Investigation of the Influence of Initial Conditions and PBL Parameterization on the Practical Predictability of Convection Initiation

Thesis Proposal

By: Bryan Burlingame (burling6@uwm.edu)

“Under the supervision of Professors Clark Evans, and Paul J. Roebber”
ABSTRACT

Convection Initiation refers to the formation of deep moist convection and is triggered by physical processes in the atmosphere. Circulations along air mass boundaries, orographically forced motions, and ducted gravity waves are amongst triggering processes that can lift parcels above their level of free convection where through buoyant motion they can rise to great heights in the atmosphere. Accurate prediction of convection initiation is a forecast problem that spans across the synoptic-, meso-, and micro- scales and is highly sensitive to the interactions between the surface and the planetary boundary layer. The varying scales of processes involved in the initiation of deep moist convection makes the forecast of it a formidable task. The Mesoscale Predictability Experiment (MPEX) aimed to test the hypothesis that the collection of additional synoptic- and mesoscale observations in the intermountain west and their assimilation into convective allowing models would significantly improve the forecasts of the timing and location of convection initiation, convective mode, and downstream evolution of deep moist convection. Utilizing these gathered observations and subsequently assimilating them into thirty-member initial analyses of the pre-convective atmospheric state, we seek to illuminate how convection initiations practical predictability is controlled by planetary boundary layer parameterization in convective allowing models.

Motivation

The predictability of weather phenomena decreases with decreasing scale, as alluded to by Lilly (1990). To explain further, as the scales decrease the temporal and spatial resolution requirements to resolve these phenomena increases. This also suggests that the processes that directly affect these phenomena are more influential at small scales given the shorter time frame in which they occur. The predictability of synoptic scale disturbances extends out to a few days whereas the predictability of sub-synoptic disturbances is on the order of 12-24 hours, and down to only a few hours for convective storms as summarized by Weismann et al. (2015). This, coupled with the notion that the forecasting skill of convection initiation (CI) has advanced at a slower rate than our ability to anticipate convective storm type, organization, and severe weather threats (Markowski and Richardson 2010), leads to us to the desire to better understand limits upon the current predictability of CI so that the forecast can be improved upon. Also, as noted by Baxter (2011), and Zhang et al. (2002 & 2003), situations involving moist convection are associated with a larger amount of intrinsic error and are inherently less predictable as a result of upscale error growth. This suggests that error initially at the smaller scales will continuously propagate upward through the larger scales. Additionally, the micro- and mesoscale processes that influence moist convection occur on shorter temporal and spatial scales increasing complexity. This further increases our motivation to reach the goal of quantifying the predictability of the forecast of CI. Moreover, we aim to understand the predictability of CI in a “pristine” environment, where pristine CI suggests a point of initiation that occurs in an environment that has not been modified by any other convection.

Methodology

The 2013 severe storm season was monitored by MPEX from 15 May to 15 June 2013. The observations were gathered using the NCAR/NSF GV Aircraft that launched 28-32 dropsondes in the morning prior to 15 different events in which DMC was forecasted across the central US. Additionally, 3-4 ground units followed convective storms in real time to gather field observations in the environments surrounding convective storms. For this study, we utilize thirty slightly different initial states of the
atmosphere to initialize ensemble forecasts for each event sampled by MPEX. These ensemble initial conditions are obtained by randomly perturbing a first-guess atmospheric analysis from the GFS model at the start of a cycled data assimilation process that began prior to MPEX and continued through the duration of the field program. These perturbations result in forecast variability (error growth) over 6-h periods as constrained by assimilating conventional data; note that the cycling interval shortens to 1-h at 0000 UTC on the day of an MPEX flight mission, and MPEX dropsondes are assimilated over the period 0900 to 1500 UTC on these days. Lateral boundary conditions are obtained by randomly perturbing GFS forecasts following the fixed covariance perturbation method Torn et al. (2006).

The ensembles will be run using the Weather Research and Forecasting (WRF) Advanced Research WRF (ARW) model version 3.4.1. The simulations will take place on the University Corporation for Atmospheric Research’s (UCAR) supercomputer, Yellowstone, which has 72,576 computational nodes and 144.6 TB of disk space. Additional to the standard configuration of WRF-ARW v3.4.1, we will be incorporating source code modifications generated by the National Severe Storms Laboratory (NSSL) to aid in calculating reflectivity derived to the -10 °C level in-line with the model code. To achieve the goal of identifying the sensitivity of the planetary boundary layer (PBL) parameterization on the forecast of CI, our thirty member ensemble complete with the MPEX observations will be run five times, with each ensemble varying only the PBL scheme and associated surface layer scheme. The five different PBL parameterizations are as follows: Mellor-Yamada-Janic (MYJ; Janic 1994), Yonsei University (YSU; Hong and Pan 1996), Quasi-Normal Scale Elimination (QNSE; Sukoriansky et al. 2005), Asymmetric Cloud Model version 2 (ACM2; Pleim 2007), and the Mellor-Yamada-Nakanishi-Niino 2.5 (MYNN; Nakanishi and Niino 2009) schemes. The remainder of the model will be configured to replicate the WRF model used

Figure 1. Domain Configuration
during the MPEX experiment (Weisman et al. 2015). That will include a two-way nested domain configuration with domain 1 encompassing the entire CONUS, and domain 2 covering the central 2/3 of the US (1046x871 gridpoints) as seen in Figure 1. Thompson’s double moment microphysical scheme (Thompson et al. 2008) will handle microphysical parameterizations, Goddard shortwave and the Rapid Radiative Transfer Model (RRTMG) longwave will control the radiation, and the NOAH land surface model will be used. Domain 1 will use Tiedtke Cumulus parameterization whereas domain 2 will resolve convection explicitly. This will be done for three events, all of which coincide with one of the fifteen research flights (RFs) that the MPEX mission conducted. The three missions will be selected based on the large scale flow regime of each event. Each regime will be subjectively chosen based on the structure of the 500mb height field, the main trigger(s) of convection, and other synoptic- and meso- scale features. The idea is to select three different large scale patterns to include variability in our examination, one chosen for example with a deep trough over the intermountain west (RF4; 19 May 2013), another with zonal flow over the central US, and another that varies from those two.

Upon completion of our simulations, we aim to identify the predictability of CI in a convectively undisturbed environment (“pristine” CI). To determine this, “pristine” CI will be identified in the observed data, and then compared to modeled CI in spatial and temporal bins to create our verification. To obtain our observed dataset, level II data from 42 radars in the Central United States were obtained from the National Climatic Data Center from 15z to 06z on the days MPEX conducted a RF (Figure 2).

These data were merged onto a uniform 0.03°x0.03° grid and then interpolated to the -10 °C level extracted from the 0 hour RAP analysis using the Warning Decision Support System – Integrated Information (WDSS-II) (Lakshmanan et al. 2007). Also using WDSS-II, CI events were identified in both the observed and modeled data using the available tracking method, w2segmotionll (Lakshmanan and Smith 2010). This was done with the definition of CI from Gremillion and Orville, 1999, which found that defining CI as reflectivity >35 dBz at the -10 °C isotherm for at least 30 minutes provided a high probability of detection and low false alarm rate for inferring the initiation and presence of a thunderstorm. These thresholds were selected to ensure that initiation occurred and was sustained, eliminating random vertical bursts that were falsely representing CI, and also to avoid the effects of bright-banding that can be seen in the reflectivity field at and near the melting level due to melting hydrometeors. Data regarding the timing and location of every CI event will afford us the opportunity to
illuminate skill in the forecast via forecast verification, statistical significance testing, and synoptic ensemble analysis.

There are many foreseeable challenges to this research, the first being identifying a verification method that can identify skill in the probabilistic forecast of rare events. CI in a pristine environment is rare in the sense that once the first CI point has initialized, the environment is no longer convectively undisturbed. Another challenge, moreover a limitation, is the number of cases we aim to examine (3), and also the number of cases we have to choose from (15). The small sample size may skew our final results. However, the conclusion of this research will provide insight into the predictability of first CI, the role that the PBL scheme has in the forecast, and will hopefully enhance the overall forecast skill.

**Hypothesis/Preliminary Findings**

At this preliminary stage, we have selected RF 4 (19 May 2013) as our first case to analyze for PBL sensitivity and the predictability of “pristine” CI. This case was chosen because it was associated with one of three large-scale flow regimes we wish to consider, with a deep trough moving over the TX panhandle, and southwesterly flow across the central and southern plains, and because it was a high-impact severe weather event. For this event, the five-thirty member ensembles described above were obtained, and CI events were identified in the same manner as the observed by linearly interpolating five minute reflectivity at the -10C to the same 0.03°x0.03° grid described above. From here, temporal and spatial distributions of the CI events were computed to better understand where and when the ensembles produced convection compared to observations. We found that the models significantly over produced the amount of CI events, as seen in Figure 3. However, the timing of CI events is in line with

![Figure 3. Temporal distribution of CI events.](image-url)
observations, ramping up around 17z and increasing steadily until around 1z. Spatial distributions show that there are times and locations in which overproduction was most prevalent. For example, between 18-21z in northeast CO and western NE, there was significant production by the ensemble members where in actuality there were very few CI events.

In terms of verifying the forecast and understanding its implications upon predictability, we have been working towards identifying an objective means of quantifying the predictability of rare forecast events. As mentioned above, we are most interested in the first, or “pristine” CI event. Separately, the over-production identified in our preliminary analysis has resulted in a desire to understand why it is occurring. Some hypotheses that have been discussed involve the amount of soil moisture in the NOAH land surface model; if there is too much soil moisture, the latent heat flux into the boundary layer may be too great, potentially better enabling thermals to become positively buoyant too easily and thus resulting in CI. Also, a lack of sufficient entrainment of drier environmental air surrounding the updraft could allow parcels to rise to heights well above what would otherwise be expected given the local environment, potentially allowing for more sustained CI events than in the real atmosphere. There is also the possibility of double counting because of the fine, 3km grid spacing we are using. This can occur because the model at this resolution is able to explicitly resolve the largest of turbulent eddies that are found in the PBL. However, the PBL parameterizations developed are designed for coarser grid spacings and thus also account for these larger turbulent eddies through parameterization. This could result in artificially-inflated values of turbulent kinetic energy, overly intense simulated vertical circulations, and, potentially, a greater likelihood for CI events to occur in the ascending branch of such circulations.


