On the Rainfall and Temperature Forecast Skill for a Tropical Andean Mountain Area in Northern South America Using Different Operational Weather Forecast Strategies: Role of the Diurnal Cycle of Rainfall on the Success of Data Assimilation

Gisel Guzmán[1], Lina Isabel Ceballos[1], Juan Camilo Hernández[1], Mauricio Zapata Henao[1], Manuela Guarin Vargas[1], and Carlos David Hoyos Ortiz[2]

[1]SIATA
[2]SIATA, Universidad Nacional de Colombia

November 24, 2022

Abstract

Numerical Weather Prediction models (NWP) have been used extensively since the '40-'50s. Despite the advances in the field, the representation and forecast of the magnitude and variability of tropical processes in models is still a challenge. One of the steps to improve the precipitation forecasts using limited-area models is to evaluate which set of physical schemes and model domain configurations represent in a better way the actual behavior observed in the tropics. We implemented, as a part of a regional risk management strategy, two different operational weather forecast strategies for a complex terrain region in the Andes mountain range in northern South America. Both strategies, together, generate a total of eleven different forecasts every day, using the Weather Research and Forecasting model (WRF) with initial and boundary conditions from the Global Forecast System (GFS). The first configuration, implemented over five years ago and referred to as SYNAPSIS, includes three nested domains (18, 6 and 2 km) and is carried out every day using the 12 UTC GFS run and three different microphysics parametrizations: Eta Ferrier scheme, Purdue Lin Scheme and Thompson Scheme. The forecast lead-time of the latter strategy is 120 hours, and it does not use data assimilation. Since December of 2017, we implemented a second configuration termed RDFS, with two nested domains (12 and 2.4 Km), which carried out four times a day using the 00, 06, 12 and 18 UTC GFS runs. This configuration has a 30-hours lead time with the Thompson microphysics scheme. In RDFS, two WRF forecast runs are performed for each start hour, one assimilating weather radar reflectivity and the other without assimilation as control run, for a total of eight forecast runs daily. In this study, we assess the rainfall and temperature forecasts for all the different configurations using precipitation derived from reflectivity from weather radar, and air temperature at 2m from a network of automatic weather stations. We use 6 hourly and monthly skill scores (RMSE, BIAS, and Correlation coefficient) to quantify the precipitation differences between the SYNAPSIS and the RDFS configurations. To evaluate the impact of data assimilation in the precipitation forecast, we aggregate the results in a region within the inner domain, and then we calculate the average precipitation forecast between 0 and 36 predicted hours for RDFS with and without data assimilation. The results suggest a strong relationship between the forecast start time and the improve of precipitation forecast accuracy using data assimilation. The diurnal cycle of precipitation in the study region has a minimum in the morning (12 UTC) and a maximum in the afternoon (00 UTC) and during the night (09 UTC). The correspondence between the forecast improvement using data assimilation and the diurnal cycle of precipitation is likely due to the amount of assimilated data. In order to quantify the precipitation differences between the diffe

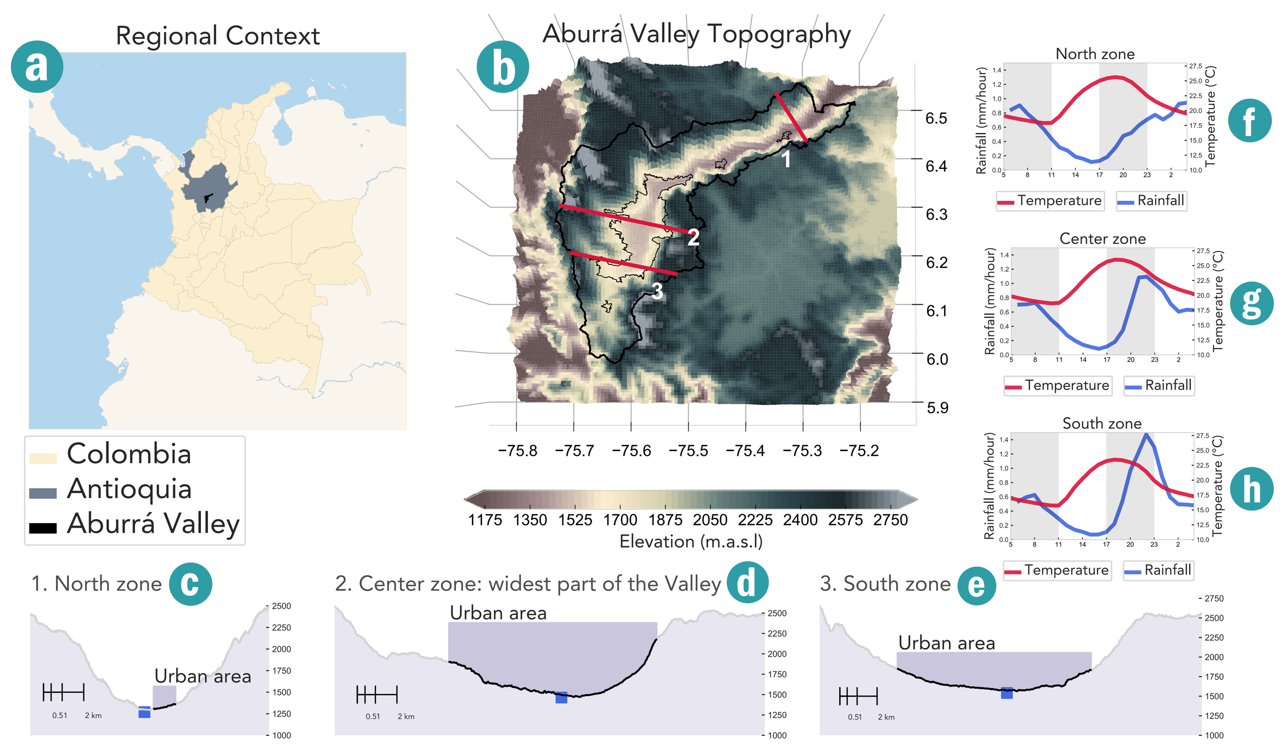
On the Rainfall and Temperature Forecast Skill for a Tropical Andean Mountainous Area in Northern South America Using Different Operational Weather Forecast Strategies:

Role of the Diurnal Cycle of Rainfall on the Success of Data Assimilation.

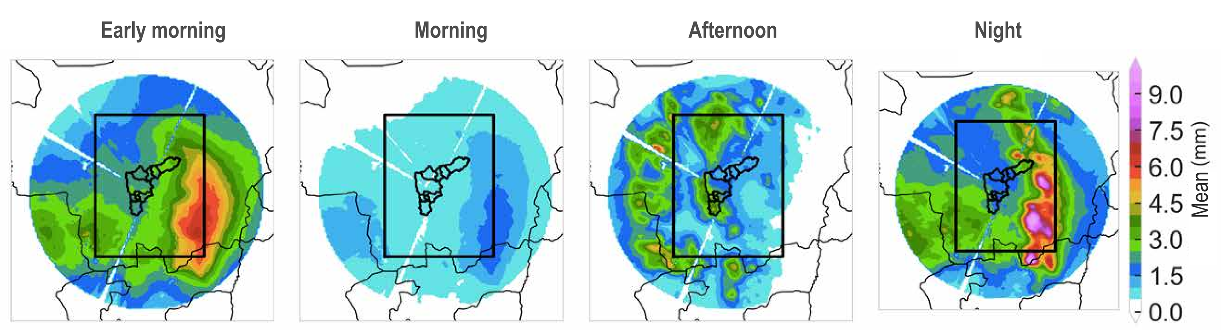
Mauricio Zapata [1,2], Gisel Guzmán-Echavarría [1], Carlos D. Hoyos [1,2], Juan C. Hernández [1], Lina I. Ceballos [1], Manuela Guarín [1]   1. Sistema de Alerta temprana de Medellín y el Valle de Aburrá   2. Universidad Nacional de Colombia, Sede Medellín.

Introduction

As a part of a regional risk management strategy, we designed two different operational weather forecast strategies for a complex terrain region in the Andes mountain range in northern South America. Both strategies, together, generate a total of eleven different forecasts every day, using the initial and boundary conditions from the GFS.



Figures (a) and (b) show the geographical context of the Aburra valley. Figures (c), (d) and (e), show a north, centre and south cross-section of the valley. Figure (f), (g) and (h) show the diurnal cycle of rainfall (Blue) and temperature (Red), obtained from the SIATA´s rain gauge and AWS network.



Diurnal cycle of precipitation inside of the study region, using the meteorological radar dBz retrievals.

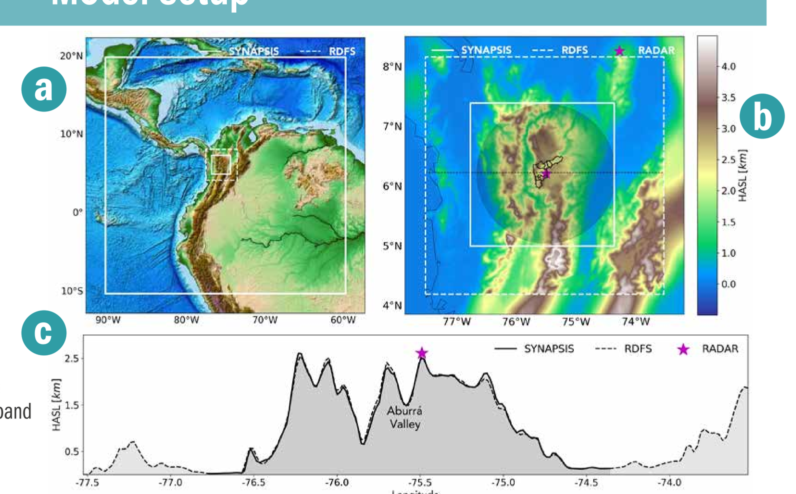
Model setup

SYNAPSIS

Medium-range forecast
Forecast lead time: 120 hours
Runs per day: 3 runs
Input data: GFS 0.5 12 UTC
Domain: Triple nested domain
Micro-physics: Purdue Lin scheme (PAR02), Eta Ferrier scheme (PAR05) and Thompson Scheme (PAR08)

RDFS

Short-range forecast
Forecast lead time: 30 hours
Runs per day: 8 runs (4 with and 4 without data assimilation)
Input data: GFS 00, 06, 12 and 18 UTC and Doppler radar C-band data for data assimilation
Domain: Double nested domain
Micro-physics: Thompson Scheme (PAR08)



Both configurations: Short-range (dashed white line) and Medium-range (solid white line), share the continental domain as can be seen in figure (a). Figure (b) shows the inner domain boundaries for both configurations. The shaded circle corresponds radar coverage. Figure (c) shows a cross-section at 6.2 north degrees. The magenta star in (b) and (c) shows the location of the radar.

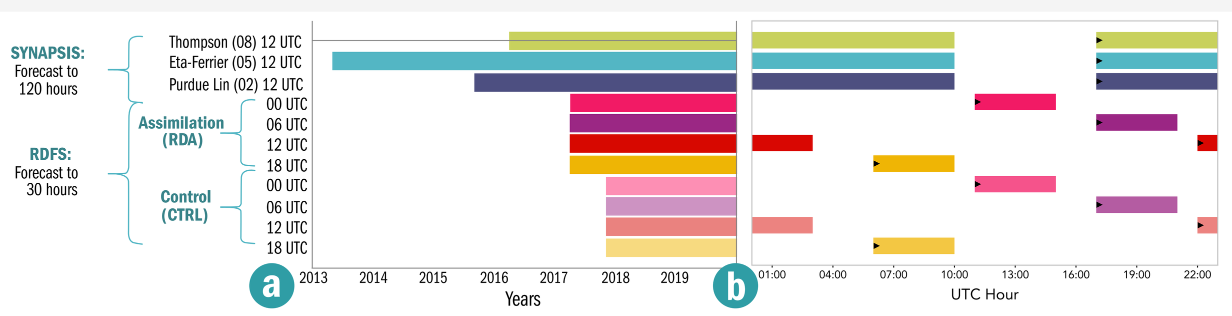
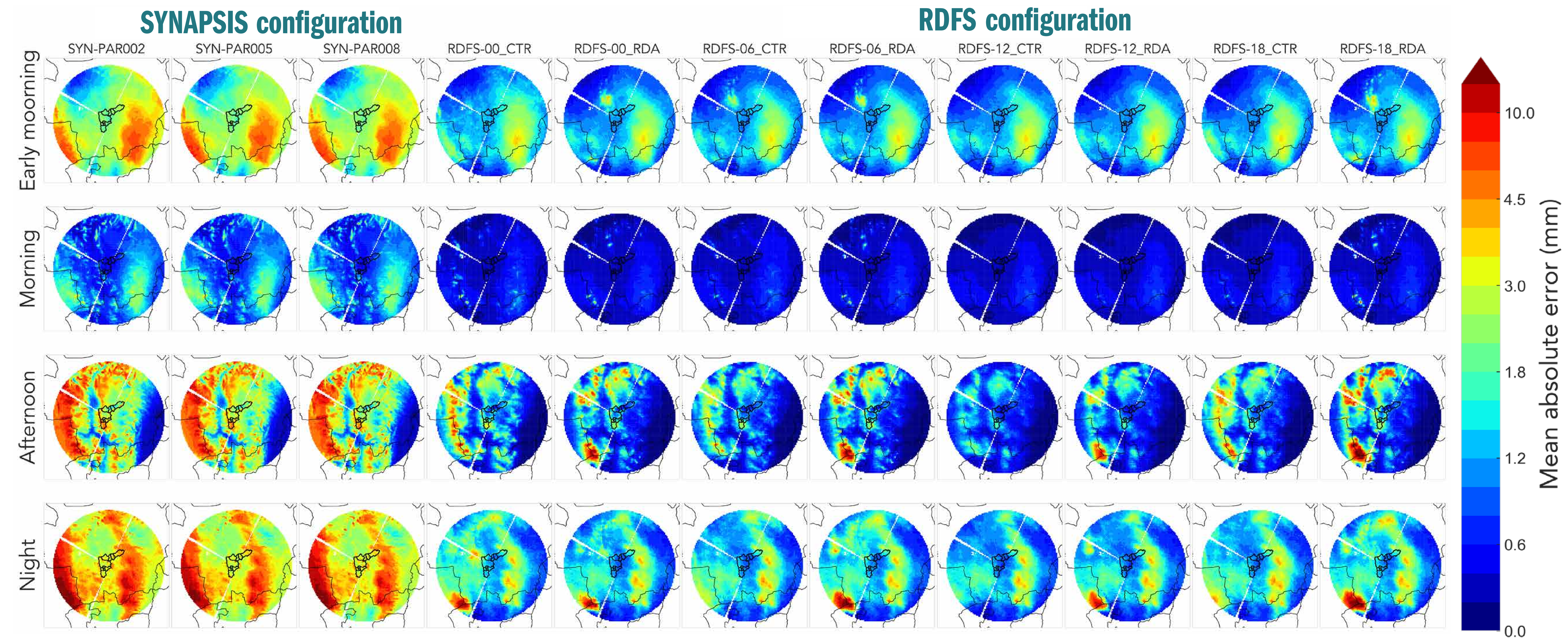


Figure (a) shows the operational forecast timeline for both configurations. Figure (b) shows the daily timeline of historical configurations (short and mid-range forecast). Short-range forecast or RDFS starts every day at eight different times, every six hours with (DA) and without data assimilation (CTRL). The mid-range forecast or SYNAPSIS starts every day with the 12 UTC boundary conditions.

Rainfall forecast verification

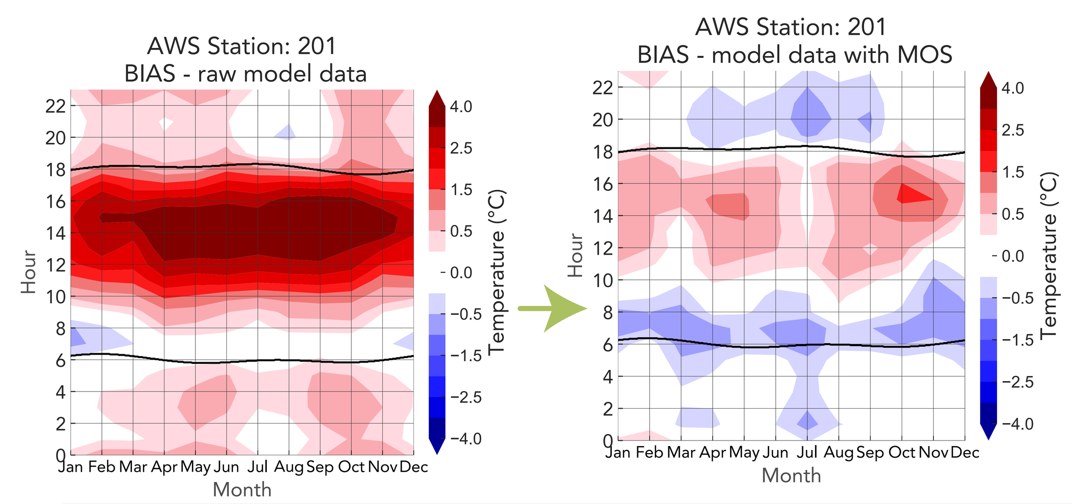
To determine the spatial differences between the rainfall forecast (SYNAPSIS and RDFS) and the rainfall observations (radar), we quantify the mean absolute error (MAE) for the cumulative rainfall during the early morning, morning, afternoon and night.



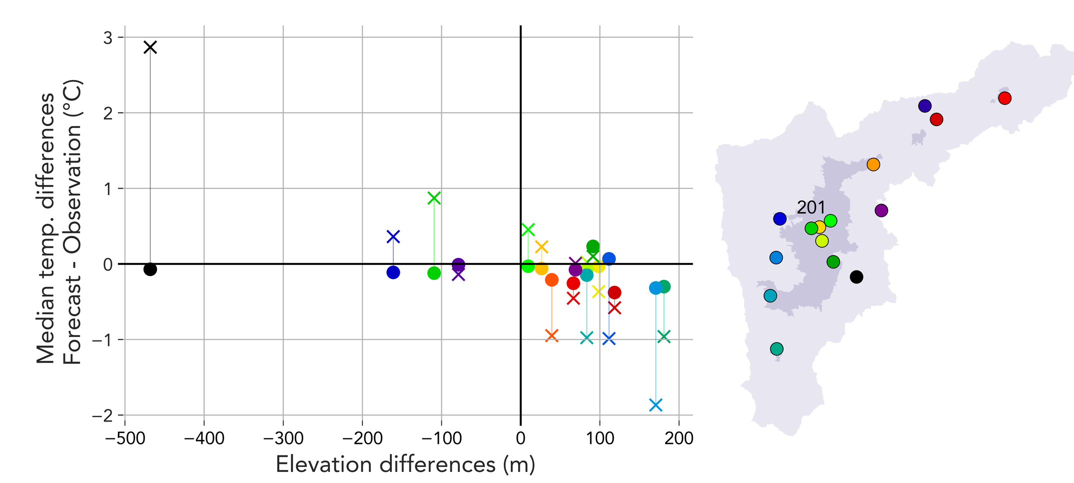
Spatial distribution of MAE. Each column corresponds to a forecast configuration. The first three columns corresponds to SYNAPSIS, RDFS CTRl and DA respectively for every initial hour (00, 06, 12 and 18 UTC). The rows represent the early morning, morning, afternoon and night.

Temperature forecast verification

There is a systematic error in the surface temperature forecast due to the difference between the real and model topography. This BIAS is corrected every day in the operational post-processing using a Model Output Statistics (MOS) scheme.



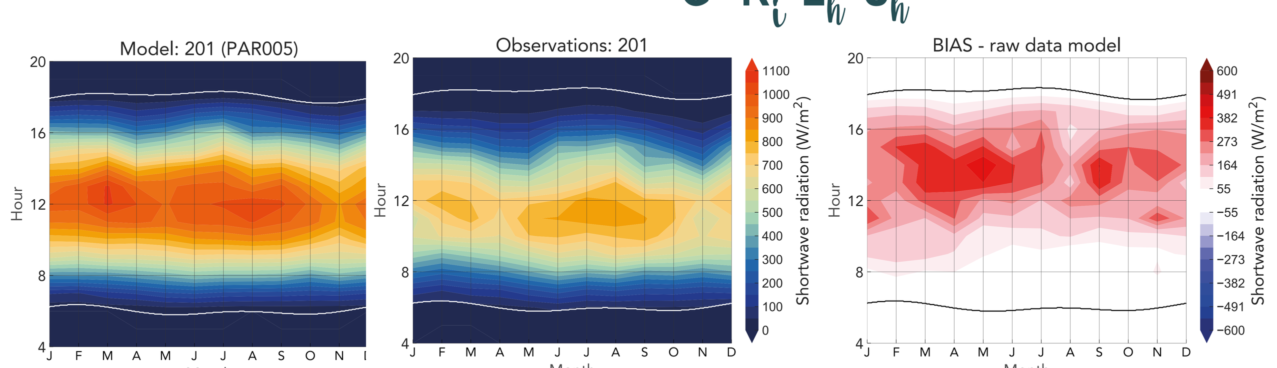
Monthly diurnal cycle of temperature differences between forecast and observations with and without MOS



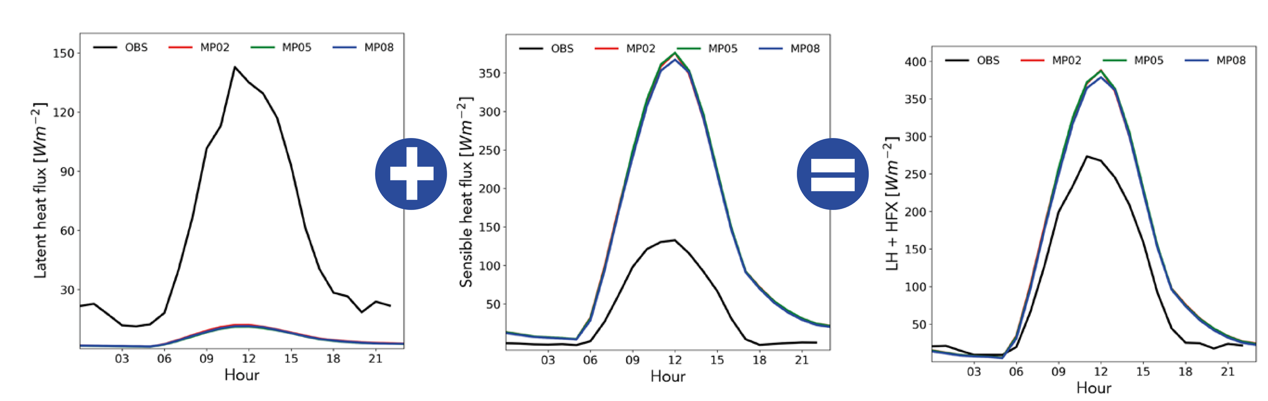
Elevation differences vs. temperature differences Crosses (points) indicate temperature differences before (after) MOS.

Surface Energy budget

$$G=R_i-L_h-S_h$$

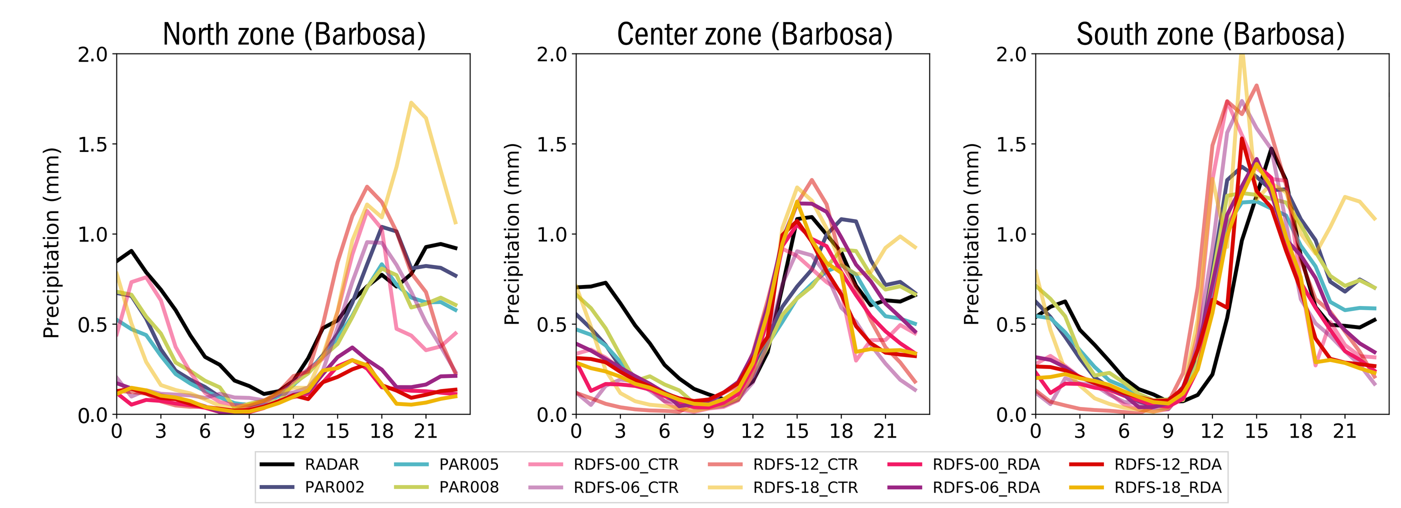


Observed, simulated and differences of surface radiation (land use = Urban)



Observed (CSAT3 Campbell) and simulated diurnal cycle of Latent and sensible heat fluxes (land use = Urban)

Diurnal cycle representation



Diurnal cycle of precipitation. Observations (radar) and the eleven WRF configurations in north, center and south of Aburrá Valley

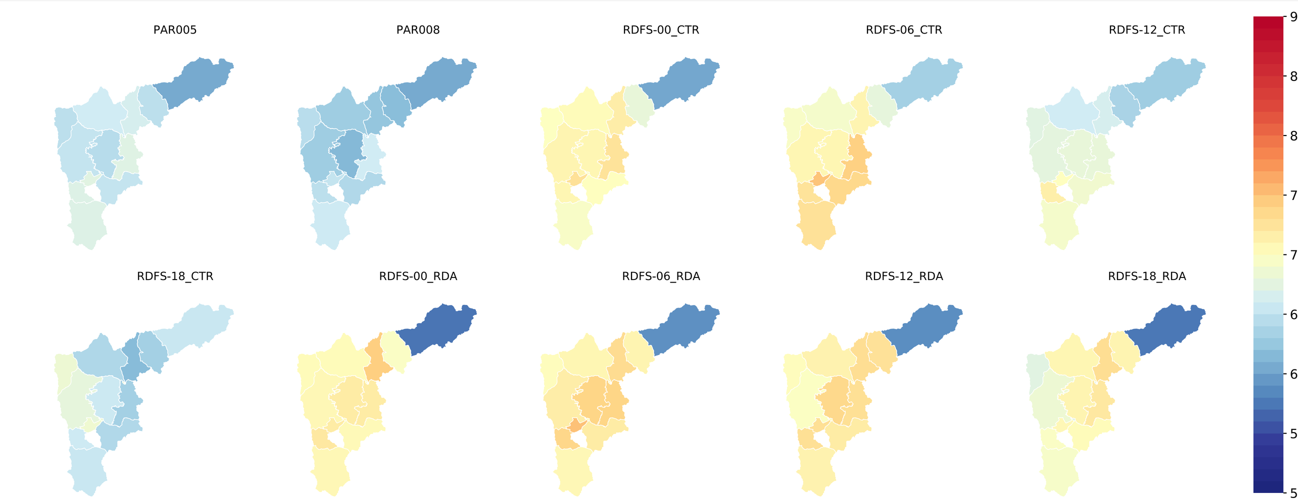
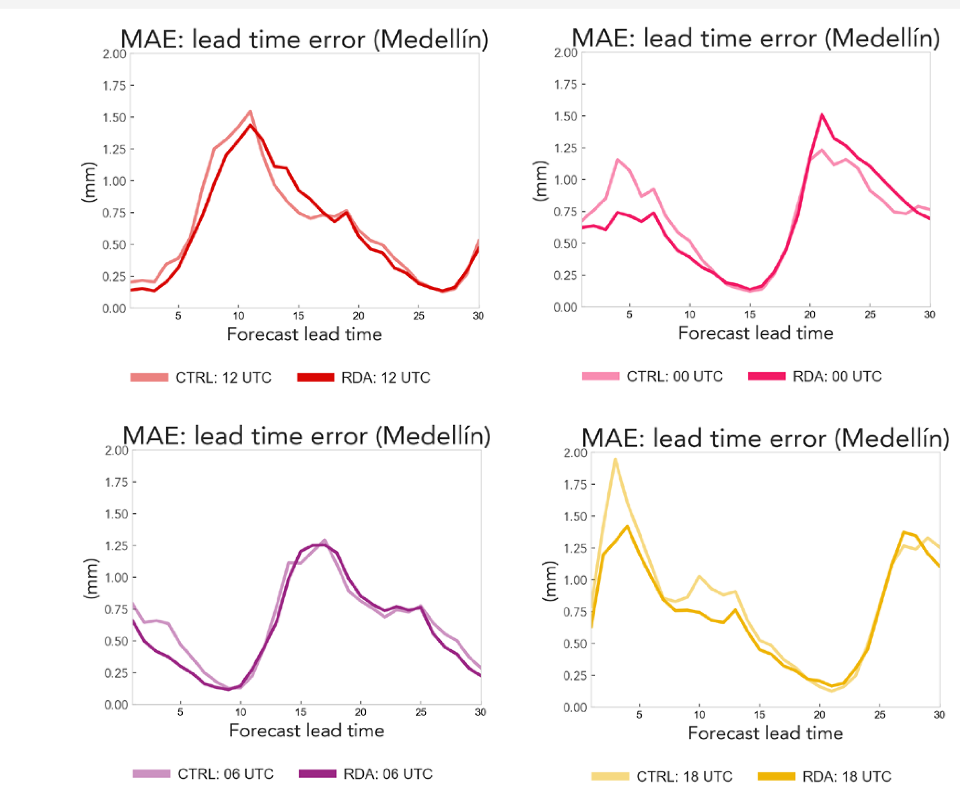


Figure shows the spatial distribution of hit rate at night for the Aburra Valley. The north zone (Barbosa) shows the worst skill, mainly in the data assimilation runs.

Effect of the data assimilation by lead time hour



Mean absolute Error (MAE), per hour in the forecast lead time on rdfs configuración. In the Center of Aburrá Valley. From left to the right 00, 06, 12, and 18 UTC.

Overall conclusions and future work

- At the inner domain, the best configuration is RDFS 12 UTC control run, but inside of Aburrá Valley, there are notable differences in the skill due to local differences in the representation of the rainfall diurnal cycle.
- Due to similarities in the performance between runs of SYNAPSIS configuration, regarding RDFS ones, we conclude that skill of rainfall forecast is highly dependent of atmospheric input data (also data assimilation) more than the choice of microphysics schema.
- The problems on the temperature representation are remarkable, due to biases in shortwave radiation and turbulent fluxes. However, temperature forecasts are accurate after a statistical correction (MOS).

Acknoledgments

This work was supported by Area Metropolitana de Medellín y del Valle de Aburrá, Municipio de Medellín, Grupo EPM, and ISAGEN under the contract CCT504 of 2019.

Alcaldía de Medellín

Contact
Gisel Guzmán gguzmane@unal.edu.co
Mauricio Zapata mzapatahe@unal.edu.co
Manuela Guarín Vargas mguarinv@unal.edu.co