Effects of Lateral Flow on the Convective Environment in a Coupled Hydrometeorological Modeling System in a Semiarid Environment

TIMOTHY M. LAHMERS,[a] CHRISTOPHER L. CASTRO, AND PIETER HAZENBERG[b]

Department of Hydrology Atmospheric Sciences, The University of Arizona, Tucson, Arizona

(Manuscript received 6 May 2019, in final form 10 January 2020)

ABSTRACT

Evidence for surface and atmosphere coupling is corroborated in both modeling and observation-based field experiments. Recent advances in high-performance computing and development of convection-permitting regional-scale atmospheric models combined with high-resolution hydrologic models have made modeling of surface–atmosphere interactions feasible for the scientific community. These hydrological models can account for the impacts of the overland flow and subsurface flow components of the hydrologic cycle and account for the impact of lateral flow on moisture redistribution at the land surface. One such model is the Weather Research and Forecasting (WRF) regional atmospheric model that can be coupled to the WRF-Hydro hydrologic model. In the present study, both the uncoupled WRF (WRF-ARW) and otherwise identical WRF-Hydro model are executed for the 2017 and 2018 summertime North American monsoon (NAM) seasons in semiarid central Arizona. In this environment, diurnal convection is impacted by precipitation recycling from the land surface. The goal of this work is to evaluate the impacts that surface runoff and shallow subsurface flow, as depicted in WRF-Hydro, have on surface–atmosphere interactions and convection in a coupled atmospheric simulation. The current work assesses the impact of surface hydrologic processes on 1) local surface energy budgets during the NAM throughout Arizona and 2) the spectral behavior of diurnally driven NAM convection. Model results suggest that adding surface and subsurface flow from WRF-Hydro increases soil moisture and latent heat near the surface. This increases the amount of instability and moisture available for deep convection in the model simulations and enhances the organization of convection at the peak of the diurnal cycle.

1. Introduction

Evidence of feedback between the lower atmosphere and the land surface, particularly in arid and semiarid environments, has been shown from analysis of near-surface fluxes and planetary boundary layer (PBL) characteristics in both modeling and observation studies (e.g., Findell and Eltahir 1997; Koster et al. 2002; Dirmeyer et al. 2009; Zeng et al. 2010; Santanello et al. 2018). The recent development of coupled mesoscale atmospheric and distributed hydrologic modeling systems (e.g., Maxwell et al. 2011; Gochis et al. 2015) now permits us to analyze the role of surface runoff and shallow subsurface hydrologic processes on surface and atmospheric coupling (e.g., Maxwell et al. 2011; Senatore et al. 2015; Arnault et al. 2016; Xiang et al. 2018).

Analysis of the impacts of surface hydrology on the lower atmosphere is relevant because of prior evidence of surface feedbacks on the atmosphere. This was first explored using relatively simple metrics with datasets like coarse-resolution GCMs and surface observations of soil moisture (e.g., Findell and Eltahir 1997; Koster et al. 2002; Dirmeyer et al. 2009; Zeng et al. 2010). Despite their simplicity, all of these studies demonstrated possible evidence for surface atmospheric coupling, including in semiarid environments in the western United States, like the Southwest and the Central Plains. More recently, Santanello et al. (2013) built upon these techniques, using a convection-permitting resolution (CPM) regional model over the Central Plains, combined with sounding and remote sensing based observations, to examine surface–atmosphere coupling. They demonstrated several metrics for coupling, based on surface–atmosphere fluxes and PBL evolution. Their work showed

Supplemental information related to this paper is available at the Journals Online website: https://doi.org/10.1175/JHM-D-19-0100.s1.

[a] Current affiliation: Hydrological Sciences Laboratory, NASA Goddard Space Flight Center, Greenbelt, Maryland.

[b] Current affiliation: Department of Hydrology, Deltares, Delft, Netherlands.

Corresponding author: Timothy Lahmers, timothy.lahmers@nasa.gov

DOI: 10.1175/JHM-D-19-0100.1

that surface coupling becomes important over this region during dry conditions.

Surface variables, discussed above, are also sensitive to surface and subsurface hydrologic processes. Maxwell et al. (2011), using the Advanced Research version of the Weather Research and Forecasting Model (WRF-ARW) coupled to the ParFlow groundwater model, suggests that movement of water across the landscape (outside of the channel network) impacts the spatial distribution of soil moisture and is thus important for the local water balance. Keune et al. (2016) showed that inclusion of 3D groundwater in a coupled simulation affected surface–atmosphere feedbacks during the 2003 European heat wave, particularly in areas with shallow groundwater. Groundwater pumping has also been shown to affect the atmospheric feedbacks over Europe (Keune et al. 2018) and the San Joaquin River Watershed in California (Gilbert et al. 2017). Anyah et al. (2008) coupled a groundwater model (Miguez-Macho et al. 2007) to a 50-km RAMS regional atmospheric model (e.g. Pielke et al. 1992; Cotton et al. 2003) and showed that shallow groundwater can enhance precipitation recycling, and subsequent convective initiation, in the semiarid southwest contiguous United States (CONUS). Barlage et al. (2015) showed similar effects with the same groundwater model coupled to WRF-ARW for the CONUS. These modeling studies differ from the present study due to their coarser resolution.

There are also several examples of recent work that consider the impacts of surface hydrology on the local atmosphere. Senatore et al. (2015) executed the WRF-ARW Model with and without coupling to WRF-Hydro for a midlatitude cyclone case over Italy. They demonstrated that coupling a calibrated WRF-Hydro simulation to WRF-ARW during the cyclone case increased model soil moisture to increase latent heating and nominally reduced precipitation overestimates by the control WRF Model. Arnault et al. (2016) coupled WRF-ARW to WRF-Hydro, and showed that overland flow from WRF-Hydro, and resultant spatial changes to ET, slightly increased precipitation and altered its distribution, within a ~100-km^2 catchment in West Africa. This was due to the rerouting of water that is otherwise removed from the Noah-MP LSM as surface flow, as lateral flow in WRF-Hydro, and these feedbacks tended to influence the spatial distribution of convection. Kerandi et al. (2018) showed that coupling WRF-Hydro to WRF had the effect of increasing surface runoff by exfiltration in some high elevations of the 3279-km^2 Tana River basin in Kenya, while decreasing runoff by increasing infiltration in other areas. Similarly, Arnault et al. (2018) and Rummler et al. (2019) both compared WRF-Hydro to WRF with otherwise identical ensemble configurations with domains in central Europe and concluded that adding surface routing to WRF, through WRF-Hydro, increased soil moisture and ET. The addition of WRF-Hydro had a small effect on precipitation, with some highly localized areas of greater changes.

Xiang et al. (2018) evaluated the sensitivity of a coupled WRF-Hydro simulation to soil moisture and vegetation fraction anomalies over the Río Sonora basin in northwest Mexico. For the 2004 and 2013 North American monsoon (NAM) seasons, coupled WRF-Hydro was able to produce consistent surface ET and soil moisture compared to an offline WRF-Hydro simulation forced with NLDAS-2 data. As this study did not consider the direct influence of surface routing caused by WRF-Hydro on the local convective environment, no work to date has considered the impacts of hydrologic routing on the local convective environment in the NAM region.

Arnault et al. (2016) analyzed the impacts of WRF-Hydro in a small domain that was not large enough to resolve convective organization. They did, however, hypothesize that WRF-Hydro, coupled to WRF-ARW, could potentially influence convective organization due to changes in soil moisture. This suggests that the impacts of WRF-Hydro may thus be important in the NAM region, where synoptic scale forcing is typically weak and diurnal convection is often initiated by local gradients caused by the terrain (e.g., Bieda et al. 2009).

To understand the impacts of surface hydrologic processes on the atmosphere in semiarid environments, we evaluate the influence of modeled surface runoff and shallow subsurface flow in the shallowest 2 m of the land surface on the surface–atmosphere feedbacks and the local convective environment in the warm season of the NAM region, across central Arizona and northern Mexico. This period is ideal for evaluating the influence of surface–atmosphere feedbacks, as its summertime convective environment is highly dependent upon local instability and orographic processes (e.g., Luong et al. 2017), but it also has sufficient surface observations (e.g., Goodrich et al. 2004) and radar data for model evaluation. This is unique compared to previous literature, as we focus on a semiarid domain with an area sufficiently large to resolve convective initiation and propagation.

Our analysis period is the 2017–18 NAM seasons, and we utilize the WRF-ARW Model configured for the southwest CONUS (e.g., Luong et al. 2017) both run in a control mode and coupled to the WRF-Hydro hydrologic model. The WRF-ARW (Skamarock et al. 2008) Model is configured as in Luong et al. (2017) with 1.5-km convection permitting resolution, and this configuration and spatial resolution is sufficient to resolve convective

organization and propagation (see section 2c for model details and section 5 for a discussion of resolution). WRF-ARW is both run in a control configuration with only Noah-MP active and with WRF-Hydro, to evaluate the impacts of lateral routing of surface flow. Unlike much of this prior work that considers deep groundwater, the scope of this research specifically considers surface flow processes, and this is appropriate for our study area, as much of Arizona has a high depth to groundwater and a hydrologic regime that is primarily associated with high volume surface runoff events driven by convective precipitation (e.g., Blasch et al. 2004; Goodrich et al. 2004; Fan et al. 2013).

The paper is organized as follows. The WRF-Hydro model structure and configuration and analysis methods are introduced in section 2, and a brief validation of the model simulations is presented in section 3. The impacts of the model coupling are presented in section 4. The discussion and conclusions are presented in sections 5 and 6, respectively.

2. WRF-Hydro modeling system and analysis metrics

a. *WRF-Hydro model structure and configuration*

WRF-Hydro (Gochis et al. 2015) is a spatially distributed hydrologic model that can be coupled to the WRF-ARW Model or forced offline (with precipitation and atmospheric forcing) as a stand-alone simulation. Vertical fluxes in WRF-Hydro are resolved using the Noah-MP LSM (Niu et al. 2011). The Noah-MP grid must be equivalent to the WRF-ARW grid used in a coupled simulation. In this study, WRF-Hydro is active in the highest-resolution nested WRF-ARW domain (see section 2c for details), with an LSM grid resolution of 1.5 km. A higher-resolution routing grid, which for this study is set to 250-m grid resolution, is also used to resolve surface runoff and subsurface flow within the lowest 2 m of the land surface. This model structure is shown in Fig. 1.

In WRF-Hydro, subsurface flow is based on changes to the water table depth of the 2-m-deep Noah-MP soil columns. Subsurface flow is computed from the hydraulic gradient, which is equivalent to the difference in the groundwater table depth along the steepest gradient (i.e., the Dupuit–Forcheimer assumption) in eight possible directions surrounding a routing grid point. Surface runoff from exfiltration is generated at a point if a grid point becomes saturated due to horizontal flow. This model configuration resolves shallow subsurface flow and does not consider deep base flow that would be captured by a groundwater model, such as ParFlow (e.g., Maxwell et al. 2011). For this simulation, groundwater is not considered, and baseflow is permitted to pass through the Noah-MP LSM grid and exit the system. Details of these routing schemes are described in Gochis et al. (2015). A diffusive wave routing scheme is used to compute surface runoff, which is also based on the steepest gradient around each grid cell (Julien et al. 1995; Ogden 1997). Diffusive wave routing, within delineated channel grid points in the 250-m routing grid is used to compute streamflow within the channels. This routing scheme can resolve the effects of backwater flow (Gochis et al. 2015). As noted earlier, streamflow in the southwest CONUS is mostly generated by surface runoff, due to the high depth to groundwater in much of the region (e.g., Blasch et al. 2004; Goodrich et al. 2004; Fan et al. 2013), so groundwater processes are of lesser importance.

In the present study hydrologic parameters for WRF-Hydro were left at defaults, as calibration is beyond the scope of this work. As there are known biases in the model water balance of WRF-Hydro (e.g., Lahmers et al. 2019), we emphasize that the results from this experiment should therefore be considered theoretical. The 250-m routing grid was generated using ArcGIS Spatial Analyst. GIS-based Routing grid variables include terrain height, flow direction, channel network, and stream order. Before terrain processing is performed, the National Hydrography Dataset Plus (NHDPlus) Version 2 channel network (McKay et al. 2012) was burned into the DEM using the AGREE DEM tool (Hellweger 1997), with a smooth drop of 5 m surrounding the channels and a sharp drop of 15 m at the channel grid points. In a few cases, deeper channels had to be carved in selected reaches to account for projection errors in narrow canyons. The Shuttle Topographic Mission (STRM) 90-m DEM was used to generate the routing grid. To generate a spatially consistent channel network, pits in the DEM were removed using a fill tool in ArcGIS. Inland sinks, defined in the NHDPlus Version 2 dataset, that were not classified as channel network ends were temporarily set to no-data values when this was performed, to prevent real inland sinks from being artificially filled. In a few cases, channel network end points were added to the sinks preserve the realism of the channel network. These points on the DEM were set back to their original values after the routing grids were computed, so that the grid did not contain missing data. After DEM processing was performed, the model channel network was computed from drainage area, and a 10-km^2 drainage area threshold was used to define channel grid points on the 250-m routing grid.

Offline WRF-Hydro simulations (i.e., WRF-Hydro as a hydrologic model with external atmospheric forcing)

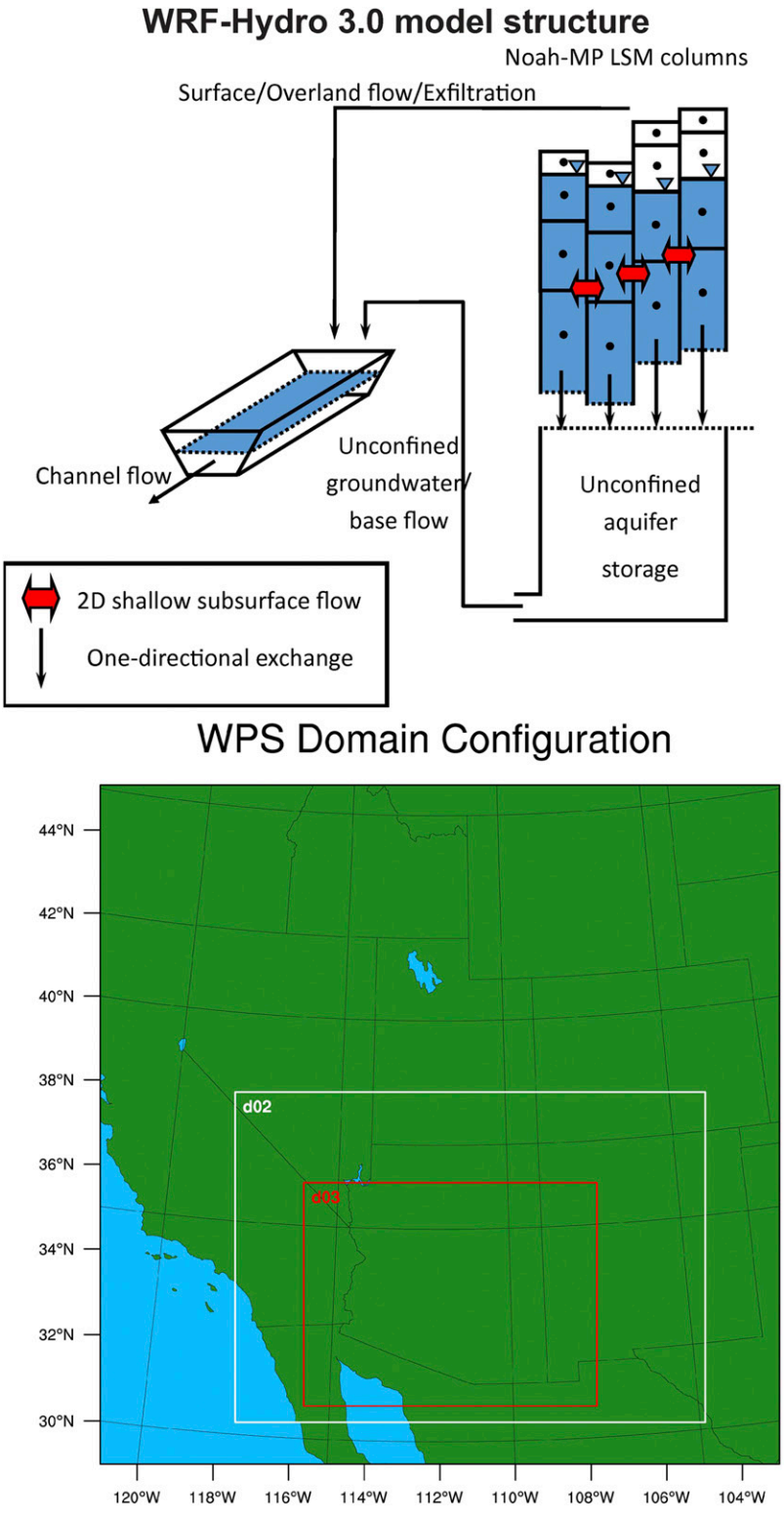

FIG. 1. (top) Illustration of the uncoupled WRF-Hydro hydrologic model structure and (bottom) analysis domain. The Noah-MP LSM and routing grid columns are shown in the top right of the figure. Water from the base flow bucket model (not used) and surface runoff from the terrain routing grid are both returned to the channel network (shown at left). Nested grids sizes are 10 km (RCM grid) for D01, 3 km for D02, and 1.5 km for D03.

were executed from 1 October 2015 until 1 October 2018 for the same model grid as the 1.5-km resolution (250-m routing grid resolution for WRF-Hydro) inner domain (domain 3) of the WRF-ARW simulations. These simulations were forced with NLDAS-2 incoming shortwave radiation, incoming longwave radiation, specific humidity, air temperature, surface pressure, and near surface wind (both u and v components) and NCEP Stage IV precipitation, which is based on a mosaic of WSR-88D radar data, with gauge correction. The soil moisture states on 1 April 2017 and 1 April 2018 from these offline simulations were used as initial surface conditions of both the coupled and uncoupled WRF-Hydro 1.5-km simulations (see next two sections for reasoning behind these start dates). These offline simulations started 1.5 and 2.5 years prior to the 2017 and 2018 WRF-ARW simulations, respectively. This was to ensure the model soil moisture states were sufficiently spun up such that the model achieved a realistic soil moisture state prior to the execution of the coupled model simulation. These model spinups were both longer than what was used on Castro et al. (2007a,b) and Liu et al. (2017), who cold-started their RCM simulations.

In some cases, the effects of the hydrologic model are difficult to distinguish because differences in modeled precipitation between the coupled and uncoupled model simulations are superimposed on differences to the model fluxes that are caused by changes to the model structure. Therefore, the offline WRF-Hydro simulations for the 2017 and 2018 NAM seasons were also compared to an equivalent WRF-Hydro simulation using only Noah-MP (Table 1). From here on, these simulations are referred to as WRF-Hydro offline and Noah-MP offline (Table 1). In other words, WRF-Hydro is run with external forcing and all hydrologic routing active (WRF-Hydro offline) and with the same external forcing but with only Noah-MP active (Noah-MP offline). The Noah-MP offline simulation is an offline LSM with the same configuration as that of the LSM for the control WRF-ARW Model simulation (i.e., WRF-Ctrl; Table 1), and the WRF-Hydro offline simulation has the same LSM and hydrologic model configuration as WRF-Hydro coupled to WRF-ARW (from here on known as WRF-Hydro; Table 1). Thus, the precise effects of surface and shallow subsurface routing without changes to precipitation can be understood from the Noah-MP offline and WRF-Hydro offline simulations, as the intention of the present study is to determine if the redistribution of water across that land surface (from the recycling of surface runoff; Arnault et al. 2016) affects surface variables in the NAM convective environment. This is necessary to fully understand the impacts of horizontal routing in a coupled environment.

Table 1. WRF-ARW and WRF-Hydro coupled and uncoupled model simulations and descriptions.

Simulation name	Description
Noah-MP offline	WRF-Hydro with only Noah-MP LSM active, forced with NLDAS-2 atmospheric variables and NCEP Stage IV precipitation
WRF-Hydro offline	WRF-Hydro and Noah-MP with surface runoff, subsurface flow, and channel routing active, forced with NLDAS-2 atmospheric variables and NCEP Stage IV precipitation
WRF-Ctrl	WRF-ARW with Noah-MP LSM parameterization
WRF-Hydro	WRF-ARW with Noah-MP LSM and WRF-Hydro modules active (surface runoff, subsurface flow, and channel routing).

b. *NAM region convective environment*

Warm season precipitation in the southwest CONUS is primarily driven by the NAM, between late June and mid-September (Douglas et al. 1993; Adams and Comrie 1997), and our model configuration is designed to capture NAM convection (e.g., Luong et al. 2017). During this time of year, the region experiences a majority of its severe weather (e.g., Maddox et al. 1995; McCollum et al. 1995), including extreme precipitation and flash flood events. The NAM convective environment is associated with a semipermanent ridge of high pressure that forms over the southwest CONUS. This pattern results in easterly flow aloft, over the southwest CONUS, which is conducive to horizontal moisture flux from the Gulf of Mexico and Gulf of California (e.g., Douglas et al. 1993; Adams and Comrie 1997). These patterns are associated with a unique convective environment that is conducive to land–atmosphere coupling, as evapotranspiration from prior convection results in precipitation recycling for convection on subsequent days (e.g., Dominguez et al. 2016). Diurnal convection is normally phase-locked to the terrain in this environment; however, atmospheric disturbances associated with synoptic-scale ascent or vertical wind shear can cause convective organization. In this situation, squall lines and mesoscale convective systems (MCSs) can propagate into the lower desert regions, often resulting in extreme precipitation and flash flooding (e.g., Pytlak et al. 2005; Bieda et al. 2009; Finch and Johnson 2010; Newman and Johnson 2012; Seastrand et al. 2014; Lahmers et al. 2016; Luong et al. 2017).

c. *WRF-ARW configuration*

To capture NAM convection, WRF was executed in three nested domains (Fig. 1). Domain 1 encompasses

the western CONUS and a small section of northwest Mexico at 12-km grid resolution. The WRF simulation is forced with version 2 of the NCEP Climate Forecast System (CFSv2) analysis data (Saha et al. 2014) with the model parameterizations shown in Table 2. The outer most domain utilized selective spectral nudging for large scale processes in the upper atmosphere (e.g., 500-hPa heights, u wind, v wind, and temperature) to preserve large-scale features yet still permit the high-resolution model to resolve mesoscale features (e.g., von Storch et al. 2000; Castro et al. 2005; Miguez-Macho et al. 2005; Rockel et al. 2008; Chang et al. 2015). Domain 2 has 3-km grid resolution, encompassing the entire lower Colorado basin. These domains are equivalent to those used in (Pal et al. 2019), and the WRF-ARW parameterizations are equivalent to Luong et al. (2017). Domain 3, with 1.5-km grid resolution, nominally encompasses the Gila basin, a main tributary of the lower Colorado River (Fig. 1). The convection parameterization is disabled in both domains 2 and 3. Domain 3 is simulated with identical initial and boundary conditions (both with and without WRF-Hydro active).

In this context, domain 3 is executed as a type-2 regional climate model (RCM), which is a limited-area regional atmospheric model that is executed on a time scale of several months, such that initial conditions are forgotten for most of the simulation. For such a simulation, lateral forcing (i.e., reanalysis) remains relevant (Castro et al. 2005). For the present study, a using a type-2 RCM permits us to analyze the impacts of surface boundary conditions from hydrologic processes (i.e., WRF-Hydro), while keeping lateral boundary conditions constant.

The NAM season, officially beginning on 15 June, is the focus of our analysis, so we execute all three domains of the RCM starting 1 April for both the 2017 and 2018 NAM seasons, to allow for model spinup, as executing the model through the dry period from April to mid-June enables the model soil columns to dry out and reach realistic premonsoon conditions. This period is consistent with Arnault et al. (2018), who also used a 3-month spinup period for their coupled RCM simulations. Both the coupled and uncoupled RCM were executed through this period. WRF-ARW simulations were executed to 15 September, the end of the NAM season, and we note that Castro et al. (2007a,b) and Liu et al. (2017) did not use any spinup for their multiyear RCM simulations. The shorter WRF-ARW periods for each NAM season are used rather than a continuous RCM simulation due to the computational requirements for the model. Due to the high computational demand of WRF-ARW, it was not feasible to execute the simulation for more than two NAM seasons.

TABLE 2. WRF-ARW parameterization settings for the RCM simulation of the 2017 and 2018 warm seasons. Note that a cumulus parameterization is not used for model domains 2 and 3, which are executed with CPM grid resolution. WRF-ARW namelist options for selected schemes are shown in parentheses.

Physics module	Scheme (option)
Microphysics	New Thompson et al. (8)
Longwave radiation	New Rapid Radiative Transfer Model (4)
Shortwave radiation	New Rapid Radiative Transfer Model (4)
Surface layer	Eta Similarity (2)
Land surface	Noah-MP (4)
Urban surfaces	Urban Canopy Model (1)
Planetary boundary layer	Mellor–Yamada–Janjić (2)
Cumulus parameterization	Kain–Fritsch (1)

d. WRF-ARW validation metrics

To evaluate our coupled model configuration, we evaluate the model compared to available SMAP–Sentinel-derived 1-km soil moisture estimates (Das et al. 2018), when available during the 2017 and 2018 NAM seasons. Both WRF-ARW simulations are also evaluated compared to NCEP Stage IV precipitation. In considering precipitation biases, we again emphasize that these results should be considered theoretical, given known biases in the Noah-MP LSM and the WRF-Hydro model. We also consider how the fully coupled WRF-Hydro model captures the diurnal precipitation cycle, in 3-h temporal resolution increments (Fig. 2). Note that the color code for these increments is used in later analysis as well. Modeled precipitation, temperature, and specific humidity were verified with station observations around Arizona, available through the National Centers for Environmental Information (NCEI). For this analysis, 32 precipitation sites and 19 sites with temperature and humidity were used.

e. Metrics for surface–atmosphere interaction

Because the purpose of this study is to consider the impacts of surface and shallow subsurface routing on RCM boundary conditions, we consider the extent of the impacts of WRF-Hydro on selected convective days [ending at 1200 UTC (0500 LST)] that have sufficient antecedent precipitation from prior days. These days were chosen, as we expect a priori that the impacts of lateral hydrologic routing will be greater on these days. Therefore, those days were selected with an antecedent precipitation index (API; Cordery 1970) of greater than 20 mm (28 days between both NAM seasons). The API is equivalent to the sum of 24-h precipitation from previous days, multiplied by a decay factor k, and is defined as

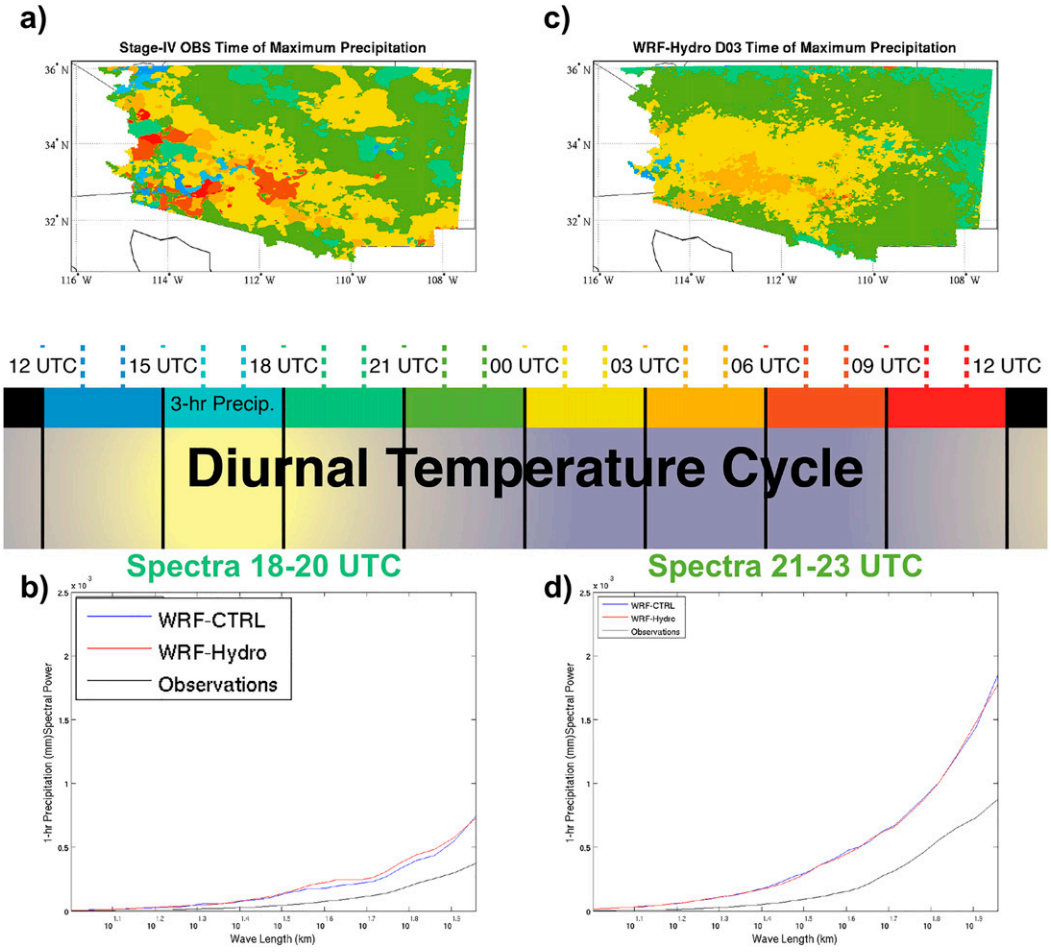

FIG. 2. The time of the peak precipitation for (a) NCEP Stage IV and (c) WRF-Hydro are plotted with the color legend for the diurnal cycle shown in the center of the graphic. Spectral power of WRF-Ctrl (blue), WRF-Hydro (red), and Stage IV observations (black) between 10 and 100 km for (b) 1800–2000 and (d) 2100–2300 UTC are shown to demonstrate the evolution of the diurnal cycle.

$$\mathrm{API}_d = P_d + kP_{d-1} + k^2P_{d-2} + \cdots = P_d + k\mathrm{API}_{d-1}, \quad (1)$$

where k defines the decay factor, indicating the impact of precipitation that occurred during previous days as being proportional to the current soil moisture. For the current study, k was assumed equal to 0.85, which has been defined as representative for arid environments (Lindsay et al. 1975). The API for the NAM season (for uncoupled WRF and coupled WRF-Hydro) for the model domain is shown in section I of the online supplemental material.

To evaluate the impacts of surface and subsurface routing on the land surface, we first consider these impacts in WRF-Hydro offline and Noah-MP offline (i.e., not coupled to WRF-ARW; Table 1). We evaluated soil moisture, latent heat, and sensible heat of both models with constant forcing. These variables are proxies for both the land surface state and surface fluxes that can impact the lower atmosphere that are not prescribed input variables for the uncoupled model. To analyze the precise impacts of coupling within WRF-Hydro on the surface and lower atmosphere, we evaluate changes in soil moisture with respect to the underlying soil conductivity, precipitation forcing, and drainage area computed from the WRF-Hydro 250-m routing grids.

We then consider the RCM atmospheric and hydrologic variables related to 1) the land surface, 2) the lower atmosphere, and 3) atmospheric instability throughout the 1.5-km resolution model domains for both the WRF and WRF-Hydro simulations. This includes soil moisture, latent heat, sensible heat, 2-m temperature, 2-m specific humidity, LCL height, and most unstable CAPE (MUCAPE). Note that LCL height, which is a useful proxy for the moisture deficit in the lower atmosphere that could inhibit convection, was also considered by

Santanello et al. (2018) to evaluate atmospheric coupling (they considered PBL height minus LCL). To evaluate the impacts of hydrologic routing on surface and atmospheric processes, we use a two-tailed Student's *t* test with a 90th percentile confidence interval, on daily averaged model data at each grid point for both the entire NAM seasons and for convective days selected based on the API threshold. A field significance test (Livezey and Chen 1983) is also performed on these composites, using a 500-member null distribution generated using a permutation. For all figures with field significance, it is shown as a percentile of the null distribution that the composite analysis in the figure exceeds.

We also compare modeled and observed mixing diagrams (e.g., Santanello et al. 2009) at six selected sites. Two sites were located in the Walnut Gulch Experimental Watershed (WGEW) in southeast Arizona, where two flux towers that represent semiarid shrubland (Lucky Hills location) and grassland (Kendall location) are maintained by the U.S. Department of Agriculture Agricultural Research Service (USDA-ARS). USDA-ARS also maintains two flux towers at its Santa Rita site, including a grassland and a mesquite savannah south of Tucson, Arizona. We also analyze surface fluxes at the Tucson and Phoenix airports. This is consistent with Santanello et al. (2013, 2018), who considered latent and sensible heating (as the evaporation fraction), LCL height, and mixing diagrams to diagnose the degree of surface coupling in multiple WRF-ARW configurations. In the present study, we are unable to compute the surface and entrainment vectors for our mixing diagrams, as in Santanello et al. (2009), due to the limited availability of upper-air data needed to compute the boundary layer height at our sites of interest.

f. Analysis of two-dimensional spectral components of propagating convection

To evaluate the impacts of the redistribution of soil moisture from horizontal routing on propagating convection in the WRF-Hydro and WRF-Ctrl simulations, we consider the convective days with sufficient atmospheric instability (i.e., Mazon et al. 2016; Luong et al. 2017) and we eliminate days that are associated with transient inverted troughs (IVs) in the CFSv2 reanalysis forcing, the forcing dataset for the RCM simulations, following the methods of Lahmers et al. (2016) to locate IVs near the model domain. Inverted troughs are tracked using potential vorticity (PV) anomalies based on a 40-yr climatology, consistent with Lahmers et al. (2016). The PV climatology used to find IVs (e.g., Bosart et al. 2011) is computed from Climate Forecast System Reanalysis (CFSR) (Saha et al. 2010) data from 1979 to 2011 and CFSv2 data from 2012 onward. Both of these models have equivalent grids, dynamic cores, and physics parameterizations. Convective days within the 50th percentile of both MUCAPE and integrated precipitable water (IPW) were selected. Both MUCAPE (e.g., Moncrieff and Miller 1976) and IPW (e.g., Lu et al. 2009) are proxies for convective development in the southwest, as they are associated with thermodynamic instability and integrated moisture, respectively. From these days, we remove "trough days," as synoptic-scale and mesoscale forcing associated with IVs (e.g., Pytlak et al. 2005; Bieda et al. 2009; Finch and Johnson 2010; Newman and Johnson 2012; Seastrand et al. 2014; Lahmers et al. 2016) may mask out the effects of surface atmospheric coupling due to their enhancement of convective propagation and organization. This left us with 55 convective days over the two NAM seasons.

Two-dimensional spectral analysis followed Castro et al. (2005) and Errico (1985) for selected convective days that are favorable for propagating convection that lack large scale forcing. In the current work, it was decided not to apply a detrending step, as the mesoscale variables we are concerned with occur on a scale where large-scale variations (such is in 500-hPa height) are not expected a priori. We consider modeled and observed NCEP Stage IV precipitation to evaluate the performance of the model relative to observations through the peak of the diurnal cycle. While not available in an observation dataset, this analysis is also performed on the modeled moisture flux convergence (e.g., Castro et al. 2007xx) and 500-hPa vertical motion (omega) to further evaluate the impacts of the land surface on convective propagation. As in the analysis of model precipitation, the diurnal cycle is broken into eight 3-h intervals (Fig. 2), so each partition of the diurnal cycle considers 55 days with three sample spectra (i.e., 165 cases). Statistical significance of spectral analysis is computed using a 500-iteration permutation method, where the control WRF-ARW and WRF-Hydro spectra are substituted into the analysis at random, and fractional differences of spectra (as in Castro et al. 2005) between two datasets beyond the 95th or 5th percentiles are considered statistically significant.

3. WRF-ARW evaluation

a. NCEP Stage IV precipitation

Figures 2 and 3 show that the WRF-Hydro simulations capture the spatial distribution and the diurnal cycle of NAM convection. Figures 2a and 2c show the diurnal cycle of convection both in the model and in the NCEP Stage IV observations, as in Fig. 5 of Luong et al. (2017), but with 3-h intervals. In this figure, the time in the diurnal

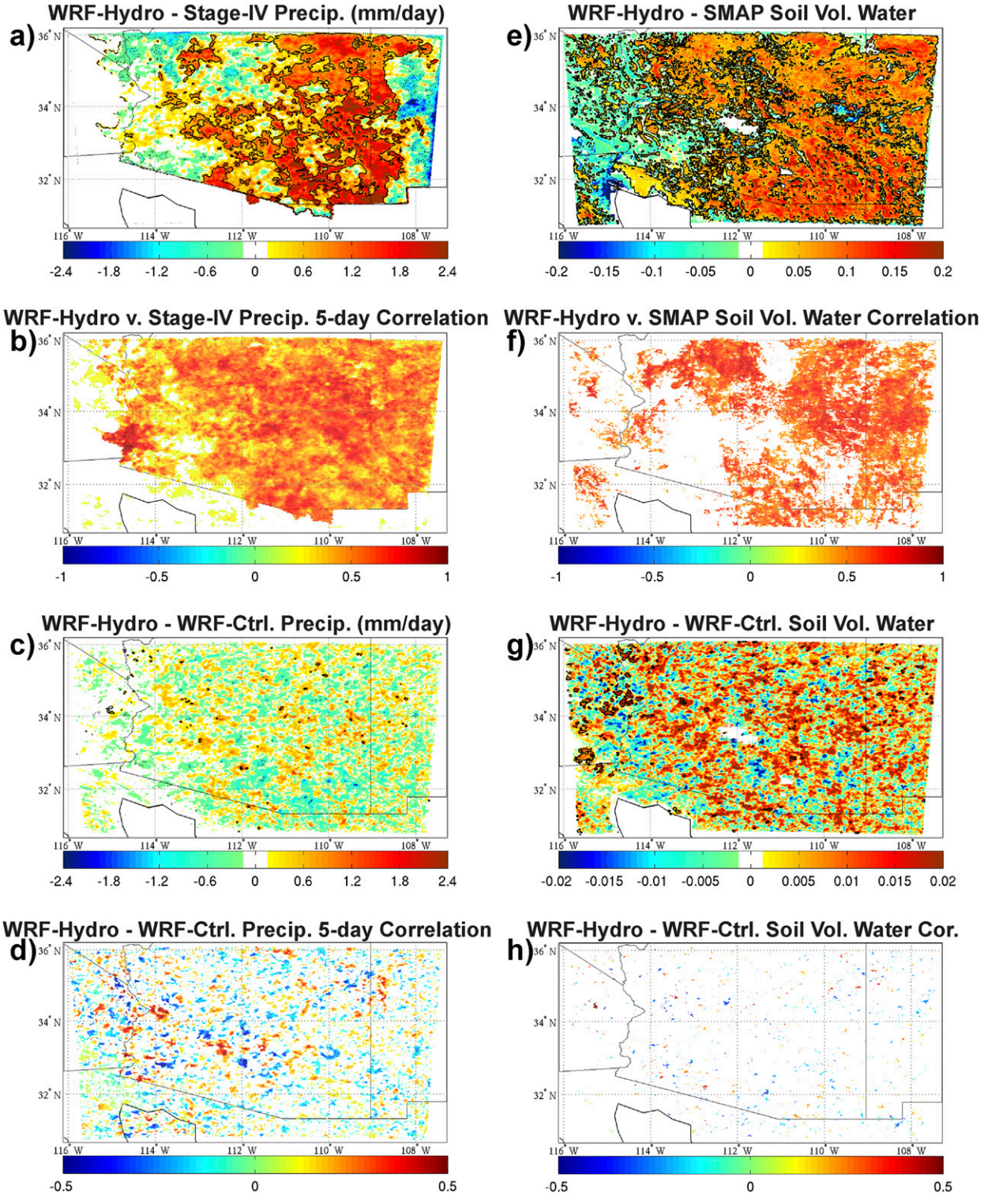

FIG. 3. (a) WRF-Hydro precipitation bias compared to NCEP Stage IV precipitation and (b) correlation coefficient between these two datasets. (c) The difference between WRF-Hydro and WRF-Ctrl precipitation. (d) The difference between the WRF-Hydro and WRF-Ctrl correlation coefficient (to observations). (e)–(h) As in (a)–(d), but with SMAP–Sentinel soil moisture 1-km data. For the difference of means in (c) and (g), solid and dashed lines show 95th and 5th percentile statistical significance from a Student's *t* test, respectively. Correlation coefficients in (b) and (f) and differences in (d) and (h) are only shown where they are statistically significant, based on a *t* test and permutation test, respectively.

cycle when the highest average precipitation occurs is denoted by color. This figure shows that the diurnal precipitation maximum on the high terrain is near the peak of the convective cycle (i.e., 2100–0000 UTC or 1400–1700 MST), where it typically initiates. This convection typically propagates into the low deserts later in the evening (i.e., 0000–0300 UTC), and often not reaching areas of western and central Arizona until 0300–0600 UTC, and sometimes even later. This pattern is captured in both the model and the observations.

Both RCM simulations capture the growth of convection upscale during the peak of the diurnal cycle as mean spectral power, excluding cases where limited coverage of precipitation causes extremely high spectral peaks (Figs. 2b,d). These data illustrate the 2D spectral power of modeled and observed precipitation, with wavelength on the x axis and spectral power on the y axis. These figures show increased power at longer wavelengths (i.e., more organization). The most important pattern from these figures is the growth in organization at all wavelengths from the early to later time steps, as Fig. 2 shows the increased convective organization between the 1800–2000 UTC periods (left) and 2100–2300 UTC periods (right), which is the peak of the diurnal heating cycle and represents the time when convective initiation most often occurs during the NAM (e.g., Luong et al. 2017). This indicates increasing convective organization, consistent in both models and the Stage IV precipitation. Both the model and observations capture these patterns in scale of convection.

WRF-Hydro has a statistically significant high precipitation bias over most of the domain, particularly the eastern sections over the high terrain (Fig. 3a) compared to the NCEP Stage IV observations. To account for timing differences of major precipitation events, in line with a type-2 RCM, where initial conditions are forgotten, as in Castro et al. (2005), the 5-day averaged correlation coefficients between modeled and observed precipitation are shown in Fig. 3b. This characteristic of an RCM can introduce some uncertainty in the precise timing of precipitation events. Both the WRF-Hydro (Fig. 3b) and the WRF-Ctrl simulations have relatively high correlation coefficients with NCEP Stage IV data when 5-day averaging is used, suggesting that the RCM is able to capture the timing of most precipitation events. There are no spatially consistent trends in the mean between the WRF-Hydro and WRF-Ctrl simulations (Fig. 3c), implying that WRF-Ctrl has similar precipitation biases. These precipitation biases may also be due to model boundary condition biases. Liu et al. (2017) found significant dry bias over much of the southwest during the NAM season in their 4-km CONUS-wide WRF-ARW simulation with ECMWF ERA-Interim forcing. As the NCEP Stage IV precipitation product is subject to beam blockage in much of the southwest United States (e.g., Zamora et al. 2014) and has low bias over the Colorado Basin River Forecast Center (RFC) domain (Nelson et al. 2016) (on the order of −15.0% and −26.0% for heavy rain), these model biases compared to the Stage IV product likely reflect uncertainties within the remote sensing product. There are multiple localized areas of statistically significant differences in correlation coefficient between WRF-Hydro and WRF-Ctrl (Fig. 3d), but there are no consistent spatial trends in these patterns. This slight high bias of precipitation from convection-permitting WRF-ARW simulations, compared with observations, is similar to that of Luong et al. (2017).

b. SMAP–Sentinel 1-km soil moisture

The right panels of Fig. 3 are the same as the left panels, but for soil moisture. They demonstrate that despite significant high bias (Fig. 3e), WRF-Hydro captures similar spatial patterns compared to observed soil moisture as well as the temporal trends, as correlation coefficients are statistically significant over much of the domain (Fig. 3f). The areas where correlation coefficients are not statistically significant are in the southern and western areas of Arizona, the parts of the state that are most reliant on MCS precipitation (e.g., Luong et al. 2017), and this suggests that the combination of the low temporal resolution of the SMAP dataset combined with uncertainty in the RCM of the placement of MCSs may degrade soil moisture results in these areas. The change to correlation coefficients between WRF-Hydro and WRF-Ctrl are not statistically significant (Fig. 3h), implying that WRF-Hydro does not add value to the temporal evolution of soil moisture, but also does no harm. WRF-Hydro increases already biased soil moisture compared to the control simulations (Fig. 3g); however, it should be emphasized that the values of soil moisture at the highest level (0–10 cm) of the Noah-MP LSM may not be one to one with near surface soil moisture measured from the SMAP–Sentinel-derived product that represents the lowest 3 cm of soil moisture (e.g., Santanello et al. 2018). The fidelity of the spatial patterns and temporal correlation of soil moisture implies the skill of our WRF-Ctrl and WRF-Hydro simulations, despite likely biases of these variables.

c. NCEI station observations

Figure 4 shows the bias and correlation coefficients for modeled and observed NCEI station data. This figure clearly shows that WRF-Hydro has a slight high precipitation bias, but these errors are less than 1.5 mm day^{-1} for all areas except the mountains in southeast Arizona.

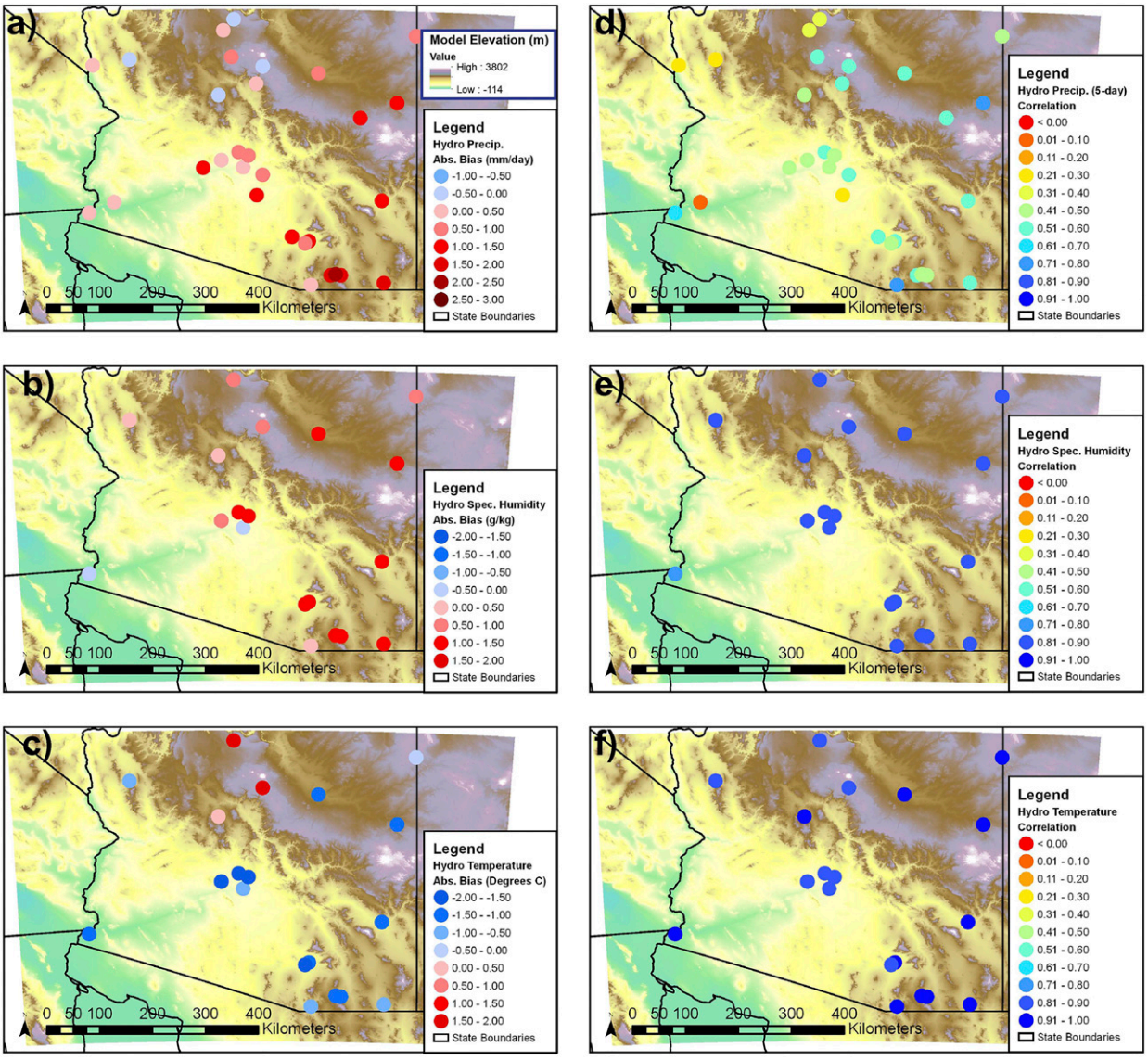

FIG. 4. (left) WRF-Hydro bias and (right) correlation coefficient between modeled and observed variables at NCEI stations throughout the model domain. Variables include (a),(d) precipitation, (b),(e) temperature, and (c),(f) specific humidity. Hourly data and observations are used for all variables, except precipitation, where 5-day smoothing is applied to daily averages when computing the correlation coefficient.

WRF-Hydro also has a cool (high) temperature (specific humidity) bias across much of the domain (Fig. 4), and these trends are the same for WRF-Ctrl (supplemental material; section 2). WRF-Hydro has correlation coefficients greater than 0.8 for nearly the entire domain for hourly specific humidity and temperature, demonstrating that it generally captures the surface environment reasonably well. As in Fig. 3b, 5-day correlation coefficients are shown in Fig. 4d. These 5-day correlation coefficients vary throughout the domain; however, they demonstrate that the model, despite timing errors of specific events, can reasonably capture the general progression of NAM convection during the 2017–18 NAM seasons.

This evaluation demonstrates that despite the biases of both the WRF-Ctrl and coupled WRF-Hydro simulations, which are consistent with previous work with a similar model configuration (i.e., Luong et al. 2017), activating hydrologic routing in WRF-ARW (i.e., coupled WRF-Hydro) generally captures the spatial and temporal patterns of NAM convection, and WRF-Hydro does no harm to the model skill metrics, compared to WRF-Ctrl. WRF-Hydro slightly increases the positive soil moisture bias of a model that has a wet

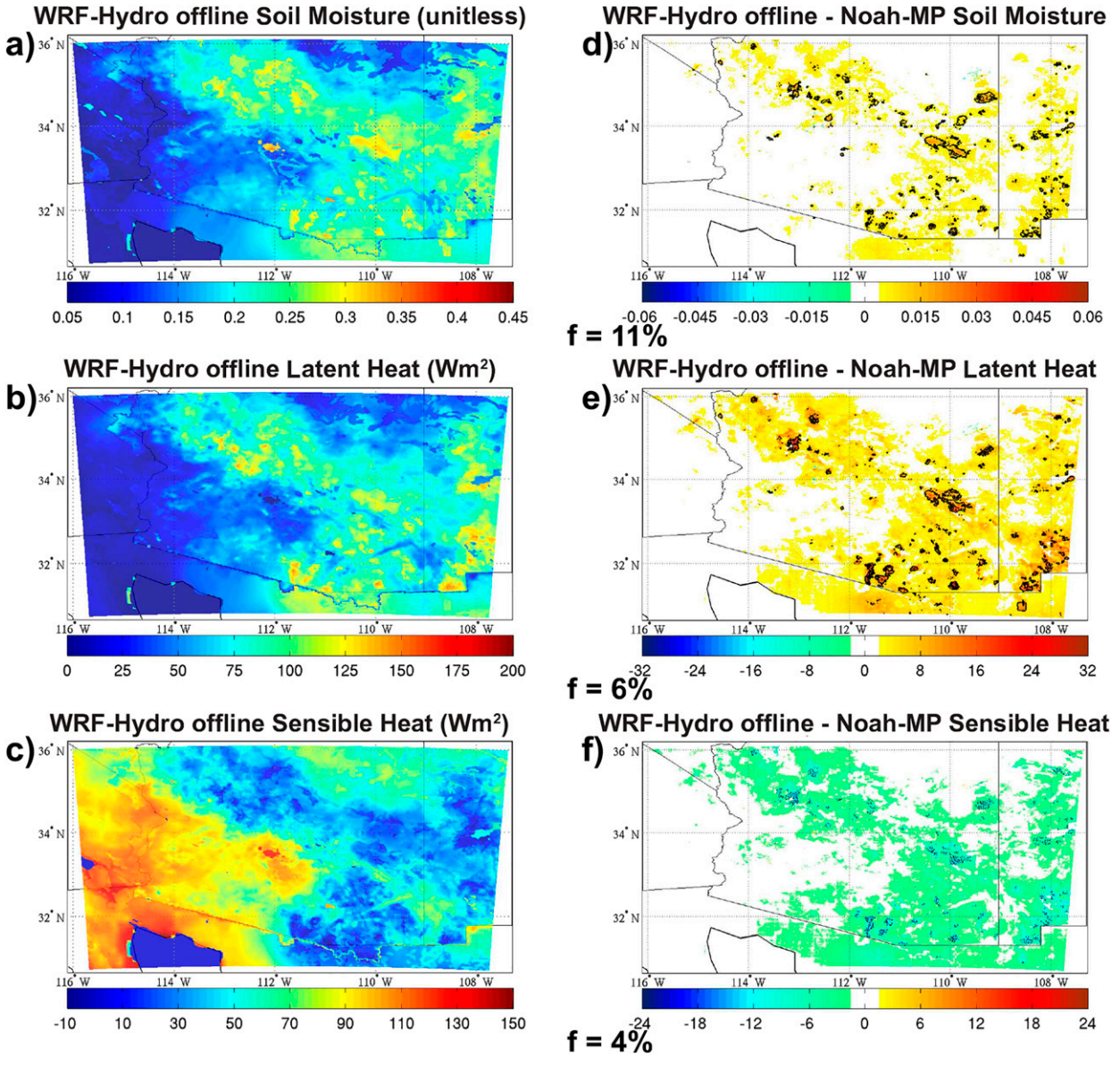

FIG. 5. WRF-Hydro offline (a) soil moisture (soil volumetric water), (b) latent heat (W m^{-2}), and (c) sensible heat (W m^{-2}) are plotted for high API days. The difference between WRF-Hydro offline and Noah-MP offline is also shown for (d) soil moisture (soil volumetric water), (e) latent heat (W m^{-2}), and (f) sensible heat (W m^{-2}) for the same high API days. For model variable differences, solid and dashed lines show 95th and 5th percentile statistical significance from a Student's *t* test, respectively. Field significance is shown in the lower left corner of these same panels.

bias. As the focus of the current work is to assess the impact of adding lateral flow through WRF-Hydro within a coupled land–atmosphere simulation, the goal within this study is not to reduce the wet bias that is present in both models through improved parameterizations (see also section 6). As such, this work is of theoretical nature and may be considered a sensitivity analysis to determine the influence of surface flow and near surface subsurface flow for an atmospheric simulation.

4. Effects of hydrologic processes in WRF-Hydro

a. Surface variables and fluxes in uncoupled simulations

Soil moisture, latent heat, and sensible heat from offline WRF-Hydro (Table 1) for the 2017–18 NAM seasons on high-API days are shown in Fig. 5. The differences between these simulations and Noah-MP offline with the same forcing (Table 1) are shown in the right panel of Fig. 5 for the same variables. These same variables

are also plotted for the full NAM season (supplemental material section 3) and show similar results, but with smaller changes. From this figure, it is clear that activating surface and subsurface routing in WRF-Hydro statistically significantly increases soil moisture, and subsequently latent heat flux, in some locations. WRF-Hydro reduces sensible heat flux in most of the model domain; however, none of these changes are field significant. Statistical significance is computed from 24-h averages from both models for the 28 high-API days.

The underlying causes of the additional soil moisture in the offline WRF-Hydro simulations, and subsequent changes to latent and sensible heat fluxes, depend on surface variables and available precipitation, which control how much additional surface runoff that WRF-Hydro can produce. When surface runoff is activated in WRF-Hydro, it redistributes existing surface runoff across the 250-m routing grid that is superimposed on the Noah-MP grid (Fig. 6). This additional runoff would otherwise be removed from the system for the Noah-MP only simulations and no longer permitted to interact with the LSM. WRF-Hydro redistributes runoff across the land surface, which subsequently increases soil moisture and latent heat (Fig. 6b).

However, for this additional runoff to affect the land surface in a significant way, several other conditions must be met. In Figs. 6b, 6c, 6e, and 6f, we plot the average change of soil moisture from WRF-Hydro offline versus Noah-MP offline (over the 2017–18 NAM seasons) for high-API days as a function of added surface runoff in WRF-Hydro (over the full model simulation including all NAM days) for several different soil conductivities. These plots show some of the more commonly occurring conductivities out of the 10 different recurring soil conductivities in our domain. This figure shows that 1) surface runoff differences between WRF-Hydro offline compared to Noah-MP offline tend to increase more for soils with lower conductivity, where runoff would already be high (not shown), and 2) soil moisture increases (in WRF-Hydro offline) with respect to increased runoff for low conductivity soils, where many of these changes are statistically significant. Thus, these results suggest that the impact of surface routing in WRF-Hydro is greatest when soil conductivity is lower.

This result is corroborated in Fig. 7, where we plot statistically significant changes in soil moisture on top of soil conductivity, and it is evident that the impacts of hydrologic routing are greatest for lower conductivities (Fig. 7a). However, soil conductivity is not the only controlling factor of surface runoff change. Figure 7d shows the same soil moisture statistical significance contours as in Fig. 7a, but plotted against precipitation. As would be expected, sufficient precipitation is also needed for the impacts of hydrologic routing to be felt in WRF-Hydro offline.

The bottom four panels of Fig. 7 show the average changes to runoff at a grid point as a function of precipitation (x axis) and drainage area (y axis; based on area averages of the 250-m routing grid interpolated to the 1-km Noah-MP grid). The surface runoff values from the different grid points are interpolated onto the grid shown in Fig. 6 for statistically significant (right) and all other points (left), based on changes to soil moisture on high-API days (Fig. 5). As we have shown previously, surface runoff increases more for lower conductivity soils as well as for increasing precipitation. For grid points without significant changes (left panels), the surface runoff difference from WRF-Hydro offline increases with increased drainage area upstream. This drainage area was computed using ArcGIS version 10 Spatial Analyst as part of the routing grid delineation, described in detail in section 2. This is directly related to the impact of lateral redistribution of water as presented in Fig. 6 (top). However, for the significant grid points (right panels), these large differences in surface runoff already occur at much smaller upstream areas for the intermediate and lowest conductivity values. Note that drainage areas included in this analysis neglect grid points with a channel drainage area greater than 10 km (one hundred sixty 250-m grid points), the threshold for defining channels, where water is removed from the WRF-Hydro channel routing scheme.

Thus, the results in Figs. 5–7 show that WRF-Hydro increases soil moisture (and subsequently latent heat) in selected areas (Fig. 5) by the following logic:

1) When WRF-Hydro is enabled, surface runoff that is otherwise removed from the system is redistributed as surface flow on the 250-m WRF-Hydro routing grid (Figs. 6a,d).
2) This additional routing subsequently increases soil moisture, and this occurs the most in areas with low soil conductivity, where we would already expect surface routing to be higher (bottom four panels of Figs. 6 and 7a).
3) For WRF-Hydro to have an appreciable impact on soil moisture, sufficiently high precipitation is also needed (Fig. 7d and Figs. 7b,c,e,f).
4) To a lesser extent, increased surface runoff and soil moisture are also favored in areas with higher drainage areas (Figs. 7b,e).

These results demonstrate that WRF-Hydro is recycling runoff from Noah-MP (particularly in areas with lower soil conductivity and high precipitation).

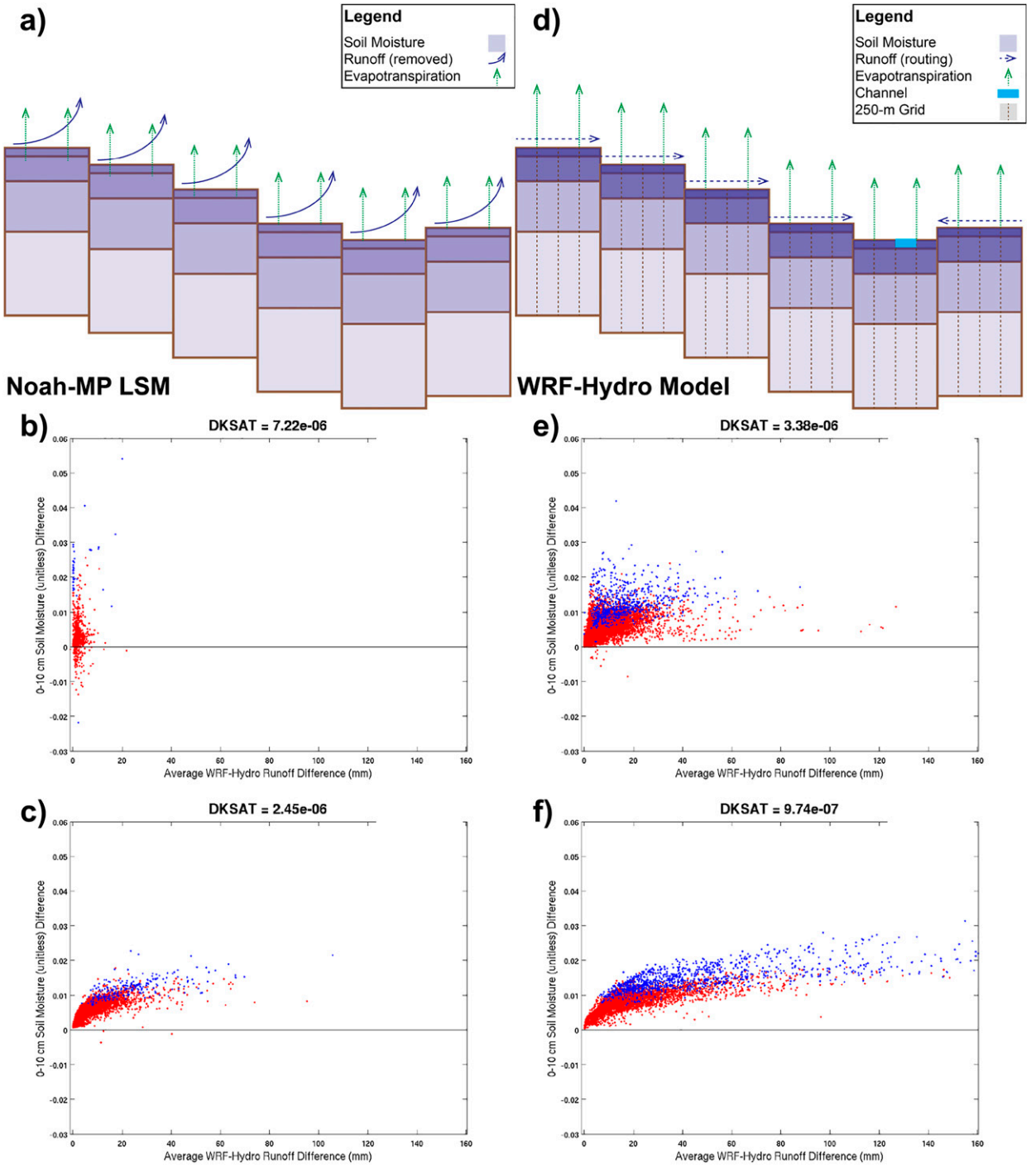

FIG. 6. Conceptual illustrations of LSM processes in (a) Noah-MP and (d) WRF-Hydro. (b),(c),(e),(f) Average soil moisture differences for WRF-Hydro offline minus Noah-MP offline vs WRF-Hydro offline minus Noah-MP runoff. Statistically significant changes in soil moisture are shown in blue, and all other changes are shown in red. Each panel shows results for a different soil conductivity.

The surface runoff option for Noah-MP in these simulations is left at the default for WRF-Hydro (runoff option 3), such that the LSM does not generate additional runoff when the depth to saturated soil in the 2-m Noah-MP columns is high. Furthermore, the infiltration equation in WRF-Hydro (Noah-MP) tends to increase infiltration with higher soil moisture (Yang et al. 2011), leading to further increases in soil

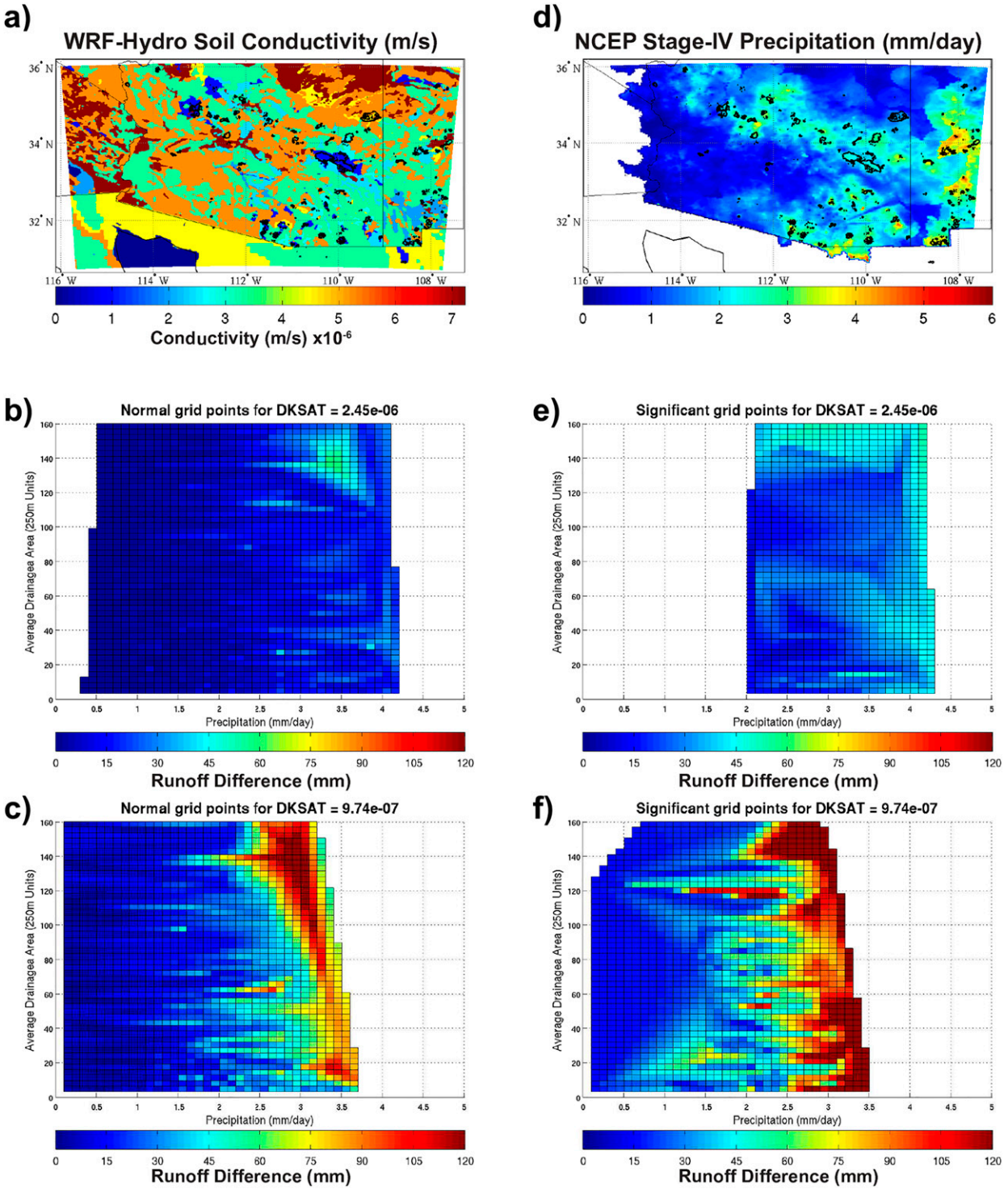

FIG. 7. (a) Soil conductivity and (d) NCEP Stage IV precipitation for offline WRF-Hydro simulations. The statistical significance for soil moisture differences (as in Fig. 5) is also shown. (b),(c),(e),(f) Average surface runoff difference for WRF-Hydro offline minus Noah-MP offline. Data are plotted as a function of precipitation (x axis) and drainage (y axis). Data are shown for grid points with 2.45×10^{-06} saturated conductivity in (b) and (e) and with 9.47×10^{-07} saturated conductivity in (c) and (f). Statistically significant grid points (for soil moisture; high API days) are shown in (e) and (f), and other points are shown in (b) and (c).

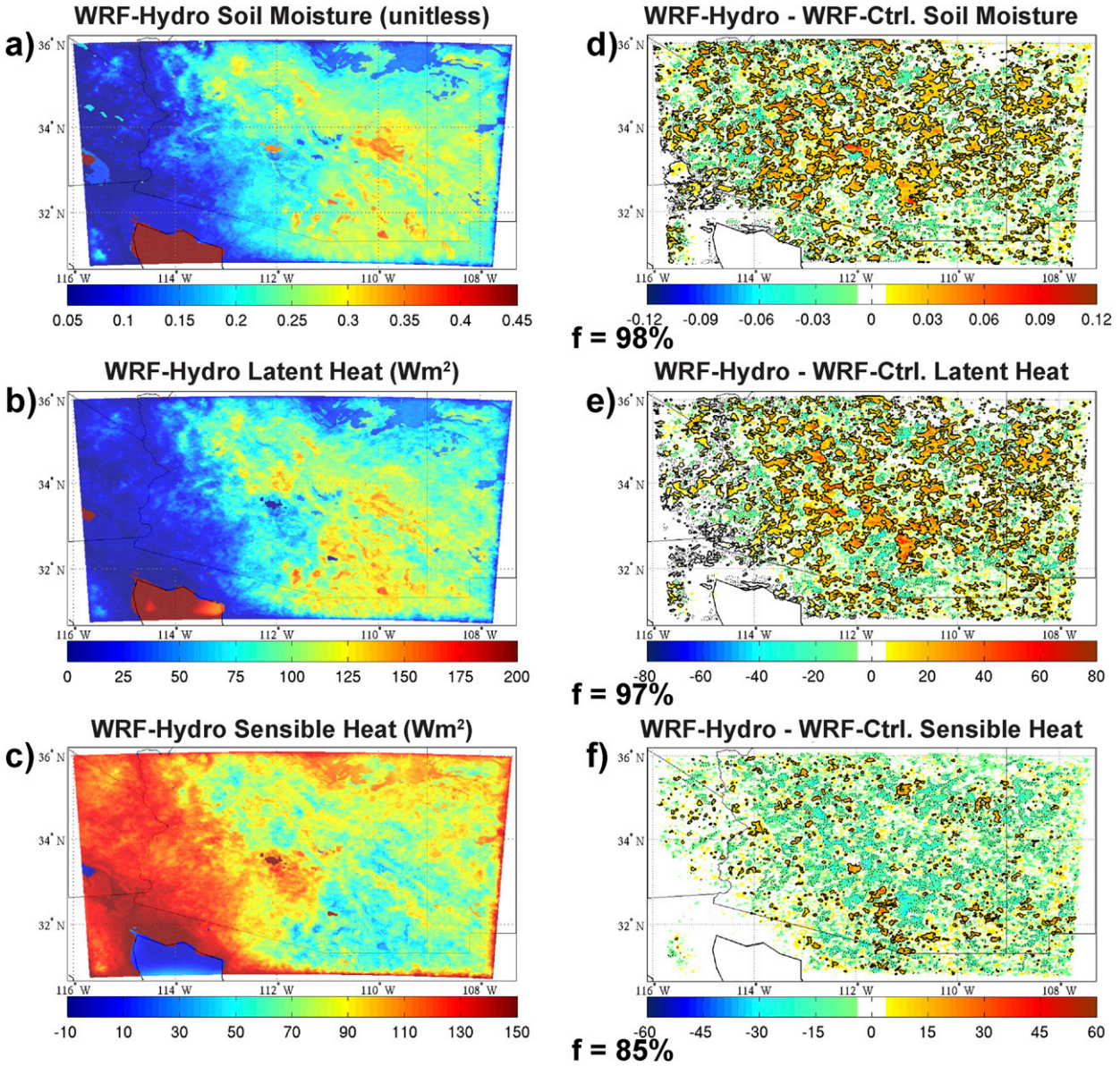

FIG. 8. As in Fig. 5, but for WRF-Hydro and WRF-Ctrl.

moisture when surface runoff occurs, an additional positive feedback.

b. *Surface variables and fluxes in coupled simulations*

Surface processes in the full atmospheric simulations (i.e., WRF-Ctrl and WRF-Hydro) are less spatially continuous because of differences in the precipitation distribution from both atmospheric simulations; however, Fig. 8 shows that WRF-Hydro consistently increases soil moisture compared to the control simulation for selected days with high API. These soil moisture changes are consistent with changes in latent heating (Fig. 8) in the WRF-Hydro simulations during the same high API days. For these days, when there is more widespread moisture in the system, latent heating increases by as much as $20\,\mathrm{W\,m^{-2}}$ across much of the model domain. Decreases in sensible heat throughout the model domain are also widespread (Fig. 8), likely due to increased partitioning of surface energy to latent heat from increased soil moisture. All of these changes are field significant, to at least 80%, implying that horizontal routing from the coupled model does impact the soil moisture and the surface energy budget.

Mixing diagrams at six sites across the domain are shown in Fig. 9, averaged across convective days (API $>$ 20 mm) at the selected stations around the

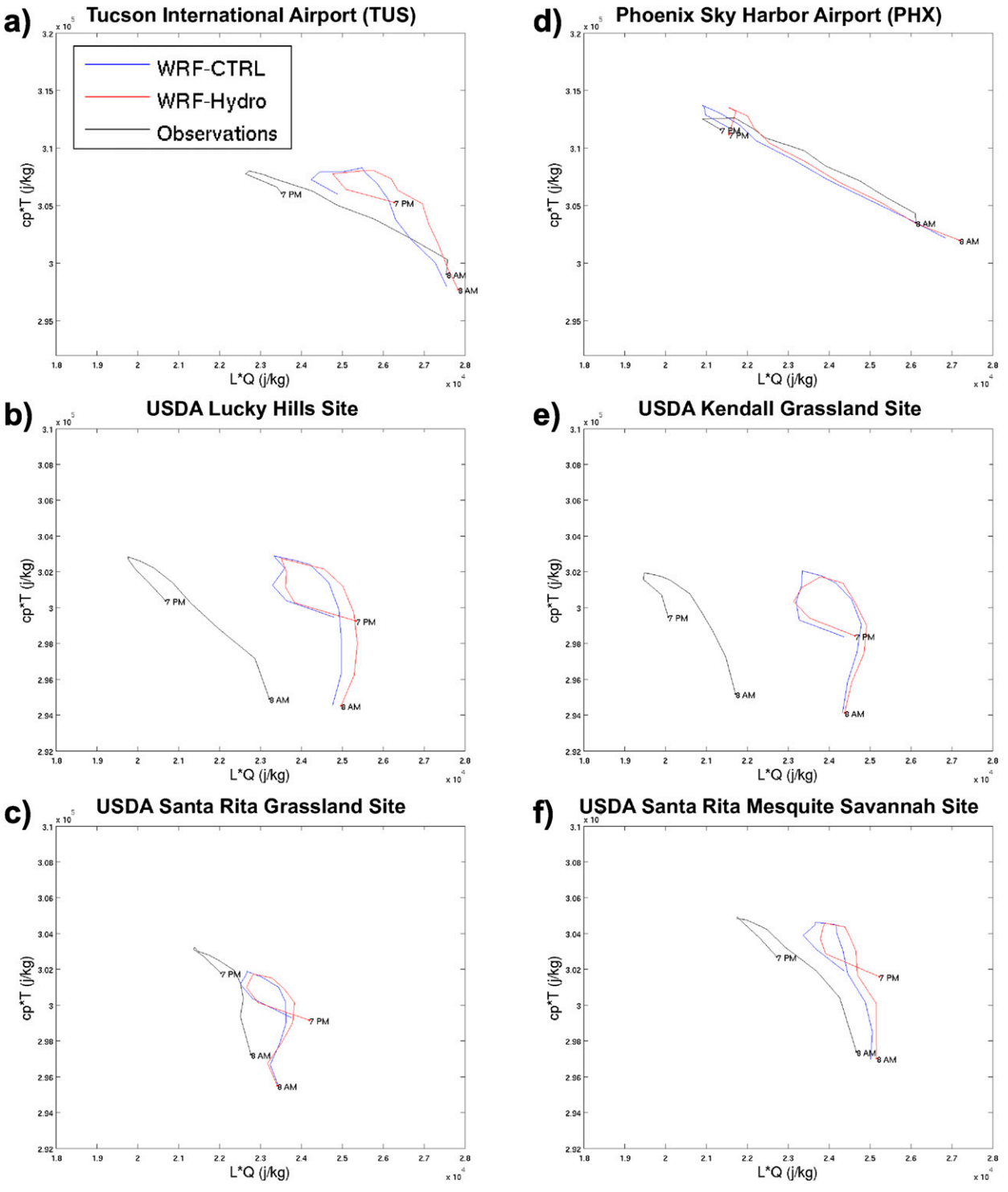

FIG. 9. WRF-Ctrl (red), WRF-Hydro (blue), and observed (black) mixing diagrams (units are J kg^{-1}) (as in Santanello et al. 2018) for (a) Tucson International Airport, (d) Phoenix Sky Harbor Airport, (b) Lucky Hills, (e) Kendall Grassland, (c) Santa Rita Grassland, and (f) Santa Rita Mesquite Savannah stations. Mixing diagrams plot the change in energy from latent heat (x axis; specific humidity multiplied by latent heat of vaporization QL_V) compared to the change in energy from sensible heat [y axis; temperature (K) multiplied by specific heat Tc_p] on days with a modeled API at the site of at least 20 mm.

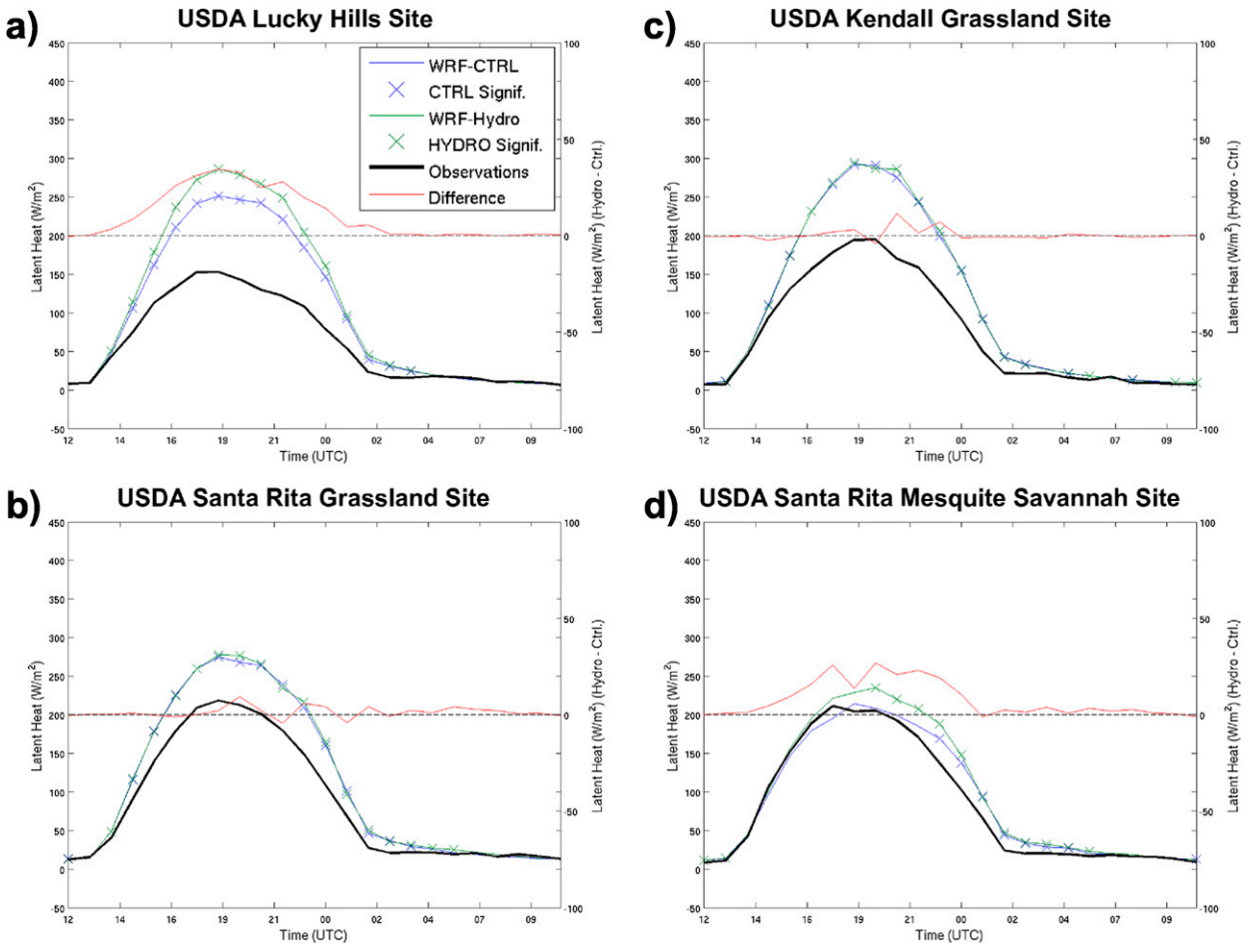

FIG. 10. The diurnal cycle for latent heat (W m^{-2}) for station API days of at least 20 mm is plotted. Stations include (a) Lucky Hills, (c) Kendall Grassland, (b) Santa Rita Grassland, and (d) Santa Rita Mesquite Savannah. Datasets include observations (black), WRF-Ctrl (blue), and WRF-Hydro (green). A red line shows the difference between WRF-Hydro and WRF-Ctrl. Model time series are marked with an additional "X" if they are statistically significantly different from the observations.

domain (different from the high API days for the full NAM domain). These figures demonstrate that WRF-Hydro is consistently wetter than the control WRF simulation. At Tucson and the WGEW flux tower sites (Kendall Grassland and Lucky Hills), both simulations produce similar diurnal cycles as shown in the mixing diagrams, with WRF-Hydro being slightly wetter, with more latent heat. The Santa Rita Grassland and Santa Rita Mesquite Savannah sites are also generally too wet. At Phoenix Sky Harbor Airport, the wetter WRF-Hydro nominally reduces a dry bias; however, WRF-Hydro increases an existing wet bias at all other locations. The magnitude of this increased bias is less than that of the bias itself. These results are consistent with the biased latent heat fluxes at all four flux tower sites through the diurnal cycle, with the exception of Santa Rita Mesquite Savannah, which is relatively close to the observations (Fig. 10). At all other sites, WRF-Ctrl and WRF-Hydro are both associated with higher latent heat flux compared to observations (Fig. 10). WRF also has a high sensible heat bias (Fig. 11), and adding WRF-Hydro increases (decreases) the latent heat (sensible heat) biases, through increases to soil moisture.

c. *Lower-atmosphere variables*

The effects on surface fluxes from adding surface routing to WRF-ARW through the WRF-Hydro model also impact the lower atmosphere, as 2-m temperature is lower across central Arizona for selected high API days (Figs. 12a,e). These changes are only statistically significant in a few areas. Similarly, specific humidity increases throughout most of the model domain (Figs. 12b,f), but these changes are only statistically significant over a small area on high API days. For high API days, temperature decreases by ~0.5°C and specific humidity increases by 0.5 g kg^{-1} across much of the

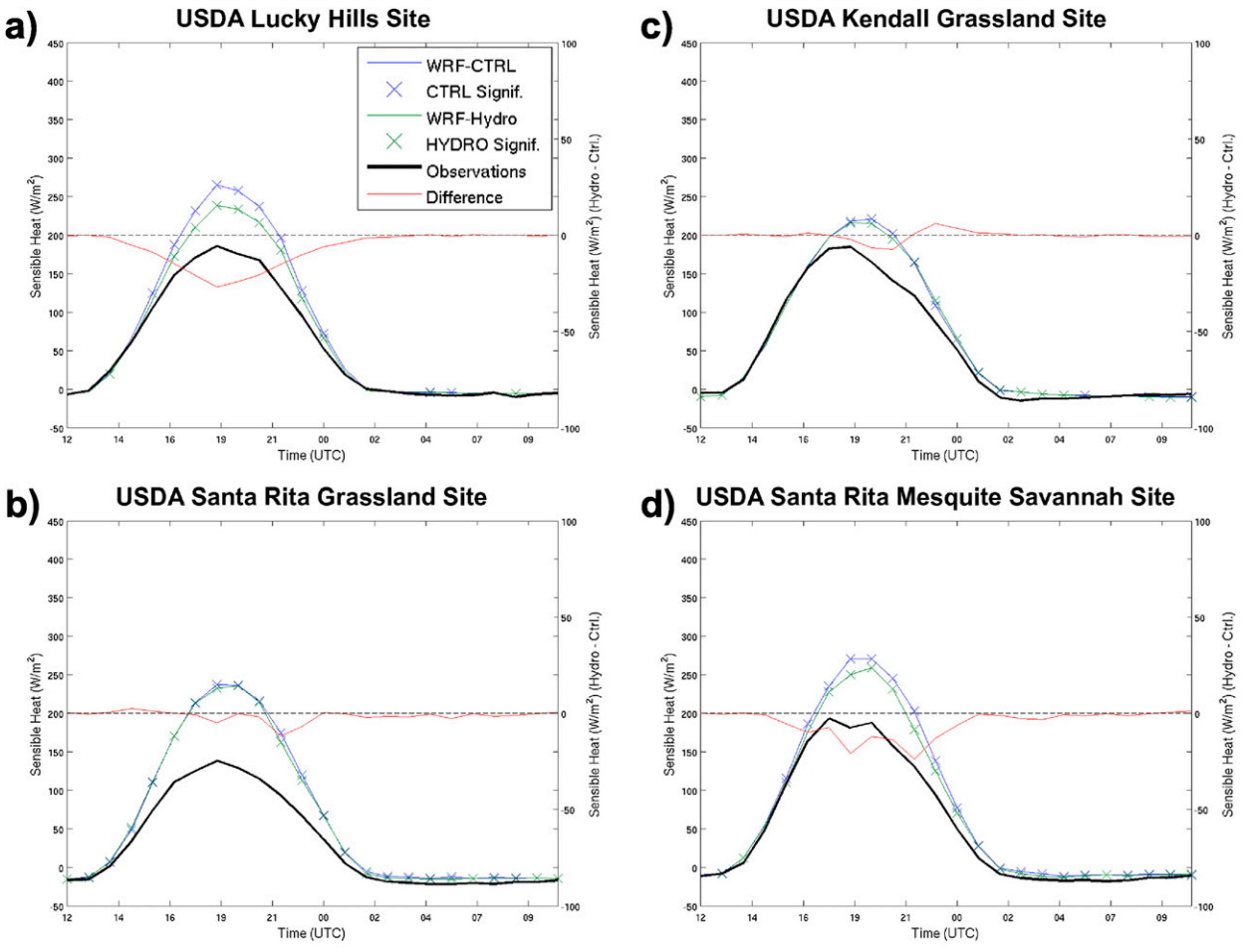

FIG. 11. As in Fig. 10, but for sensible heat.

domain. This suggests that redistribution of water across the land surface may be important for the lower atmosphere in this semiarid environment, as was shown in the uncoupled simulations in Figs. 5–7.

d. Local convective environment

Increases to the modeled specific humidity from surface routing subsequently reduce the LCL height in the coupled simulation (Figs. 12c,g). The changes are somewhat spatially consistent on selected high API days, with decreases to the LCL height on the order of 100 m across much of the domain, including central Arizona near the Phoenix metro area. Despite the small decreases in temperature, MUCAPE increases across much of the model domain, on the order of over ~50 J kg^{-1} over much of Arizona (Figs. 12d,h). These changes are statistically significant over a small area near the Phoenix Metro. These results suggest that surface routing can potentially influence the NAM convective environment, increasing the thermodynamic instability of the atmosphere, particularly on days when antecedent soil moisture from prior precipitation is greater. These changes occur because increased soil moisture in WRF-Hydro subsequently increases latent heating and specific humidity (Fig. 8). This added moisture makes the atmosphere more convectively unstable, reducing the LCL height and increasing MUCAPE. Due to the high variance in MUCAPE from day to day, these changes are not statistically significant over most of the domain.

e. Effects on convective propagation

Figure 13 shows the fractional differences of WRF-Hydro compared to WRF-Ctrl spectra for moisture flux convergence, as in Castro et al. (2005), where fractional difference ΔS_{frac} of two spectra S for a given wavenumber k is $\Delta S(k) = S(k)_{\text{hydro}}/S(k)_{\text{ctrl}} - 1$. Fractional differences between WRF-Ctrl and WRF-Hydro for the 3-h intervals, starting at 1200, 1500, 1800, 2100, 0000, and 0300 UTC are shown in Fig. 13, along with the bounds for statistical significance. Titles in this plot feature the same colors as the diurnal maxima shown in Fig. 2. Also as in Fig. 2, these plots show wavelength on

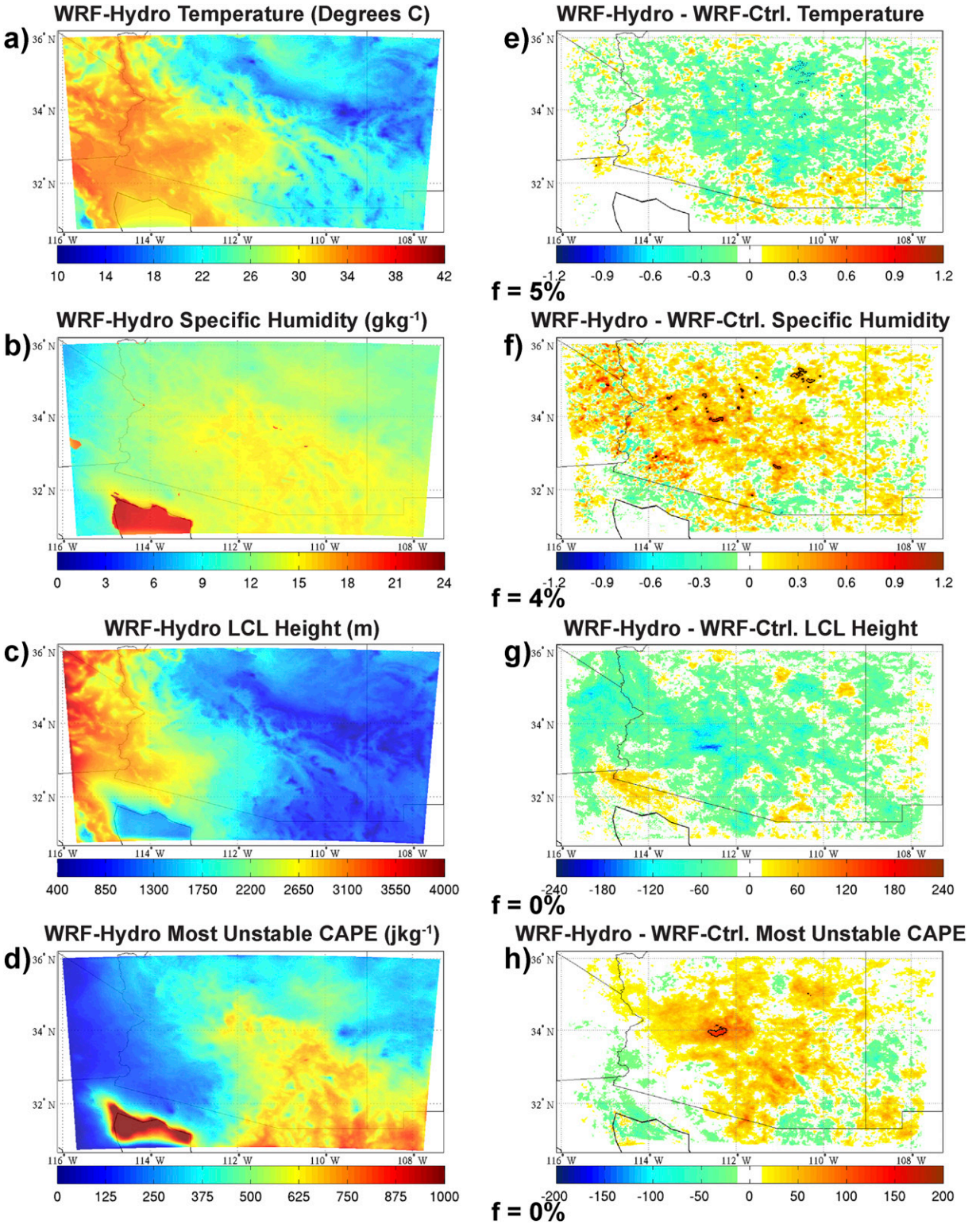

FIG. 12. WRF-Hydro averages for high API (API > 20 mm) days for (a) 2-m temperature (°C), (b) specific humidity (g kg^{-1}), (c) LCL height (m), and (d) MUCAPE (J kg^{-1}). (e)–(h) As in (a)–(d), but for differences between WRF-Hydro and WRF-Ctrl. For model variable differences, solid and dashed lines show 95th and 5th percentile statistical significance from a Student's *t* test, respectively. Field significance is plotted in the lower-left corner of these plots.

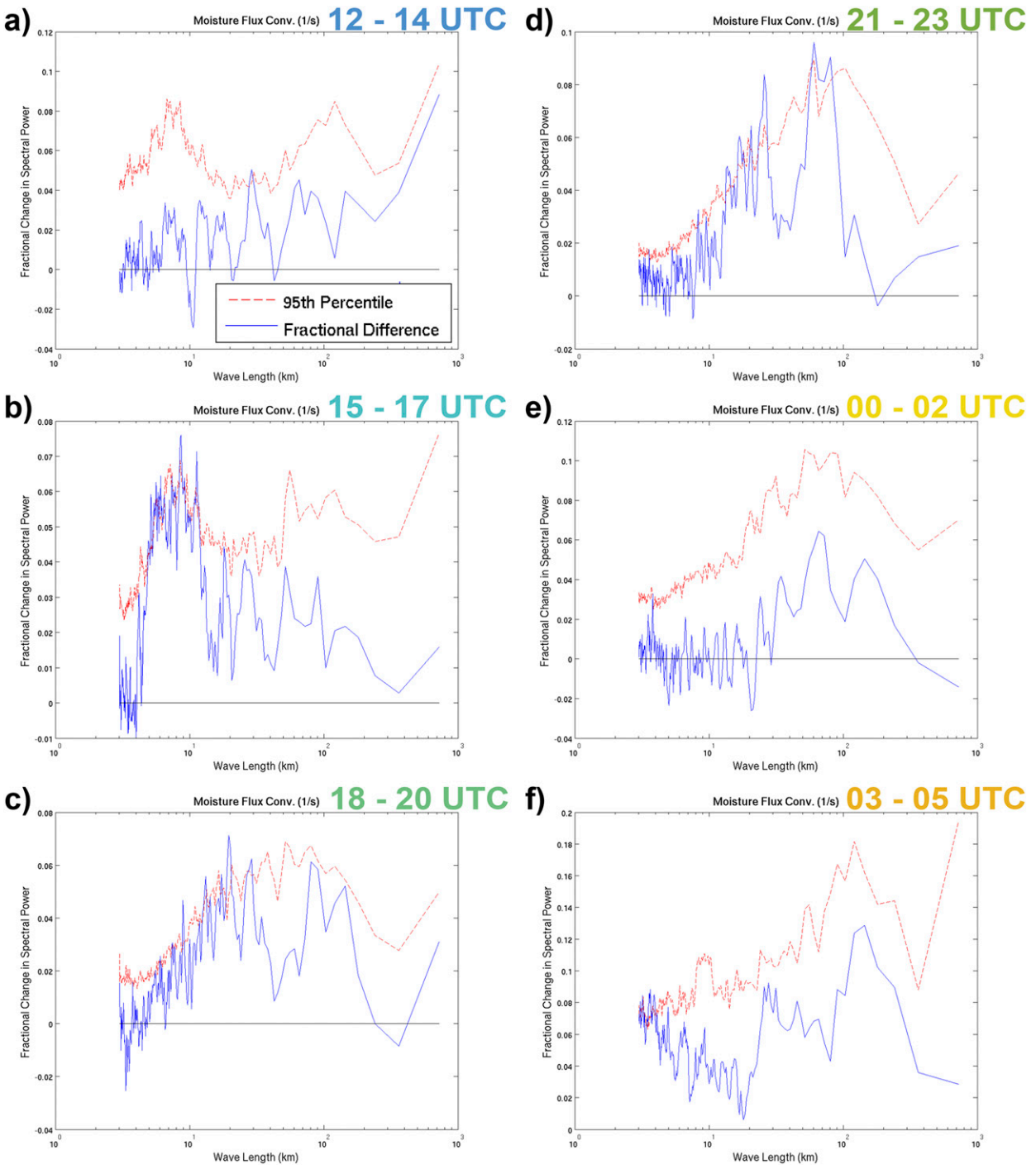

FIG. 13. Fractional difference of moisture flux convergence spectra (WRF-Hydro vs WRF-Ctrl) (blue) and 95th (red) percentile bounds for statistical significance, computed with Monte Carlo substitution. The 5th percentile bounds are omitted. Shown are the (a) 1200–1400, (b) 1500–1700, (c) 1800–2000, (d) 2100–2300, (e) 0000–0200, and (f) 0300–0500 UTC time steps. The time steps for each panel are given in the top-right corners, with coloring consistent with the diurnal cycle colors used in Fig. 2.

the *x* axis, but change in organization on the *y* axis, such that positive values indicate more organization with WRF-Hydro, compared to WRF-Ctrl. Thus, higher values in Fig. 13 conceptually indicate times and scales where WRF-Hydro enhances convective organization. The results from this figure suggest that the predominant scales of moisture flux convergence organization (on days with high MUCAPE and PW and no synoptic

scale influences) are greater in WRF-Hydro compared to WRF-Ctrl, on the scale of 5–10 km, starting during the 1500–1700 UTC (0800–1000 LST) interval. This grows up in scale, with some statistically significant increases in moisture flux convergence at scales of ~100 km by the 2100–2300 UTC (1400–1600 LST) interval. Spectral analysis of 500-hPa omega, another proxy for vertical motion, also shows enhanced convective organization in the WRF-Hydro simulations at the same times of the diurnal cycle, but to a lesser extent (supplemental material section 4). This suggests that the thermodynamic effects of hydrologic processes in the WRF-Hydro hydrologic model results in a higher propensity for convection to become more organized during the peak of the diurnal cycle, despite little change to the net amounts of precipitation in the system (Fig. 3). We note that while the increases to MUCAPE shown in Fig. 12 are not statistically significant for most of the domain, according to these simulations, they are sufficient to increase the spatial organization of convection at the peak of the diurnal cycle.

Figure 14 suggests that these changes to convective organization associated with WRF-Hydro might impact the location of MCS precipitation later in the diurnal cycle. Figure 14 shows the change in precipitation between WRF-Ctrl and WRF-Hydro at 3-h intervals of the diurnal cycle (using the same times and colors as Fig. 2). There is little consistent change from WRF-Hydro early in the convective cycle (i.e., before 0000 UTC), but by 0000–0300 UTC and 0300–0600 UTC (after the changes in convective organization are seen in the model spectra), WRF-Hydro favors more precipitation farther west in the domain, and is associated with a decrease in some areas farther east. Many of these changes are not statistically significant, and consequently not field significant (partly due to the high temporal variability of precipitation in both systems), but a longer simulation with more MCS events may be able to further verify these spatial patterns.

5. Discussion

We have used a configuration of the WRF-ARW model that has been validated for research and quasi-operational forecasts for the NAM region (e.g., Luong et al. 2017) and executed it as a type-2 RCM (e.g., Castro et al. 2005) with constant boundary conditions for two NAM seasons, both in a control configuration and coupled with the WRF-Hydro hydrologic model, to include surface and subsurface routing (top panels of Fig. 6) in the 2-m soil column closest to the surface. In this way, we are varying the surface boundary conditions of the atmospheric model through the inclusion of surface hydrologic processes. Despite both RCM simulations exhibiting similar temperature, specific humidity, and precipitation biases, compared to observations, to those documented in Luong et al. (2017), the models capture the diurnal cycle of organized propagating convection, as well as the spatial and temporal patterns of temperature, specific humidity, and soil moisture. Precipitation biases have been found in other RCM simulations as well (e.g., Liu et al. 2017) and likely also reflect uncertainties in the NCEP Stage IV precipitation product (e.g., Zamora et al. 2014; Nelson et al. 2016). Thus, we can establish both the WRF-Ctrl and WRF-Hydro RCM simulations as realistic representations of the NAM system.

Using these model configurations, we have shown that including lateral surface flow in the WRF-ARW model through the use of WRF-Hydro, for a number of locations results in increased soil moisture, and therefore increased latent heating and atmospheric instability in the simulations of the NAM season. Based on spectral analysis, this leads to enhanced convection, as shown by the statistically significantly greater organization of moisture flux convergence at the peak of the diurnal convective cycle (Fig. 13). Conceptually, these processes are shown in Fig. 15, where we illustrate the WRF-Ctrl and WRF-Hydro model states and processes. We emphasize that these results should be considered a sensitivity analysis, with an LSM that was not calibrated. We also acknowledge that there is a slight positive precipitation bias in the WRF-ARW configuration, which likely contributes to the model system being slightly too moist. This soil moisture bias is greater (by at least an order of magnitude) than the differences in moisture between WRF-Ctrl and WRF-Hydro (Fig. 3); however, this also reflects differences between the nature of the model state variables and surface measurements (e.g., Santanello et al. 2018). Increased soil moisture occurs in regions of the model domain where surface runoff increases from WRF-Hydro, suggesting that additional surface runoff (from Noah-MP surface runoff being rerouted with WRF-Hydro) is the main cause of this change (Fig. 6). This occurs in areas of the domain with high precipitation and low soil conductivity, and to a lesser extent a high drainage area.

This moistening of the land surface from WRF-Hydro is consistent with past studies that showed coupling WRF-Hydro to WRF-ARW also increased soil moisture, and subsequently latent heat (e.g., Arnault et al. 2016; Arnault et al. 2018; Rummler et al. 2019). As in these previous studies, we find that WRF-Hydro has little effect on total precipitation, averaged across the domain. Arnault et al. (2016) performed their simulations in a small domain in West Africa, another semiarid region,

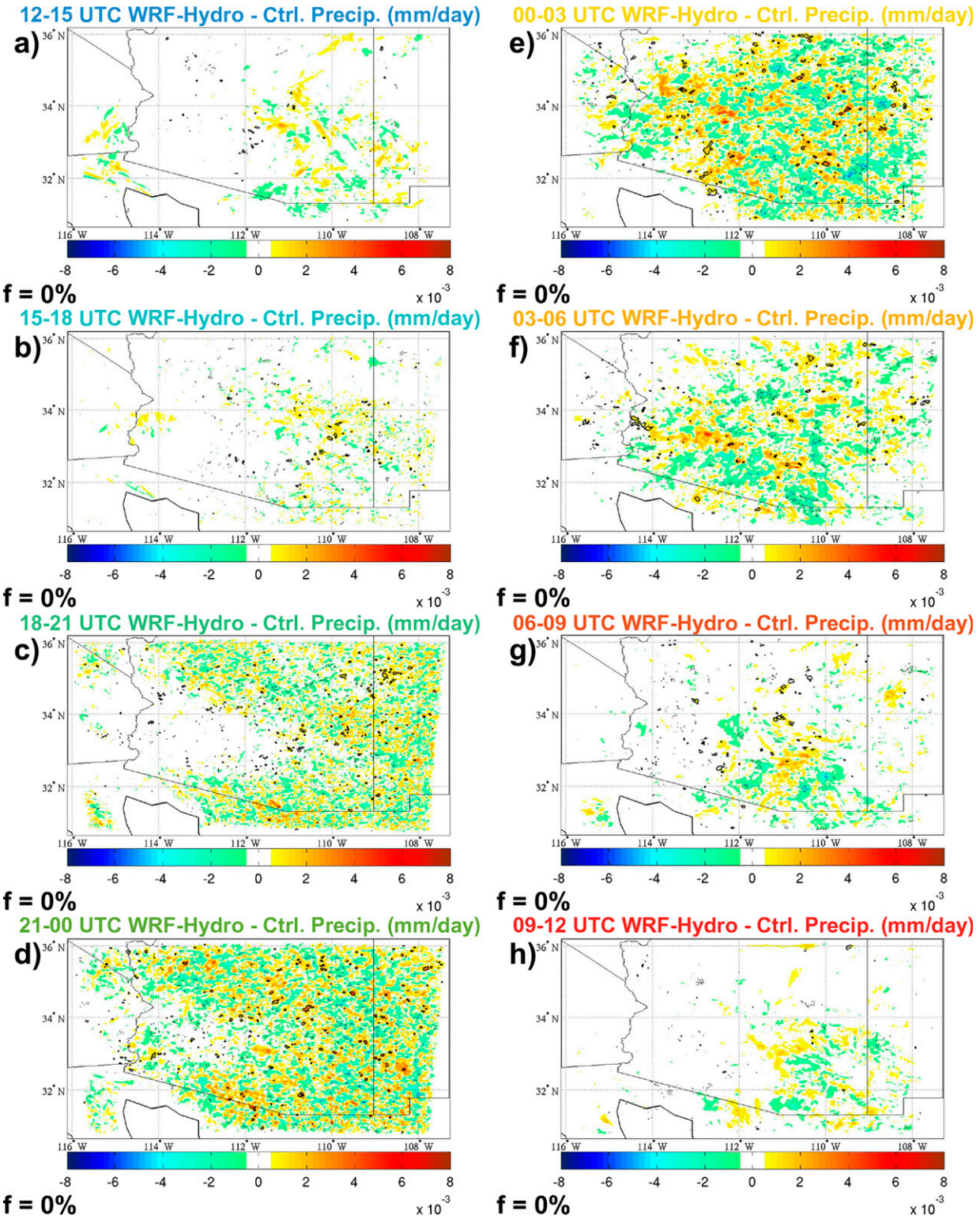

FIG. 14. The difference between WRF-Hydro and WRF-Ctrl precipitation, with 90% statistical significance, for the (a) 1200–1500, (b) 1500–1800, (c) 1800–2100, (d) 2100–0000, (e) 0000–0300, (f) 0300–0600, (g) 0600–0900, and (h) 0900–1200 UTC time steps is plotted. For all plots, solid and dashed lines show 95th and 5th percentile statistical significance from a Student's *t* test, respectively. Field significance is plotted in the lower-left corner of these plots. Titles use coloring consistent with the diurnal cycle colors used in Fig. 2.

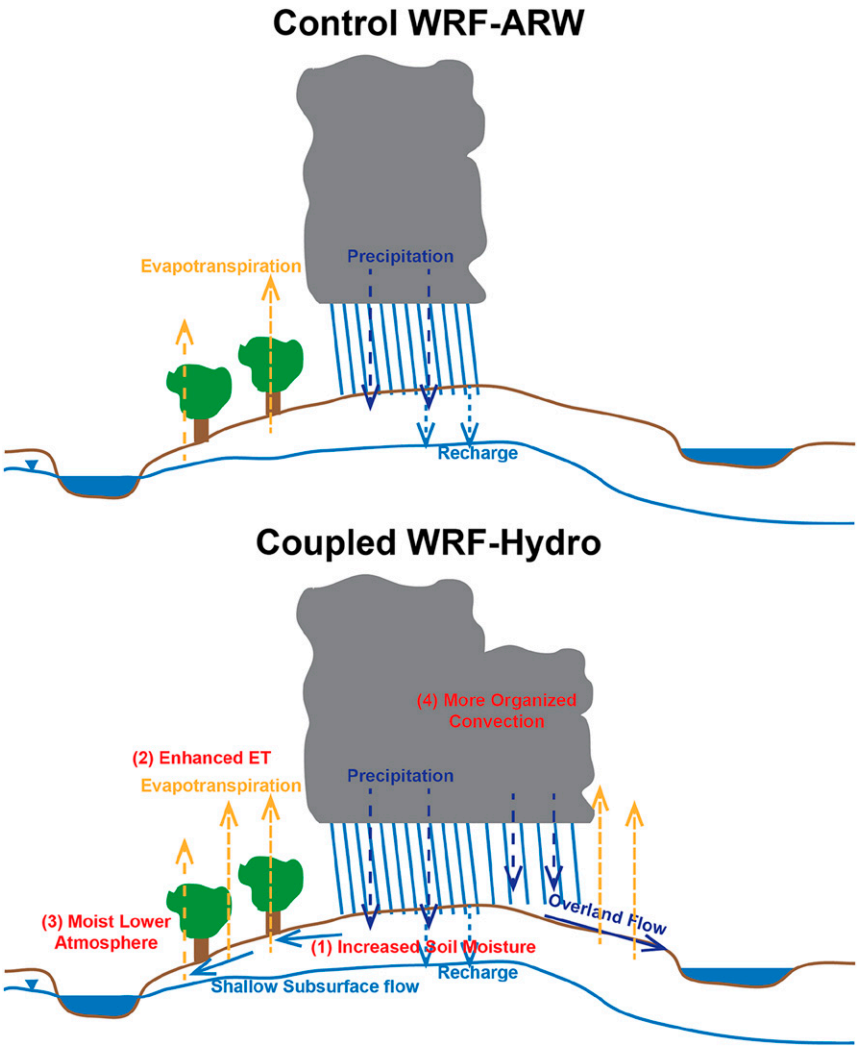

FIG. 15. Conceptual illustration of the effects of (bottom) WRF-Hydro compared to (top) WRF-Ctrl on surface–atmosphere exchanges and convection. Note, that in WRF-Hydro, deep groundwater recharge is approximated when Noah-MP drainage is removed from the system.

and also showed that changes to soil moisture influenced convective initiation and propagation. Accordingly, we have shown that the addition of WRF-Hydro increases (decreases) surface specific humidity (temperature), and results in a more unstable convective environment that is more favorable for propagating convection (Fig. 15).

As discussed above, the impact of WRF-Hydro on surface routing as shown here is theoretical, and its impact on real world conditions should be interpreted with caution. To quantify the precise impact of surface routing in a coupled WRF-Hydro simulation, calibration of the WRF-Hydro hydrologic model, including soil moisture (e.g., Xiang et al. 2018) and the LSM configuration (e.g., Santanello et al. 2013) is needed. However, despite these suggested limitations of our model configuration and the theoretical nature of the work presented here, these results still suggest that surface routing has some impact on the distribution of soil moisture in a semiarid environment, which in turn impacts the lower atmosphere (e.g., Arnault et al. 2016).

Moistening of the convective environment shown here in our mixing diagrams could be further analyzed with a more robust observation record (as in Santanello et al. 2013, 2018), to compute the surface and entrainment vector components in the mixing diagrams (Fig. 9). More complete boundary layer data would also permit us to analyze LCL height versus boundary layer height and its influence on the convective environment. A field campaign, such as the 2004 North American Monsoon Experiment (NAME) (Higgins et al. 2006), including both surface radar data to improve precipitation estimates, combined with frequent upper air observations

based at a USDA site such as WGEW could potentially be used to investigate the influences of surface hydrology during the NAM. This is in addition to improvements to the model configuration to better address soil moisture and surface variable biases.

The effects of groundwater coupled back to WRF-Hydro may also have an influence (e.g., Maxwell et al. 2011), particularly in areas like the southern plains where groundwater is of greater importance for surface hydrology and surface–atmosphere interactions (e.g., Santanello et al. 2018). While our model configuration is a reasonable approximation to hydrologic processes across much of southern Arizona (e.g., Blasch et al. 2004), future releases of the WRF-Hydro model that utilize a more advanced groundwater scheme that can be coupled to the surface (e.g., Maxwell et al. 2015) may help to further investigate this question.

Furthermore, we caution that a longer analysis period would be more useful to evaluate the effects of hydrologic routing in the system, as they are superimposed on climate variability and change (e.g., Castro et al. 2007a). This would ameliorate the limitations of our small sample of 28 high-API days. For such an analysis, a lower resolution coupled WRF-Hydro simulation would not have been useful, as water moving horizontally on and within the land surface must be reaggregated whenever Noah-MP is called (every hour in our simulations). In a coarser domain, this would lessen the impacts of hydrologic routing from WRF-Hydro. A coarser RCM domain (i.e., greater than 5-km grid resolution) would have decreased runtime considerably, but would not have been able to capture convective initiation and organization in the NAM region (e.g., Luong et al. 2017; Prein et al. 2015) as well as small-scale horizontally moving processes at the land surface. The computational cost of executing a longer-range RCM simulation over a large regional domain at high resolution prohibits a longer analysis in the current work. However, the results shown here suggest that such an analysis would be valuable in the future to fully realize the impacts of a coupled atmospheric and hydrologic simulation.

6. Conclusions

We have shown that coupling a WRF-ARW configuration, for the NAM convective environment to the uncalibrated WRF-Hydro (version 3.0) hydrologic model influences the model simulations by the following effects (Fig. 15):

1) WRF-Hydro is associated with increased soil moisture caused by water that is otherwise removed from Noah-MP as surface runoff being rerouted through the surface runoff module of WRF-Hydro. This added surface runoff occurs in some favorable locations in the model domain.
2) These widespread increases to soil moisture in WRF-Hydro increase latent heating and reduce sensible heating.
3) Despite lower temperatures at the land surface, WRF-Hydro is associated with a moister surface and a more unstable convective environment.
4) These changes to the WRF-ARW configuration cause convection to become more organized at the peak of the diurnal cycle.

Despite these impacts, further analysis is needed to address several caveats in our methodology, including the following:

1) A long-range RCM simulation (~10 years) is needed to capture a larger dataset of convective events and determine how they might vary as a function of climate change and variability (e.g., Luong et al. 2017; Castro et al. 2007b).
2) Given the biases in model soil moisture, a calibrated hydrologic model is needed to further verify and quantify the influences of WRF-Hydro in our system. Spatial regularization (e.g., Dugger et al. 2017) that is currently being tested with the WRF-Hydro model for CONUS-scale forecasts could be used to generate a physically realistic parameter set for such a configuration, as well as evaluating the LSM configuration (e.g., Santanello et al. 2013). Advanced parameter regionalization methods that can produce regionally continuous parameter surfaces (e.g., Samaniego et al. 2010; Kumar et al. 2013; Mizukami et al. 2017) would also be desirable for such simulations. A more physically consistent model architecture to simulate groundwater (e.g., Maxwell et al. 2015) could also potentially improve future simulations.
3) A field campaign over the study region (e.g., Higgins et al. 2006) that includes frequent upper-atmosphere profiles (e.g., Santanello et al. 2013) could be used to further investigate the impacts of surface water routing on atmospheric fluxes and the planetary boundary layer.
4) Future work could include dynamically downscaling the Liu et al. (2017) or similar RCM datasets to the ~1-km scale to evaluate these trends. Such a dataset could also be used for climate change impacts assessment (e.g., Luong et al. 2017).

Acknowledgments. Observation data, including NCEP Stage IV Precipitation, SMAP–Sentinel soil moisture, and surface observations are available online. Surface station observations and CFSR forcing are available through the

National Centers for Environmental Information (NCEI), and flux tower observations are available through AmeriFlux. All model outputs are permanently archived on a server maintained by The University of Arizona Department of Hydrology and Atmospheric Sciences are available upon request. This work was supported by NOAA Joint Technology Transfer Initiative (JTTI) Federal Grant Number: NA17OAR4590183. We acknowledge Russ Scott (USDA-ARS Southwest Watershed Research Center) for providing WGEW flux tower data records from the 2017 and 2018 NAM seasons.

REFERENCES

Adams, D. K., and A. C. Comrie, 1997: The North American monsoon. *Bull. Amer. Meteor. Soc.*, **78**, 2197–2213, https://doi.org/10.1175/1520-0477(1997)078<2197:TNAM>2.0.CO;2.

Anyah, R. O., C. P. Weaver, G. Miguez-Macho, Y. Fan, and A. Robock, 2008: Incorporating water table dynamics in climate modeling: 3. Simulated groundwater influence on coupled land-atmosphere variability. *J. Geophys. Res.*, **113**, D07103, https://doi.org/10.1029/2007JD009087.

Arnault, J., S. Wagner, T. Rummler, B. Fersch, J. Bliefernicht, S. Andresen, and H. Kunstmann, 2016: Role of runoff–infiltration partitioning and resolved overland flow on land–atmosphere feedbacks: A case study with the WRF-Hydro coupled modeling system for West Africa. *J. Hydrometeor.*, **17**, 1489–1516, https://doi.org/10.1175/JHM-D-15-0089.1.

——, and Coauthors, 2018: Precipitation sensitivity to the uncertainty of terrestrial water flow in WRF-Hydro: An ensemble analysis for central Europe. *J. Hydrometeor.*, **19**, 1007–1025, https://doi.org/10.1175/JHM-D-17-0042.1.

Barlage, M., M. Tewari, F. Chen, G. Miguez-Macho, Z.-L. Yang, and G.-Y. Niu, 2015: The effect of groundwater interaction in North American regional climate simulations with WRF/Noah-MP. *Climatic Change*, **129**, 485–498, https://doi.org/10.1007/s10584-014-1308-8.

Bieda, S. W., C. L. Castro, S. L. Mullen, A. C. Comrie, and E. Pytlak, 2009: The relationship of transient upper-level troughs to variability of the North American monsoon system. *J. Climate*, **22**, 4213–4227, https://doi.org/10.1175/2009JCLI2487.1.

Blasch, K., T. P. Ferré, J. Hoffmann, D. Poll, M. Baily, and J. Cordova, 2004: Processes controlling recharge beneath ephemeral streams in Southern Arizona. *Groundwater Recharge in a Desert Environment: The Southwestern United States*, J. F. Hogan, F. M. Phillips, and B. R. Scanlon, Eds., Water Science and Application Series, Vol. 9, Amer. Geophys. Union, 69–76.

Bosart, L., T. J. Melino, S. R. Sukup, E. S. Pytlak, J. E. Matusiak, S. J. Weiss, J. Racy, and R. S. Schneider, 2011: Potential vorticity disturbances as a trigger of southwest U.S. severe weather. *24th Conf. on Weather and Forecasting*, Seattle, WA, Amer. Meteor. Soc., 11A.3, https://ams.confex.com/ams/91Annual/webprogram/Paper181548.html.

Castro, C. L., R. A. Pielke Sr., and G. Leoncini, 2005: Dynamical downscaling: Assessment of value retained and added using the Regional Atmospheric Modeling System (RAMS). *J. Geophys. Res.*, **110**, D05108, https://doi.org/10.1029/2004JD004721.

——, R. A. Pielke, and J. O. Adegoke, 2007a: Investigation of the summer climate of the contiguous United States and Mexico using the Regional Atmospheric Modeling System (RAMS). Part I: Model climatology (1950–2002). *J. Climate*, **20**, 3844–3865, https://doi.org/10.1175/JCLI4211.1.

——, R. A. Pielke Sr., J. O. Adegoke, S. D. Schubert, and P. J. Pegion, 2007b: Investigation of the summer climate of the contiguous United States and Mexico using the Regional Atmospheric Modeling System (RAMS). Part II: Model climate variability. *J. Climate*, **20**, 3866–3887, https://doi.org/10.1175/JCLI4212.1.

Chang, H.-I., C. L. Castro, C. M. Carrillo, and F. Dominguez, 2015: The more extreme nature of U.S. warm season climate in the recent observational record and two "well-performing" dynamically downscaled CMIP3 models. *J. Geophys. Res. Atmos.*, **120**, 8244–8263, https://doi.org/10.1002/2015JD023333.

Cordery, I., 1970: Initial loss for flood estimation and forecasting. *J. Hydraul. Div.*, **96**, 2447–2466.

Cotton, W. R., and Coauthors, 2003: RAMS 2001: Current status and future directions. *Meteor. Atmos. Phys.*, **82**, 5–29, https://doi.org/10.1007/s00703-001-0584-9.

Das, N., D. Entekhabi, R. S. Dunbar, S. Kim, S. Yueh, A. Colliander, P. E. O'Neill, and T. Jackson, 2018: SMAP/Sentinel-1 L2 Radiometer/Radar 30-Second Scene 3 km EASE-Grid Soil Moisture, version 2. NASA National Snow and Ice Data Center Distributed Active Archive Center, accessed 9 October 2018, https://doi.org/10.5067/KE1CSVXMI95Y.

Dirmeyer, P. A., C. A. Schlosser, and K. L. Brubaker, 2009: Precipitation, recycling, and land memory: An integrated analysis. *J. Hydrometeor.*, **10**, 278–288, https://doi.org/10.1175/2008JHM1016.1.

Dominguez, F., G. Miguez-Macho, and H. Hu, 2016: WRF with water vapor tracers: A study of moisture sources for the North American monsoon. *J. Hydrometeor.*, **17**, 1915–1927, https://doi.org/10.1175/JHM-D-15-0221.1.

Douglas, M. W., R. A. Maddox, K. Howard, and S. Reyes, 1993: The Mexican monsoon. *J. Climate*, **6**, 1665–1677, https://doi.org/10.1175/1520-0442(1993)006<1665:TMM>2.0.CO;2.

Dugger, A. L., and Coauthors, 2017: Learning from the National Water Model: Regional improvements in streamflow prediction through experimental parameter and physics updates to the WRF-Hydro Community Model. *31st Conf. on Hydrology*, Seattle, WA, Amer. Meteor. Soc., 6A.3, https://ams.confex.com/ams/97Annual/webprogram/Paper314352.html.

Errico, R. M., 1985: Spectra computed from a limited area grid. *Mon. Wea. Rev.*, **113**, 1554–1562, https://doi.org/10.1175/1520-0493(1985)113<1554:SCFALA>2.0.CO;2.

Fan, Y., H. Li, and G. Miguez-Macho, 2013: Global patterns of groundwater table depth. *Science*, **339**, 940–943, https://doi.org/10.1126/science.1229881.

Finch, Z. O., and R. H. Johnson, 2010: Observational analysis of an upper-level inverted trough during the 2004 North American monsoon experiment. *Mon. Wea. Rev.*, **138**, 3540–3555, https://doi.org/10.1175/2010MWR3369.1.

Findell, K. L., and E. A. Eltahir, 1997: An analysis of the soil moisture-rainfall feedback, based on direct observations from Illinois. *Water Resour. Res.*, **33**, 725–735, https://doi.org/10.1029/96WR03756.

Gilbert, J. M., R. M. Maxwell, and D. J. Gochis, 2017: Effects of water-table configuration on the planetary boundary layer over the San Joaquin River Watershed, California. *J. Hydrometeor.*, **18**, 1471–1488, https://doi.org/10.1175/JHM-D-16-0134.1.

Gochis, D. J., W. Yu, and D. N. Yates, 2015: The WRF-Hydro model technical description and user's guide, version 3.0. NCAR Tech. Doc., 120 pp., http://www.ral.ucar.edu/projects/wrf_hydro/.

Goodrich, D. C., and Coauthors, 2004: Comparison of methods to estimate ephemeral channel recharge, Walnut Gulch San Pedro River Basin, Arizona. *Recharge and Valdose Zone Processes: Alluvia Basins of the Southwestern United States*, F. M. Phillips, J. F. Hogan, and B. Scanlon, Eds., Water Science and Application Series, Vol. 9, Amer. Geophys. Union, 77–99.

Hellweger, F., 1997: AGREE - DEM Surface Reconditioning System. http://www.ce.utexas.edu/prof/maidment/GISHYDRO/ferdi/research/agree/agree.html.

Higgins, W., and Coauthors, 2006: The NAME 2004 field campaign and modeling strategy. *Bull. Amer. Meteor. Soc.*, **87**, 79–94, https://doi.org/10.1175/BAMS-87-1-79.

Julien, P. Y., B. Saghafian, and F. L. Ogden, 1995: Raster-based hydrological modeling of spatially-varied surface runoff. *J. Amer. Water Resour. Assoc.*, **31**, 523–536, https://doi.org/10.1111/j.1752-1688.1995.tb04039.x.

Kerandi, N., J. Arnault, P. Laux, S. Wagner, J. Kitheka, and H. Kunstmann, 2018: Joint atmospheric-terrestrial water balances for East Africa: A WRF-Hydro case study for the upper Tana River basin. *Theor. Appl. Climatol.*, **131**, 1337–1355, https://doi.org/10.1007/s00704-017-2050-8.

Keune, J., F. Gasper, K. Goergen, A. Hense, P. Shrestha, M. Sulis, and S. Kollet, 2016: Studying the influence of groundwater representations on land surface-atmosphere feedbacks during the European heat wave in 2003. *J. Geophys. Res. Atmos.*, **121**, 13 301–13 325, https://doi.org/10.1002/2016JD025426.

——, M. Sulis, S. Kollet, S. Siebert, and Y. Wada, 2018: Human water use impacts on the strength of the continental sink for atmospheric water. *Geophys. Res. Lett.*, **45**, 4068–4076, https://doi.org/10.1029/2018GL077621.

Koster, R. D., P. A. Dirmeyer, A. N. Hahmann, R. Ijpelaar, L. Tyahla, P. Cox, and M. J. Suarez, 2002: Comparing the degree of land–atmosphere interaction in four atmospheric general circulation models. *J. Hydrometeor.*, **3**, 363–375, https://doi.org/10.1175/1525-7541(2002)003<0363:CTDOLA>2.0.CO;2.

Kumar, R., B. Livneh, and L. Samaniego, 2013: Toward computationally efficient large-scale hydrologic predictions with a multiscale regionalization scheme. *Water Resour. Res.*, **49**, 5700–5714, https://doi.org/10.1002/wrcr.20431.

Lahmers, T. M., C. L. Castro, D. K. Adams, Y. L. Serra, J. J. Brost, and T. Luong, 2016: Long-term changes in the climatology of transient inverted troughs over the North American monsoon region and their effects on precipitation. *J. Climate*, **29**, 6037–6064, https://doi.org/10.1175/JCLI-D-15-0726.1.

——, H. V. Gupta, C. L. Castro, D. J. Gochis, D. N. Yates, A. L. Dugger, D. C. Goodrich, and P. Hazenberg, 2019: Enhancing the structure of the WRF-Hydro hydrologic model for semi-arid environments. *J. Hydrometeor.*, **20**, 691–714, https://doi.org/10.1175/JHM-D-18-0064.1.

Lindsay, R. K., M. A. Kohler, and J. L. H. Paulhus, 1975: *Hydrology for Engineers*. 2nd ed. McGraw-Hill, 482 pp.

Liu, C., and Coauthors, 2017: Continental-scale convection-permitting modeling of the current and future climate of North America. *Climate Dyn.*, **49**, 71–95, https://doi.org/10.1007/s00382-016-3327-9.

Livezey, R. E., and W. Y. Chen, 1983: Statistical field significance and its determination by Monte Carlo techniques. *Mon. Wea. Rev.*, **111**, 46–59, https://doi.org/10.1175/1520-0493(1983)111<0046:SFSAID>2.0.CO;2.

Lu, E., X. Zeng, Z. Jiang, Y. Wang, and Q. Zhang, 2009: Precipitation and precipitable water: Their temporal-spatial behaviors and use in determining monsoon onset/retreat and monsoon regions. *J. Geophys. Res.*, **114**, D23105, https://doi.org/10.1029/2009JD012146.

Luong, T. M., C. L. Castro, H. Chang, T. Lahmers, D. K. Adams, and C. A. Ochoa-Moya, 2017: The more extreme nature of North American monsoon precipitation in the southwestern United States as revealed by a historical climatology of simulated severe weather events. *J. Appl. Meteor. Climatol.*, **56**, 2509–2529, https://doi.org/10.1175/JAMC-D-16-0358.1.

Maddox, R. A., D. M. McCollum, and K. W. Howard, 1995: Large-scale patterns associated with severe summertime thunderstorms over central Arizona. *Wea. Forecasting*, **10**, 763–778, https://doi.org/10.1175/1520-0434(1995)010<0763:LSPAWS>2.0.CO;2.

Maxwell, R. M., J. K. Lundquist, J. D. Mirocha, S. G. Smith, C. S. Woodward, and A. F. B. Tompson, 2011: Development of a coupled groundwater–atmosphere model. *Mon. Wea. Rev.*, **139**, 96–116, https://doi.org/10.1175/2010MWR3392.1.

——, L. E. Condon, and S. J. Kollet, 2015: A high-resolution simulation of groundwater and surface water over most of the continental US with the integrated hydrologic model ParFlow v3. *Geosci. Model Dev.*, **8**, 923–937, https://doi.org/10.5194/gmd-8-923-2015.

Mazon, J. J., C. L. Castro, D. K. Adams, H. Chang, C. M. Carrillo, and J. J. Brost, 2016: Objective climatological analysis of extreme weather events in Arizona during the North American monsoon. *J. Appl. Meteor. Climatol.*, **55**, 2431–2450, https://doi.org/10.1175/JAMC-D-16-0075.1.

McCollum, D. M., R. A. Maddox, and K. W. Howard, 1995: Case study of a severe mesoscale convective system in central Arizona. *Wea. Forecasting*, **10**, 643–665, https://doi.org/10.1175/1520-0434(1995)010<0643:CSOASM>2.0.CO;2.

McKay, L., T. Bondelid, T. Dewald, J. Johnston, R. Moore, and A. Rea, 2012: NHDPlus Version 2: User Guide. U.S. EPA Tech. Doc., 182 pp., ftp://ftp.horizon-systems.com/nhdplus/NHDPlusV21/Documentation/NHDPlusV2_User_Guide.pdf.

Miguez-Macho, G., G. L. Stenchikov, and A. Robock, 2005: Regional climate simulations over North America: Interaction of local processes with improved large-scale flow. *J. Climate*, **18**, 1227–1246, https://doi.org/10.1175/JCLI3369.1.

——, Y. Fan, C. P. Weaver, R. Walko, and A. Robock, 2007: Incorporating water table dynamics in climate modeling: 2. Formulation, validation, and soil moisture simulation. *J. Geophys. Res.*, **112**, D13108, https://doi.org/10.1029/2006JD008112.

Mizukami, N., M. P. Clark, A. J. Newman, A. W. Wood, E. D. Gutmann, B. Nijssen, O. Rakovec, and L. Samaniego, 2017: Towards seamless large-domain parameter estimation for hydrologic models. *Water Resour. Res.*, **53**, 8020–8040, https://doi.org/10.1002/2017WR020401.

Moncrieff, M., and M. Miller, 1976: The dynamics and simulation of tropical cumulonimbus and squall lines. *Quart. J. Roy. Meteor. Soc.*, **102**, 373–394, https://doi.org/10.1002/qj.49710243208.

Nelson, B. R., O. P. Prat, D. Seo, and E. Habib, 2016: Assessment and Implications of NCEP Stage IV quantitative precipitation estimates for product intercomparisons. *Wea. Forecasting*, **31**, 371–394, https://doi.org/10.1175/WAF-D-14-00112.1.

Newman, A., and R. H. Johnson, 2012: Mechanisms for precipitation enhancement in a North American monsoon upper-tropospheric trough. *J. Atmos. Sci.*, **69**, 1775–1792, https://doi.org/10.1175/JAS-D-11-0223.1.

Niu, G. Y., and Coauthors, 2011: The community Noah land surface model with multiparameterization options (Noah-MP): 1. Model description and evaluation with local-scale measurements. *J. Geophys. Res.*, **116**, D12109, https://doi.org/10.1029/2010JD015139.

Ogden, F. L., 1997: CASC2D reference manual, version 2.0. Department of Civil and Evnironmental Engineering U-37, University of Connecticut, 106 pp.

Pal, S., H.-I. Chang, C. L. Castro, and F. Dominguez, 2019: Credibility of convection-permitting modeling to improve seasonal precipitation forecasting in the southwestern United States. *Front. Earth Sci.*, **7**, 11, https://doi.org/10.3389/FEART.2019.00011.

Pielke, R. A., and Coauthors, 1992: A comprehensive meteorological modeling system—RAMS. *Meteor. Atmos. Phys.*, **49**, 69–91, https://doi.org/10.1007/BF01025401.

Prein, A. F., and Coauthors, 2015: A review on regional convection-permitting climate modeling: Demonstrations, prospects, and challenges. *Rev. Geophys.*, **53**, 323–361, https://doi.org/10.1002/2014RG000475.

Pytlak, E., M. Goering, and A. Bennett, 2005: Upper-tropospheric troughs and their interaction with the North American monsoon. *19th Conf. on Hydrology*, San Diego, CA, Amer. Meteor. Soc., JP2.3, https://ams.confex.com/ams/Annual2005/techprogram/paper_85393.htm.

Rockel, B., C. L. Castro, R. A. Pielke Sr., H. von Storch, and G. Leoncini, 2008: Dynamical downscaling: Assessment of model system dependent retained and added variability for two different regional climate models. *J. Geophys. Res.*, **113**, D21107, https://doi.org/10.1029/2007JD009461.

Rummler, T., J. Arnault, D. Gochis, and H. Kunstmann, 2019: Role of lateral terrestrial water flow on the regional water cycle in a complex terrain region: Investigation with a fully coupled model system. *J. Geophys. Res. Atmos.*, **124**, 507–529, https://doi.org/10.1029/2018JD029004.

Saha, S., and Coauthors, 2010: The NCEP Climate Forecast System Reanalysis. *Bull. Amer. Meteor. Soc.*, **91**, 1015–1057, https://doi.org/10.1175/2010BAMS3001.1.

——, and Coauthors, 2014: The NCEP Climate Forecast System version 2. *J. Climate*, **27**, 2185–2208, https://doi.org/10.1175/JCLI-D-12-00823.1.

Samaniego, L., R. Kumar, and S. Attinger, 2010: Multiscale parameter regionalization of a grid-based hydrologic model at the mesoscale. *Water Resour. Res.*, **46**, W05523, https://doi.org/10.1029/2008WR007327.

Santanello, J. A., C. D. Peters-Lidard, S. V. Kumar, C. Alonge, and W.-K. Tao, 2009: A modeling and observational framework for diagnosing local land–atmosphere coupling on diurnal time scales. *J. Hydrometeor.*, **10**, 577–599, https://doi.org/10.1175/2009JHM1066.1.

——, ——, A. Kennedy, and S. V. Kumar, 2013: Diagnosing the nature of land–atmosphere coupling: A case study of dry/wet extremes in the U.S. Southern Great Plains. *J. Hydrometeor.*, **14**, 3–24, https://doi.org/10.1175/JHM-D-12-023.1.

——, and Coauthors, 2018: Land–atmosphere interactions: The LoCo perspective. *Bull. Amer. Meteor. Soc.*, **99**, 1253–1272, https://doi.org/10.1175/BAMS-D-17-0001.1.

Seastrand, S., Y. Serra, C. L. Castro, and E. Ritchie, 2014: The dominant synoptic-scale modes of North American monsoon precipitation. *Int. J. Climatol.*, **35**, 2019–2032, https://doi.org/10.1002/joc.4104.

Senatore, A., G. Mendicino, D. J. Gochis, W. Yu, D. N. Yates, and H. Kunstmann, 2015: Fully coupled atmosphere-hydrology simulations for the central Mediterranean: Impact of enhanced hydrological parameterization for short and long time scales. *J. Adv. Model. Earth Syst.*, **7**, 1693–1715, https://doi.org/10.1002/2015MS000510.

Skamarock, W. C., and Coauthors, 2008: A description of the Advanced Research WRF version 3. NCAR Tech. Note NCAR/TN-475+STR, 113 pp., http://doi.org/10.5065/D68S4MVH.

von Storch, H., H. Langenberg, and F. Feser, 2000: A spectral nudging technique for dynamical downscaling purposes. *Mon. Wea. Rev.*, **128**, 3664–3673, https://doi.org/10.1175/1520-0493(2000)128<3664:ASNTFD>2.0.CO;2.

Xiang, T., E. R. Vivoni, and D. J. Gochis, 2018: Influence of initial soil moisture and vegetation conditions on monsoon precipitation events in northwest Mexico. *Atmósfera*, **31**, 25–45, https://doi.org/10.20937/ATM.2018.31.01.03.

Yang, Z.-L., X. Cai, G. Zhang, A. A. Tavakoly, Q. Jin, L. H. Meyer, and X. Guan, 2011: The Community Noah Land Surface Model with Multi-Parameterization Options (Noah-MP): Technical description. Tech. rep., Center for Integrated Earth System Science, The University of Texas at Austin, 72 pp., http://www.jsg.utexas.edu/noah-mp/files/Noah-MP_Technote_v0.2.pdf.

Zamora, R. J., E. P. Clark, E. Rogers, M. B. Ek, and T. M. Lahmers, 2014: An examination of meteorological and soil moisture conditions in the Babocomari River basin before the flood event of 2008. *J. Hydrometeor.*, **15**, 243–260, https://doi.org/10.1175/JHM-D-12-0142.1.

Zeng, X., M. Barlange, C. Castro, and K. Fling, 2010: Comparison of land–precipitation coupling strength using observations and models. *J. Hydrometeor.*, **11**, 979–994, https://doi.org/10.1175/2010JHM1226.1.