ENSEMBLE VARIABILITY IN RAINFALL FORECASTS OF HURRICANE IRENE (2011)

by

Molly B. Smith

A Thesis

Submitted to the University at Albany, State University of New York

in Partial Fulfillment of

the Requirements for the Degree of

Master of Science

College of Arts & Sciences

Department of Atmospheric & Environmental Sciences

2017

Abstract

As tropical cyclones (TCs) move into the midlatitudes, they are often associated with extensive heavy precipitation. This precipitation can lead to widespread flooding events, such as occurred with Hurricane Irene (2011) over the northeastern United States. Despite the high-impact nature of these events, there are relatively few studies that explore the sensitivity of precipitation forecasts to model initial conditions, instead focusing on the variability in TC track.

The goal of this work is to understand what modulates precipitation forecasts over the northeastern United States during Hurricane Irene. This is investigated using the Global Forecasting System (GFS) ensemble prediction system, initialized at 0000 UTC 27 August 2011. The ensemble members that forecast the largest precipitation totals (i.e., wet members) over the Catskill Mountains of New York (where over 15" of rain were observed) are then compared to the members that predicted the least precipitation (i.e., dry members), to diagnose the processes that lead to the rainfall differences. Results indicate that the amount of rainfall is tied to storm track, with wetter members clustered on the western side of the track envelope, and drier members on the east. Variability in storm track is associated with variability in each member's potential vorticity field at model initialization.

The 0.5° GFS ensemble is then downscaled to 3 km using the Weather Research and Forecasting (WRF) model, in order to better simulate the effects of mesoscale processes and terrain on precipitation distribution. This ensemble features a more complex relationship between precipitation and storm track than a simple east–west correlation. The wettest members are characterized by lower-tropospheric winds that are directed perpendicular to the eastern face of the Catskills (where the heaviest rainfall occurred), allowing maximum upslope forcing during the period of highest rainfall rates, as well as by greater horizontal moisture flux convergence. It

is interesting to note, however, that the drier eastern members have the greatest synoptic forcing for ascent in the Catskills. Their lower rainfall totals may be due to the eastern members positioning the bulk of the available moisture too far east for that forcing to produce large rainfall amounts.

1. Introduction

1.1 Precipitation in landfalling tropical cyclones

As tropical cyclones (TCs) move from the tropics into the midlatitudes and interact with synoptic weather systems, they are often associated with extensive regions of heavy precipitation (e.g., DiMego and Bosart 1982a; Atallah et al. 2007), which can lead to widespread, devastating flood events. Chen et al. (2010) sought to uncover the dynamics driving these flood events by performing a comprehensive analysis of precipitation characteristics in landfalling TCs, and found that as a TC's regular spiral rainband structure is disrupted during landfall, other factors determine the distribution of precipitation over land. On the synoptic scale, moisture availability was determined to be the dominant control of rainfall amount, with landfalling TCs producing much greater amounts of rain when there was strong water vapor transport into the region. Likewise, the process of extratropical transition (ET; discussed in more detail in section 1.2) alters the distribution of TC precipitation by introducing cold air (and thus asymmetric baroclinicity) into the system (Klein et al. 2000). The highest rainfall rates are found in the vicinity of this baroclinicity (Bosart and Dean 1991), where frontogenesis (and thus synoptic forcing for ascent) occurs, although if the cold air manages to permeate the center of the transitioning TC, rainfall rates for the entire system can decrease (Ritchie and Elsberry 2007).

Topography and mesoscale processes also play a large role in the horizontal distribution of precipitation in landfalling TCs (Chen et al. 2010). The upslope flow generated by a TC's winds blowing orthogonal to a topography gradient provides enough ascent to produce locally elevated rainfall totals (e.g., Lin et al. 2001; Wu et al. 2002; Lin et al. 2011). In addition, mesoscale vortices can form within the larger TC circulation (Harr and Elsberry 1996), again leading to locally increased ascent and higher rainfall amounts (Wu and Kuo 1999). Therefore,

precipitation forecasting for landfalling tropical cyclones can be difficult, as it is necessary to consider both the synoptic aspects of the flow, such as moisture transport and structural changes within the TC circulation, and mesoscale processes, such as terrain interactions and embedded vortices, to make an accurate prediction.

1.2 Extratropical transition

As discussed in the previous section, extratropical transition is one of the major influences on the horizontal precipitation distributions of many landfalling TCs, as the heaviest rainfall rates in a transitioning cyclone tend to emerge in regions of maximum baroclinicity (Bosart and Dean 1991). Hart and Evans (2001) estimated that half of all TCs that make landfall on the east coast of the United States and Canada undergo ET, so an understanding of ET is necessary in order to create accurate precipitation forecasts for landfalls in this region. ET occurs when a decaying TC becomes baroclinic in nature and reintensifies into a cold-core extratropical cyclone (Jones et al. 2003). The definition of ET presented in Klein et al. (2000) includes a transformation stage, in which a TC transitions into an extratropical cyclone, and a reintensification stage, in which that extratropical cyclone either strengthens or dissipates. The transformation stage itself is further divided into three steps. During step one, the TC begins to move into a baroclinic zone, placing the TC near a supply of cold, dry air to its northwest. This cold air begins to be entrained into the TC, decreasing convection and precipitation on its western side. In step two, the TC proceeds even farther into the baroclinic zone. The cyclonic circulation of the TC, coupled with the cold air to the north and the warm air to the south, gives rise to substantial cold air advection (CAA) on its western side, and warm air advection (WAA) to its east. The WAA is associated with rising air on the eastern side of the TC, which flows into the jet and creates a large asymmetric cloud shield aloft. The WAA also transports a substantial

amount of moisture into the system, much as a warm conveyor belt does in an extratropical cyclone (Eckhardt et al. 2004), increasing precipitation rates to the northeast of the storm center (Atallah and Bosart 2003). The CAA helps to further entrain dry air to the west, which ultimately ends the convection in the storm core. In step three, the eyewall dissipates, as do the last remnants of the TC warm core. The new dominant circulation is defined by the rising air occurring within the WAA to the east of the low, and the sinking air occurring within the CAA to the west. The precipitation structure within the storm now strongly resembles that of an extratropical cyclone, with the highest rainfall rates located in the regions of maximum baroclinicity, and the system's transition to a cold core low is complete.

As the process of ET is underway, the transitioning TC's outflow can become confluent with the midlatitude jet, which can be associated with a substantial increase of the jet's kinetic energy (Palmén 1958; DiMego and Bosart 1982b). This interaction is associated with an increasingly meridional flow pattern downstream of the TC (Archambault et al. 2013). Through the use of idealized simulations, Riemer et al. (2008) identified the primary mechanism by which TCs induce Rossby wave amplification in the extratropical flow: advection of potential vorticity (PV) by the divergent outflow of the TC. As a TC undergoes ET, its cyclonic circulation will advect low-PV air from the tropics towards the midlatitude PV gradient. The low-PV air impinges on the jet, bowing it northwards and leading to ridge amplification. This bowing also increases the gradient of PV in the baroclinic zone, leading to the strengthening of its associated winds and the formation of a jet streak. The exit region of this jet streak can provide a favorable environment for extratropical cyclogenesis, allowing a new cold-core cyclone to form. The cyclonic flow around this new low can further amplify the downstream ridge (Riemer et al. 2008).

Davis et al. (2008) found that as a TC becomes extratropical, its outflow at upper levels will increase rapidly, which speeds the advection of low-PV air and amplifies the downstream response. Furthermore, Archambault et al. (2013) tested the importance of three main factors that modulate the extratropical flow response: the characteristics of the large-scale flow pattern, the characteristics of the TC, and the strength of the interaction between the TC and the extratropical flow. The sensitivity of the downstream response to the characteristics of either the large-scale flow pattern or the TC was found to be small, while the strength of the interaction between the TC and the extratropical flow was recognized as the dominant control on the magnitude of downstream impacts.

Archambault et al. (2015) builds on the work of Archambault et al. (2013) by performing an in-depth analysis of the dynamics behind strong and weak interactions between TCs and extratropical flow. Cases with strong interactions (ETs that feature greater advection of PV by the TC's irrotational outflow) produce greater ridge amplification downstream of the TC, leading to the formation of a strong high-latitude ridge that can persist for several days. Cases with weak interactions, however, feature the excitation of a weak wave train that soon decays. Strong cases are able to produce the large amount of divergent outflow needed to maintain an amplified wave train because they tend to occur in environments that feature greater midlevel forcing for ascent and frontogenesis. These environments also tend to have a greater moisture supply available for latent heat release, which additionally strengthens a TC's divergent outflow.

1.3 Numerical weather prediction

Although ET plays a large role in determining the precipitation distribution of landfalling TCs, there are as a variety of synoptic features which also modulate rainfall. These additional features increase the difficulty of creating accurate landfall precipitation forecasts, leading

forecasters to numerical weather models for guidance. Global models, such as the National Centers for Environmental Prediction's (NCEP) Global Forecast System (GFS) and the European Centre for Medium-Range Weather Forecasts' (ECMWF) model, use a global grid to simulate atmospheric dynamics, incorporating numerical solutions for dynamical equations and parameterizing smaller-scale processes (Krishnamurti 1995). Regional models, such as the North American Mesoscale Forecast System and the Weather Research and Forecasting (WRF) model, have more localized domains, but use a higher resolution, are more easily customizable, and present the opportunity to dynamically calculate precipitation (Michalakes et al. 2001). Both types of models assimilate observational data in an attempt to improve accuracy (Rabier 2005). However, a single, deterministic model simulation is prone to increasing error with time, due to error introduced both by a lack of continuous initial observations and by assumptions made within the model physics, which grows with time due to the chaotic nature of the atmosphere (Lorenz 1963).

Probabilistic weather forecasting has been increasingly emphasized in recent years as a way to determine the range of solutions and account for the chaotic nature of the atmosphere. Many authors have advocated the use of probabilistic forecasts as a way to account for some of the model error arising from initial condition and model physics inaccuracies (e.g., Epstein 1969; Gleeson 1970), often focusing on model ensembles as the best way to generate these probabilistic forecasts (e.g., Krishnamurti et al. 2000). Ensembles attempt to replicate a realistic probability density function through individual forecasts, with the assumption that the observed field will fall somewhere within the distribution of the members.

Multiple methods have been employed over the years to deduce initial condition uncertainty. The ECMWF Ensemble Prediction System (EPS) uses singular vectors, representing

the most unstable phase space directions early on in the simulation, linearized around a short-term forecast trajectory to dynamically add perturbations to an initial analysis (Molteni et al. 1996), generating multiple ensemble members. The United States' National Meteorological Center created an early operational model ensemble that used the breeding of growth nodes method, in which perturbations were added to a higher-resolution control simulation by computing the difference between two nonlinear forecasts, and applying the uncertainty implied by that difference to the control (Tracton and Kalnay 1993; Toth and Kalnay 1997), thereby producing a number of ensemble members, although this method has been replaced in the modern GFS.

The Ensemble Kalman Filter (EnKF; Whitaker and Hamill 2002) is another popular data assimilation technique used to generate initial condition perturbations. The EnKF uses a Monte Carlo statistical method to provide each ensemble member with unique initial conditions based on available observations. The EnKF is employed by initializing and running a short-term ensemble for a given period of time. Then, observations are assimilated to create a new ensemble analysis, and the short-term ensemble is re-initialized with this analysis and run for another time step. This cycle of short-term data assimilation can continue for any length of time, and then the ensemble can be run forward from the last analysis to produce a forecast (Evensen 2003). The current operational version of the GFS uses another type of Kalman filter called the ensemble transform Kalman filter (ETKF), which uses ensemble transformation to more quickly assimilate data (Bishop et al. 2001).

Statistical data assimilation methods such as the EnKF are mainly intended to solve the problem of incomplete observational data and its resulting error. However, models also gain error by relying on physics parameterizations, which save computational power and replicate

physical processes that occur on scales that models can't resolve. This parameterization error can become especially pronounced in the case of a complicated forecast scenario such as ET (Davis et al. 2008). For example, Bassill (2014) examined the differences in forecast track for Hurricane Sandy (2012) predicted by the GFS and ECMWF. This study found that, although the ECMWF was able to accurately predict the track that Sandy would take days before the GFS, the GFS could produce a correct track with similar lead time as the ECMWF when it was run with the same cumulus parameterization as was utilized in the ECMWF. This cumulus parameterization was better able to simulate the increase in vertical mass flux typically observed in TCs undergoing ET, producing a more accurate representation of divergent outflow aloft and advection of low-PV air into an upstream trough. Because of this accuracy, the ECMWF predicted that the trough would slow and steer Sandy to the west, correctly placing its precipitation swath over the Northeast.

Although physics parameterizations can introduce error into forecasts, this work will focus more on the effects that initial condition errors in particular can have on rainfall forecasts by using Hurricane Irene (2011) as a case study. The analysis will examine Irene in terms of precipitation variability between members of an ensemble forecast, and will test whether some of the variability in rainfall is due to differences in the way that the various members simulate the TC–jet interactions that occur during ET. Even though the members all use the same physics, small differences in the orientation and intensity of Irene, and the midlatitude jet, can be magnified by the process of ET, eventually producing much different forecasts.

1.4 Synoptic history of Hurricane Irene (2011)

Hurricane Irene formed from a tropical wave that moved off the coast of Africa on 15 August 2011. Although the disturbance was associated with relatively strong convection, it took

several days for a closed circulation to develop, preventing the system from being classified as a tropical depression. Irene was finally named as a tropical storm late on 20 August 2011 while located to the east of Martinique, and began to move west-northwest through the Caribbean Sea. It reached hurricane strength on 22 August, while located just off the coast of eastern Puerto Rico, causing extensive flood damage in that region before moving northwestward past Hispaniola and strengthening to a category 3 hurricane. On 24 August, Irene turned northward towards the Bahamas and mid-Atlantic states. After passing over the Bahamas as a category 2 storm, Irene made landfall near Cape Lookout, North Carolina at 1200 UTC 27 August as a category 1 hurricane, before re-emerging over the Atlantic and continuing on a northward track. The hurricane's final landfall as a category 1 hurricane occurred just east of New York City shortly after 1200 UTC 28 August, and Irene was declared extratropical at 0000 UTC 29 August (Avila and Cangialosi 2011; Figure 1).

As Irene moved north from the tropics to the midlatitudes, the storm's motion was strongly influenced by the larger-scale synoptic pattern of the atmosphere, most notably by two separate troughs situated over the continental United States. As troughs deepen into the midlatitudes and subtropics, they bring colder, higher-PV air equatorward (Hoskins et al. 1985). High-PV perturbations in the atmosphere are associated with cyclonic winds (Davis 1992), so the introduction of a high-PV trough into the vicinity of a TC can substantially alter the environmental steering flow acting upon that cyclone (e.g., Galarneau and Davis 2013). On 24 August, as Irene moved over the southern Bahamas, it began to interact with a weak cyclonic upper-level PV perturbation over the southeastern United States and Gulf of Mexico (Figure 2a). Because this trough was situated to the west of Irene, the cyclonic circulation around it placed the TC in a region of southerly flow, turning Irene north towards the United States (Figure 2b–d).

By the time Irene reached the Carolinas, a second, stronger trough had formed over the central and eastern United States, once again placing the TC in a region of southerly steering flow, sending the storm farther north towards New York and the northeastern United States (Figure 3a–e). This trough is discussed further in section 2.2.

Hurricane–trough interactions such as this have been responsible for some very large precipitation totals in landfalling tropical cyclones. In October 1954, for example, Hurricane Hazel was situated off of the coast of Florida. Thanks to a strong trough situated over the central and eastern United States, Hazel accelerated to the north-northwest and made landfall in the Carolinas (Palmén 1958), interacting with a pre-existing frontal boundary and depositing nearly a foot of rain across the region (Collins 1954). In 1972, Hurricane Agnes re-emerged over the Atlantic, off the coast of Virginia, on a trajectory out to sea. However, a Canadian trough developed a strong negative tilt over the Great Lakes region, placing Agnes in southeasterly environmental steering flow and turning it northwestward into upstate New York and Pennsylvania, where it was associated with widespread rainfall totals of 8–16 in (DiMego and Bosart 1982a). Hurricane Floyd (1999) had a storm track remarkably similar to Irene's, and a very similar rainfall distribution (although Floyd was shifted slightly to the east), again the result of a deep trough situated over the central and eastern United States (Atallah and Bosart 2003). More recently, Hurricane Sandy (2012) was also steered inland due to a strong negatively-tilted trough over the central United States, leading to severe coastal flooding across the Northeast (Blake et al. 2013).

Like these other TCs that interacted with upstream troughs, Irene was responsible for substantial flooding and damage across the Northeast. As the storm moved inland, it deposited widespread rainfall totals of 4–7 in, with locally higher amounts. The Catskill region of New

York in particular received very large amounts of precipitation, with some locations seeing up to a foot of rain in less than 24 h. Irene directly caused 41 deaths and $15.8 billion in damages in the United States, with three towns in the Catskills deemed uninhabitable due to flooding after Irene had passed (Avila and Cangialosi 2011).

As Irene approached the eastern United States, the National Hurricane Center (NHC) recognized it as a heavy rainfall threat, and put a high priority on producing accurate precipitation forecasts. This goal was facilitated by obtaining a substantial number of supplemental observations for assimilation into NCEP's real-time GFS forecasts (Majumdar et al. 2013). Ten aircraft missions were flown between 23 and 27 August to obtain dropsonde data from Irene, while upper-air sites in the National Weather Service (NWS) Southern and Eastern Regions launched additional rawinsondes at 0600 and 1800 UTC beginning at 1800 UTC 22 August. At 0600 UTC 25 August, the region of supplemental rawinsondes was expanded to include all stations east of the Rocky Mountains. The dropsonde data proved to significantly increase forecast accuracy at all initialization times, while the rawinsonde data significantly improved forecast accuracy for GFS simulations initialized at 0600 and 1800 UTC. Majumdar et al. (2013) further explained that if the GFS is run without these additional data, it allows the upstream trough to progress too quickly to the east, which displaces Irene too far offshore, thereby understating the potential impacts on land. With the supplemental observations, however, the GFS was able to simulate a slowly-moving trough and produce a relatively accurate track forecast for Irene. This success helps to illustrate that more continuous initial conditions do result in a quantifiable decrease in forecast error.

Like the other TCs discussed in previous sections, Irene underwent ET around the time of landfall, meaning that its precipitation field was influenced by ET dynamics. Using a cyclone

phase space diagram (Hart 2003; Figure 4), it can be seen that Hurricane Irene was undergoing ET for about 36 h prior to the NHC's official declaration of its status as an extratropical cyclone at 0000 UTC 29 August (Figure 3e). As Irene moved nearer to the trough situated over the eastern United States, the processes described in Klein et al. (2000) became apparent: WAA increased on the eastern side of the TC and CAA to the west, while outflow aloft accelerated (not shown, but can be inferred from the increasing PV gradient to the north of Irene illustrated in Figure 3). A jet streak formed to the north of Irene, indicating that the kinetic energy of the midlatitude jet was locally increasing. Finally, Irene's circulation became fully coupled to the jet and its associated upstream trough, producing an extratropical cyclone. Therefore, the modeling challenges associated with ET are most likely present in simulations of Irene during this time period, making the storm a good candidate for an ET modeling case study.

1.5 Goals of this work

Rappaport (2000) estimated that, between 1970 and 1999, half of the TC related deaths that occurred in the United States were due to freshwater flooding; thus, accurate forecasts of such events are vital to prepare for, and to mitigate, the negative effects of TC landfall on communities. Tropical cyclone track and intensity forecasts have improved substantially in recent years (Yamaguchi et al. 2015), largely due to advances in numerical weather prediction, although models still retain some biases and inaccuracies (e.g., Davis et al. 2016). However, despite the high-impact nature of heavy precipitation events associated with landfalling TCs, relatively few studies have explored the sensitivity of TC-related precipitation forecasts to model initial conditions beyond looking at the variability in TC track. The goal of this work is to understand what modulates precipitation forecasts over the northeastern United States during Hurricane Irene by examining the synoptic and mesoscale dynamics of an ensemble initialized

36 h before Irene made landfall over the Northeast. section 2 will focus on determining the synoptic sources of precipitation variability in an 80-member GFS ensemble, while section 3 will focus on the mesoscale causes of precipitation variability in an 80-member WRF ensemble.

2. GFS ensemble rainfall variability

2.1 GFS 0.5° model description

An 80-member ensemble of the 2014 experimental version of the GFS was initialized at 0000 UTC 27 August 2011 (when Hurricane Irene was off the coast of North Carolina; Figure 3a) and run through 0000 UTC 29 August 2011 (when Irene was declared extratropical over New England; Figure 3e). The 80-member ensemble was based off of a high-resolution (~27 km) control forecast, which was perturbed using the EnKF (Whitaker and Hamill 2002) and hybrid 3D-variational data assimilation (Wu et al. 2002b). This statistical assimilation method combines initial static background error with the 80 perturbed members produced by the EnKF, and adds the resulting perturbations back to the initial high-resolution control forecast. The resulting 80 high-resolution initial analyses are used as initial conditions for the ensemble member forecasts. These forecasts were output with 0.5° spatial resolution and three-hourly temporal resolution.

2.2 GFS 0.5° Results

2.2.1 Analysis of ensemble forecasts

The main purpose of this research is to determine why some ensemble members predict large precipitation totals at a given location, while others predict much less. As discussed in section 1, this work uses Hurricane Irene as a case study to examine ensemble variability in rainfall forecasts for the Catskill region of New York (Figure 5). The Catskills were one of the hardest hit areas by Irene, receiving up to 305 mm of rain in some areas (Figure 6b), making the region an ideal metric for evaluating the accuracy of each ensemble member's precipitation forecast. Sets of the wettest and the driest of the GFS ensemble members (in terms of Catskills

precipitation) are therefore compared in order to deduce the dynamical causes behind the variability in their precipitation.

2.2.2 General overview of ensemble member rainfall characteristics

The performance of the ensemble as a whole was assessed by comparing the 48 h ensemble mean and standard deviation of precipitation to observations for the same time period, obtained from the Earth Observing Laboratory's NCEP / Environmental Modeling Center 4 km Stage IV precipitation data (Figure 6; Lin 2011). Overall, the GFS ensemble mean did well at predicting the spatial extent of Irene's rainfall, with a rain swath about 5° of longitude wide extending from North Carolina up through New England present in both the ensemble mean rainfall (Figure 6a) and the observed rainfall (Figure 6b). The highest rain totals were also correctly located over the Mid-Atlantic and Carolina coastlines. However, the ensemble mean underpredicted the actual amount of rainfall received nearly everywhere over land, and did not reproduce the greater than 254 mm accumulations seen over North Carolina, Delaware, or the Catskill region of New York (Figure 6c). This difference is perhaps to be expected, as the process of calculating an ensemble mean will damp out any local maxima seen in individual members, resulting in a smooth, lower-magnitude distribution of values. The high standard deviation (Figure 6d) seen over Virginia and North Carolina allows the possibility that some members produced locally higher totals, and an examination of individual ensemble members showed that many of these simulations did produce maxima of greater than 254 mm there, albeit in differing locations. However, it is notable that while the ensemble standard deviation was elevated over the Catskills (~50 mm), its magnitude was less than that produced farther south (>65 mm), suggesting that fewer individual members predicted high enough rainfall totals

(possibly due to errors in the GFS's inability to simulate terrain effects and mesoscale processes; see section 3 for details).

The GFS ensemble mean (Fig. 6a) also displays a sharp gradient on the western edge of the accumulated precipitation swath. This implies that track variability between the individual ensemble members could have been a driver of precipitation variability, as a small east-west shift of the rain swath could substantially alter the amount of precipitation delivered to a specific location. To test this hypothesis, the predicted storm track for each ensemble member was derived by finding the latitude and longitude of the minimum value in the mean sea level pressure field on a defined eastern United States domain (31–50°N, 60–82°W) at three-hourly intervals, and comparing these tracks to Irene's observed storm track, obtained from the revised Atlantic hurricane database (HURDAT2) best track analysis (Landsea and Franklin 2013) (Figure 7). In the case of Irene, 1–2° of logitudinal variability was present between the storm tracks produced by each ensemble member. When combined with the sharp gradient in precipitation that occurred on the western edge of the GFS ensemble mean rain swath (Figure 6a), this track variability makes it likely that storm position was a major control on the amount of precipitation delivered to certain areas (such as the Catskills).

To test this hypothesis, the total 48 h precipitation received by the Catskills (defined here as 41.5–42.5°N, 73.5–75°W) was compared with the average storm track longitude for each ensemble member (Figure 8). The data show a strongly linear relationship between these two variables ($r = 0.77$), with precipitation amount increasing for a more westward average storm longitude. The strong linearity implies that storm position was the dominant factor controlling Catskills rainfall amount. Furthermore, Irene's observed track fell about one third of a degree of longitude to the west of the ensemble mean track at landfall (Figure 7), suggesting that the

westward-tracking members may be more realistic than the eastward-tracking ones, assuming the model's precipitation is correct.

Figure 9 takes a closer look at this implication by comparing a westward-tracking ensemble member with an eastward-tracking one, both in terms of total rainfall amount and of three-hourly rainfall rate during the interval of maximum precipitation at the latitude of the Catskills (36–39 h). The westward-tracking member features a 48 h rain swath with its highest values over eastern New York, while the eastern one has its highest values over New England. Likewise, the western member placed the largest 36–39 h rainfall rates directly over the Catskills (centered on ~74.5°W), with the eastern one placing the largest rates over eastern Massachusetts and New Hampshire (centered on ~71.5°W). Even though the specific rainfall amounts generated were visually similar between the two members, the large difference in position results in very different rainfall totals for specific regions of the Northeast. Again, the precipitation produced by the westward-tracking member most closely resembles the observed precipitation (Figure 6b).

Figure 9 also hints at some differences in timing between the wet western members and the dry eastern ones, in addition to the previously examined spatial differences. The eastern member placed the 36–39 h precipitation farther north than the western one, suggesting that Irene may have moved more quickly in that simulation. A time series of mean precipitation rate over the Catskills for the 10 wettest ensemble members, the 10 driest, and the remaining 60 members (Figure 10) clearly shows that the drier members produced a peak rainfall rate for this region three hours earlier than the wetter ones, although the two composites had similarly-shaped curves. Figures 8–10 imply Irene took a more westward, slowly progressing track in the wetter simulations, allowing the area of maximum rainfall to move closer to the Catskills and linger

there, while in the drier simulations, the region of maximum rainfall occurred farther to the east and moved away from the Catskills more quickly.

2.2.3 Comparison of the 10 wettest and 10 driest ensemble members

The GFS ensemble forecasts show that the farther west a particular member tracked Hurricane Irene, the more rain that member delivered to the Catskills. Therefore, the remainder of this analysis will focus on two scenarios to determine the causes of precipitation variability in the GFS ensemble: 1) wet, westward-tracking, and 2) dry, eastward-tracking simulations. Composite fields were created for the ten wettest ensemble members and the ten driest for comparison. As discussed in section 2.2.2, there are both spatial and temporal differences in how Irene is positioned in the wet members as opposed to the dry members. The speed and direction of a hurricane's propagation is often influenced by the timing and orientation of upstream features (Carr and Elsberry 2000), so a working hypothesis for the differences observed in this ensemble is that differences in track between the wet and dry members are due to differences in the timing of upstream features. This hypothesis was tested by comparing the composited synoptic-scale features of the 10 wettest ensemble members with those of the 10 driest, using the method described below. If the hypothesis is true, significant positional differences should be observed both in Irene and in upstream features between the two composites.

The synoptic-scale features of the wet and dry member composites were primarily compared using composite difference plots of various fields. Composite difference plots display the ensemble mean of a particular field, the normalized difference between the wet and dry scenarios, and the statistical significance of that difference. The normalized difference for each field in question was computed with the formula

$$\Delta x_i = \frac{\bar{x}_i^{wet} - \bar{x}_i^{dry}}{\sigma_{x_i}} \quad (1)$$

where $\bar{x}_i^{wet}$ represents the mean of the *i*th field for the wet members, $\bar{x}_i^{dry}$ represents the mean of the *i*th field for the dry members, and σ_{x_i} is the ensemble standard deviation of the *i*th field (Torn et al. 2015). Normalizing by the standard deviation allows disparate fields and times to be compared to one another, even if their actual magnitudes are on different scales. Statistical significance was computed at $\alpha = 0.05$ using a Student's t-test (Ruxton 2006).

The expected positional differences between synoptic-scale features are immediately apparent when examining composite difference plots of 300 hPa area-averaged circulation at 00 h, 06 h, 18 h, and 36 h (Figure 11; the area average for a particular gridpoint was calculated by taking the average of all neighboring gridpoints within a 200 km radius). 300 hPa circulation was selected because TCs are areas of strong cyclonic circulation, so any position difference in that circulation would show up very well in a composite difference plot. In addition, midlatitude features (such as an upstream trough) likely to influence the TC are maximized in the upper levels of the troposphere. Although differences between the two composites are of relatively low magnitude at the time of model initialization (Figure 11a), a statistically significant positional difference co-located with Irene arose by 06 h (Figure 11b), as evidenced by the +/- 0.8 difference centered on Irene off the coast of the Carolinas. The warm colors to Irene's west indicate that the wetter members positioned Irene's circulation west of the ensemble mean, while the dry colors to its east show that the drier members positioned it east of the ensemble mean. This positional difference grew in magnitude through 18 h (Figure 11c) and 36 h (Figure 11d), when Irene made landfall in New York. The fact that positive differences arise to the west of the ensemble mean and negative differences to the east supports the previous conclusion that the wetter members are taking a westward track and the drier members an eastward one.

In addition to the differences associated with the position of Irene, another strong positional difference can be observed in the placement of the upstream trough located over the Great Lakes. At 06 h, or six hours into the simulation, the positive circulation connected to the leading edge of the trough had progressed slightly farther east in the drier members (as evidenced by the -0.8 differences in place over Pennsylvania and Virginia), although this perturbation is not statistically significant. By 18 h, the upstream trough in the drier members had progressed significantly farther to the east (as evidenced by statistically significant -1.6 anomalies over western New York and Pennsylvania, than the trough in the wetter members, which remained over the Great Lakes. When Irene made landfall at 36 h, the positional differences in the trough were especially pronounced over the eastern Great Lakes and western New York and Pennsylvania, with differences greater than +/- 2. The timing differences in this upstream trough in the 18 hours preceding landfall were likely responsible for the differences in storm position at the time of landfall. The drier members featured a more progressive trough, blocking Irene from moving inland and accelerating it off to the east. The trough in the wetter members, on the other hand, remained farther to the west, allowing Irene to maintain a predominately northward heading into eastern New York, placing the maximum rainfall rates over the Catskills.

It is important to note, however, that Irene developed statistically significant positional differences before the upstream trough, implying that the forecast variability that ultimately developed originated in the vicinity of the TC. As tropical cyclones are capable of modifying their environments, especially during extratropical transition (e.g., Harr et al. 2008; Torn 2010; Archambault et al. 2013), which Irene was undergoing at the time, it is plausible that processes within Irene itself were responsible for the timing differences that developed in the synoptic-

scale flow. Therefore, the next step is to determine which dynamic processes within Irene could be responsible for the modeled position differences in the upstream trough.

First, it is important to confirm that Irene had a mechanism for altering its environment. As discussed in section 1, divergent outflow can allow forecast error from within a TC to propagate to the surrounding midlatitude flow, so divergence is a likely candidate for such a mechanism. Figure 12 shows that from 06 h onwards, the wetter members featured substantially stronger upper-level divergence (as indicated by the positive 1.2–2 difference values oriented NE-SW over the east coast) on the northwestern side on Irene, indicating that there was a mechanism in place for halting the progression of the upstream trough. Using this information, we proposed two hypotheses: first, that the wetter members featured a greater magnitude of divergence on Irene's western side because they were simulating greater mid-level latent heat release on that side of the storm, and this divergence redistributed the PV field around Irene in the wetter members, allowing the TC to take a more westward track. The second hypothesis focuses less on the magnitude of the divergence aloft than on its position relative to the trough, and asserts that the initial differences in storm track between the wet and dry members were due to differences in the broader steering flow which began in the vicinity of Irene. Once set on a more westward path, presumably by stronger easterly steering flow, the wetter members placed the divergent outflow from Irene closer to the upstream trough than the drier members, allowing the outflow to more efficiently slow the trough's progress. The first hypothesis will be proven correct if the wetter members featured greater latent heat release on the western side of Irene than the drier members during the first 18 h of the model run. The second will be verified if the wetter members were initialized with greater easterly steering flow, which shifted the divergence aloft closer to the upstream trough.

A composite time series of latent heat release at 500 hPa (500 hPa was selected for analysis because Irene displayed large amounts of variability in latent heating in between members at that level) on the western side of Irene (defined here as the area encompassed by +/- 2°N of latitude and +4/+1°W of longitude from the storm center for each ensemble member, chosen to correspond with the maximum positive values seen on the divergence composite difference plots) for the 10 wettest members, the 10 driest, and the remaining 60 members. The time series (shown in Figure 13), show that the first of the two hypotheses is unlikely to be correct. The three composites exhibit similar magnitudes of latent heat release (ranging from -1 to +7 K/3hr) throughout the 48 h simulation, with the wetter members producing less latent heat than the drier members at all times except 9–15 h, and to a lesser extent 27–33 h. The only interval in which the wetter members were consistently associated with the most latent heat release, and the driest members the least, is between 09 h and 15 h. If latent heat release is related to the linear correlation between storm position and rainfall total, it would likely be the latent heat release from this time interval.

Figure 14 displays vertical profiles of latent heat release for each of the three composites at 12-h (the time of maximum positive difference between the wet and dry members). The wetter members do appear to display elevated latent heat throughout the mid-troposphere at this time, although the magnitude of the difference is very small. However, when 12 h latent heat release at 500 hPa is compared to total precipitation received by the Catskills (the metric by which wet and dry members are defined) there is almost no correlation ($r = 0.19$; Figure 15). Therefore, we conclude that mid-level latent heat release is not a good predictor of whether a particular ensemble member will be wet or dry over the Catskills, and was not the driving mechanism by which the upstream trough was slowed and Irene allowed to track farther west.

With the first hypothesis disproved, the second hypothesis (that the initial differences in storm track between the wet and dry members were due to differences in the broader steering flow) is tested by examining composite difference plots of 250–850 hPa layer zonal steering flow at 00 h, 06 h, 18 h, and 36 h (Figure 16). Zonal steering flow was calculated by removing the irrotational and nondivergent wind vectors from the total wind within a 3° radius of the TC's center following Galarneau and Davis (2013). At the time of model initialization, the wetter members placed Hurricane Irene within a broad region of statistically significant perturbation easterly steering flow, as evidenced by -1.2 standardized zonal flow anomalies maximized to the northeast of Irene (Figure 16a). By 06 h, the magnitude of the difference between the wetter and the drier members increased, with Irene situated at the center of a strong easterly perturbation (Figure 16b). By 18 h, the easterly perturbations strengthened and grew in area (Figure 16c), persisting through the time of Irene's landfall (Figure 16d). Thus, from the time of model initialization, the wetter members possessed the necessary synoptic setup to steer Irene closer to the approaching trough than the dry members.

Torn et al. (2015) performed a similar verification of GFS ensemble forecasts on Hurricane Sandy (2012), and also found that ensemble members that produced westward tracks were characterized by easterly perturbations in the steering flow surrounding the storm. In Sandy, the easterly perturbations were determined to have been the result of a negative PV anomaly associated with an anticyclone located poleward of Sandy. The clockwise flow around the negative PV anomaly resulted in increased easterly flow to its south, steering Sandy farther west.

To test whether a similar PV anomaly was responsible for steering Irene, composite differences are displayed for the 350 K PV field surrounding Irene at 00 h, 06 h, 18 h, and 36 h

(Figure 17). At the time of model initialization (Figure 17a), the wetter members featured a positive PV anomaly on the southern wave guide, to the southwest of Irene, just off the coast of northeastern Florida. The cyclonic flow associated with this positive PV anomaly would be expected to generate easterly steering flow perturbations in the vicinity of Irene (Figure 16a), yielding a perturbation westerly motion. The cyclonic PV anomaly persisted through 06 h (Figure 17b), maintaining the increased easterly steering flow in the wet members (Figure 16b). By 18 h (Figure 17c), the initial PV perturbation has disappeared; however, by this time, a significant difference had developed between the wet and dry members' placement of both the TC and the upstream trough, as evidenced by the < -2 magnitude perturbations over the Appalachians (Figure 17a), indicating that the high-PV trough had not yet arrived in that region. The anticyclonic flow anomalies associated with this perturbation placed Irene once again in a region of strong easterly steering flow (Figure 16c), which persisted through the storm's landfall at 36 h (Figure 17d).

Many studies have examined the mechanism by which tropical cyclones undergoing ET affect PV distributions along the midlatitude wave guide (e.g., Riemer et al. 2008; Torn 2010; Archambault et al. 2013). A primary mechanism by which energy can be transferred from a tropical cyclone to the jet is through negative PV advection by the storm's irrotational outflow. As per the second hypothesis above, the wetter members are the ones in which easterly steering flow perturbations started Irene on a more westward course, bringing its divergent outflow in closer proximity to the upstream trough in the midlatitude flow. The outflow then advected more low-PV air poleward into the jet, slowing the forward progression of the high-PV trough. This interaction set up a feedback mechanism between Irene and the midlatitude wave guide, wherein Irene advected low-PV air poleward and westward, slowing the trough, which allowed Irene to

track even farther west, further advecting low-PV air and slowing the trough even more. This feedback was largely absent in the dry, eastward-tracking members, which never got close enough to the trough to effectively slow its progress. Figure 18 displays a composite difference plot of 250 hPa irrotational outflow at 15 h. Irene's upper-level outflow was strongly divergent at this time, and is represented in the ensemble mean as a starburst pattern of outward-pointing vectors radiating from the TC's center. A strong easterly wind perturbation can be seen on the leading edge of the upstream trough (represented by the -1.2 differences over Ohio), indicating that the wetter members had significantly higher magnitude irrotational wind advecting low-PV values in that location.

Altogether, the information provided in Figs. 16–18 illustrates the mechanism by which the wetter ensemble members were able to take a more westerly track. The analysis supports the following steering flow hypothesis: the differences in Irene's track were due to differences in the timing of upstream features, which in turn were due to differences in steering flow at the time of model initialization. This sequence of events is consistent with each ensemble member being initialized with a different PV field, in order to compensate for the discontinuous nature of real-time weather observations.

3. WRF ensemble rainfall variability

Although the GFS is a good synoptic-scale model, it can have difficulty producing accurate forecasts over regions of complex terrain (Carpenter et al. 2004). Lower-resolution terrain input files don't resolve individual peaks and valleys, and thus don't accurately reflect the mesoscales forcings or hydrology of complex areas. In order to examine the role that these types of complex terrain and mesoscale processes played in the evolution of precipitation in Hurricane Irene, the 0.5° GFS ensemble discussed in the previous section was downscaled to 15 km and 3 km using the Weather Research and Forecasting (WRF) model. The methods and results associated with these downscaled ensembles are discussed in this section.

3.1 WRF 15 km model description

The 0.5° GFS ensemble was initially downscaled to 15 km over the eastern United States and western Atlantic (25–50°N, 60–95°W) using WRF version 3.6, with physics comparable to those employed in the High-Resolution Rapid Refresh model (HRRR): Thompson microphysics (Thompson et al. 2004), the Rapid Radiative Transfer Model for Global Climate Models (GCMs) (RRTMG) longwave and shortwave radiation (Mlawer et al. 1997), the Mellor-Yamada-Nakanishi-Niino (MYNN) planetary boundary layer (PBL) and surface scheme (Nakanishi and Niino 2009), the Rapid Update Cycle (RUC) land surface model (LSM; Benjamin et al. 2004), and the Kain-Fritsch cumulus parameterization (CP; Kain 2004). This setup allowed for improved representation of terrain and mesoscale processes, although it should be noted that 15 km terrain is still too coarse to capture individual peaks and valleys (thereby retaining inaccuracies in simulated terrain processes such as upslope flow).

3.2 WRF 15 km Results

With the GFS ensemble, the major factor that determined how much precipitation the Catskills received from a particular ensemble member was Irene's storm track. Members that tracked Irene farther west produced higher Catskills rainfall totals, while members that tracked Irene farther east produced lower Catskills rainfall totals. However, this simple east-west dependency does not exist in the WRF 15 km ensemble. When the total precipitation received by the Catskills in the WRF 15 km ensemble is plotted against the average storm track longitude for each member (Figure 19), the plot reveals that there is very little correlation between those two variables. Instead, members on the eastern and western edges of the distribution produced the lowest rainfall totals, while members with central tracks produced the highest. These results indicate that there is still a relationship between storm track and Catskills rainfall in the WRF 15 km distribution, but that it is a much more complex and nonlinear.

This nonlinearity in the relationship between track and rainfall makes the composite differencing used in the GFS section an ineffective analysis technique, as the 10 wettest members will all be centrally located, but the 10 driest will be dispersed east and west (thereby cancelling out each other's vortex signatures when composited), creating a false signal that wetter members have a stronger (and smaller) vortex circulation. Indeed, when a composite difference plot of 300 hPa circulation was produced for the time of Irene's landfall (36 h into the model run; Figure 20) this exact pattern is seen: positive differences in the center of the hurricane (represented by the 0.8–2 magnitude positive differences off the coast of New Jersey, and negative differences in the surrounding environment (represented by the -0.8 to -1.6 differences across most of the northeastern US).

Because of the false signal produced by composite differencing, an alternative analysis technique is needed. In addition, the fact that there is clearly a more complex relationship

between storm track and rainfall total suggests that terrain and mesoscale processes play an important role in determining the distribution of precipitation; thus, it is vital to simulate both as accurately as possible. As mentioned earlier, while 15 km resolution is a vast improvement over 0.5°, it is still too coarse to capture the nuances of very complex terrain, such as exists in the Catskills. The WRF 15 km ensemble also relies on a CP, so mesoscale processes (specifically convective precipitation), too, are not being simulated as accurately as they could be. In order to address all of these issues, the 15 km WRF domain was further downscaled to 3 km. 3 km terrain is finely resolved enough to capture individual peaks and valleys of mountain ranges, and at 3 km, convective precipitation is directly calculated, instead of being parameterized.

3.3 WRF 3 km model description

A portion of the WRF 15 km domain spanning the eastern United States from Michigan to Maine (29–46°N, 68–87°W) was further downscaled to 3 km, once again using WRF version 3.6, with physics comparable to those employed in the HRRR. As discussed previously, the 15 km simulation relied on a CP to approximate convective precipitation, but the 3 km grid was high-resolution enough that it was not necessary to use the cumulus scheme. In addition, new boundaries were defined for the Catskills in the WRF 3 km output, to better reflect the actual location of the mountain range in the higher-resolution topography. These new bounds (41.75–42.75°N, 73.9–75.25°W, as opposed to the GFS's 41.5–42.5°N, 73.5–75°W) are illustrated in Figure 21.

3.4 WRF 3 km Results

3.4.1 General overview of ensemble member rainfall characteristics

When the mean precipitation of the 80-member WRF 3 km ensemble is compared to observed precipitation over the eastern United States (Figure 22a–b), it can be seen that this ensemble mean is much better than the GFS ensemble mean (Figure 6a) at predicting rainfall for this time period in the Northeast. The WRF 3 km ensemble mean accurately forecasts accumulations of greater than 254 mm over the peaks of the Catskills, and widespread totals between 152 and 254 mm across southeastern New York, western Massachusetts and Connecticut, and southern Vermont, while the GFS ensemble mean failed to predict anything greater than 152 mm outside of extreme southeastern New York, western Massachusetts, and Connecticut (Figure 6a).

The WRF 3 km ensemble, like the GFS ensemble, features a high standard deviation over the Mid-Atlantic states (Figure 22d), but while the GFS featured a fairly uniform standard deviation of 30–50 mm over the Northeast, the WRF 3 km displays values exceeding 65 mm in regions with complex terrain. Most notably for this study, there is high variability on the *eastern* slopes of the Catskills, suggesting that the steep terrain of these eastern slopes are a source for some of the large variability that occurs in Catskills rainfall forecasts between ensemble members. The WRF 3 km ensemble mean also underpredicted rainfall totals on the eastern slopes (Figure 22c).

One major conclusion of section 2 was that Catskills rainfall variability in the GFS ensemble was mainly due to how far east or west individual members tracked Irene. The WRF 3 km storm tracks do feature longitudinal variability (Figure 23), and individual members have roughly the same average storm track longitude in the WRF 3 km ensemble as they do in the GFS ensemble (Figure 23a); but, as was seen in the WRF 15 km ensemble, the WRF 3 km ensemble does not feature a linear relationship between Catskills precipitation and storm track

longitude (Figure 23b). Members with storm tracks near the center of the longitude distribution have generally higher precipitation amounts, although there is still ~60 mm of spread in precipitation for those members.

3.4.2 Objective clustering methods

The 3 km WRF ensemble features a more complex relationship between precipitation and storm track than a simple east–west correlation (as was seen with the GFS). Therefore, composite difference plots of the ten wettest and ten driest ensemble members were not effective analysis tools for this ensemble, and a new technique was needed to stratify members and address the question of what modulates rainfall variability over the Catskills. Objective clustering of ensemble members based on the precipitation field proved to be an effective method.

For this analysis, the k-means objective clustering algorithm (Hartigan and Wong 1979) was used to group the 80 ensemble members into three clusters (Figure 24a), based on the distribution of rainfall over the domain 41.5–43.5°N, 73–76.5°W at 39 h into the forecast. 39 h was selected because it featured the highest precipitation rates over the Catskills, and members with the greatest rainfall at this time generally produced the greatest total accumulations over the region. The domain 41.5–43.5°N, 73–76.5°W was subsequently chosen because it encompassed the extent of the members' rainfall swaths over New York at this time. This clustering methodology was proven to be robust, as very similar groupings were produced when the members were clustered based on the horizontal distribution of precipitation over the entire duration of the simulation, instead of just at 39 h. In addition, using more than three clusters proved to be redundant, indicating that three is sufficient to accurately portray the variability present in the ensemble (Figure 24d).

3.4.3 Objective clustering results

The first cluster is comprised of members that tracked Irene to the center and west of the track distribution (Figure 24b; purple), the second consists of members with central and easterly tracks (Figure 24b; red), and the third is a residual cluster, encompassing members that were either too slow or too far on the edges of the track distribution to bring much rainfall to the Catskills (Figure 24b; blue). The western cluster brought the highest rainfall totals to the Catskills, while the eastern cluster brought reduced, although still substantial, accumulations (Figure 24c; purple and red lines). The residual cluster brought the least amount of rain, as its members placed Irene's rain swath too far away from the region. The remainder of this thesis will focus on comparisons of the wetter, western cluster and drier, eastern cluster as two scenarios that resulted in substantially different rainfall rates, while still tracking near enough to the Catskills to produce significant accumulations.

Three hypotheses are proposed to explain the variability between the wetter and drier clusters: 1) wetter members feature the greatest upslope forcing over the Catskills, 2) wetter members have increased moisture flux convergence over the Catskills, and 3) wetter members position the region of maximum Q-vector convergence over the Catskills, thus producing greater synoptic forcing for ascent. It should be noted that the first two of these hypotheses are related to mesoscale dynamics and terrain effects, and are not entirely independent of one another (as upslope flow in a moist environment will cause moisture convergence), while the third hypothesis deals with synoptic dynamics.

Figure 25 tests the first of the three hypotheses (that the wetter cluster had stronger upslope forcing) by comparing the composited 900 hPa winds and three-hourly precipitation rate in relation to terrain for the wetter cluster (Figure 25a) and the drier cluster (Figure 25b) at the

time of maximum precipitation over the Catskills. The wetter members feature easterly low-level winds perpendicular to the steep eastern slopes of the Catskills, where the highest rainfall rates were seen. The drier members, on the other hand, have a more northerly flow (with a wind angle nearly 90° less than that of the wet members at 39 h), impacting shallower terrain gradients at the northern side of the Catskills. When an upslope metric ($V_{10} \bullet \nabla Z_s$, where V_{10} is the 10-meter horizontal wind vector and ∇Z_s is the gradient of the terrain height; Tang et al. (2016)) is examined for the Catskills locations that received the top 20th percentile of accumulated rainfall (the top 20th percentile was chosen in order to provide a clearer signal for comparison between the two clusters), it can be seen that the wind setup in the wetter members provides the wetter members with an additional band of strong positive upslope near the southeast part of the Catskills (near 42°N, 74° 15'W; Figure 26a), which is absent in the drier members (Figure 26b). An examination of Figure 26a reveals that the southeast part of the Catskills (near 42°N, 74° 15'W) had one of the largest differences in precipitation rate between the western and eastern clusters, with the former receiving over 102 mm per 3 h, and the latter receiving only 38–64 mm per 3 h (Figure 25a–b). In addition, the drier members have much stronger downslope values over much of the 20th percentile region (as evidenced by the higher prevalence of dark blues in Figure 26b). These patterns persist throughout most of the time interval when rain was impacting the Catskills. This evidence supports the first hypothesis, which stated that wetter members feature the greatest upslope forcing over the Catskills.

Figure 27 tests the second hypothesis (that the wetter cluster had stronger moisture convergence) by comparing the composited 100–1000 hPa layer mean winds, 100–1000 hPa integrated moisture transport by the southeasterly wind (which was the dominant direction of moisture transport for both clusters), and convergence of the 100–1000 hPa integrated moisture

transport (including all wind directions) for the wetter, western cluster (Figure 27a) and the drier, eastern cluster (Figure 27b), at the time of maximum precipitation over the Catskills (the 100–1000 hPa layer was chosen as a means to approximate the moisture content of the whole troposphere). Not only do wetter members have greater moisture transport by the southeasterly wind into the Catskills (as can be seen by the larger contour values in Figure 27a), they also have greater moisture convergence over the mountain range (with a maximum $6*10^3$ kg/m^2/s, as opposed to a maximum of $3*10^3$ kg/m^2/s in the drier members), which would lead to increased precipitation. The drier members actually have moisture divergence over the Catskills (shown by the blue shading in southeastern New York), especially over that southeast part of the Catskills where the wetter members have an additional band of strong positive upslope. This evidence supports the second hypothesis, which stated that wetter members have increased moisture flux convergence over the Catskills. It should be noted again, however, that the first two hypotheses are not independent of one another, as upslope forcing can lead to moisture convergence. Indeed, the greater upslope values of the wetter members are possibly responsible for those members' greater moisture convergence, although the wetter cluster does have larger actual values of moisture being transported into the Catskills (contours in Figure 27a). These patterns persist throughout most of the time interval when rain was impacting the Catskills.

Finally, Figure 28 tests the final hypothesis (wetter members position the region of maximum Q-vector convergence over the Catskills) by comparing the composited 500–800 hPa layer mean Q-vectors and isotherms, and the Q-vector convergence for the wetter cluster (Figure 28a) and the drier cluster (Figure 28b), at the time of maximum precipitation over the Catskills. The 500–800 hPa layer was selected because it represented the layer of strongest forcing for ascent in this system. Q-vectors are defined by the equation

$$\vec{Q} = \left(-\frac{R}{\sigma p}\left[\frac{\partial u_g}{\partial x}\frac{\partial T}{\partial x} + \frac{\partial v_g}{\partial x}\frac{\partial T}{\partial y}\right], -\frac{R}{\sigma p}\left[\frac{\partial u_g}{\partial y}\frac{\partial T}{\partial x} + \frac{\partial v_g}{\partial y}\frac{\partial T}{\partial y}\right] \right) \quad (2)$$

where $\vec{Q}$ is the total Q-vector, R is the ideal gas constant for dry air, σ is static stability, p is pressure, u_g and v_g are the zonal and meridional components of the geostrophic wind, and T is temperature (Sanders and Hoskins 1990). Physically, Q-vectors represent the change with time of the potential temperature gradient by the geostrophic wind. Convergence of Q-vectors indicates synoptic forcing for ascent (Hoskins and Pedder 1980), which is why the third hypothesis posed in this section is that wetter members position the region of maximum Q-vector convergence over the Catskills, thus creating greater quasigeostrophic forcing for ascent. Surprisingly, Figure 28 shows that the drier, eastern members actually feature much stronger Q-vector convergence (and therefore synoptic forcing for ascent) than the wetter members ($5*10^{13}$ m2/kg/s of convergence in the dry members, as opposed to around $1*10^{13}$ m2/kg/s of convergence in the wet members), disproving the third hypothesis. The drier members have much stronger synoptic forcing for ascent, and thus for precipitation, throughout most of the interval when rain occurred over the Catskills.

The main drivers for precipitation in the wetter and drier clusters can be more clearly seen by comparing composited time series of all the metrics discussed in this section (precipitation, upslope, moisture convergence, and synoptic forcing; Figure 29). The wetter members feature large upslope (in red) and moisture convergence (in blue) values from around 24 h into the forecast (when precipitation began over the Catskills) to the end of the model run, with peaks at 36 and 39 h, which correspond with the interval of maximum precipitation over the region (Figure 29a). The drier members have some weak upslope forcing initally, but it quickly disappears, while moisture convergence values remain moderate throughout the duration of the model run (Figure 29b). The precipitation curve (in green) in the dry members closely follows

the shape of the Q-vector convergence curve (in orange), with both peaking at 39 h and then decreasing to the end of the model run. Q-vector convergence does exist in the wetter members, but its magnitude is much lower. It should be noted that all four time series are scaled to fit on the same graph, in order to allow for easy comparison of curve shape. In addition, the synoptic forcing and precipitation rate are calculated as area averages over the Catskills domain, while upslope and moisture convergence are averaged only over the gridpoints in that domain that received the top 20th percentile of precipitation values, in order to display a clearer signal.

To summarize the results of these cluster comparisons, precipitation in the wetter, western cluster was driven primarily by mesoscale processes and terrain effects (strong upslope forcing and moisture convergence), while precipitation in the drier, eastern cluster was driven primarily by synoptic forcing for ascent. The wetter cluster featured easterly low-level flow directly into the sharp terrain gradient of the eastern Catskills, creating large upslope magnitudes, while the drier cluster had more northerly low-level flow, which failed to impact as large a terrain gradient. Likewise, the wetter cluster simulated stronger moisture transport into the region than the drier cluster, and greater convergence of that moisture (the larger upslope magnitudes likely contributed to this convergence). The drier cluster, on the other hand, had stronger Q-vector convergence than the wetter cluster, which generally featured either weak or negative Q-vector convergence (and thus reduced synoptic forcing for ascent). These results illustrate the importance of terrain and mesoscale processes in producing large rainfall totals, as the cluster with a favorable mesoscale setup was able to produce substantially more precipitation than the synoptically forced cluster.

4. Conclusions

As TCs track into the midlatitudes, they are often associated with extensive regions of heavy precipitation. This precipitation can lead to devastating, widespread floods, such as occurred with Hurricane Irene (2011) over the Northeast, and particularly the Catskill region of New York. The goal of this work is to understand what modulated precipitation forecasts over the northeastern United States during Hurricane Irene by examining the synoptic and mesoscale dynamics of an ensemble initialized 36 h before Irene made landfall over the Northeast. In particular, the goal is to determine why some ensemble members predict large precipitation totals over the Catskills, while others predict much less.

A large amount of variability exists in an 80-member GFS ensemble in terms of precipitation in the Catskills. The differences that exist between the rainfall forecasts of the ten wettest and ten driest GFS ensemble members are almost entirely due to the forecast position of Irene. The wetter members tracked Irene farther to the west, while the drier members featured a more eastward path. Ultimately, differences in storm track appear to be related to differences in the 0 h potential vorticity (PV) to the southwest of Irene. Wetter members were characterized by greater cyclonic PV to the southwest of the storm center, in a tropopause-based trough over the southeastern U.S., which placed the hurricane in a region of anomalously easterly steering flow. This steering flow started Irene on a more westward track, enabling its upper-level outflow to affect the midlatitude jet and causing a Great Lakes trough to slow and deepen. With the trough remaining well upstream, Irene was able to track even farther to the west, positioning the region of maximum rainfall directly over eastern New York and Vermont. These results imply that synoptic-scale interactions are the dominant source of variability in GFS forecasts of Irene, and

that variability forecasted precipitation distribution between ensemble members is largely controlled by storm position.

The 0.5° GFS output was then downscaled to 3 km using the Weather Research and Forecasting (WRF) model, in order to allow a better representation of mesoscale processes and the effects of terrain on the precipitation distribution. Contrary to the GFS-based results, the 3 km WRF output does not show a linear relationship between storm track and precipitation over the Catskills. Instead, precipitation in the wetter members was driven primarily by mesoscale processes and terrain effects (strong upslope forcing and moisture convergence), while precipitation in the drier members was driven primarily by quasigeostrophic forcing for ascent. The wetter members featured easterly low-level flow directly into the sharp terrain gradient of the eastern Catskills, creating large upslope magnitudes, while the drier members had more northerly low-level flow, encountering more shallow slopes. Likewise, the wetter members simulated stronger moisture transport into the region than the drier members, and greater convergence of that moisture (the larger upslope magnitudes likely contributed to this convergence). The drier members, on the other hand, had stronger Q-vector convergence than the wetter members, and therefore greater quasigeostrophic forcing for ascent. All of this illustrates the importance of terrain and mesoscale processes in producing large rainfall totals, as the members with a favorable upslope and moisture convergence setup were able to produce substantially more precipitation than the quasigeostophically forced cluster.

In summary, when a synoptic-scale GFS ensemble is used to forecast precipitation in Irene, variability between members is predominantly due to differences in storm position, which arise through interactions with midlatitude synoptic features. The mesoscale WRF model, on the other hand, produces variations that are more due to differences in terrain effects and simulated

mesoscale processes brought about by differences in storm track. With both models, a small east-west deviation in a member's storm track can have a huge effect on the amount of rain received by a particular location. The largest precipitation accumulations in an ensemble are obtained by positioning the storm track such that high rainfall rates are already in place over a region, and then enhancing those rainfall rates with strong upslope forcing. When forecasting precipitation in landfalling tropical cyclones, therefore, a forecaster must take into consideration whether a TC's track will place it in a region where upslope forcing can locally enhance its synoptic rain swath. The worst flooding can result when a storm's low-level winds are perpendicular to a steep terrain gradient in an already moist environment, and the enhanced precipitation over that terrain can later overwhelm downstream watersheds. In addition, clustering ensemble members into specific forecast scenarios can reveal more information than just using the ensemble mean and standard deviation, as it allows forecasters to evaluate several possible forecasts that appear in an ensemble.

Acknowledgements

I would like to thank my two advisors, Dr. Ryan Torn and Dr. Kristen Corbosiero, for providing me with data, support, and guidance over the past three years, and for helping me to develop some semblance of research skills during my time at the University at Albany. I would also like to thank Philip Pegion from the Cooperative Institute for Research in Environmental Sciences (CIRES) for providing me with my GFS ensemble. Thanks to Sarah Ditchek, Rosimar Rios-Berrios, Hannah Huelsing, Kaitlyn Krzyzaniak, Jeremy Berman, the population of office ES234, and the rest of my fellow graduate students for the moral support and help when I was stuck, the faculty and staff of the UAlbany Department of Atmospheric and Environmental Sciences for their work in creating an excellent atmospheric science graduate program, and my family for 25 years of encouragement.

References

Archambault, H. M., L. F. Bosart, D. Keyser, and J. M. Cordeira, 2013: A climatological analysis of the extratropical flow response to recurving Western North Pacific tropical cyclones. *Mon. Wea. Rev.*, **141**, 2325–2346, doi:10.1175/MWR-D-12-00257.1.

——, ——, ——, C. A. Davis, and J. M. Cordeira, 2015: A composite perspective of the extratropical flow response to recurving Western North Pacific tropical cyclones. *Mon. Wea. Rev.*, **143**, 1122–1141, doi:10.1175/MWR-D-14-00270.1.

Atallah, E. H., and L. F. Bosart, 2003: The extratropical transition and precipitation distribution of Hurricane Floyd (1999). *Mon. Wea. Rev.*, **131**, 1063–1081, doi:10.1175/1520-0493(2003)131<1063:TETAPD>2.0.CO;2.

——, ——, and A. R. Aiyyer, 2007: Precipitation distribution associated with landfalling tropical cyclones over the eastern united states. *Mon. Wea. Rev.*, **135**, 2185–2206, doi:10.1175/MWR3382.1.

Avila, L. A., and J. P. Cangialosi, 2011: *Tropical cyclone report Hurricane Irene (AL092011) 21 – 28 August 2011*. 1-45 pp.

Bassill, N. P., 2014: Accuracy of early GFS and ECMWF Sandy (2012) track forecasts: Evidence for a dependence on cumulus parameterization. *Geophys. Res. Lett.*, **41**, 3274–3281, doi:10.1002/2014GL059839.

Benjamin, S. G., and Coauthors, 2004: An Hourly Assimilation–Forecast Cycle: The RUC. *Mon. Wea. Rev.*, **132**, 495–518, doi:10.1175/1520-0493(2004)132<0495:AHACTR>2.0.CO;2.

Bishop, C. H., B. J. Etherton, and S. J. Majumdar, 2001: Adaptive sampling with the ensemble transform Kalman filter. Part I: Theoretical aspects. *Mon. Weather Rev.*, **129**, 420–436, doi:10.1175/1520-0493(2001)129<0420:ASWTET>2.0.CO;2.

Blake, E. S., T. B. Kimberlain, R. J. Berg, Cangia, J. P. Losi, and J. L. Beven II, 2013: *Tropical cyclone report Hurricane Sandy (AL182012) 22 – 29 October 2012*. 1-157 pp.

Bosart, L. F., and D. B. Dean, 1991: The Agnes rainstorm of June 1972 - surface-feature evolution culminating in inland storm redevelopment. *Wea. Forecasting*, **6**, 515–537, doi:10.1175/1520-0434(1991)006<0515:TAROJS>2.0.CO%3B2.

Carpenter, R. L., and Coauthors, 2004: A globally relocatable numerical weather prediction system based on WRF and ADAS. *20th Conference on Weather Analysis and Forecasting and 16th Conference on Numerical Weather Prediction*, Seattle, WA.

Carr, L. E., and R. L. Elsberry, 2000: Dynamical tropical cyclone track forecast errors. Part II: Midlatitude circulation influences. *Wea. Forecasting*, **15**, 662–681, doi:10.1175/1520-0434(2000)015<0662:DTCTFE>2.0.CO;2.

Chen, L. S., Y. Li, and Z. Q. Cheng, 2010: An overview of research and forecasting on rainfall associated with landfalling tropical cyclones. *Adv. Atmos. Sci.*, **27**, 967–976, doi:10.1007/s00376-010-8171-y.

Collins, C., Hurricane Hazel, October 15, 1954. *NWS Newport/Morehead*,. http://www.weather.gov/mhx/Oct151954EventReview (Accessed December 29, 2016).

Davis, C. A., 1992: Piecewise potential vorticity inversion. *J. Atmos. Sci.*, **119**, 1397–1411, doi:10.1175/1520-0469(1992)049<1397:PPVI>2.0.CO;2.

——, S. C. Jones, and M. Riemer, 2008: Hurricane vortex dynamics during atlantic extratropical transition. *J. Atmos. Sci.*, **65**, 714–736, doi:10.1175/2007JAS2488.1.

——, D. A. Ahijevych, W. Wang, and W. C. Skamarock, 2016: Evaluating medium-range tropical cyclone forecasts in uniform- and variable-resolution global models. *Mon. Wea. Rev.*, **144**, 4141–4160, doi:10.1175/MWR-D-16-0021.1.

DiMego, G. J., and L. F. Bosart, 1982a: The transformation of Tropical Storm Agnes into an extratropical cyclone. Part I: The observed fields and vertical motion computations. *Mon. Wea. Rev.*, **110**, 385–411, doi:10.1175/1520-0493(1982)110<0385:TTOTSA>2.0.CO;2.

——, and ——, 1982b: The transformation of Tropical Storm Agnes into an extratropical cyclone. Part II: Moisture, vorticity and kinetic energy budgets. *Mon. Wea. Rev.*, **110**, 412–433, doi:10.1175/1520-0493(1982)110<0412:TTOTSA>2.0.CO;2.

Eckhardt, S., A. Stohl, H. Wernli, P. James, C. Forster, and N. Spichtinger, 2004: A 15-year climatology of warm conveyor belts. *J. Climate.*, **17**, 218–237, doi:10.1175/1520-0442(2004)017<0218:AYCOWC>2.0.CO;2.

Epstein, E. S., 1969: A scoring system for probability forecasts of ranked categories. *J. Appl. Meteor.*, **8**, 985–987, doi:10.1175/1520-0450(1969)008<0985:ASSFPF>2.0.CO;2.

Evensen, G., 2003: The Ensemble Kalman Filter: Theoretical formulation and practical implementation. *Ocean Dyn.*, **53**, 343–367, doi:10.1007/s10236-003-0036-9.

Galarneau, T. J., and C. A. Davis, 2013: Diagnosing forecast errors in tropical cyclone motion. *Mon. Wea. Rev.*, **141**, 405–430, doi:10.1175/MWR-D-12-00071.1.

Gleeson, T. A., 1970: Statistical-dynamical predictions. *J. Appl. Meteor.*, **9**, 333–344, doi:10.1175/1520-0450(1970)009<0333:SDP>2.0.CO;2.

Harr, P. A., and R. L. Elsberry, 1996: Structure of a mesoscale convective system embedded in Typhoon Robyn during TCM-93. *Mon. Wea. Rev.*, **124**, 634–652, doi:10.1175/1520-0493(1996)124<0634:SOAMCS>2.0.CO;2.

——, D. Anwender, and S. C. Jones, 2008: Predictability associated with the downstream impacts of the extratropical transition of tropical cyclones: Methodology and a case study of Typhoon Nabi (2005). *Mon. Wea. Rev.*, **136**, 3205–3225, doi:10.1175/2008MWR2248.1.

Hart, R. E., 2003: A cyclone phase space derived from thermal wind and thermal asymmetry. *Mon. Wea. Rev.*, **131**, 585–616, doi:10.1175/1520-0493(2003)131<0585:ACPSDF>2.0.CO;2.

——, and J. L. Evans, 2001: A climatology of the extratropical transition of Atlantic tropical cyclones. *J. Climate.*, **14**, 546–564.

Hartigan, J. A., and M. A. Wong, 1979: Algorithm AS 136: A K-means clustering algorithm. *J. R. Stat. Soc. Ser. C (Applied Stat.*, **28**, 100–108, doi:10.2307/2346830.

Hoskins, B. J., and M. A. Pedder, 1980: The diagnosis of middle latitude synoptic development. *Quart. J. Roy. Meteor. Soc.*, **106**, 707–719, doi:10.1002/qj.49710645004.

——, M. E. McIntyre, and A. W. Robertson, 1985: On the use of isentropic potential vorticity maps. *Quart. J. Roy. Meteor. Soc.*, **111**, 877–946.

Jones, S. C., and Coauthors, 2003: The extratropical transition of tropical cyclones: Forecast challenges, current understanding, and future directions. *Wea. Forecasting*, **18**, 1052–1092, doi:10.1175/1520-0434(2003)018<1052:TETOTC>2.0.CO;2.

Kain, J. S., 2004: The Kain–Fritsch Convective Parameterization: An update. *J. Appl. Meteor.*, **43**, 170–181, doi:10.1175/1520-0450(2004)043<0170:TKCPAU>2.0.CO;2.

Klein, P. M., P. A. Harr, and R. L. Elsberry, 2000: Extratropical transition of Western North Pacific tropical cyclones: An overview and conceptual model of the transformation stage. *Wea. Forecasting*, **15**, 373–395, doi:10.1175/1520-0434(2000)015<0373:ETOWNP>2.0.CO;2.

Krishnamurti, T. N., 1995: Numerical weather prediction. *Annu. Rev. Fluid Mech.*, **27**, 195–224.

——, C. M. Kishtawal, Z. Zhang, T. Larow, D. Bachiochi, and E. Williford, 2000: Multimodel ensemble forecasts for weather and seasonal climate. *J. Climate.*, **13**, 4196–4216, doi:10.1175/1520-0442(2000)013<4196:MEFFWA>2.0.CO;2.

Landsea, C. W., and J. L. Franklin, 2013: Atlantic hurricane database uncertainty and presentation of a new database format. *Mon. Wea. Rev.*, **141**, 3576–3592, doi:10.1175/MWR-D-12-00254.1.

Lin, C. Y., H. M. Hsu, Y. F. Shengl, C. H. Kuo, and Y. A. Liou, 2011: Mesoscale processes for super heavy rainfall of typhoon Morakot (2009) over southern Taiwan. *Atmos. Chem. Phys.*, **11**, 345–361, doi:10.5194/acp-11-345-2011.

Lin, Y., 2011: GCIP/EOP Surface: Precipitation NCEP/EMC 4KM Gridded Data (GRIB) Stage IV Data, Version 1.0. *UCAR/NCAR - Earth Obs. Lab.*, doi:http://data.eol.ucar.edu/dataset/21.093.

Lin, Y.-L., S. Chiao, T.-A. Wang, M. L. Kaplan, and R. P. Weglarz, 2001: Some common ingredients for heavy orographic rainfall. *Wea. Forecasting*, **16**, 633–660, doi:10.1175/1520-0434(2001)016<0633:SCIFHO>2.0.CO;2.

Lorenz, E. N., 1963: Deterministic nonperiodic flow. *J. Atmos. Sci.*, **20**, 130–141, doi:10.1175/1520-0469(1963)020<0130:DNF>2.0.CO;2.

Majumdar, S. J., M. J. Brennan, and K. Howard, 2013: The impact of dropwindsonde and supplemental rawinsonde observations on track forecasts for Hurricane Irene (2011). *Wea. Forecasting*, **28**, 1385–1403, doi:10.1175/WAF-D-13-00018.1.

Michalakes, J., S. Chen, J. Dudhia, L. Hart, J. Klemp, J. Middlecoff, and W. Skamarock, 2001: Development of a next-generation regional weather research and forecasting model. *Developments in Teracomputing: Proceedings of the Ninth ECMWF Workshop on the use of high performance computing in meteorology, Vol. 1*, World Scientific, 269–276.

Mlawer, E. J., S. J. Taubman, P. D. Brown, M. J. Iacono, and S. A. Clough, 1997: Radiative transfer for inhomogeneous atmospheres: RRTM, a validated correlated-k model for the longwave. *J. Geophys. Res.*, **102**, 16663, doi:10.1029/97JD00237.

Molteni, F., R. Buizza, T. N. Palmer, and T. Petroliagis, 1996: The ECMWF Ensemble Prediction System: Methodology and validation. *Quart. J. Roy. Meteor. Soc.*, **122**, 73–119, doi:10.1002/qj.49712252905.

Nakanishi, M., and H. Niino, 2009: Development of an Improved Turbulence Closure Model for the Atmospheric Boundary Layer. *J. Meteorol. Soc. Japan*, **87**, 895–912, doi:10.2151/jmsj.87.895.

Palmén, E., 1958: Vertical circulation and release of kinetic energy during the development of Hurricane Hazel into an extratropical storm. *Tellus, Ser. A A Q. J. Geophys.*, **10**, 1–23, doi:10.1111/j.2153-3490.1958.tb01982.x.

Rabier, F., 2005: Overview of global data assimilation developments in numerical weather-prediction centres. *Quart. J. Roy. Meteor. Soc.*, **131**, 3215–3233, doi:10.1256/qj.05.129.

Rappaport, E. N., 2000: Loss of life in the United States associated with recent atlantic tropical cyclones. *Bull. Amer. Meteor. Soc.*, **81**, 2065–2073, doi:10.1175/1520-0477(2000)081<2065:LOLITU>2.3.CO;2.

Riemer, M., S. C. Jones, C. A. Davis, M. Riemer, S. C. Jones, and C. A. Davis, 2008: The impact of extratropical transition on the downstream flow: An idealized modelling study with a straight jet. *Quart. J. Roy. Meteor. Soc.*, **134**, 69–91, doi:10.1002/qj.189.

Ritchie, E. A., and R. L. Elsberry, 2007: Simulations of the extratropical transition of tropical cyclones: Phasing between the upper-level trough and tropical cyclones. *Mon. Wea. Rev.*, **135**, 862–876, doi:10.1175/MWR3303.1.

Ruxton, G. D., 2006: The unequal variance t-test is an underused alternative to Student's t-test and the Mann-Whitney U test. *Behav. Ecol.*, **17**, 688–690, doi:10.1093/beheco/ark016.

Sanders, F., and B. J. Hoskins, 1990: An easy method for estimation of Q-vectors from weather maps. *Wea. Forecasting*, **5**, 346–353, doi:DOI: 10.1175/1520-0434(1990)005<0346:AEMFEO>2.0.CO;2.

Tang, B., M. Vaughan, R. Lazear, K. Corbosiero, L. Bosart, T. Wasula, I. Lee, and K. Lipton, 2016: Topographic and boundary influences on the 22 May 2014 Duanesburg, New York, tornadic supercell. *Wea. Forecasting*, **31**, 107–127, doi:10.1175/WAF-D-15-0101.1.

Thompson, G., R. M. Rasmussen, and K. Manning, 2004: Explicit forecasts of winter precipitation using an improved bulk microphysics scheme. Part I: Description and sensitivity analysis. *Mon. Wea. Rev.*, **132**, 519–542, doi:10.1175/1520-0493(2004)132<0519:EFOWPU>2.0.CO;2.

Torn, R. D., 2010: Diagnosis of the downstream ridging associated with extratropical transition using short-term ensemble forecasts. *J. Atmos. Sci.*, **67**, 817–833, doi:10.1175/2009JAS3093.1.

——, J. S. Whitaker, P. Pegion, T. M. Hamill, and G. J. Hakim, 2015: Diagnosis of the source of GFS medium-range track errors in Hurricane Sandy (2012). *Mon. Wea. Rev.*, **143**, 132–152, doi:10.1175/MWR-D-14-00086.1.

Toth, Z., and E. Kalnay, 1997: Ensemble forecasting at NCEP and the breeding method. *Mon. Wea. Rev.*, **125**, 3297–3319, doi:10.1175/1520-0493(1997)125<3297:EFANAT>2.0.CO;2.

Tracton, M. S., and E. Kalnay, 1993: Operational ensemble prediction at the National Meteorological Center: Practical aspects. *Wea. Forecasting*, **8**, 379–398, doi:10.1175/1520-0434(1993)008<0379:OEPATN>2.0.CO;2.

Whitaker, J. S., and T. M. Hamill, 2002: Ensemble data assimilation without perturbed observations. *Mon. Wea. Rev.*, **130**, 1913–1924, doi:10.1175/1520-0493(2002)130<1913:EDAWPO>2.0.CO;2.

Wu, C.-C., T.-H. Yen, Y.-H. Kuo, and W. Wang, 2002a: Rainfall simulation associated with Typhoon Herb (1996) near Taiwan. Part I: The topographic effect. *Wea. Forecasting*, **17**, 1001–1015, doi:10.1175/1520-0434(2003)017<1001:RSAWTH>2.0.CO;2.

Wu, C. C., and Y. H. Kuo, 1999: Typhoons affecting Taiwan: Current understanding and future challenges. *Bull. Amer. Meteor. Soc.*, **80**, 67–80, doi:10.1175/1520-0477(1999)080<0067:TATCUA>2.0.CO;2.

Wu, W.-S., R. J. Purser, and D. F. Parrish, 2002b: Three-dimensional variational analysis with spatially inhomogeneous covariances. *Mon. Wea. Rev.*, **130**, 2905–2916, doi:10.1175/1520-0493(2002)130<2905:TDVAWS>2.0.CO;2.

Yamaguchi, M., F. Vitart, S. T. K. Lang, L. Magnusson, R. L. Elsberry, G. Elliott, M. Kyouda, and T. Nakazawa, 2015: Global distribution of the skill of tropical cyclone activity forecasts on short- to medium-range time scales. *Wea. Forecasting*, **30**, 1695–1709, doi:10.1175/WAF-D-14-00136.1.

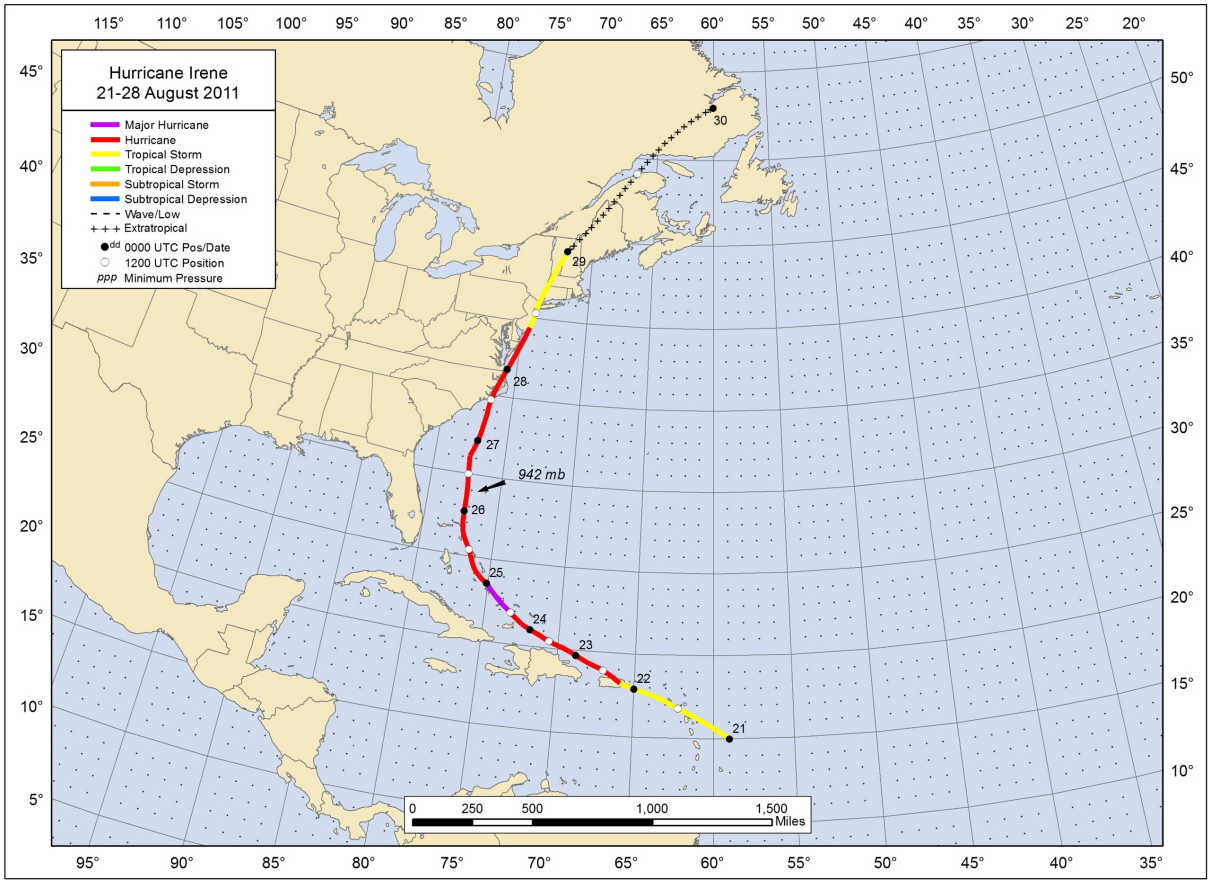

Fig. 1: Track and intensity of Hurricane Irene 0000 UTC 21–29 August 2011. Source: NOAA NHC (http://www.noaanews.noaa.gov/stories2012/images/irene_track_full.jpg). Image is open source.

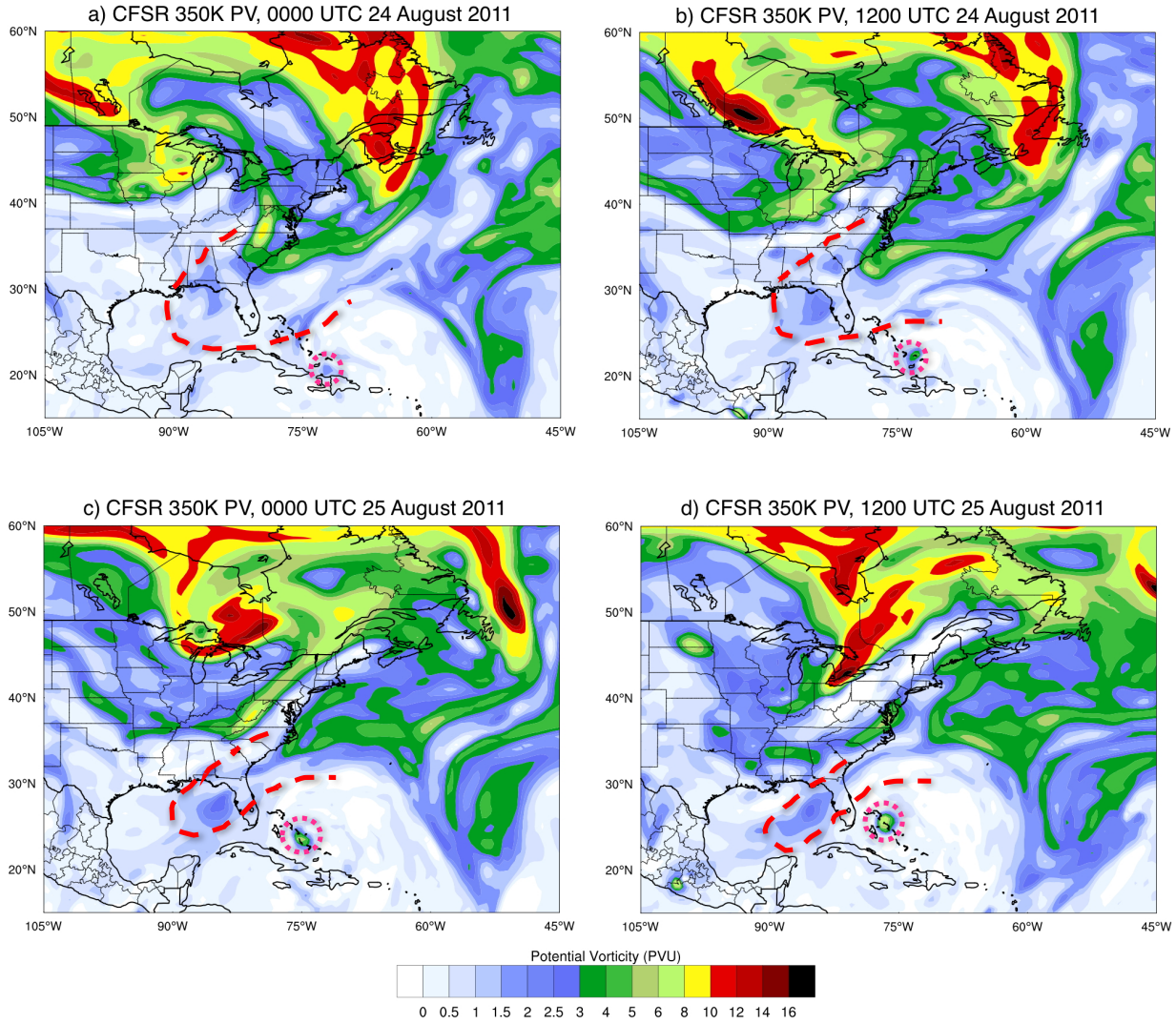

Fig. 2: Climate Forecast System Reanalysis (CFSR) 350-K PV analysis at a) 0000 UTC 24 August, b) 1200 UTC 24 August, c) 0000 UTC 25 August, and d) 1200 UTC 25 August. Hurricane Irene is outlined in pink and the first upstream trough interacting with it is outlined in red.

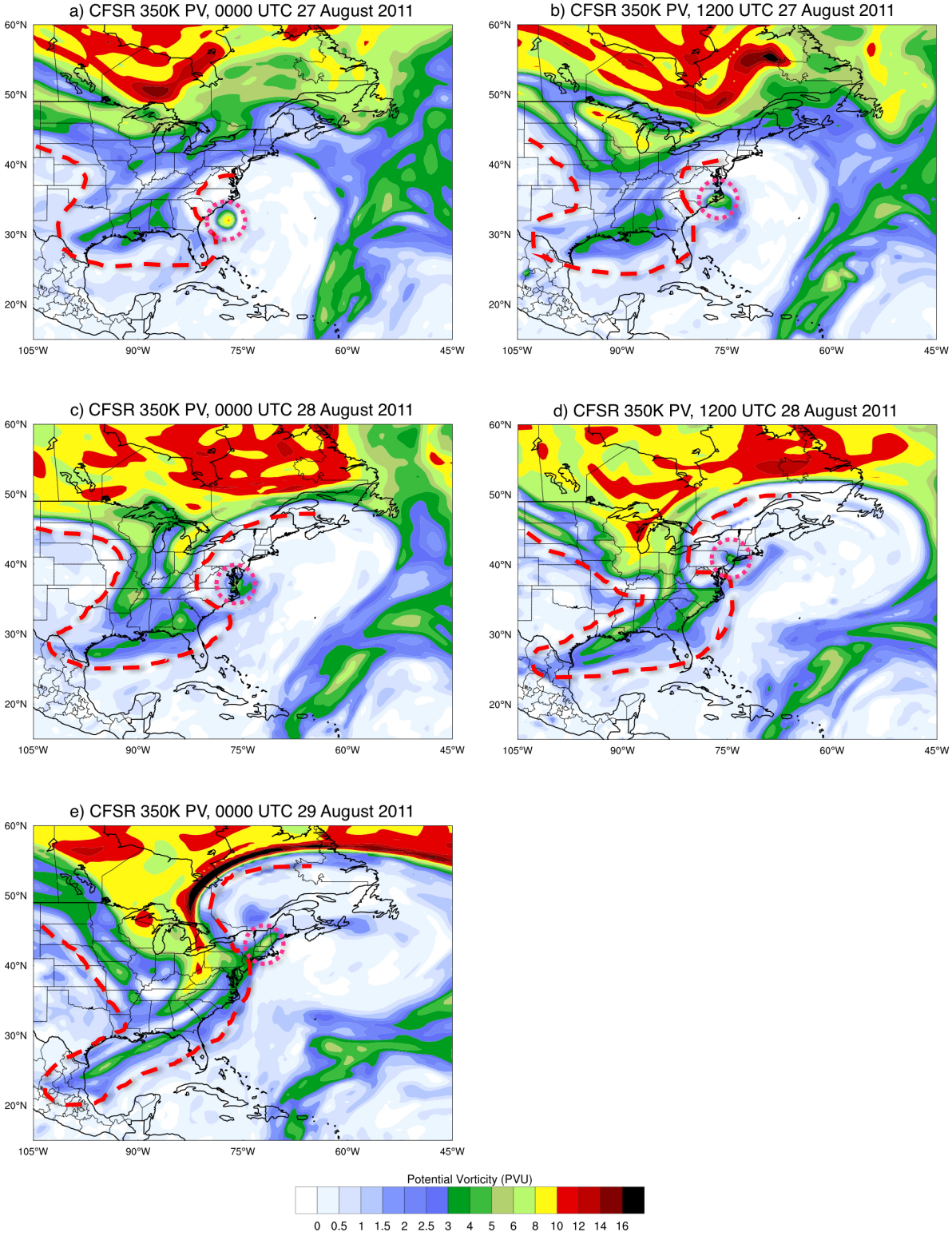

Fig. 3: CFSR 350-K PV analysis at a) 0000 UTC 27 August, b) 1200 UTC 27 August, c) 0000 UTC 28 August, d) 1200 UTC 28 August, and e) 0000 UTC 29 August. Hurricane Irene is outlined in pink and the second upstream trough interacting with it is outlined in red.

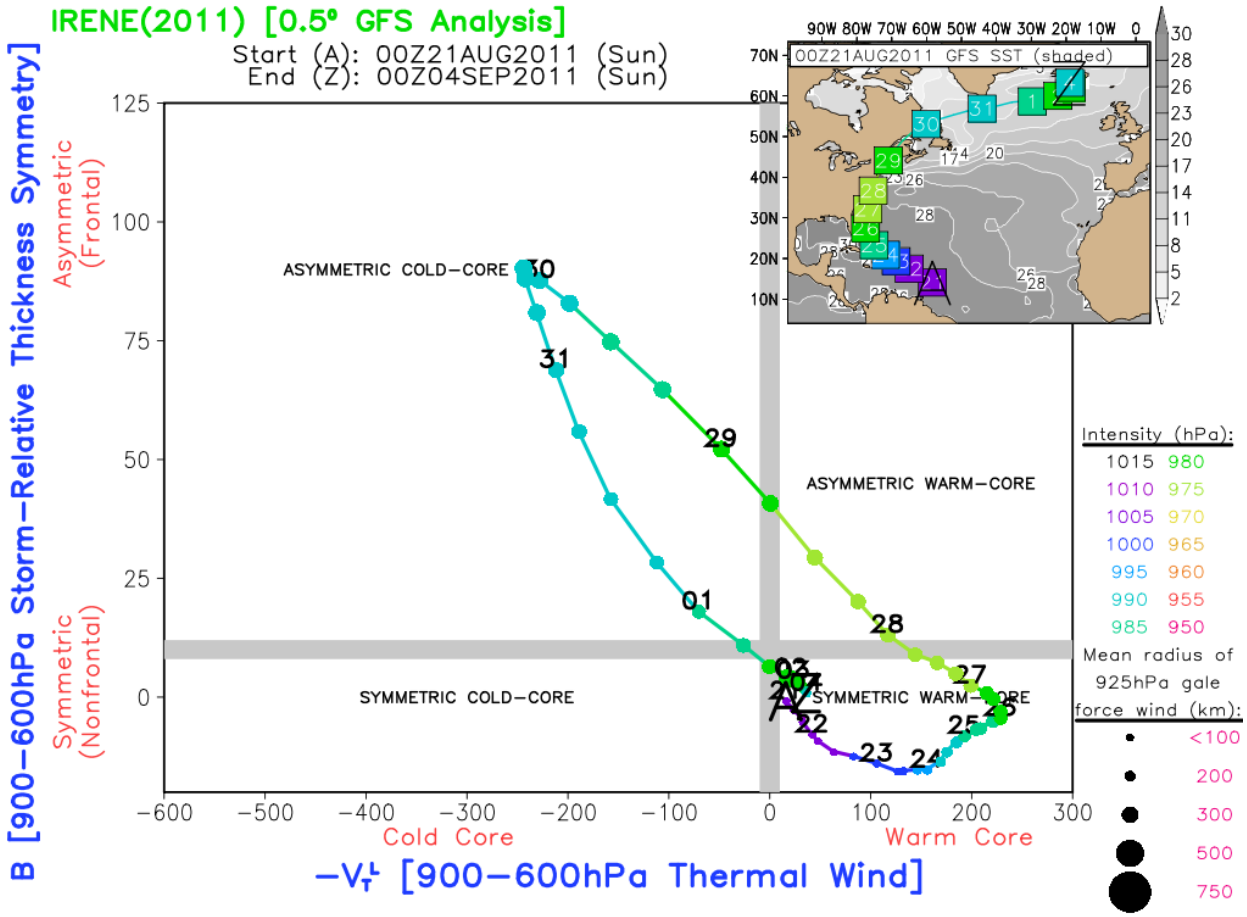

Fig. 4: Hurricane Irene cyclone phase space diagram, showing the storm's transition from a symmetric warm-core system to an asymmetric cold-core system. Source: Florida State University cyclone phase evolution page (http://moe.met.fsu.edu/cyclonephase/archive/2011/gifs/irene2011/irene2011.a.gfs.57.png). Image is open source.

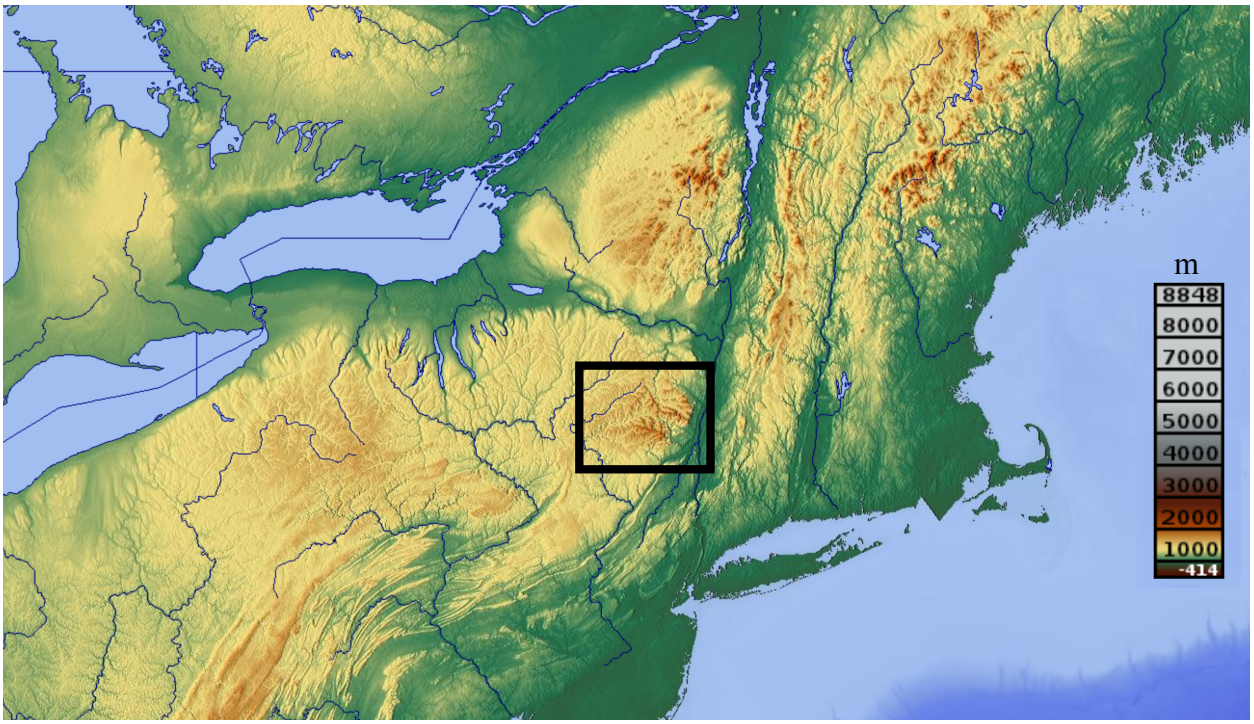

Fig. 5: Topographic map of New York and most of the Northeast. The Catskills are indicated by the black box. Source: maps-for-free.com. Image is open source.

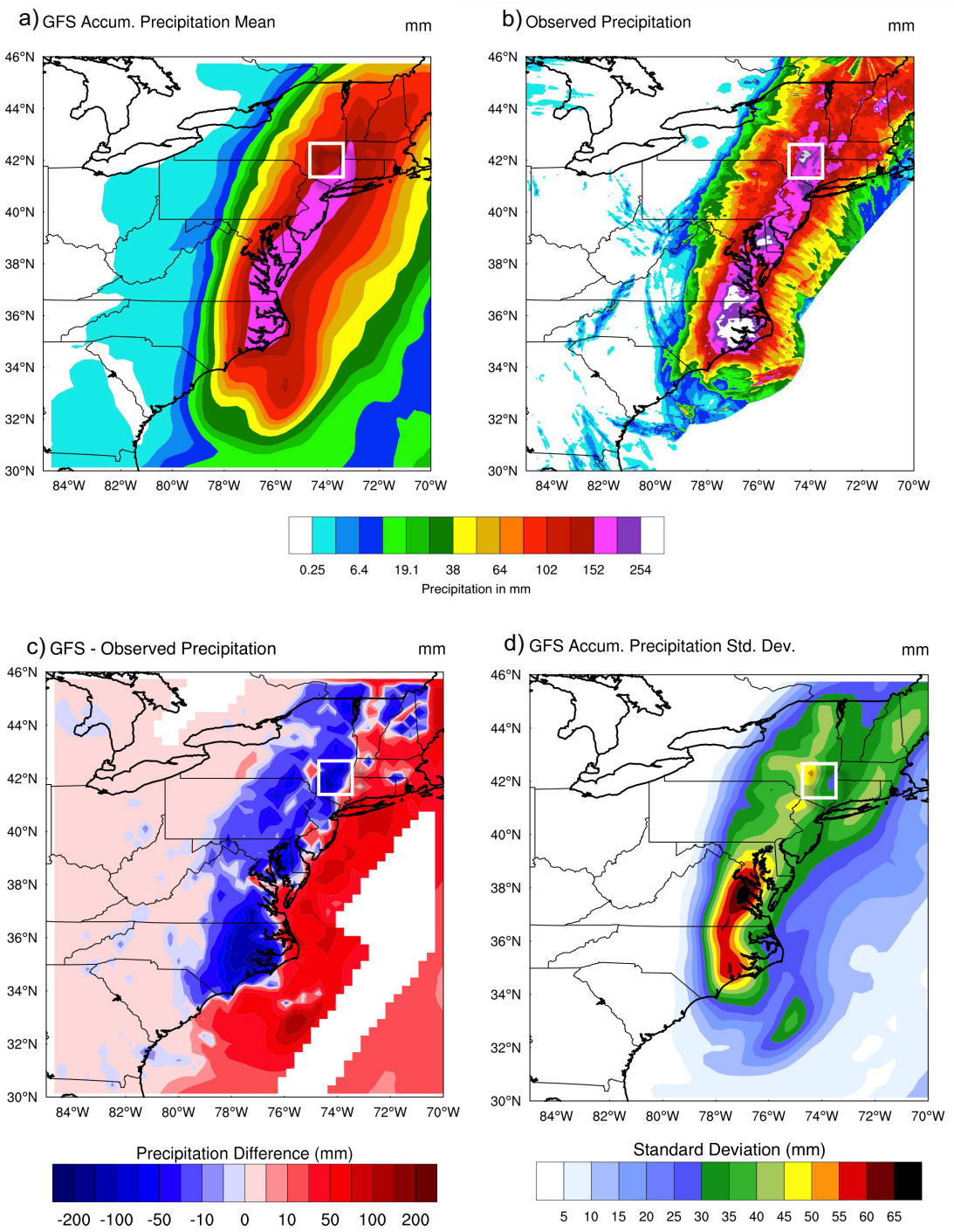

Fig. 6: GFS 27–29 August 2011 a) ensemble mean total accumulated precipitation (mm), b) observed accumulated precipitation (mm), c) difference between the ensemble mean and observed accumulated precipitation (mm), and d) ensemble standard deviation (mm). The Catskills domain is indicated by the white box.

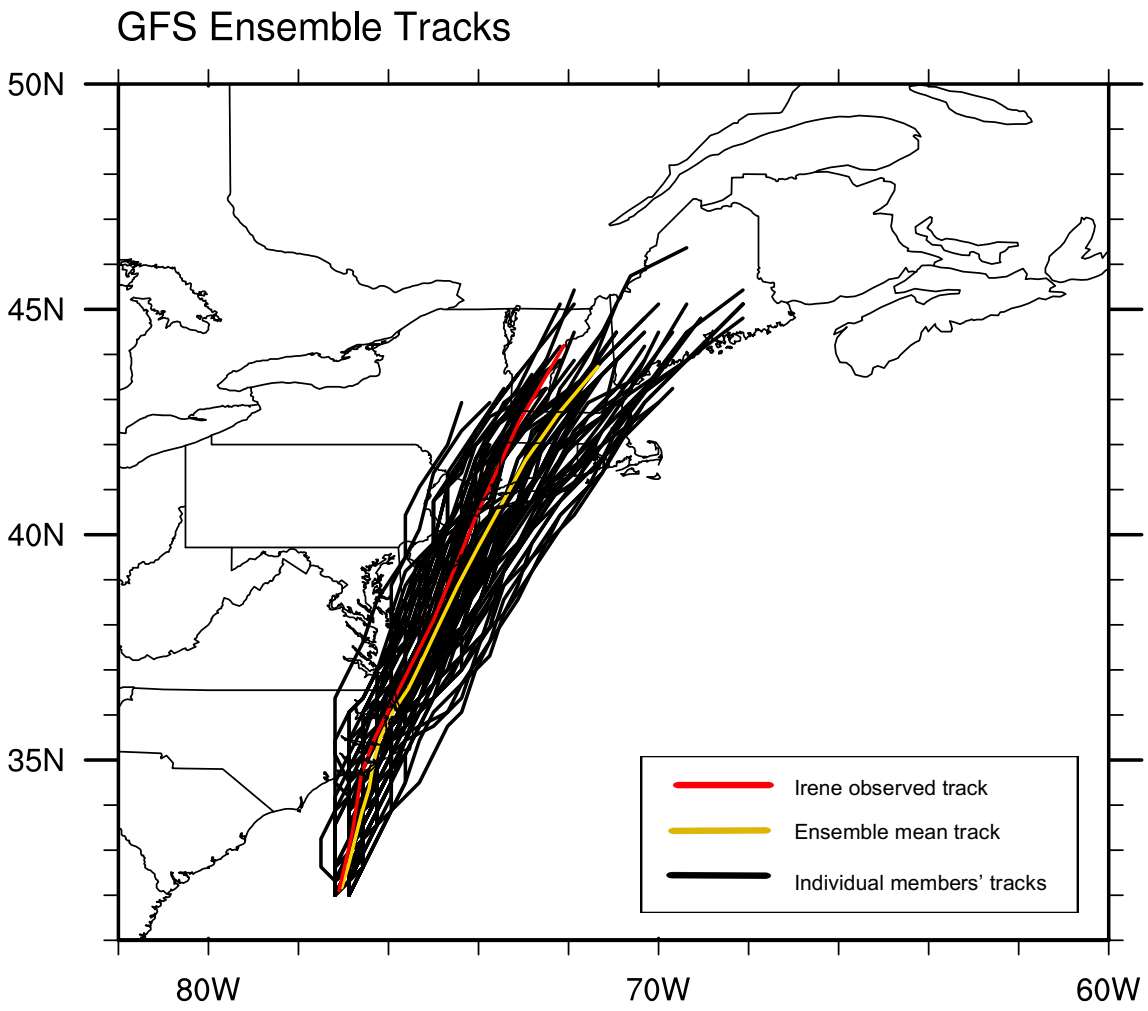

Fig. 7: All 80 storm tracks for Irene produced by the GFS ensemble (black), as well as the ensemble mean track (yellow) and the observed storm track (red).

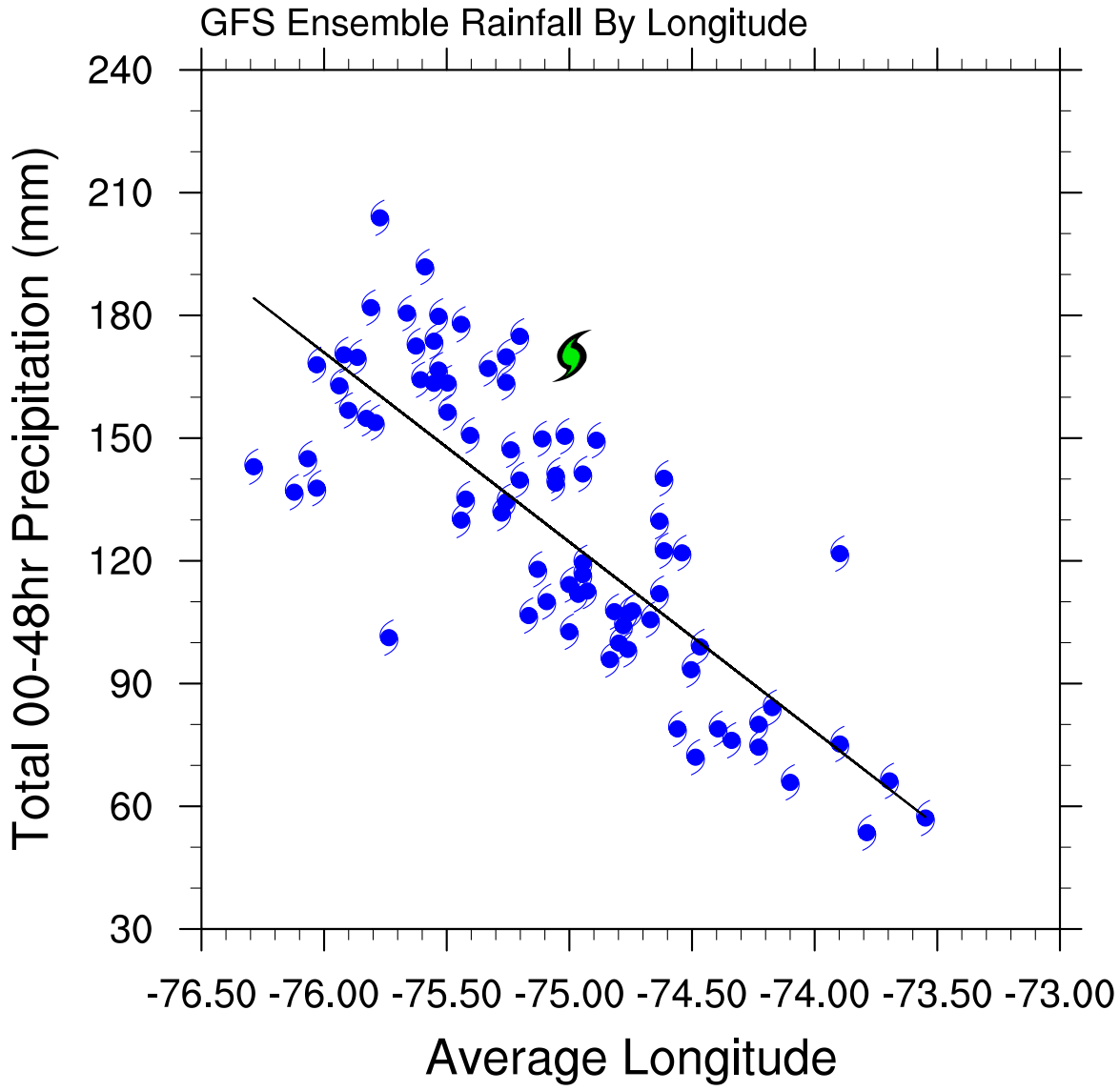

Fig. 8: 27–29 August 2011 accumulated Catskills precipitation (mm) as a function of Irene's average storm track longitude (degrees east) for each GFS ensemble member (blue). Irene's observed position and precipitation is indicated in green.

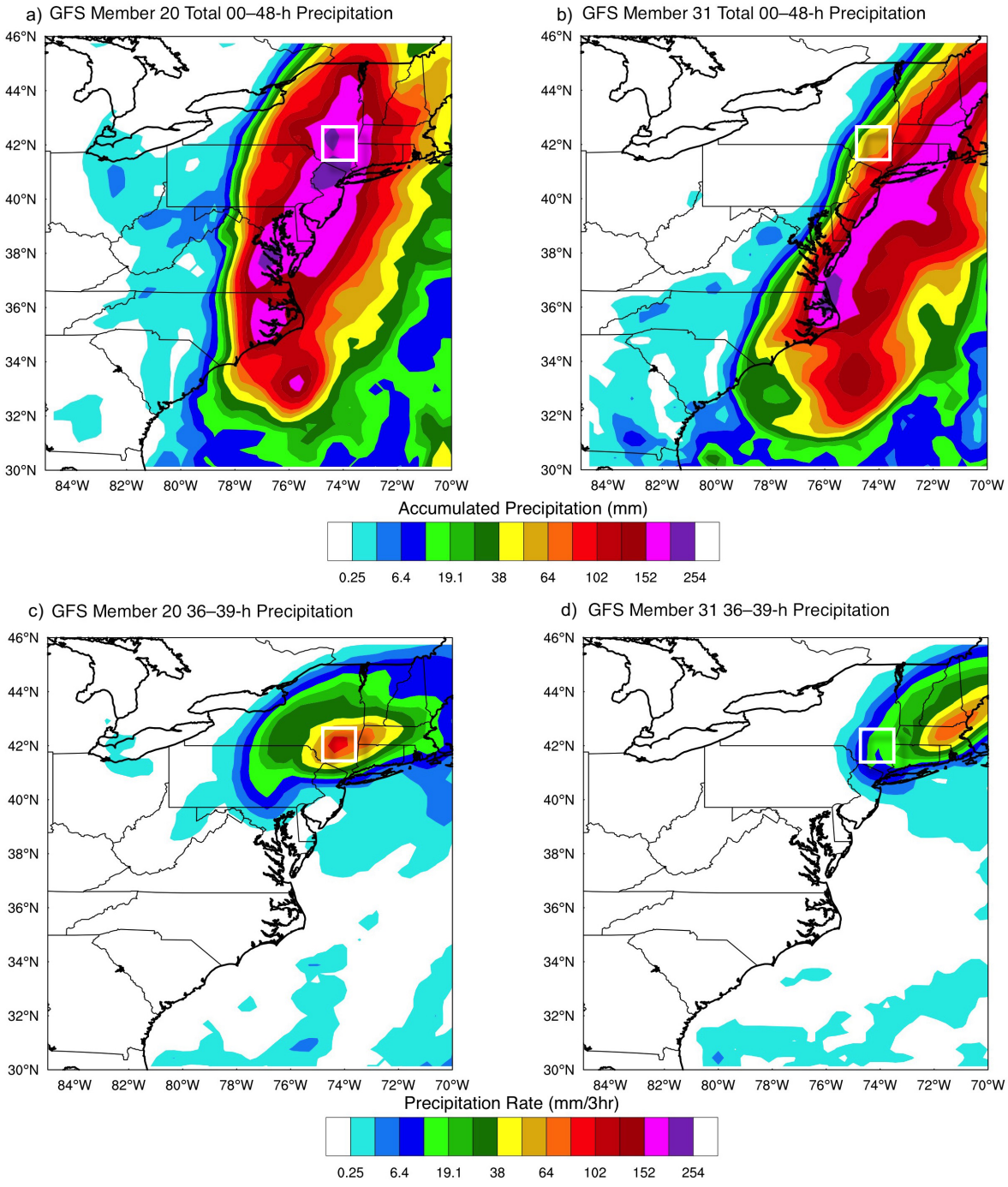

Fig. 9: Total accumulated precipitation (mm; a,b) and 36–39 h rainfall rate (mm/3hr; c,d) for a westward-tracking member (20; a,c) and an eastward-tracking member (31; b,d). The Catskills domain is indicated by the white box.

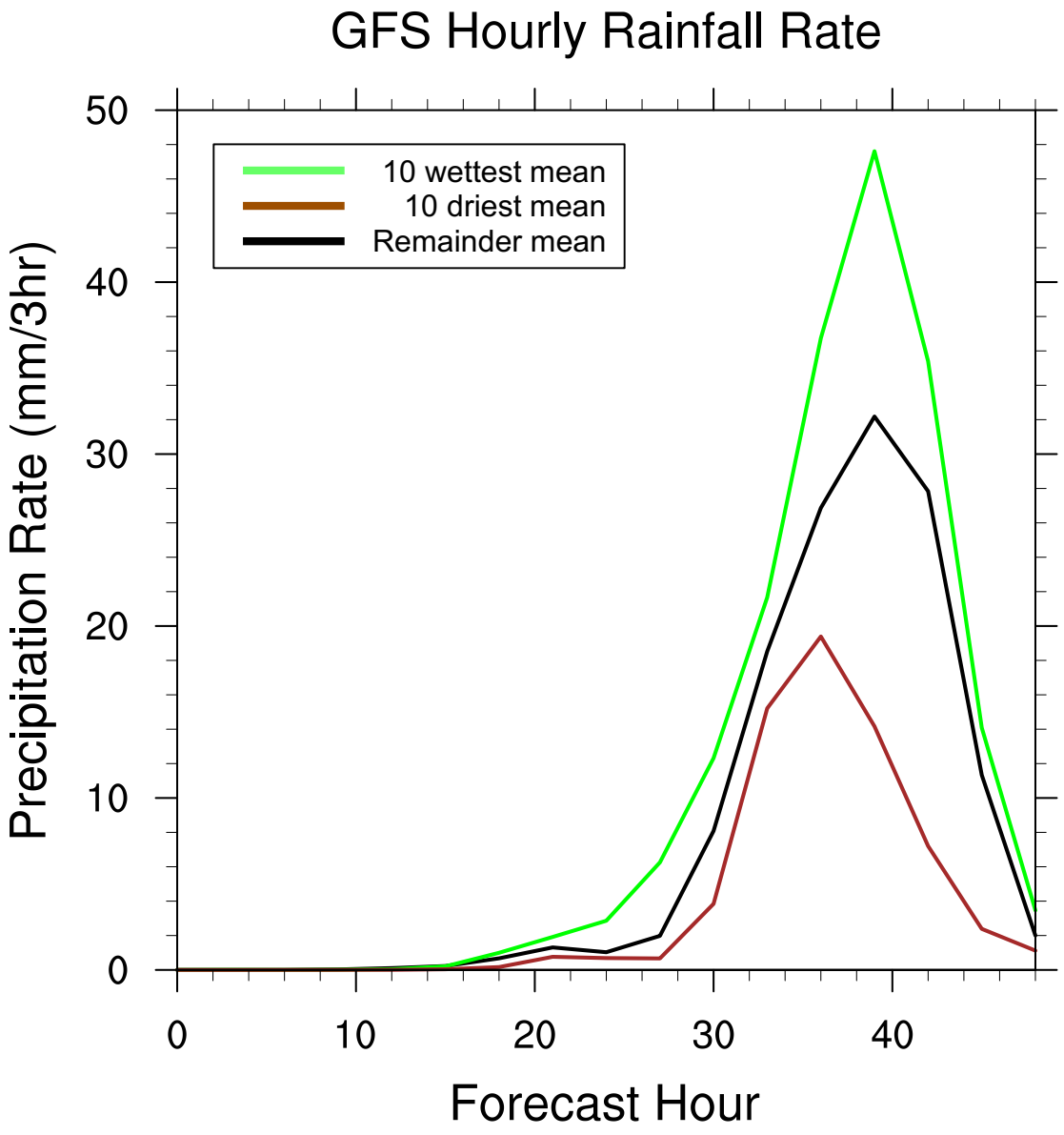

Fig. 10: Rainfall rates (mm/3hr) for the ten wettest (green) and ten driest (brown) GFS ensemble members over the Catskills domain, with the remaining 60 members (black).

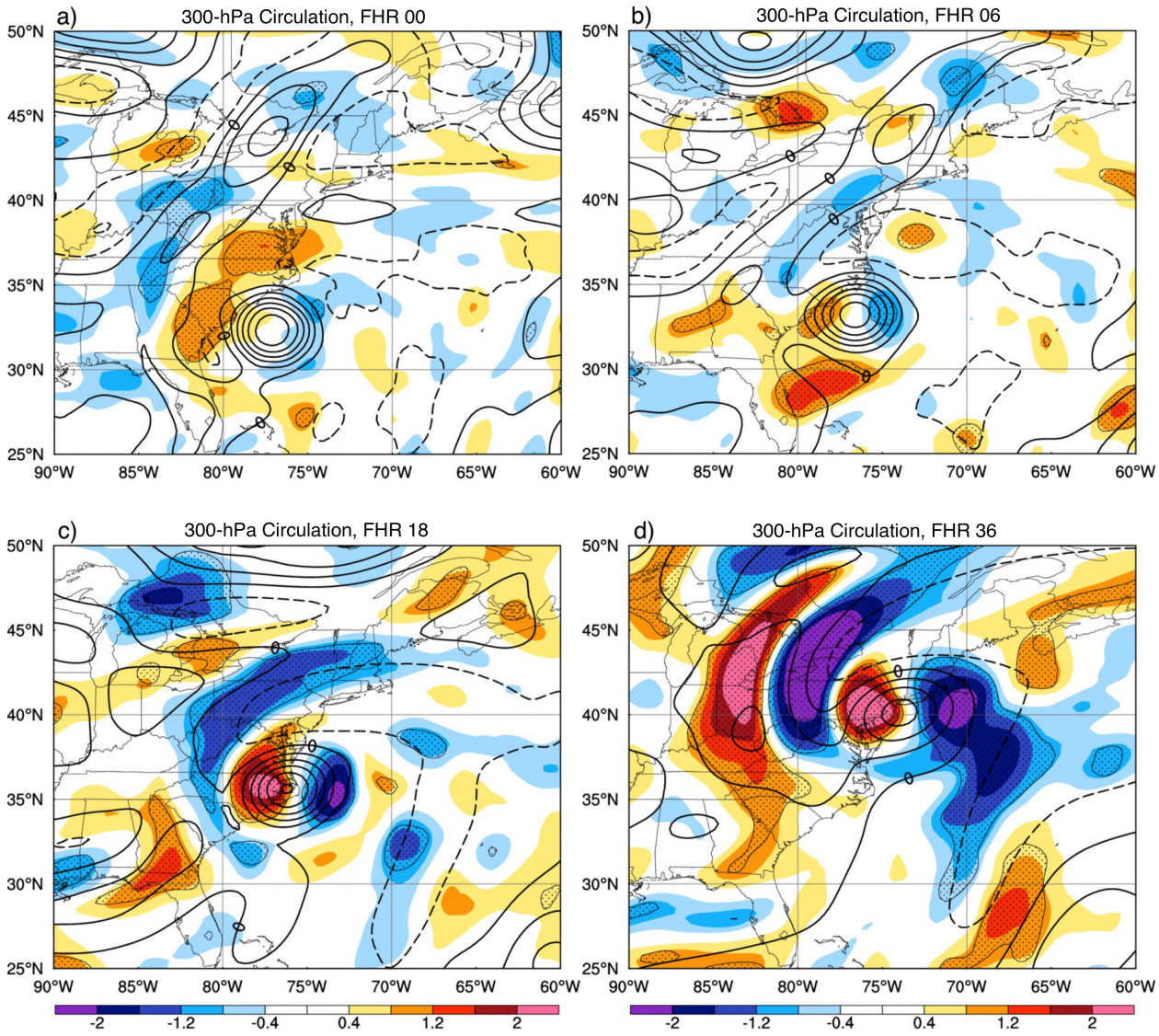

Fig. 11: 300-hPa circulation differences between the 10 wettest and 10 driest members at a) 00-h, b) 06-h, c) 18-h, and d) 36-h. In this plot, contours represent the ensemble mean circulation (s^{-1}), color shading represents the standardized difference between the wet member composite field and the dry member composite field, and stippling represents regions of statistical significance.

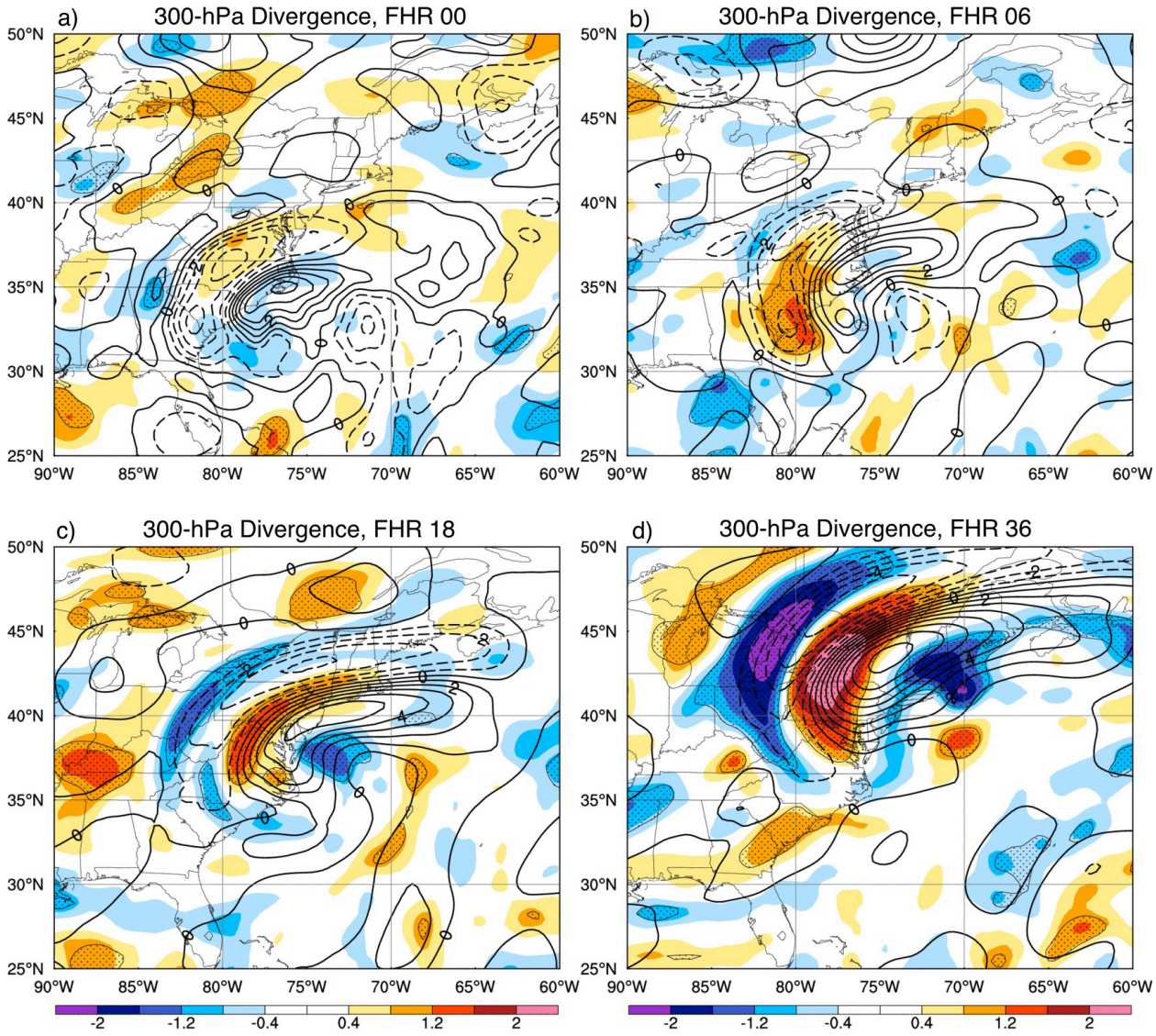

Fig. 12: Same as Fig. 10, but for 300-hPa area-averaged divergence (s^{-1}).

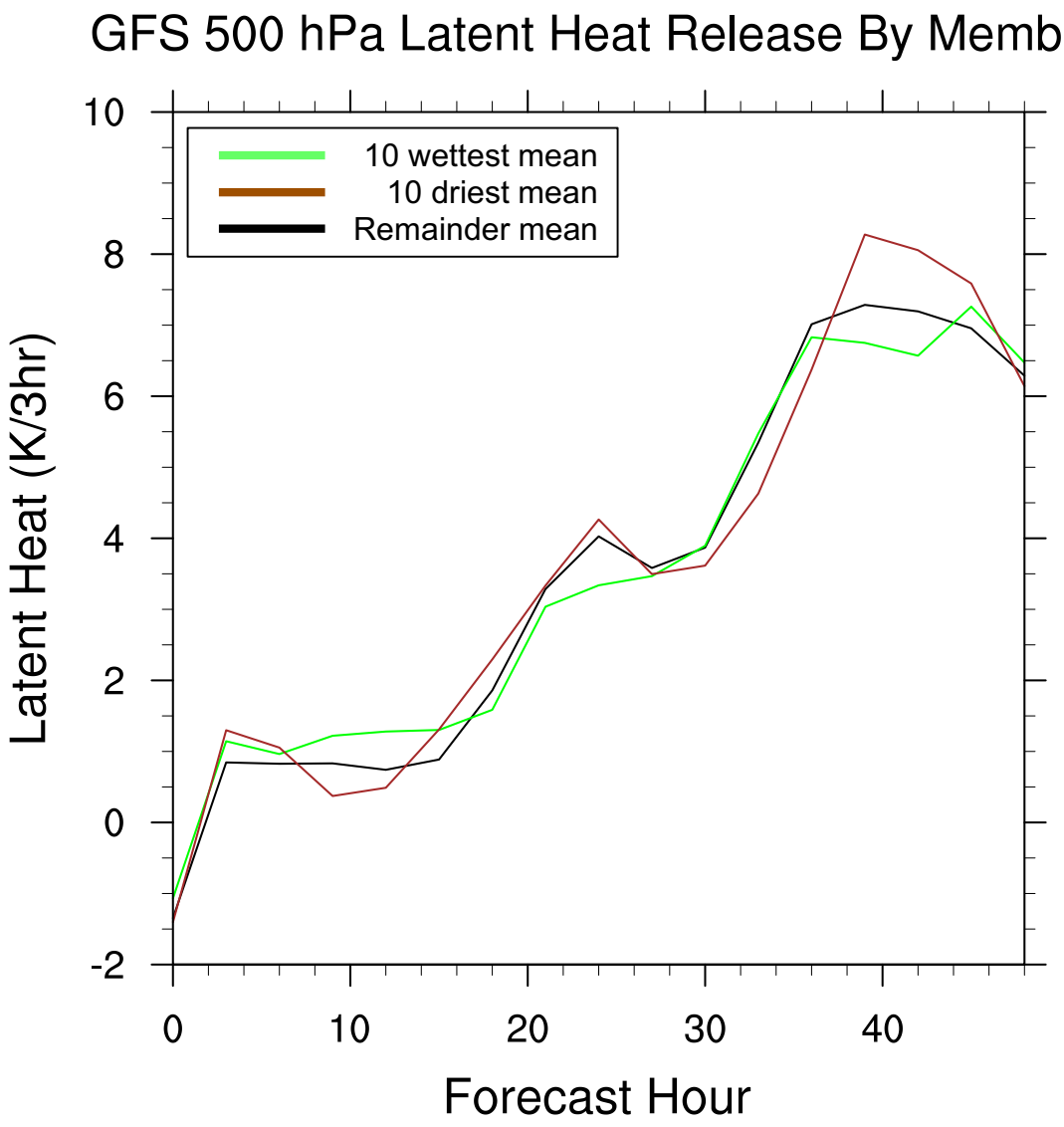

Fig. 13: Time series of latent heat release (K/3hr) at 500-hPa on the western side of Irene.

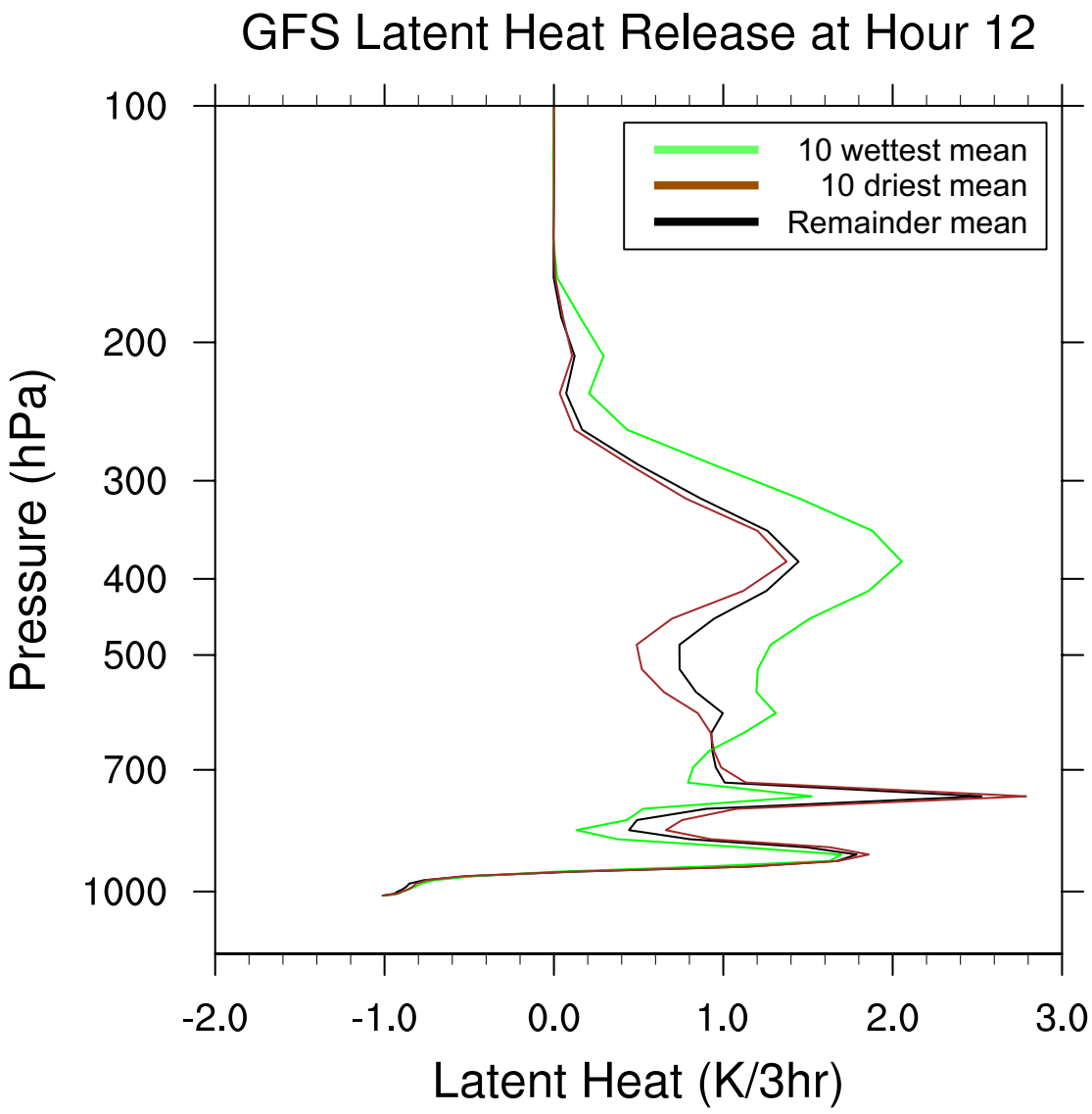

Fig. 14: Vertical profile of latent heat release (K/3hr) on the western side of Irene at 12-h.

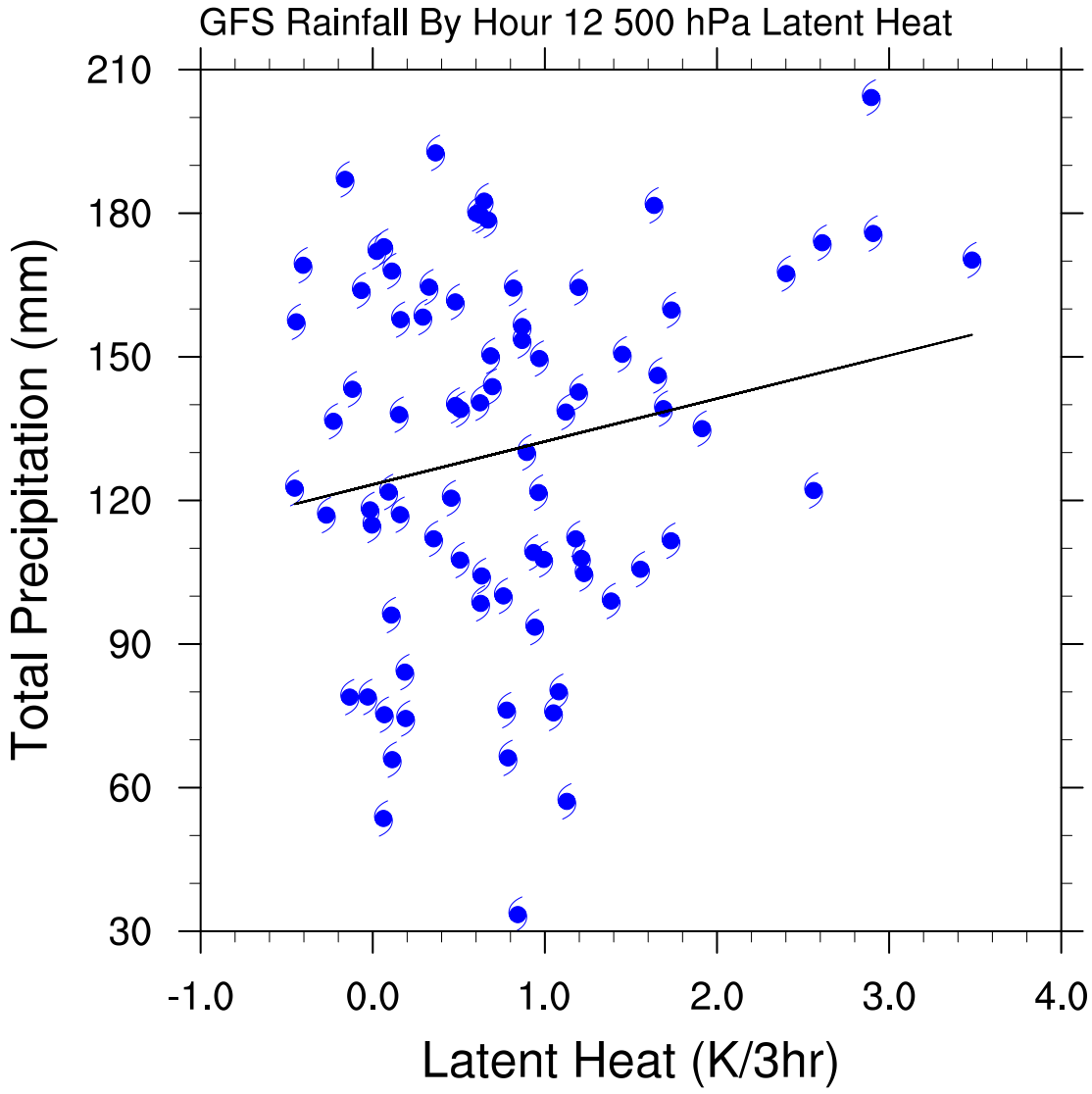

Fig. 15: 00–48 h Catskills precipitation (mm) as a function of 12 h 500-hPa latent heat release (K/3hr) on the western side of Irene.

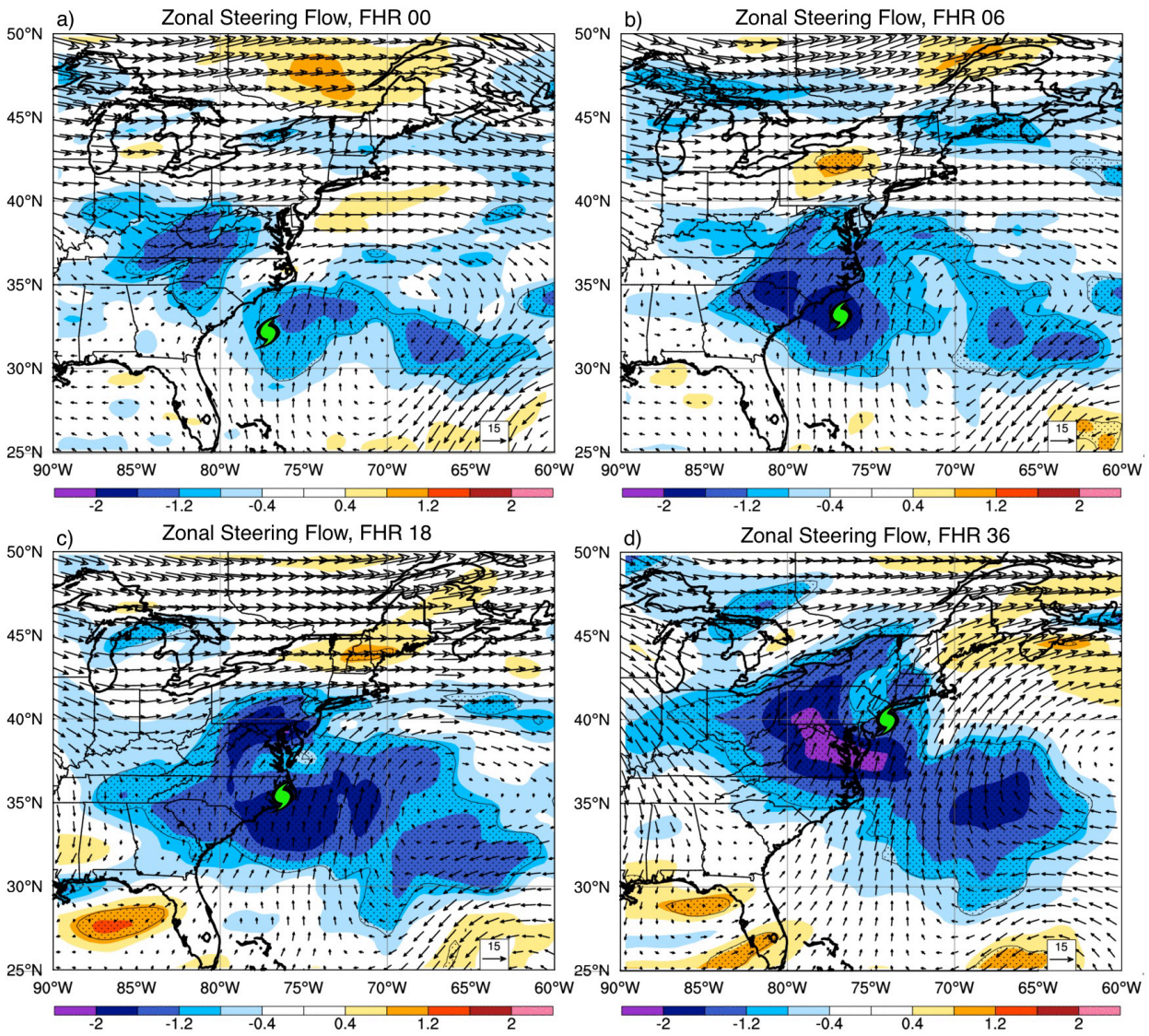

Fig. 16: Same as Fig. 10, but for zonal steering flow (m/s). The green cyclone marker represents the position of Irene at each time.

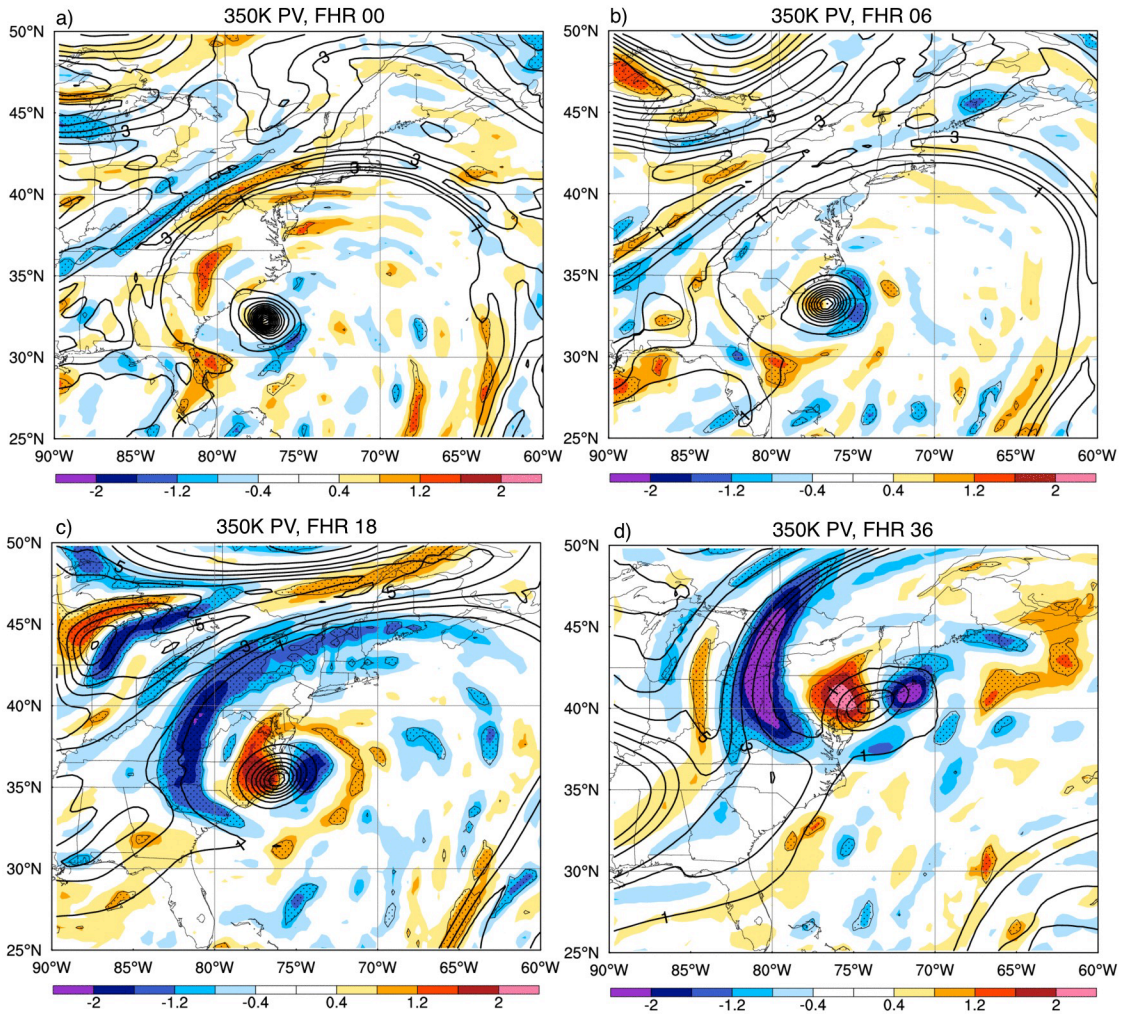

Fig. 17: Same as Fig. 10, but for 350-K PV (units are PVU).

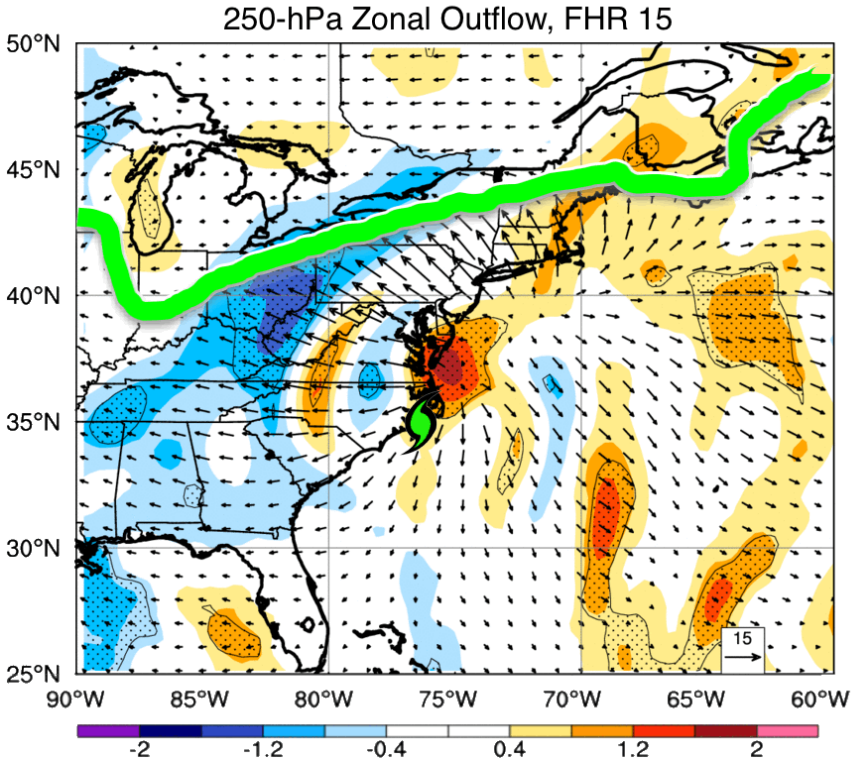

Fig. 18: Same as Fig. 10, but for 250-hPa divergent zonal wind at 15-h. The green line marker represents the position of Irene, and the green contour denotes the 2 PVU contour on this surface at this time.

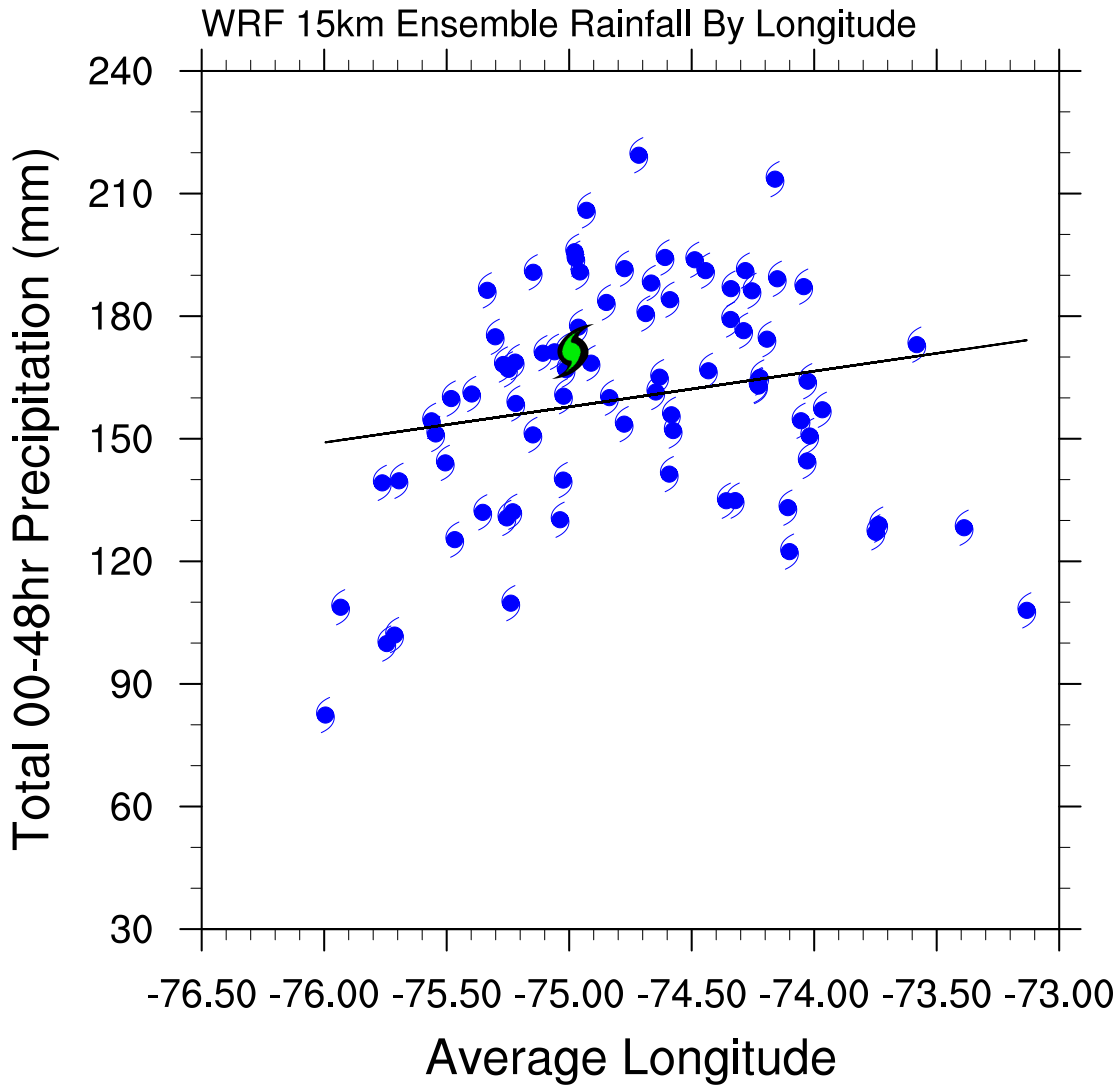

Fig. 19: 27–29 August 2011 accumulated Catskills precipitation (mm) as a function of Irene's average storm track longitude (degrees east) for each WRF 15 km ensemble member (blue). Irene's observed position and precipitation is indicated in green.

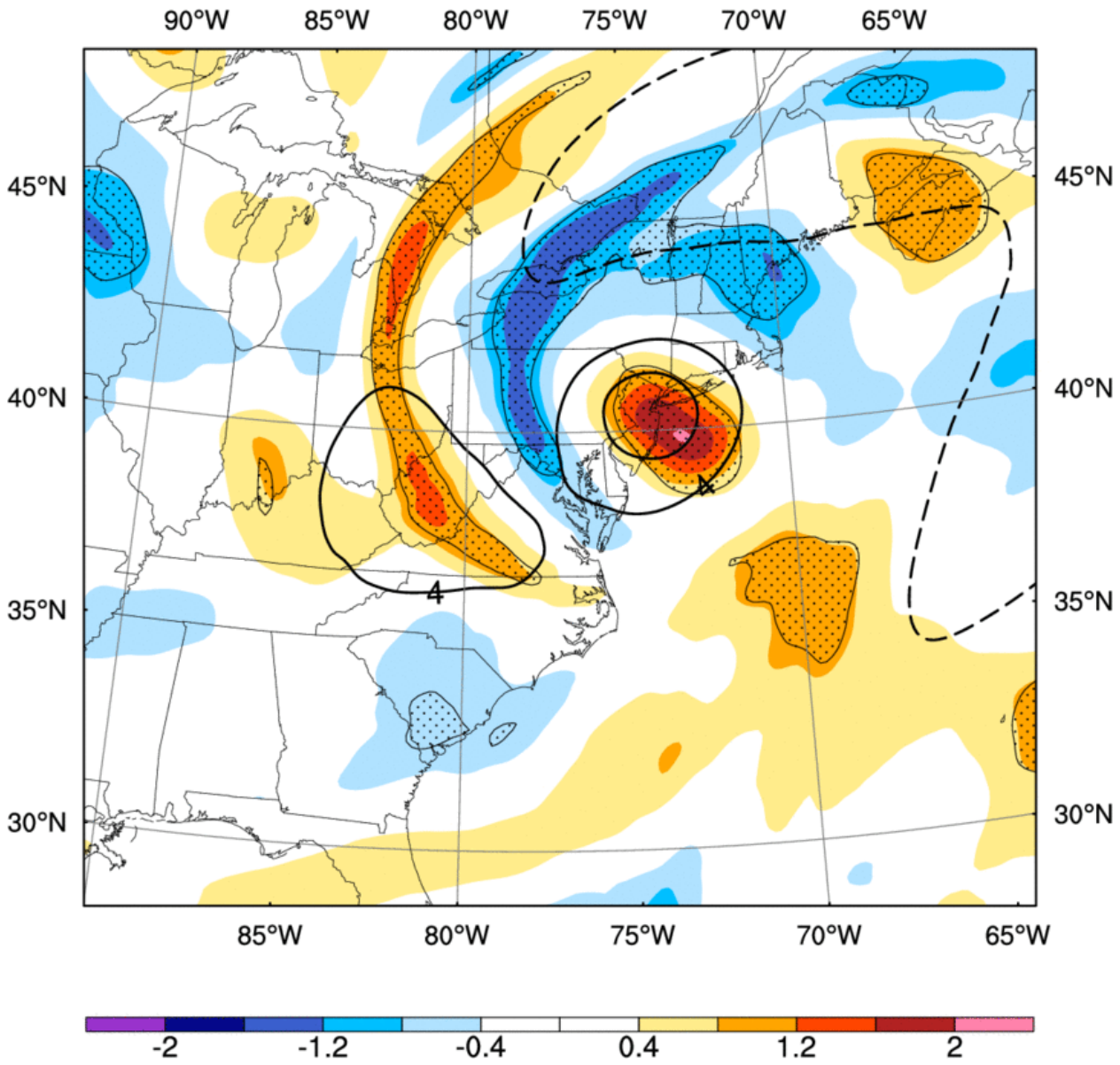

Fig. 20: WRF 15 km 300-hPa standardized circulation (s^{-1}) differences between the 10 wettest and 10 driest members at 36 h into the forecast. In this plot, contours represent the ensemble mean circulation (s^{-1}), color shading represents the standardized difference between the wet member composite field and the dry member composite field, and stippling represents regions of statistical significance.

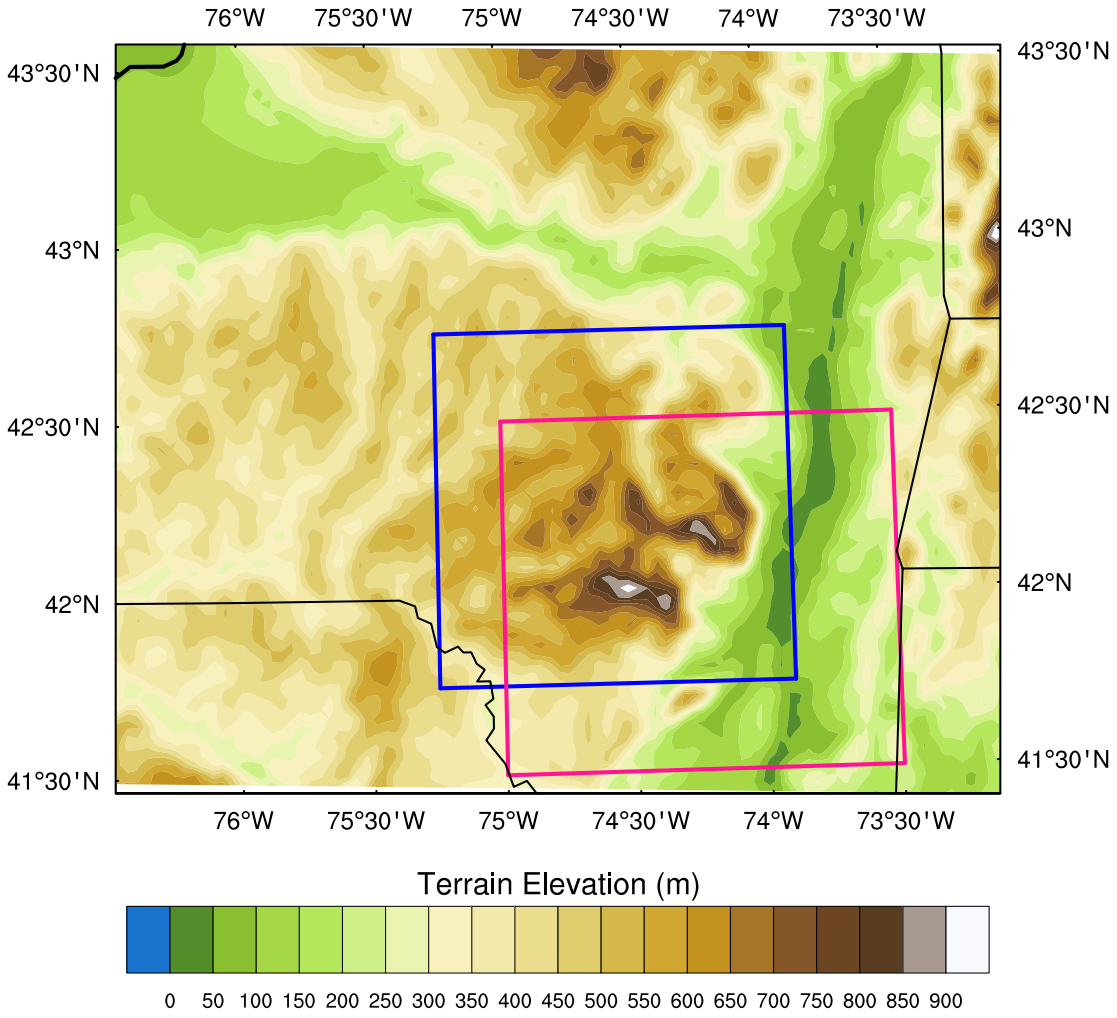

Fig. 21: Defined Catskills bounds for the GFS and WRF 15 km (pink box) and WRF 3 km (blue box) analyses.

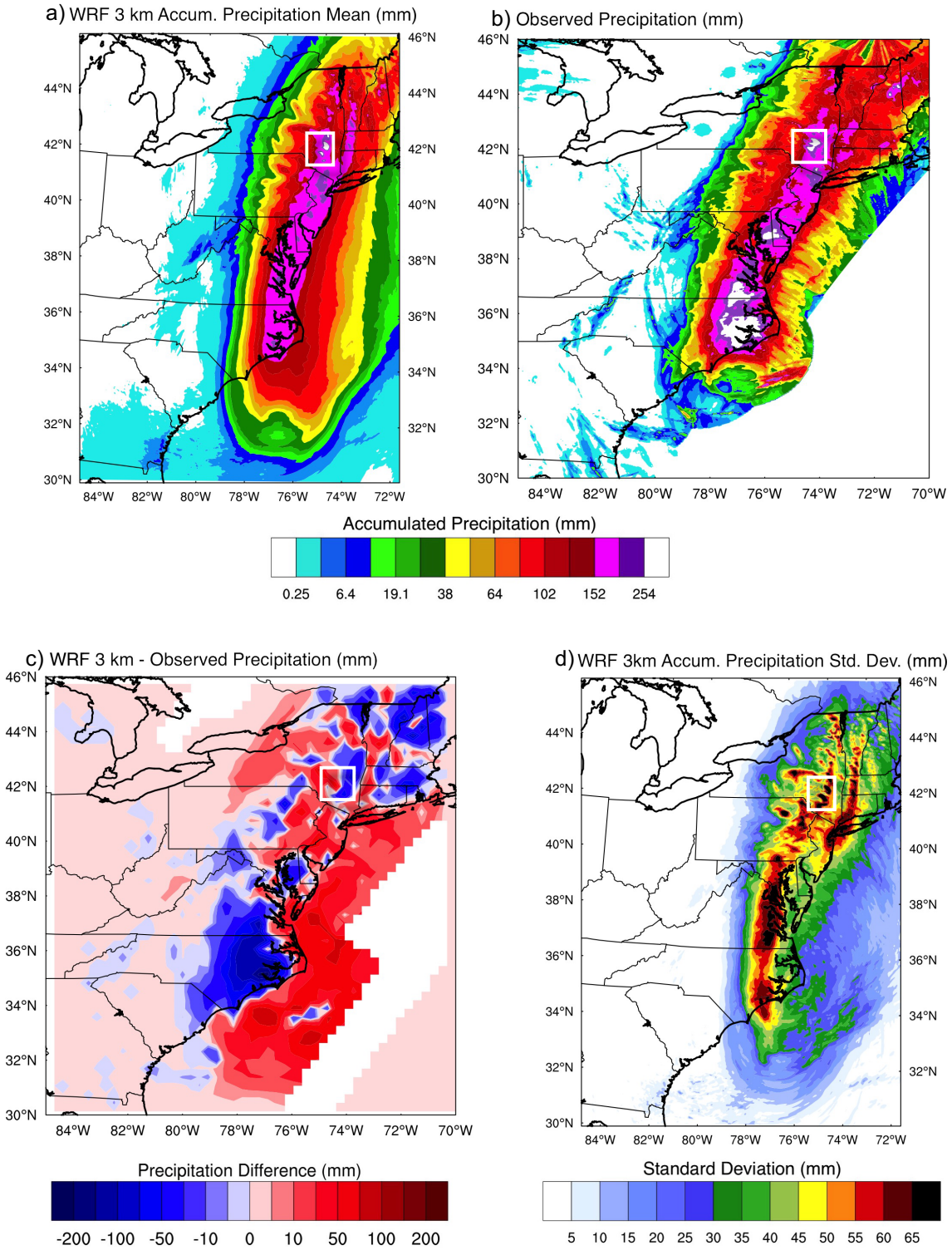

Fig. 22: WRF 3 km 27–29 August 2011 a) ensemble mean total accumulated precipitation (mm), b) observed accumulated precipitation (mm), c) difference between the ensemble mean and observed accumulated precipitation (mm), and d) ensemble standard deviation (mm). The Catskills domain is indicated by the white box.

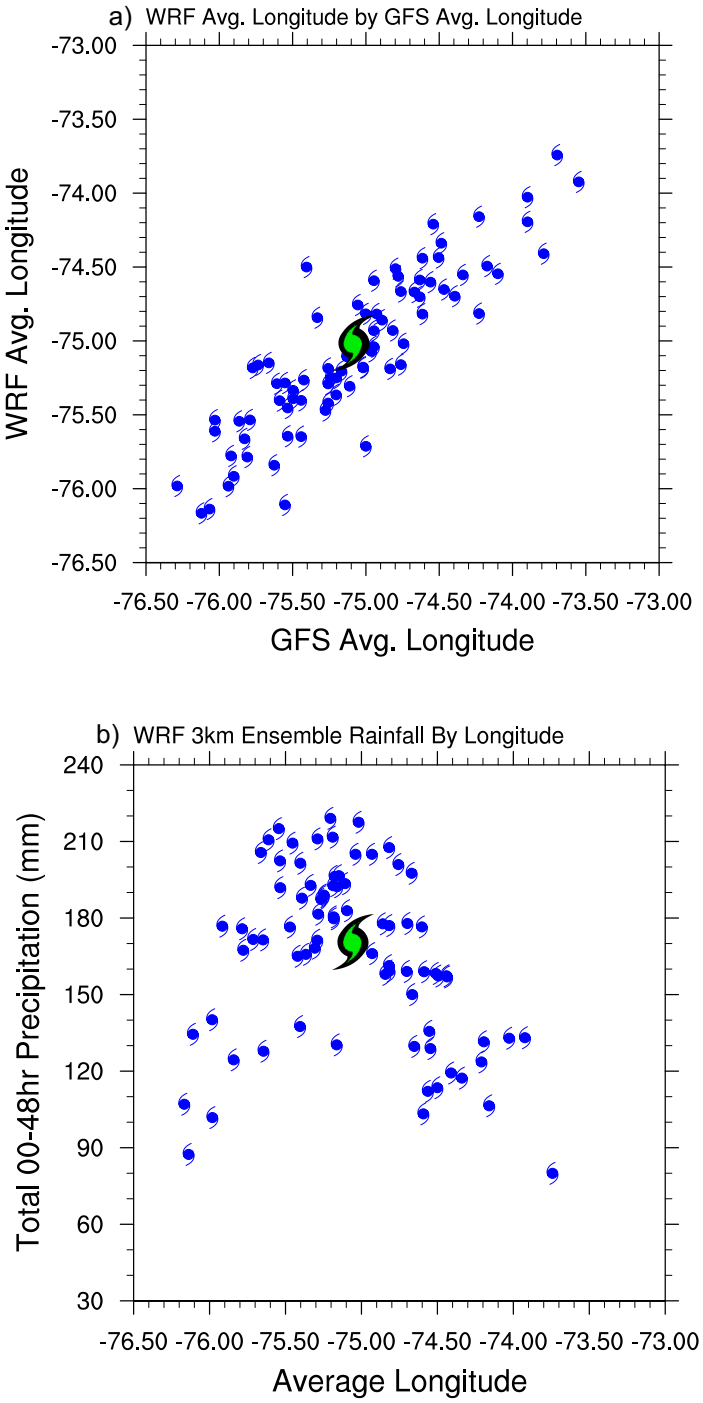

Fig. 23: a) GFS average storm track longitude (degrees east) compared to WRF 3 km average storm track longitude for each ensemble member, and b) WRF 3 km average storm track longitude compared to total Catskills rainfall (mm) for each ensemble member. Irene's observed position and precipitation is indicated in green.

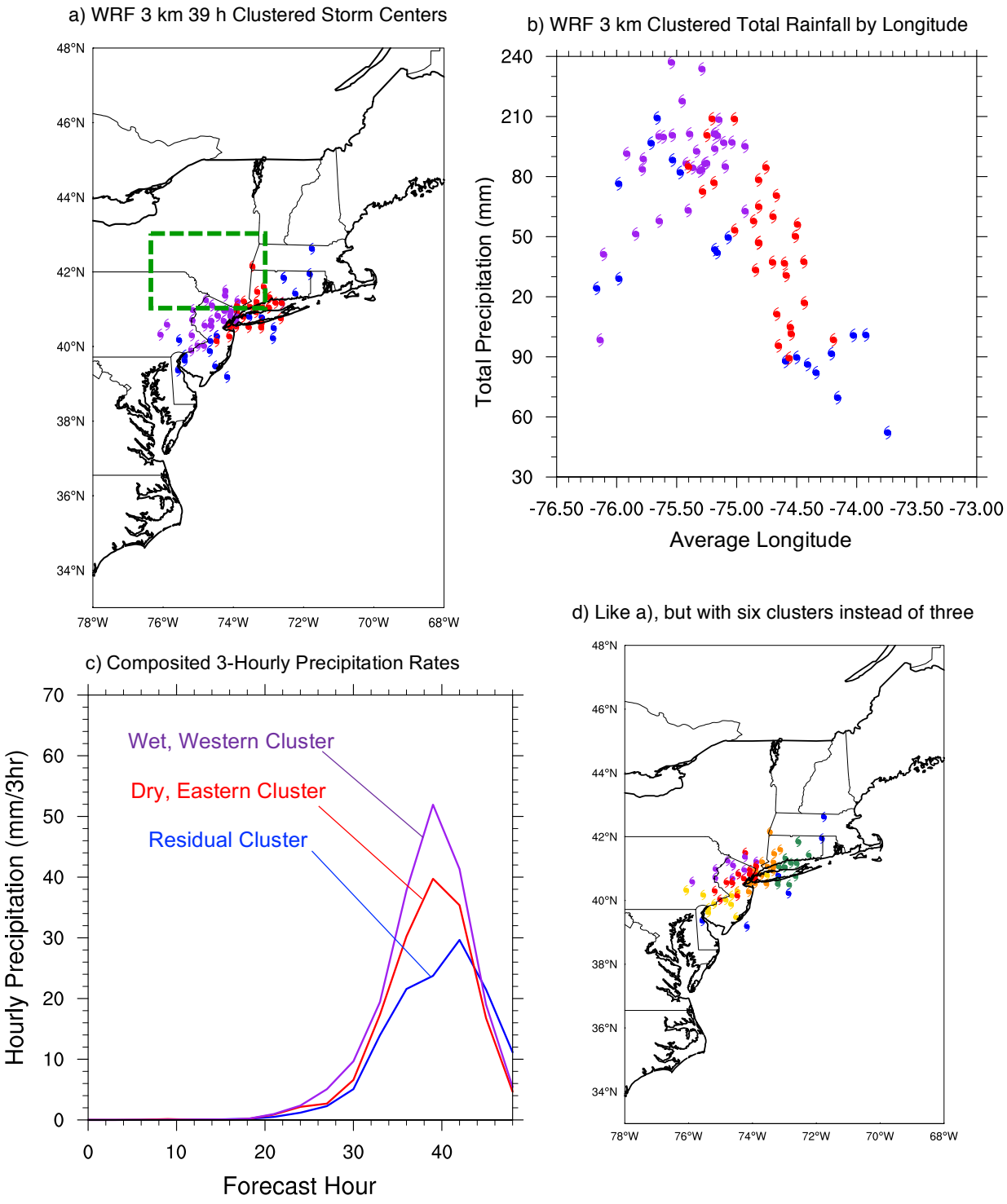

Fig. 24: a) Storm center positions at 39 h into the simulation sorted into three clusters (represented here by color). Clusters are based on the horizontal distribution of precipitation at this time over the domain outlined in green. b) Like Fig. 23b, with each member color-coded by cluster. c) Composited time series of precipitation rate (mm/3hr) over the Catskills for each cluster. d) Like a), but with six clusters instead of three.

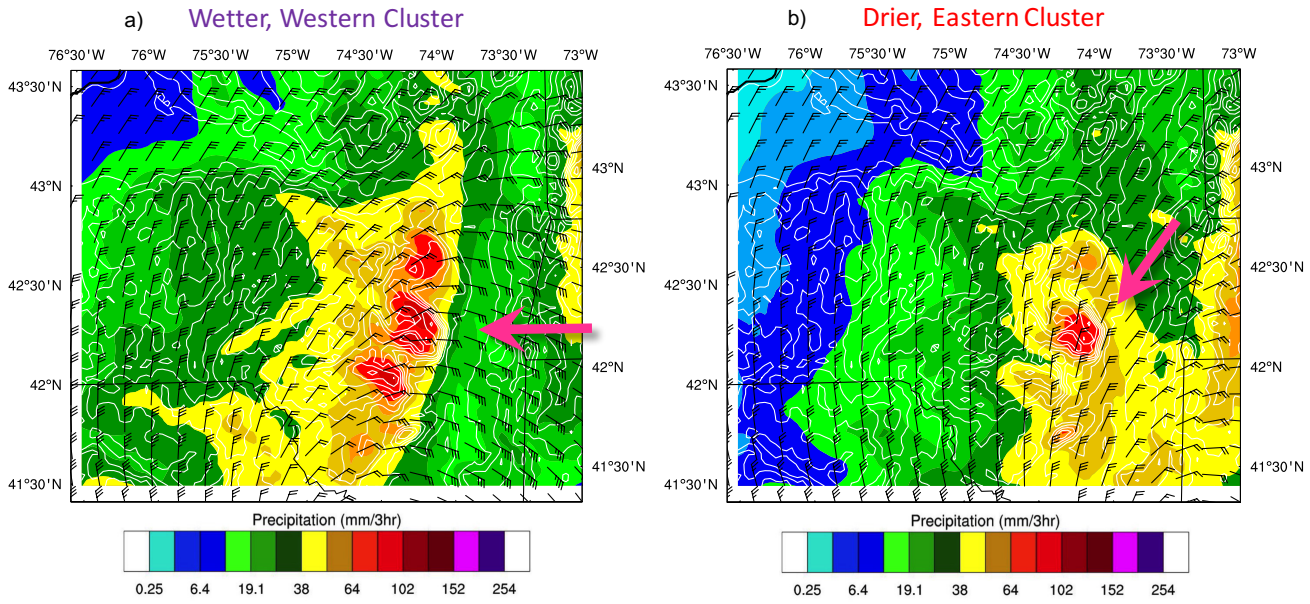

Fig. 25: 39 h precipitation (shaded; mm/3hr), 900-hPa winds (barbs; m/s), and topography (white contours every 100 m) for a) the wet, western cluster and b) the dry, eastern cluster. The pink arrow indicates the predominant wind direction into the Catskills.

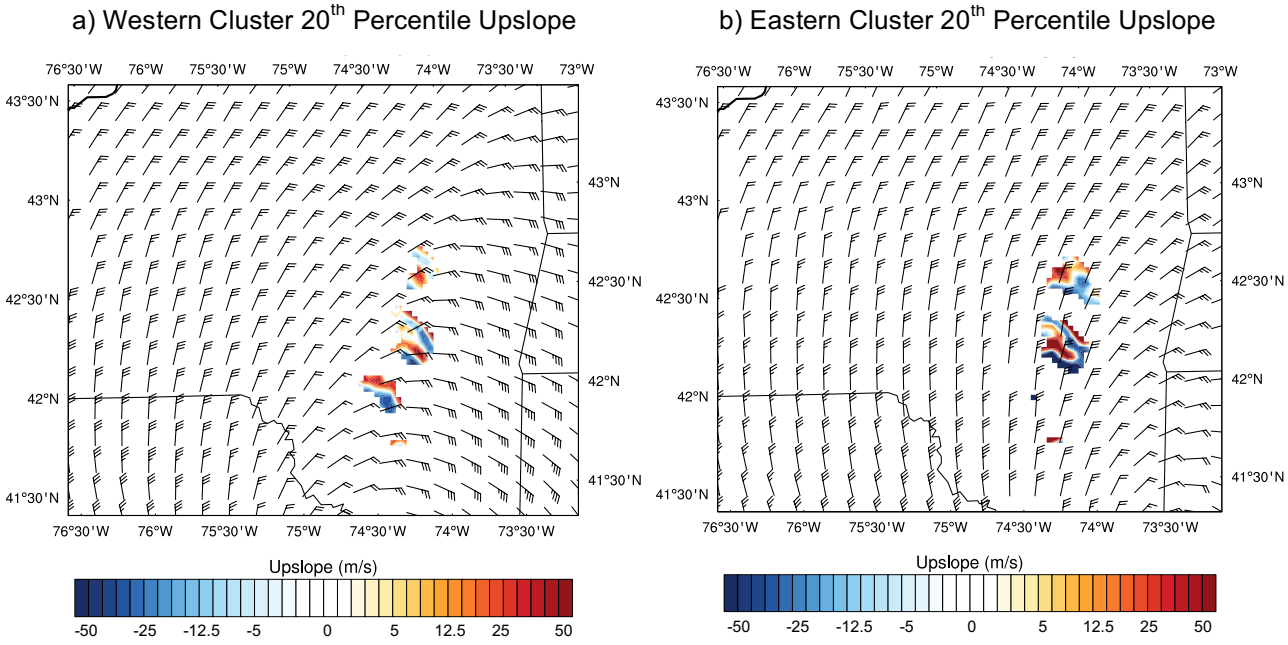

Fig. 26: 39 h surface upslope velocity (shaded; m/s) and 900-hPa winds (barbs; m/s) over Catskills locations that received the top 20th percentile of precipitation accumulations for a) the wet, western cluster and b) the dry, eastern cluster.

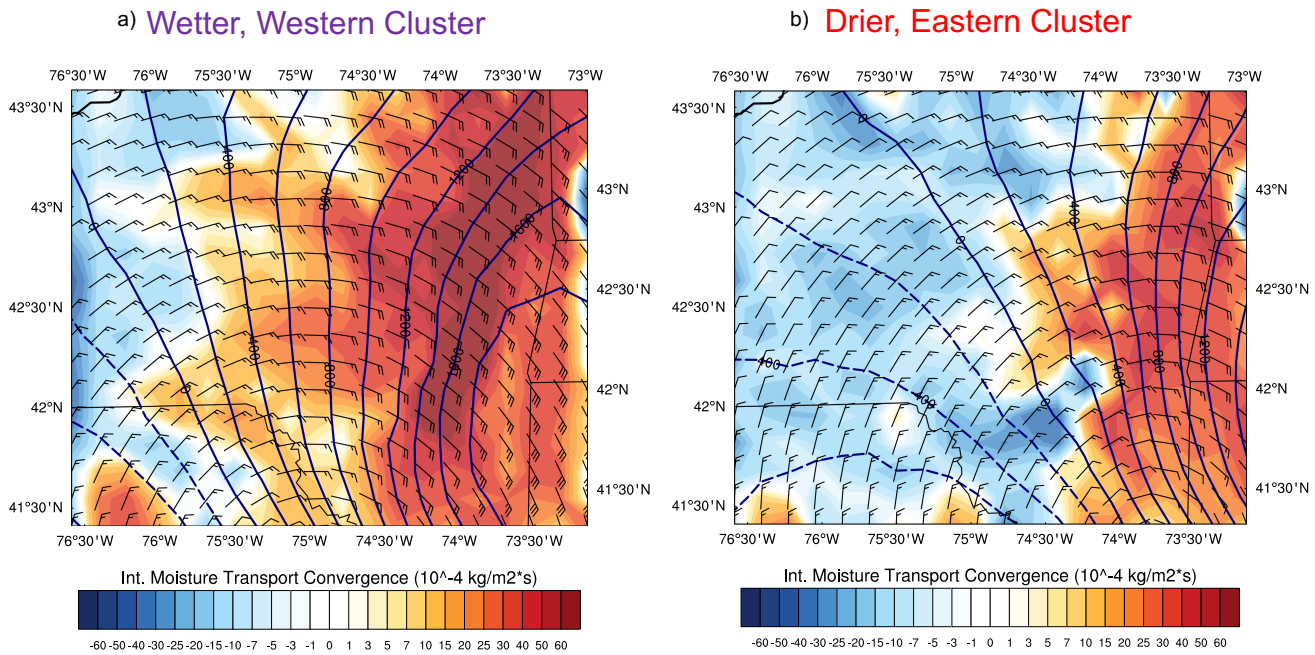

Fig. 27: 39 h 100–1000-hPa layer mean winds (barbs; m/s), integrated moisture transport by the southeasterly wind (contours; every 100 * 10^{-4} kg/m/s), and integrated moisture transport convergence (shaded; 10^{-4} kg/m2/s) for a) the wet, western cluster and b) the dry, eastern cluster.

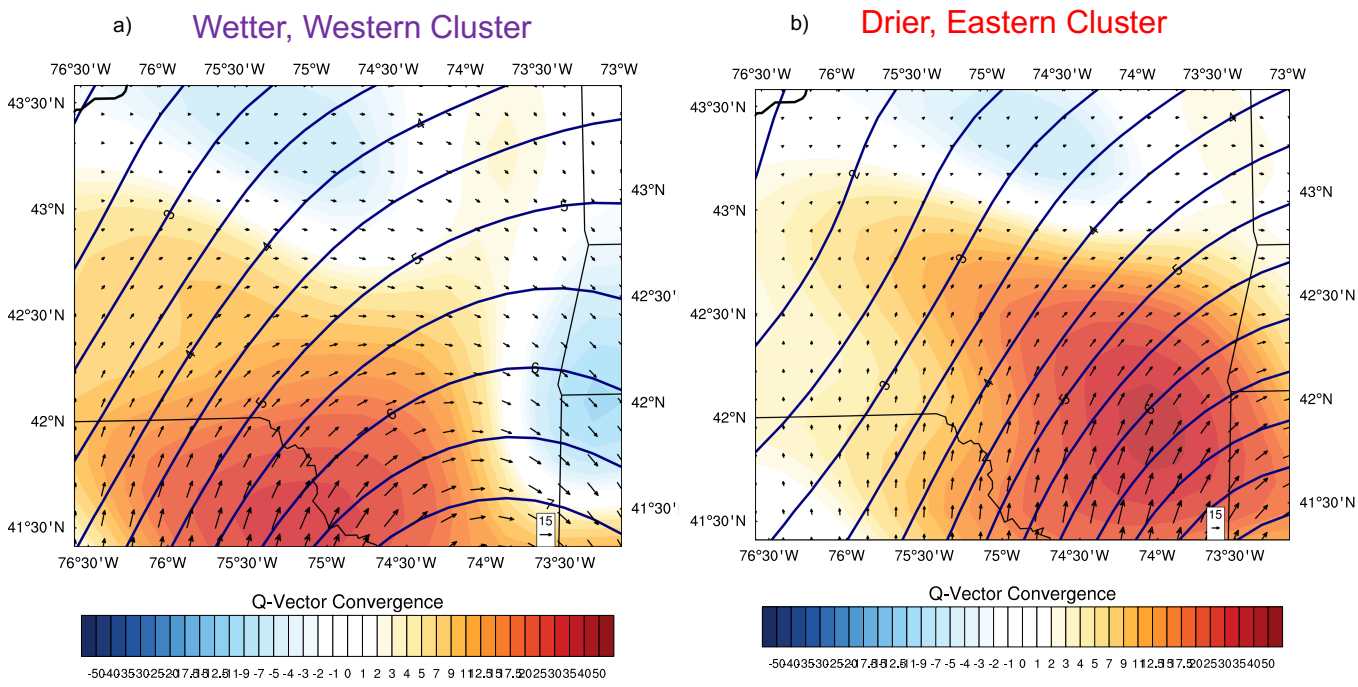

Fig. 28: 39 h 500–800-hPa layer mean Q-vectors (arrows; 10^{12} m2/kg/s), isotherms (contoured; every 0.5 K), and Q-vector convergence (shaded; 10^{12} m/kg/s) for a) the wet, western cluster and b) the dry, eastern cluster.

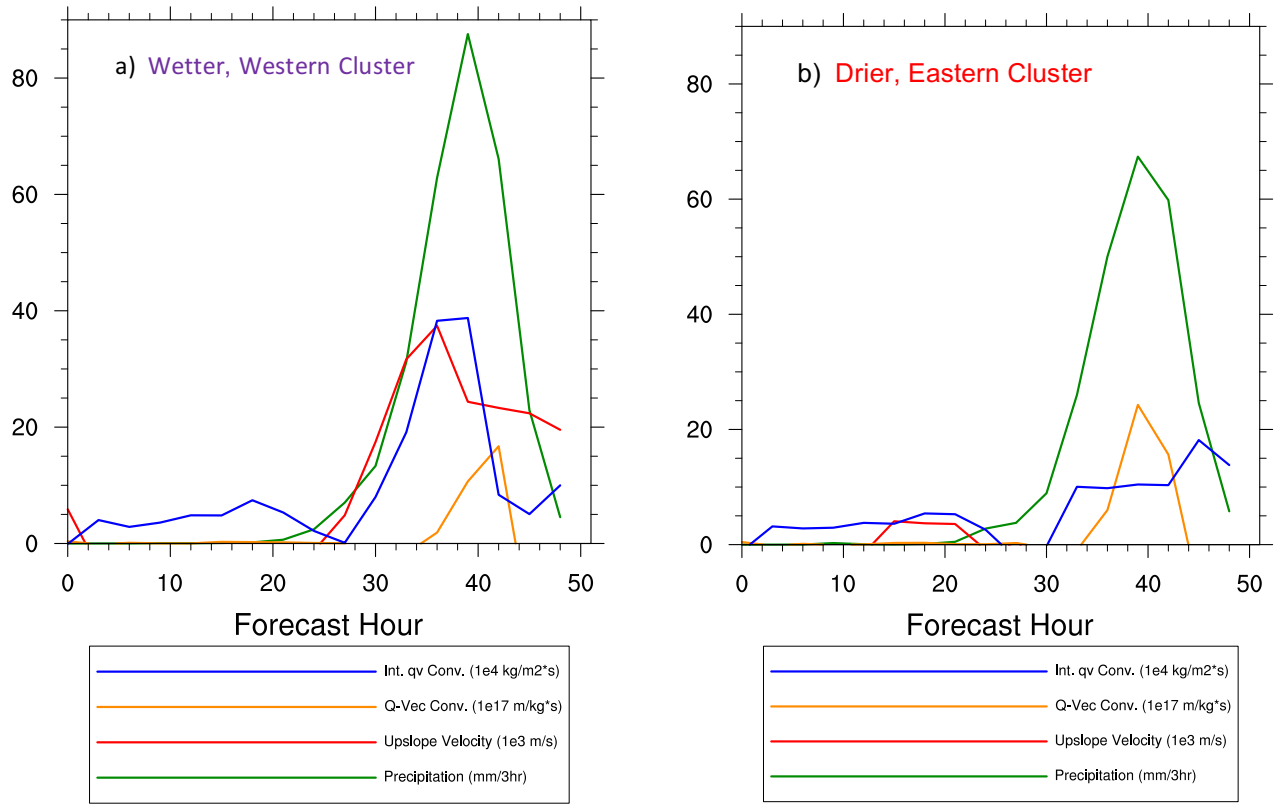

Fig. 29: Scaled time series of Catskills precipitation (green), upslope (red), and moisture convergence (blue) over Catskills locations that received the top 20th percentile of precipitation accumulations, and Q-vector convergence (orange) over the domain 41.5–43.5°N, 73–74.8°W.