
THOMAS T. WARNER
National Center for Atmospheric Research,*
and Program in Atmospheric and Oceanic Sciences, University of Colorado, Boulder, Colorado

EDWARD A. BRANDES AND JUANZHEN SUN
National Center for Atmospheric Research, Boulder, Colorado

DAVID N. YATES
National Center for Atmospheric Research,
and Department of Civil Engineering, University of Colorado, Boulder, Colorado

CYNTHIA K. MUELLER
National Center for Atmospheric Research, Boulder, Colorado

(Manuscript received 18 February 1999, in final form 27 July 1999)

ABSTRACT

Operational prediction of flash floods caused by convective rainfall in mountainous areas requires accurate estimates or predictions of the rainfall distribution in space and time. The details of the spatial distribution are especially critical in complex terrain because the watersheds generally are small in size, and position errors in the placement of the rainfall can distribute the rain over the wrong watershed. In addition to the need for good rainfall estimates, accurate flood prediction requires a surface-hydrologic model that is capable of predicting stream or river discharge based on the rainfall-rate input data. In part 1 of this study, different techniques for the estimation and prediction of convective rainfall are applied to the Buffalo Creek, Colorado, flash flood of July 1996, during which over 75 mm of rain from a thunderstorm fell on the watershed in less than 1 h. The hydrologic impact of the rainfall was exacerbated by the fact that a considerable fraction of the watershed experienced a wildfire approximately two months prior to the rain event.

Precipitation estimates from the National Weather Service Weather Surveillance Radar-1988 Doppler and the National Center for Atmospheric Research S-band, dual-polarization radar, collocated east of Denver, Colorado, were compared. Very short range simulations from a convection-resolving dynamic model that was initialized variationally using the radar reflectivity and Doppler winds were compared with simulations from an automated algorithmic forecast system that also employs the radar data. The radar estimates of rain rate and the two forecasting systems that employ the radar data have degraded accuracy by virtue of the fact that they are applied in complex terrain. Nevertheless, the dynamic model and automated algorithms both produce simulations that could be useful operationally for input to surface-hydrologic models employed for flood warning. Part 2 of this study, reported in a companion paper, describes experiments in which these radar-based precipitation estimates and dynamic model- and automated algorithm-based precipitation simulations are used as input to a surface-hydrologic model for simulation of the stream discharge associated with the flood.

1. Introduction

Flash floods are defined as events in which stream discharge rises from normal to flood level within 6 h (National Research Council 1996). Such events often occur in mountainous terrain because of the prevalent short response times of watersheds and the fact that the irregular topography forces atmospheric circulations that can initiate heavy convective precipitation. Flash floods represent a major threat to life for two reasons, one meteorological and one related to the surface-hydrological conditions. First, the small spatial scales and short lifetimes of the convective precipitation events make them difficult to predict explicitly with atmospheric forecast models. At this time, only the probability of occurrence of the events can be estimated,
based on operational model predictions of large-scale conditions. Even when increased computing power, better meteorological observing systems, better techniques for taking advantage of the current observing systems, and improved knowledge of thunderstorm dynamics, make deterministic prediction of thunderstorms more tractable, the inherent low predictability of such short timescale and space-scale phenomena will render the predictions accurate for perhaps an hour at most. Second, even when short-range, accurate convective rainfall forecasts are attainable, the short hydrologic response times of typical mountain watersheds will exacerbate the problem of providing timely warnings to the public.

There are a number of different approaches that are being developed in the atmospheric and hydrologic sciences for providing information on convective precipitation amounts in complex terrain that can be used as input to surface-hydrologic models for flood prediction. Some techniques are diagnostic and provide analyses of current and antecedent precipitation; others attempt to provide short-range convective precipitation predictions. With the diagnostic approach, the response time of the watershed defines the lead time of the forecast, whereas the short-range precipitation forecasts add an additional (albeit generally short) increment to the time between when a forecast is issued and the time of flooding discharge.

In part 1 of this study, reported here, three techniques for the estimation and prediction of thunderstorm precipitation in very complex terrain were applied to the convective system that caused the Buffalo Creek, Colorado, flash flood of 12 July 1996. Part 2 (Yates et al. 2000) involved testing the relative utility of these rainfall estimates and simulations as input to a surface-hydrologic model that produces simulations of the flash-flood discharge. Two months prior to the rainfall event, a considerable area of the watershed was burned by a wildfire, which drastically increased the ratio of precipitation runoff to infiltration. Thus, a qualitative metric of rainfall simulation accuracy is the degree to which the simulated rainfall is located correctly over the watershed and burn area.

The precipitation estimation and forecast techniques applied in part 1 were 1) estimation of the precipitation using the National Weather Service (NWS) Weather Surveillance Radar-1988 Doppler (WSR-88D) and the National Center for Atmospheric Research (NCAR) S-band, dual-polarization (S-Pol) radars, 2) prediction of the precipitation utilizing a convection-resolving dynamic model, and 3) prediction of the precipitation using an automated algorithmic system. Other possible approaches for providing the precipitation input to surface-hydrologic models will be summarized in the next section. The three methods tested here already have demonstrated considerable potential for application to thunderstorm-related precipitation in areas where the terrain is not complex. For example, numerous studies have used radar to estimate rainfall successfully over relatively flat terrain [see, e.g., summaries given by Wilson and Brandes (1979) and Joss and Waldvogel (1990)].

Recent applications with WSR-88D include Bauer-Messmer et al. (1997) and Baeck and Smith (1998). Roberts et al. (1999) describe the application of the automated algorithmic system over flat terrain. Tests of the convection-resolving dynamic model over flat terrain are described in Sun and Crook (1998). Thus, it is important to document the performance of these techniques in mountainous regions because they represent examples of perhaps the best approaches for eventually diagnosing and predicting flash flood–producing convective rainfall. The discussions here, however, will document that all three of the techniques inherently are handicapped in one respect or another by their application in complex topography. Nevertheless, they are the most promising available with which to work, and it is appropriate that their strengths and shortcomings be investigated and summarized through application to a common test case. Thus, these results represent a benchmark for future work and serve as an illustration of the challenges associated with flash-flood prediction in mountainous areas.

Section 2 provides a background summary of the various general approaches for diagnosing and predicting precipitation, especially in complex terrain, for use as input to flash-flood models. Section 3 describes the meteorological conditions associated with the Buffalo Creek, Colorado, flash flood. A description of the three different convective precipitation estimation and prediction techniques employed in this study is provided in section 4. In section 5, the precipitation estimates from the three techniques are compared, and a summary is provided in section 6.

2. Summary of available techniques for providing diagnoses and predictions of precipitation in complex terrain

A variety of different precipitation estimation and prediction techniques are available or are under development for possible use in providing operational input to surface-hydrologic models for flood or flash-flood prediction. To place the techniques employed in this study in their proper context among the range of approaches available, a brief summary of the latter is provided in this section. In addition, their appropriateness for use in flash-flood prediction in mountainous terrain is commented upon, and estimates are provided of their degree of maturity for operational use. Note that specific techniques for diagnosing and predicting convective precipitation often use a combination of radar data, dynamic-model products, and other input, and thus it is difficult to categorize them in a simple way.

a. Radar estimation of precipitation

Diagnoses of convective precipitation have been made possible by the availability of high-resolution re-
filtering creates a $K$.

Other recent research (Ryzhkov and Zrnic 1995a,b, watershed with an area of 926 km$^2$. The diagnosis of activity—inferred rainfall amounts as input to a discharge and Hirano (1991) who utilized gauge and radar re-areas. An exception to this is the work of Moriyama 1991), but few have been performed for mountainous precipitation rates as input to surface-runoff models algorithms more difficult.

Last, the paucity of rain gauge data in mountainous areas is problematic. Locating them in the valleys results away from the mountains. The specific differential propagation phase ($K_{dp}$) is the range derivative of $\phi_{dp}$. Other recent research (Ryzhkov and Zrnic 1995a,b, 1996; Ryzhkov et al. 1997) shows that rainfall estimates derived from $K_{dp}$ are relatively insensitive to radar calibration error, attenuation, beam blockage, and anomalous propagation. The parameter also is relatively insensitive to dry tumbling hail (Balakrishnan and Zrnić 1990). The $\phi_{dp}$ measurement has a standard error of 3°–4°, which is reduced by filtering. The filtering procedure yields a $K_{dp}$ with an accuracy of $\pm0.3$ km$^{-1}$. Filtering creates a $K_{dp}$ rain field that is slightly smoother than that derived from radar reflectivity alone but does not introduce a bias. The current operational WSR-88D radar does not have dual-polarization capability. An additional difficulty is that siting radars in mountainous areas is problematic. Locating them in the valleys results in beam blockage; locating them on the ridges means they do not observe the lower atmosphere and there are extended ground echoes. The result of this siting dilemma is that radars simply might not be installed in mountainous areas, leaving a data void where it is least tolerable from the standpoint of flash-flood prediction. Last, the paucity of rain gauge data in mountainous areas makes the local calibration of radar-reflectivity/rain-rate algorithms more difficult.

Studies have investigated the use of radar-inferred precipitation rates as input to surface-runoff models (e.g., Peters and Easton 1997; Engdahl and McKim 1991), but few have been performed for mountainous areas. An exception to this is the work of Moriyama and Hirano (1991) who utilized gauge and radar reflectivity—inferrned rainfall amounts as input to a discharge model in tests of flash prediction over a mountainous watershed with an area of 926 km$^2$. The diagnosis of the Buffalo Creek convective precipitation by the operational Denver WSR-88D is one of the approaches tested in this study for providing rainfall estimates that can be used as input to a flash-flood model. In addition, rainfall estimates are employed from the experimental dual-polarization S-Pol radar operated by NCAR.

b. Three-dimensional dynamic meteorological models for precipitation prediction

Numerical weather prediction (NWP) models have been employed for four decades to produce research simulations and operational forecasts of precipitation. Even though the skill of the operational-model predictions of large-scale circulations has improved markedly during the period, progress has been especially slow in improving summer-season quantitative precipitation forecasts (QPF) (Fritsch et al. 1998; Georgakakos and Hudlow 1984). One of the impediments has been the lack of conventional in situ data on the thunderstorm scale. Thus, the development of techniques for the variational initialization of mesoscale models with WSR-88D radar reflectivity and Doppler winds holds promise of contributing significantly toward progress in this area (Xu 1996; Sun and Crook 1994, 1997). Radar-beam blockage in complex terrain, however, unfortunately means that boundary layer winds generally cannot be retrieved, and there is no reflection from precipitation at low beam angles. The lack of this information makes the variational-initialization process much less skillful than when it is applied over flat terrain. In addition, full-physics adjoints of convection-resolving models are not available yet. The model adjoints are very computationally intensive to use, but, by the time they become sufficiently mature for operational application, it is probable that the necessary computing power will be available.

In general, NWP models do not have a significant history of use for providing input to surface-hydrologic models. Perhaps one of the first examples was the use by Warner et al. (1991) of a mesoscale model, with a horizontal resolution approximately that of the current NCEP Eta Model, to provide precipitation input to the Army Corps of Engineers Hydrologic Engineering Center HEC-1 runoff model for a flood simulation. More recently, Westrick et al. (1998) have employed a much higher resolution mesoscale atmospheric model coupled with a surface-hydrologic model for the study of a flood event in complex terrain. Both of these studies, however, employed winter precipitation events that were much more predictable than are summer convective events (even though winter runoff predictions may be more difficult than summer ones if snowmelt contributes).

The Sun and Crook model (1994, 1997) was employed in this study as an example of the use of a convection-resolving NWP model for the explicit prediction of convection in complex terrain. It was initialized variationally using WSR-88D reflectivity and winds.

c. Automated algorithmic systems for precipitation prediction

Because NWP models and their initialization techniques have not been sufficiently mature to produce very
short–range operational convective-scale QPF, a number of automated algorithmic techniques have been developed for the “nowcasting” of convective precipitation, where the term nowcasting implies forecasts of up to 3-h duration. These systems vary widely in their constructs and complexity, but, in their most general form, they utilize information from radar, profilers, surface observations, and models to produce very short–range forecasts of thunderstorm initiation, development, movement, and decay. The reader should refer to Conway et al. (1996), Cluckie and Collier (1991), and Wilson et al. (1998) for descriptions of the variety of techniques that are used for thunderstorm nowcasting. Two of the more-complex nowcasting systems are the Generating Advanced Nowcasts for Deployment in Operational Land Surface Flood Forecast system developed by the United Kingdom Meteorological Office (Hand and Conway 1995; Hand 1996) and the thunderstorm Autonowcaster developed by NCAR (Wilson and Mueller 1993; Wilson et al. 1994, 1998). Often these systems employ a variety of data types. The paucity and limited accuracy of some data in mountains can greatly limit the accuracy of the forecasts, however.

A component of the Autonowcaster system that can be employed in complex terrain will be used in this study as an example of an automated algorithmic system. As with the above-described use of the radar and dynamic model, the terrain effects cause serious limitations here also. Specifically, much of the skill of the Autonowcaster is associated with its use of radar-inferred boundary layer winds, which are not available in this case, as noted above, because of the terrain blocking.

d. One-dimensional cloud models for precipitation prediction

This approach first was described by Georgakakos and Bras (1984a,b) as a means for providing short-range rainfall forecasts for flash-flood prediction. After a number of improvements, Lee and Georgakakos (1996) describe its use for very short–range prediction of convective precipitation events in Oklahoma. The basic concept behind the technique involves embedding a very fine grid within the coarser grid of a complete NWP model and solving a cloud model on the fine grid. Both radar- and gauge-inferred precipitation estimates can be used to update, in real time, the state variables of the cloud model. The model also produces estimates of the forecast precipitation variance in real time. The precipitation forecasts outperformed persistence for forecast periods in excess of 3 h for two storms in Oklahoma (Lee and Georgakakas 1996). The improved model has not been tested yet over mountainous terrain.

e. Mass continuity–based orographic precipitation models for precipitation estimation and prediction

This general class of techniques for estimation of the modulation of precipitation by complex orography relies upon mass-continuity constraints to estimate the uplift and condensation associated with the impingement of the large-scale atmospheric flow on complex mountainous terrain whose structure can be defined with very high resolution. Barros and Lettenmaier (1994) contrast this type of model with dynamic NWP and other models. If the large-scale atmospheric state is based on observations, the estimation of precipitation is diagnostic. If the large-scale atmospheric state is predicted by a coarse-resolution (in comparison with the terrain dataset resolution) model, the precipitation-rate product is prognostic. Leung et al. (1996), Leung and Ghan (1995, 1996), Hay (1998), and Rhea (1977) are examples of the numerous studies that demonstrate success in simulating such orographic modulation of precipitation.

These approaches for representing the dynamic effects of orography on the precipitation are applied most appropriately and reliably for conditions wherein large-scale, stable, upslope flow (i.e., not thunderstorms) produces the rainfall. In addition, much of this work has involved long-term simulations for production of seasonal climate descriptions, and it has been found that the models verify best for these types of simulations rather than for simulations of individual rain events. Nevertheless, with further development, these methods likely have considerable potential applicability to individual nonconvective storms that produce flooding in complex terrain.

f. Statistically based orographic precipitation estimation

In this approach, statistical relationships between precipitation and elevation are used to interpolate precipitation between rain gauge observations in complex terrain for use in rainfall-runoff models. For example, Hay et al. (1998) use a digital elevation model to define various characteristics of the orography for an array of grid cells that covers an area of complex terrain. The terrain characteristics (including mean elevation, range of elevations within the cell, slope, aspect, and others) then are used in the statistical interpolation between the relatively coarse array of precipitation observations. In the example used in the Hay et al. (1998) study, the interpolated precipitation was used as input to a stream-discharge model. As with the mass-continuity models, 1) most of the validation of these techniques has involved longer-period precipitation totals rather than rates for a single event, 2) the logical application would be for diagnosis of precipitation rather than for prediction (even though one could envision using this statistical interpolation between gridpoint precipitation values from a coarse-grid model), and 3) the success of the technique would seem to be greatest for wintertime stable precipitation events rather than for summer convective events.
3. Description of the meteorological conditions of the Buffalo Creek flash-flood event

During the day prior to the storm, a weak front moved from the northeast into Colorado, turning the winds to upslope easterlies in the Eastern Plains as it advanced (Fig. 1). Surface dewpoint temperatures in advance of the front and upstream of the mountains were between 55°F and 65°F. Upper-air flow was from the northwest, with a jet approaching the Rockies from Canada. Ageostrophic circulations in the exit region of the jet could have been responsible for large-scale vertical motions over Colorado. A 500-hPa temperature wave that produced a cold anomaly over the northern Rockies at 1200 UTC 12 July and 0000 UTC 13 July 1996 could represent evidence of upward vertical motion in this area. Figure 2 shows the Denver sounding for 0000 UTC 13 July. It is relatively unstable, and the boundary layer moisture has increased and deepened since the previous sounding. The 10–15 m s⁻¹ northerly boundary layer winds likely result from an outflow boundary. Thus, the combination of high surface dewpoints, weak frontal forcing, low-level upslope flow, and the possible upper-level support from the jet produced an environment that was conducive to convection over the Front Range of the Colorado Rockies.

The precipitation that caused the flood was associated with a thunderstorm that moved from west to east during the evening (local time) of 12 July 1996. There was other scattered convection throughout the Front Range of the Rockies to the west of Denver and to the east over the plains. Some of these events produced precipitation that is comparable to that associated with the Buffalo Creek thunderstorm. Figure 3 shows the total precipitation accumulation along the Front Range and neighboring High Plains for the period of the convection that evening between 0000 UTC and 0423 UTC 13 July 1996 [1700 and 2123 Local Time (LT) 12 July 1996], as estimated from the NCAR S-Pol φdp measurements (Vivekanandan et al. 1999). The outline of the Buffalo Creek watershed is shown on the figure. Figure 4 shows the temporal distribution of the radar-estimated precipitation associated with the thunderstorm over the watershed. The rainfall rates are in terms of 6-min accumulations for the period ending at the local time indicated. Also shown on the figure is the outline of the watershed and the area burned by the wildfire.

Because the Buffalo Creek and surrounding watersheds are small and were not instrumented well with rain gauges, less-exact methods need to be employed to corroborate the radar-inferred estimates. R. D. Jarrett and T. W. Browning (1999, manuscript submitted to J. Hydrol.) utilized “bucket-survey” and paleohydrologic-estimation techniques to produce the storm-total analysis shown in Fig 5. “Paleohydrologic estimates” were based on calibration of hillslope erosion where bucket data were available. Hillslope erosion alone then was used where bucket data were not available. R. D. Jarrett and T. W. Browning (1999, manuscript submitted to J. Hydrol.) suggest that the paleohydrologic estimates are accurate to within ±20%. The general patterns of the precipitation distribution from the radar and paleohydrologic/bucket estimates are similar, but the latter reflects higher storm totals. There is also a considerable difference in the orientation of the axes of the elliptical rainfall patterns. The largest amount recorded in the bucket survey was 80 mm, which was reported to have accumulated in less than 1 h.
4. Description of the different convective precipitation estimation and prediction techniques employed in this study

a. Radar estimation using the WSR-88D and the dual-polarization S-Pol radars

During the summer of 1996, NCAR conducted a field study in the Front Range of the Rocky Mountains in which a polarimetric radar (S Pol) was deployed in a program designed to improve NWS WSR-88D reflectivity-based algorithms for estimating rainfall and to determine the benefits that might be gained if these radars were modified with a dual-polarization capability (Brandes et al. 1999). The WSR-88D radar currently has only horizontal polarization. The S-Pol radar is located within 2 km of the Denver NWS WSR-88D, approximately 35 km east of downtown Denver and 60 km east of the Colorado foothills. The Buffalo Creek watershed is about 75 km from the radars at an azimuth of 233° (see Fig. 3).

Radar-reflectivity measurements $Z_H$ were converted to rain rates $R$ using the WSR-88D default relationship $R = 0.017Z_H^{0.714}$, where a lower bound of 25 dBZ was used to remove the effects of clear-air returns. All returns greater than 51 dBZ were assumed to have been contaminated with hail, and the values were reset to 51 dBZ. Copious hail fell with the storm. In fact, damage-survey crews noted that hail still could be found the following morning. (The hail was detected readily with the differential reflectivity measurement of the S-Pol radar.) Radar-reflectivity estimates of rainfall are sensitive to the hail threshold used. For example, the maximum rainfall from radar reflectivity increased by almost 60% when the threshold was not used. Additional error sources in this experiment are thought to be associated with shielding by intervening obstructions (Harrold et al. 1974), the vertical profile of reflectivity (Kitchen and Jackson 1993), smoothing of reflectivity gradients (Zawadzki 1984), and spatial and temporal variations in the relationships between the radar parameters $Z_H$ and $K_{DP}$ and the actual rainfall rate (Harrold et al. 1974; Wilson and Brandes 1979; Ryzhkov and Zrnić 1996). The differential phase-based estimate was based on the relationship $R = 40.56K_{DP}^{0.87}$ (Sachidananda and Zrnić 1987), and rainfall rates were computed only if the reflectivity was greater than 25 dBZ.

Rainfall accumulations were defined on polar grids at the basic measurement resolution of each radar (1° × 0.15 km for S-Pol and 1° × 1 km for WSR-88D). Sampling intervals were 110 s with S-Pol and 5 min with WSR-88D. Accumulations were based on measurements from the 0.5° antenna elevation. Beam blockage can be reduced by incorporating measurements from higher angles. Such procedures, however, introduce other problems because the distance increases between the elevated radar beam and the ground. Hence, this procedure was not employed here.

b. The rule-based thunderstorm Autonowcaster

Even though only one component of the Autonowcaster system can be applied here because of the existence of complex terrain, all components will be described briefly to emphasize the limitations imposed by the terrain. The Autonowcaster system combines feature-detection algorithms and thunderstorm extrapolation/trending software in an “expert” system to produce short-term thunderstorm forecasts (nowcasts). Outputs from the component algorithms are used as inputs to the Autonowcaster algorithm, which predicts the evolution of the reflectivity field. For this application, the simulated new reflectivity fields are converted to rain rates using the WSR-88D algorithm. As with virtually all automated algorithmic systems, the existence of mountains unfortunately imposes serious limitations on the use of most of the components of the current version of the Autonowcaster. The only component that remains operative is the Thunderstorm Identification, Tracking, Analysis, and Nowcasting (TITAN) algorithm (Dixon and Wiener 1993) that is used to detect and forecast automatically the location and size of existing thunderstorms. The thunderstorm detections and forecasts are based on three-dimensional Cartesian radar data. A method that matches storms at one radar volume time to those at a subsequent time is used to track the storms. Based on past storm trends, TITAN predicts future storm location and size. Products from TITAN are input into
The components of the Autonowcaster that could not be used in this complex terrain application are 1) a Convergence Line Detection algorithm that provides forecasts of radar-detected convergence lines located near the surface, 2) a satellite-based cumulus cloud-detection algorithm that is used to locate clouds in the vicinity of a surface convergence boundary to identify areas with a high likelihood that vigorous convection may occur, and 3) a Tracking Radar Echoes by Correlation algorithm (Tuttle and Foote 1990) that retrieves the three-dimensional wind speed and direction within the clear-air planetary boundary layer from single Doppler radar data.

Outputs from all of the various algorithms described above normally are combined by the Autonowcaster system to produce forecasts of thunderstorm initiation, growth, decay, and movement. As noted, because the Autonowcaster was not designed for use in complex terrain, the TITAN algorithm is the only component that can be employed for this case and for similar ones in mountainous terrain. Additional rules that employ satellite data could be developed to improve predictive skills in these areas where the boundary layer is not observable by radar.
FIG. 4. Temporal distribution of WSR-88D radar-estimated precipitation associated with the Buffalo Creek thunderstorm. The rainfall rates are 6-min accumulations for the period ending at the time indicated (LT). The watershed is shown in black and the burn area is outlined in red.
c. The dynamic model

A nonhydrostatic, storm-scale, anelastic numerical model (Sun and Crook 1997) is used in both the assimilation of the radar observations and the subsequent numerical simulation of the Buffalo Creek storm. The prognostic variables include the three velocity components; the perturbation, liquid-water potential temperature; rainwater mixing ratio; and total-water mixing ratio. The temperature and cloud water mixing ratio are diagnosed by assuming that all vapor in excess of the saturation value is converted to cloud water. The microphysical processes represented in this model are condensation and evaporation of cloud water, evaporation of raindrops in subsaturated air, autoconversion of cloud to rain, accretion of cloud by rain, and sedimentation of rain. Ice processes are not considered. Kessler explicit warm-rain microphysics (Kessler 1969) is employed.

Both the radar-data assimilation and the subsequent simulation experiments are performed on a domain with dimensions of 65 km × 65 km × 13 km (Fig. 6). The southwest corner of this domain has east–west and north–south coordinates of \( x = -113 \) km and \( y = -78 \) km, respectively, relative to the radar location. The grid increment is 2 km in the horizontal and 500 m in the vertical. The lateral boundary conditions are open such that the inflow is prescribed by a mean profile from a coarser-resolution model and the outflow is extrapolated using the closest inner two grid points. The top and bottom boundary conditions are set to zero for vertical velocity; for all other variables, they are defined such that their derivatives vanish.

This storm-scale numerical model is used as the constraint in the Variational Doppler Radar Analysis System (VDRAS). VDRAS was developed to assimilate a time series of radar observations (radial velocity and reflectivity) from single or multiple Doppler radars. By fitting the model to observations over a specified time period, a set of optimal initial conditions for the constraining numerical model can be obtained. A cost function that measures the fit between the model variables and the observations is defined, based on the difference between the radar-inferred and model-simulated radial velocity and rainwater mixing ratio. The background term in the cost function is based on a horizontally averaged mean profile obtained from a coarser-resolution model simulation of this case.

As stated earlier, current-generation prototype numerical models and their adjoints used for four-dimensional variational assimilation of radar data and forecasting on the convective scale generally have limitations. Specifically, full-physics adjoints of convection-resolving models are not available yet. In this case, the model does not consider the surface topographic variation in spite of the fact that the altitude over the domain varies by about 1700 m. As a result, the dynamic effects of the orographic forcing are not accounted for. Furthermore, there is no parameterization of the planetary boundary layer turbulence, nor are there surface fluxes of heat and moisture. Also, as noted earlier, because of the distance of the storm from the radar and the storm’s location in the mountains, boundary layer winds are not available for assimilation.

Thus, the skill obtained here represents a lower limit to what will be available in the future. Even when such modeling methodologies become more mature, however, there still will be inherent limitations associated with their use in complex terrain. Specifically, the terrain blockage of the radar beam and the distance of the storms from the radar will mean that boundary layer winds will not be available for assimilation, which represents a significant limitation.

5. Quantitative precipitation estimates and simulations

This section discusses both the estimated and simulated precipitation that will be used in the second part of this study as input to the surface-hydrologic model. Radar data are used in all three approaches discussed here for the estimation and prediction of the rainfall for the Buffalo Creek flood. The radar-inferred precipitation data themselves will be used as input to the surface-hydrologic model, and the radar data are used as input to the two techniques in which very short range precipitation simulations are used by the surface-hydrologic model.

For the simulations, the radar data will be used for verification. To quantify the position error of simulated precipitation, the threat score (TS), described by Anthes (1983), is employed. It is a measure of the quality of the forecasts in terms of predicting the area of precipitation amounts over any given threshold and is defined as

\[
TS = \frac{CFA}{FA + OA - CFA},
\]

where CFA is the correctly forecast area, FA is the forecast area, and OA is the observed area. The areas are defined in terms of a threshold that has been specified. A perfect score is unity. If there is no overlap in the FA and the OA for a given threshold, the CFA is zero and the threat score is zero. When there is no observed or forecast precipitation above a threshold, the threat score is undefined. Fritsch et al. (1998) report that the threat scores for 0–24-h operational U.S. precipitation forecasts have ranged between about 0.1 and 0.3 during the last 30 yr for thresholds between 0.5 and 2.0 inches.

a. Radar-based estimates

Because the S-Pol \( K_{DP} \) estimates of rainfall are less affected by beam blockage, they should more-correctly reflect actual amounts than reflectivity-based estimates from the WSR-88D do (Vivekanandan et al. 1998). Figure 7 shows the cumulative rainfall volume as a function of time within the Buffalo Creek watershed boundary,
based on dual-polarization measurements by S-Pol and reflectivity measurements from WSR-88D. The S-Pol storm-total, watershed, rainfall volume estimate from $K_{dp}$ at 0300 UTC is about 15% greater than the WSR-88D reflectivity estimate.

The temporal distribution of the watershed-average precipitation estimates from the WSR-88D, S-Pol differential phase, and S-Pol reflectivity measurements are shown in Fig. 8. Reflectivity estimates (from either S-Pol or WSR-88D) produce peak 6-min rainfall rates that are approximately 20% lower than those from the S-Pol $K_{dp}$ estimate. Also note that the WSR-88D reflectivity tends to give higher estimates at the onset of the storm (between 1900 and 1930 LT), and the S-Pol estimates give slightly higher estimates after 1930 LT. The lower reflectivity-based quantitative precipitation estimates during the most intense periods of the storm are attributable largely to beam blockage by complex terrain. Figure 9, based on Vivekanandan et al. (1999), shows the spatial distribution of estimated storm-total rainfall based on the WSR-88D and S-Pol $K_{dp}$ radar estimates. Note that the accumulations differ in amount and in spatial detail; for example, the rainfall maxima are not coincident. The differences are attributed to blockage of the WSR-88D radar beam and sampling. The S-Pol $K_{dp}$ and WSR-88D rainfall estimates were interpolated to a rain gauge in Buffalo Creek that recorded 68.1 mm (Fig. 5). The S-Pol and WSR-88D amounts were 70.9 and 50.6 mm, respectively. We believe this result is more than fortuitous and is indicative of the superiority of the $K_{dp}$ rainfall estimate in complex terrain. Differences between the radar estimates of the positions of the rainfall maxima and the amounts both could contribute to differences in the computed flood intensity.

These radar-estimated precipitation rates at 6-min intervals were inserted into the surface-hydrologic model that calculates stream discharge (Yates et al. 2000). The

---

Fig. 6. The computational domain for the dynamic model retrieval and simulation. The solid lines are terrain contours labeled in meters and plotted at an interval of 100 m. The outline of the Buffalo Creek watershed is shown.

Fig. 5. Bucket-survey and paleohydrologic estimate of storm-total precipitation. Isohyets (solid lines) are labeled in millimeters and are plotted at an irregular interval. The shaded area reflects the area burned by the wildfire. The "$\times$" symbols and associated numbers reflect the locations and amounts (mm) of bucket estimates. Location 1 (triangle) is the site of the rain gauge measurement noted in the text (from Jarrett and Browning 1999, manuscript submitted to J. Hydrol.). The color bands correspond to the precipitation banding in Figs. 3 and 9, and facilitate comparison with those figures of radar-derived accumulations.
advantage of using these radar data to force a surface-hydrologic model, relative to using output from a forecast algorithm, is that the errors are perhaps more predictable and smaller. The drawback to using the observations is that the time between a warning and the flood-level discharge is essentially the response time of the watershed.

b. Dynamic-model and automated algorithmic simulations

Results of the dynamic-model simulations will be discussed first, and then comparisons will be made of the skill of these simulations with those from the automated algorithmic system and those based on persistence. Three pairs of simulations using both the dynamic model and the automated algorithmic system were conducted, for which the differences among the experiments are the time period spanned by the simulations during the life cycle of the Buffalo Creek thunderstorm. Two radar-based datasets of reflectivity and radial velocity, separated by approximately 6 min, were used in the retrieval for the dynamic-model initialization. Table 1 shows the retrieval period and the simulation period for each of the experiments. Experiments 1 and 2 involve contiguous simulation periods of approximately 1 h each. Experiment 3 is a longer simulation that is initialized 12 min after experiment 1 is initialized and runs for 105 min until the time of the end of simulation 2. The 105-min period of experiment 3 is challenging given the expected skill of both the dynamic model and the automated algorithm in mountainous terrain. For this reason, the experiment-3 simulations were initiated 12 min later than those of experiment 1 in order for there to be a stronger signature of the rainfall at the initial time. The WSR-88D data rather than the S-Pol data were used as input to the dynamic model and the automated algorithm, because S-Pol data were not acquired above the lowest scan elevations. In addition, employing WSR-88D data is consistent with the concept of using data that currently are available operationally rather than using experimental data.

1) Dynamic-model simulations

Figure 10 depicts the domain-total rainfall as simulated by the dynamic model for all three periods and as inferred from the WSR-88D radar retrieval. The WSR-88D data are used for comparison because they were...
Fig. 9. Storm-total rain accumulations for the Buffalo Creek storm based on the WSR-88D reflectivity and S-Pol differential-phase measurements.
TABLE 1. Retrieval period and simulation period for the dynamic model simulations of precipitation (13 Jul 1996). Local standard time at the watershed is UTC minus 7 h.

<table>
<thead>
<tr>
<th>Experiment</th>
<th>Start of retrieval</th>
<th>End retrieval/start simulation</th>
<th>End of simulation</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0122 UTC</td>
<td>0128 UTC</td>
<td>0227 UTC</td>
</tr>
<tr>
<td>2</td>
<td>0221 UTC</td>
<td>0227 UTC</td>
<td>0325 UTC</td>
</tr>
<tr>
<td>3</td>
<td>0134 UTC</td>
<td>0140 UTC</td>
<td>0325 UTC</td>
</tr>
</tbody>
</table>

used for model initialization and thus should be the standard against which to gauge the model performance. The experiments initialized near the genesis of the storm (experiments 1 and 3) correctly simulate the intensification in the rain rate by a factor of 4–8 during approximately the first hour, but the simulated timing of the intensification is approximately 20 min late. The experiment initialized near the middle of the event (experiment 2) shows an erroneous decay in the precipitation intensity for about 30 min after the initialization, before the rates recover to more-realistic values. These curves, of course, only reflect biases in the domain-total, model-simulated precipitation and contain no information about possible critical position errors.

As a metric of position error, Fig. 11 shows the dynamic-model simulations’ threat scores for a range of thresholds for the precipitation accumulated at various times in the simulation period. For experiment 1, the threat scores improve with time because, as shown in Fig. 10, the simulated rain rate is slow to develop and does not approach the observed value until late in the simulation. For experiment 2, the threat scores increase late in the period as the simulated precipitation recovers from its anomalous minimum. The scores for experiment 3 are low initially because, as with experiment 1, the simulated rainfall is slow to develop. Because the threat scores pertain to accumulated precipitation, the scores, especially for higher thresholds, tend to increase with time as patterns of heavier accumulations become more organized.

In all three of the experiments, the dynamic model—simulated, domain-averaged rainfall rate decreases with time in the first 12–24 min and then increases to near the observed rate (Fig. 10). All of the limitations of the model and the data assimilation system described earlier can contribute to the errors in the assimilation and simulation. These initial adjustments can result from at least two problems. First, the radar cannot discriminate between liquid and solid precipitation, nor does the model represent the ice phase, so that all the radar-observed

![Figure 10](image1.png)  
**Fig. 10.** Storm-total average rain rate as a function of time based on WSR-88D data and rain rates from the three dynamic-model simulations.

![Figure 11](image2.png)  
**Fig. 11.** Threat scores for cumulative precipitation as a function of time based on the WSR-88D data and rain rates from the three dynamic-model simulations. Different curves are labeled in terms of the precipitation thresholds.
water is assumed to be liquid. This overestimation of the rainwater mixing ratio in upper levels, and the associated vertical velocity, can result in a delay in the precipitation. Second, the retrieved initial conditions are not perfectly balanced because of the lack of prior information, so the model undergoes some initial adjustments that affect the rainfall rates.

2) COMPARISON OF DYNAMIC-MODEL SIMULATIONS WITH THOSE FROM THE AUTOMATED ALGORITHM AND FROM PERSISTENCE

Figure 12 shows threat scores for the two types of precipitation simulations (dynamic model and automated algorithm based) and for “persistence” for the rainfall totals for each of the three simulation periods. Persistence corresponds to the assumption that the rainfall-rate pattern does not change with time. The abscissa in the figure represents the threshold for the threat score. To help in understanding these plots, Fig. 13 shows the total observed rainfall and the simulated rainfall for the dynamic model, the automated algorithm, and persistence for each of the three simulation periods.

Experiment 1 spans the development period and the early part of the mature period of the storm, as shown in Fig. 10. Figures 12 and 13 show that the persistence simulation does not perform well, because the observed rain rates are small at the beginning of the period and increase rapidly (Fig. 10). The fact that the pattern simulated by persistence does not move is reflected in the localization of the greatest rainfall accumulation. The TITAN algorithm of the Autonowcaster caused the two maxima in the initial pattern (as reflected in the double maxima in the persistence simulation) to divide and to intensify during the period. The eastern precipitation area verifies somewhat favorably against the observations, but the western maximum is completely erroneous. The area of the dynamic model–simulated rainfall is reasonable, but the accumulation is less than that inferred by radar, as seen in Fig. 10. In summary, based on Figs. 12 and 13, the model simulation is best, with good 1-h threat scores that far exceed persistence. The automated algorithm also has threat scores that greatly exceed persistence and places considerable rainfall over the Buffalo Creek watershed, but the existence of the erroneous precipitation over a watershed to the west is a major error.

The period covered by experiment 2 includes the latter part of the mature phase and roughly the first half of the decay phase of the storm. The precipitation at the initial time is near its maximum intensity and is widespread. Persistence performs reasonably well during this period because the rain rates and storm position do not change greatly. In terms of the threat scores, the automated-algorithm and persistence forecasts have similar quality, and both produce good simulations in terms of generating rainfall over the watershed. The skill of the dynamic-model simulation is greater than that of the other two techniques for the low-to-moderate thresholds.

Experiment 3 spans most of the development phase, all of the mature phase, and most of the decay phase of the storm. Both the dynamic model and the automated algorithm showed considerable skill for the 105-min simulation in terms of the threat scores for the lower thresholds. A comparison with persistence probably is not reasonable because of the longer simulation period.
6. Summary

Precipitation estimates for the Buffalo Creek storm that are based on WSR-88D and S-Pol radars are impacted by beam blockage at the lower elevation angles. Vivekanandan et al. (1999) describe the superiority of S-Pol rainfall estimates derived from differential propagation-phase measurements in complex terrain. The Buffalo Creek storm total maximum rainfall based on S-Pol is about 81 mm, and, based on the WSR-88D, it is about 61 mm. Maxima in both analyses are over the northeast corner of the watershed and in the wildfire burn area.

In a review of the techniques that are available or are under development for the operational very short range prediction of rainfall, it was pointed out that virtually all are not ideal for application to convective weather in complex terrain. Nevertheless, it is essential that the best available techniques be identified and refined for use in providing rainfall forecasts for input to surface-hydrologic models that can predict stream discharge.

Rainfall simulations produced by the dynamic model were generally superior to those from the automated algorithmic system. Even though the dynamic-model simulations did show considerable erroneous temporal variation in the rain rate, there was sufficient overall success to warrant optimism about potential future operational storm-scale, model-based prediction.

In spite of some problems, the automated algorithm did produce simulations that had threat scores that often were not greatly less than the dynamic model’s threat scores. Also, a subjective evaluation is that these automated algorithmic simulations would have been useful as operational forecasts for flash-flood prediction. An important point of comparison between the techniques is that the variationally initialized dynamic model will not be usable operationally for at least a few years because of the computational requirements associated with the use of its adjoint and the fact that a full-physics, convective-scale adjoint has not been developed and tested yet. In contrast, automated algorithmic systems currently are being used operationally for hydrologic applications, and the one employed in this case has demonstrated sufficient skill to give us hope that such computationally undemanding systems could be used operationally in complex terrain. Certainly when there is an area such as Buffalo Creek that is extremely vulnerable to large precipitation rates because of the burned surface, even a technique with modest quantitative skill would be valuable in defining the first-order characteristics of thunderstorm movement and growth.

In the companion paper (Yates et al. 2000) the radar-estimated precipitation and the precipitation simulated by the dynamic model and the automated algorithm will be used as input to a runoff/stream-discharge model. Based on paleohydrologic estimates of the peak flood discharge, a tentative evaluation will be made of the relative utility of these precipitation estimates and simulations for use in flash-flood prediction in mountainous areas.

Acknowledgments. This research was supported through special funds from the National Science Foundation that have been designated for the U.S. Weather Research Program activities at NCAR. Robert Jarrett of the U.S. Geological Survey, Denver, assisted in the interpretation of the paleohydrologic estimates of the precipitation and the discharge maxima.

REFERENCES
——, and M. D. Hudlow, 1984: Quantitative precipitation forecast

Fig. 13. Total rain (mm) for the simulation period based on the WSR-88D radar data and as simulated by the dynamic model, the automated algorithmic system (Autonowcaster), and persistence for the three simulation periods of the different experiments. The area of the watershed is shaded. The isopleth interval for the isohyets is uneven, with values greater than 40 not plotted.


