The Kain–Fritsch Convective Parameterization: An Update

JOHN S. KAIN*

Cooperative Institute for Mesoscale Meteorological Studies, University of Oklahoma, Norman, Oklahoma

(Manuscript received 23 December 2002, in final form 16 July 2003)

ABSTRACT

Numerous modifications to the Kain–Fritsch convective parameterization have been implemented over the last decade. These modifications are described, and the motivating factors for the changes are discussed. Most changes were inspired by feedback from users of the scheme (primarily numerical modelers) and interpreters of the model output (mainly operational forecasters). The specific formulation of the modifications evolved from an effort to produce desired effects in numerical weather prediction while also rendering the scheme more faithful to observations and cloud-resolving modeling studies.

1. Introduction

Convective parameterization continues to be one of the most challenging aspects of numerical modeling of the atmosphere, especially for numerical weather prediction and global climate prediction. A number of convective parameterization schemes (CPSs) have been developed over the years (e.g., Manabe et al. 1965; Ooyama 1971; Kuo 1974; Arakawa and Schubert 1974; Fritsch and Chappell 1980; Bougeault 1985; Betts 1986; Frank and Cohen 1987; Tiedtke 1989; Gregory and Rowntree 1990; Emanuel 1991; Grell 1993), and many of these schemes continue to be used and modified (e.g., Janjić 1994; Cheng and Arakawa 1997; Emanuel and Zivkovic-Rothman 1999; Gregory et al. 2000; Grell and Devenyi 2002). One such parameterization is the Kain–Fritsch (KF) scheme (Kain and Fritsch 1990, 1993, hereinafter KF90, KF93, respectively), which has been used successfully for many years in the Pennsylvania State University–National Center for Atmospheric Research Mesoscale Model (Wang and Seaman 1997; Kuo et al. 1996, 1997; Cohen 2002) and has been incorporated more recently into experimental versions of the National Centers for Environmental Prediction (NCEP) Eta Model (Black 1994), the new Weather Research and Forecasting model (Skamarock et al. 2001), and various other models (e.g., Bechtold et al. 2001).

Testing of the scheme within the Eta Model has been unique in that it has been carried out in close collaboration with forecasters at the National Oceanic and Atmospheric Administration (NOAA)/National Weather Service (NWS)/Storm Prediction Center (SPC). A modified configuration of the Eta Model, including the KF scheme, has been run at the NOAA/OFFICE OF Oceanic and Atmospheric Research/National Severe Storms Laboratory (NSSL) in a semioperational mode since 1998. In particular, this configuration (locally known as the EtaKF run) has been run 2 times per day in parallel with the operational Eta Model. Output from these forecasts arrives typically about 1–3 h after the operational guidance, well within the window of time during which it has potential to be useful for daily forecasts at SPC. Scientists from NSSL and the Cooperative Institute of Mesoscale Meteorological Studies (CIMMS) University of Oklahoma have striven to make this output useful to SPC forecasters, and feedback from the forecasting unit has played a significant role in assessing model performance (Kain et al. 2003a). This feedback has ultimately led to modifications of the KF scheme and improved forecasts.

The purpose of this paper is to document and describe these modifications as a resource for users of the KF scheme. In the next section, the original version of the scheme is described briefly. The next section describes modifications to the scheme and the motivations for making these changes. The last section provides a summary.

2. The “original” Kain–Fritsch scheme

The KF scheme was derived from the Fritsch–Chappell CPS, and its fundamental framework and closure assumptions are described by Fritsch and Chappell (1980). KF90 modified the updraft model in the scheme
and later introduced numerous other changes, so that it eventually became distinctly different from the Fritsch–Chappell scheme. It was distinguished from its parent algorithm by referring to the more elaborate code as the KF scheme, beginning in the early 1990s (KF93).

These early papers documented many details of the code. Additional details can be found in Bechtold et al. (2001); although this paper describes a significantly modified version of the KF scheme, it documents some sections of the KF code that are not available in print elsewhere; it thus provides a valuable additional reference. Furthermore, a less quantitative description of the code was recently presented in a paper that describes one of its unique applications (Kain et al. 2003b). Here, a brief overview of the “old” versions of the code is presented to provide the context for the description of the recent modifications.

The KF scheme is a mass flux parameterization. It uses the Lagrangian parcel method (e.g., Simpson and Wiggert 1969; Kreitzberg and Perkey 1976), including vertical momentum dynamics (Donner 1993), to estimate whether instability exists, whether any existing instability will become available for cloud growth, and what the properties of any convective clouds might be. For the sake of this discussion, it is convenient to compartmentalize the KF scheme into three parts: 1) the convective trigger function, 2) the mass flux formulation, and 3) the closure assumptions. Each of these is discussed briefly below.

a. The trigger function

The first task of the scheme is to identify potential source layers for convective clouds, that is, updraft source layers (USLs). Beginning at the surface, vertically adjacent layers in the host model are mixed until the depth of the mixture is at least 60 hPa. This combination of adjacent model layers composes the first potential USL. The mean thermodynamic characteristics of this mixture are computed, along with the temperature and height of this “parcel” at its lifting condensation level (LCL). As a first measure of the likelihood of convective initiation, parcel temperature $T_{\text{LCL}}$ is compared with the ambient temperature $T_{\text{ENV}}$ at the parcel LCL. The parcel will typically be colder than its environment, that is, negatively buoyant. Based on observations suggesting that convective development tends to be favored by background vertical motion (Fritsch and Chappell 1980), the parcel is assigned a temperature perturbation linked to the magnitude of grid-resolved vertical motion. The specific formula for this perturbation $\delta T_{vv}$ is

$$\delta T_{vv} = k[w_g - c(z)]^{1/3}, \quad (1)$$

where $k$ is a unit number with dimensions K s $^{1/3}$ cm$^{-1/3}$, $w_g$ is an approximate running-mean grid-resolved vertical velocity at the LCL (cm s$^{-1}$), and $c(z)$ is a threshold vertical velocity given by

$$c(z) = \begin{cases} w_0(Z_{\text{LCL}}/2000), & Z_{\text{LCL}} \leq 2000 \\ w_0, & Z_{\text{LCL}} > 2000. \end{cases} \quad (2)$$

where $w_0 = 2$ cm s$^{-1}$ and $Z_{\text{LCL}}$ is the height of the LCL above the ground (m). For example, this equation yields a temperature perturbation of 1 K for a background vertical velocity of 1 cm s$^{-1}$ above the threshold value and just over 2 K when $w_g$ is 10 cm s$^{-1}$ above the threshold value.

Use of this perturbation term allows us effectively to eliminate most parcels as candidates for deep convection, which is important for computational efficiency. The elimination process involves adding the computed temperature perturbation (typically 1–2 K, e.g., in environments with weak to moderate upward motion) to the parcel temperature at the LCL. If the resulting temperature is still less than the environmental value (i.e., $T_{\text{LCL}} + \delta T_{vv} < T_{\text{ENV}}$), then this parcel is eliminated from consideration, the base of the USL is moved up one model level, and the above test is repeated for a new potential USL. If, however, the perturbed parcel is warmer than its environment, it is allowed to proceed as a candidate for deep convection. At this stage, the parcel is released at its LCL with its original (unperturbed) temperature and moisture content and a vertical velocity derived from the perturbation temperature. To be specific, its initial vertical velocity $w_{p0}$ is loosely based on the parcel buoyancy equation and is given by

$$w_{p0} = 1 + 1.1[(Z_{\text{USL}} - Z_{\text{LCL}})\delta T_{vv}/T_{\text{ENV}}]^{1/2}, \quad (3)$$

where $Z_{\text{USL}}$ is the height at the base of the USL. This formula yields starting vertical velocities of up to several meters per second.

Above the LCL, parcel vertical velocity is estimated at each model level using the Lagrangian parcel method, including the effects of entrainment, detrainment, and water loading (Frank and Cohen 1987; Bechtold et al. 2001). If vertical velocity remains positive over a depth that exceeds a specified minimum cloud depth (typically 3–4 km), deep convection is activated using this USL. If not, the base of the potential USL is moved up one model layer and the procedure is repeated. This process continues until either the first suitable source layer is found or the sequential search has moved up above the lowest 300 hPa of the atmosphere, where the search is terminated. This complete set of criteria composes the trigger function, but note that the updraft model described in the next section plays an important role in determining cloud depth and, as a consequence, whether the parameterization is activated.

b. Mass flux formulation

Convective updrafts in the KF scheme are represented using a steady-state entraining–detraining plume model, where equivalent potential temperature ($\theta_e$) and water vapor ($q_e$) are both entrained and detrained, and detrainment also includes various hydrometeors, as de-
scribed in detail in KF90. In this model, entrainment and detrainment rates are inversely proportional, with high entrainment (detrainment) rates being favored by high (low) parcel buoyancy and moist (dry) environments. In practice, the distinction between the updraft and the trigger function can become blurred because the specific formulation of the updraft can determine whether the specified minimum cloud depth for deep convection is achieved.

Convective downdrafts are fueled by evaporation of condensate that is generated within the updraft. A fraction of this total condensate is made available for evaporation within the downdraft, based on empirical formulas for precipitation efficiency as a function of vertical wind shear and cloud-base height (Zhang and Fritsch 1986). This fraction effectively dictates the relative wind shear and cloud-base height (Zhang and Fritsch 1986). This fraction effectively dictates the relative magnitudes between downdraft and updraft mass fluxes once other critical downdraft parameters are specified. These other parameters include the downdraft starting and ending levels, its relative humidity profile, and the characteristics and amounts of entrained air. The downdraft is specified to start at the level of minimum saturation equivalent potential temperature \( \theta_e \) in the cloud layer with a mixture of updraft and environmental air. It is moved downward in a Lagrangian sense, with a specified entrainment rate (entraining environmental air only) and a fixed relative humidity of 100% above cloud base and 90% below cloud base. The downdraft is terminated if it becomes warmer than its environment or if it reaches the surface. It is forced to detrain into the environment within and immediately above the termination level, such that the minimum depth of the detrainment layer is the same as the minimum depth of the USL, 60 hPa.

Environmental mass fluxes are required to compensate for the upward and downward transports in the updrafts and downdrafts, so that the net convective mass flux at any level in the column is zero. This general framework for computing convective effects has been used for many years (e.g., Johnson 1976); the specific formulation for the KF scheme is described in KF93.

c. Closure assumptions

The method by which the KF scheme satisfies its closure assumptions is described in Bechtold et al. (2001). In fundamental terms, the KF scheme rearranges mass in a column using the updraft, downdraft, and environmental mass fluxes until at least 90% of the convective available potential energy (CAPE) is removed. CAPE is computed in the traditional way, using \( \text{undilute} \) parcel ascent, with the parcel characteristics being those of the USL. CAPE is removed by the combined effects of lowering \( \theta_e \) in the USL and warming the environment aloft. The convective time scale, or relaxation period, is based on the advective time scale in the cloud layer, with an upper limit of 1 h and a lower limit of 0.5 h. The scheme feeds back convective tendencies of temperature, water vapor mixing ratio, and cloud water mixing ratio. By default, convective precipitation particles simply accumulate at the surface rather than being introduced aloft, but the code has a “switch” to activate feedback of precipitation at the level it is formed. The switch can be set to any value from 0 (no feedback) to 1 (100% feedback).

3. Recent modifications to the KF scheme

Several different components of the KF scheme have been changed in recent years. They are described individually below.

a. Updraft formulation

The algorithm for the KF updraft has been modified with a specified minimum entrainment rate and formulations to allow variability in the cloud radius and cloud-depth threshold for deep (precipitating) convection. Furthermore, the effects of shallow (nonprecipitating) convective clouds are now included as well. These changes and the motivations for making them are discussed below.

1) Minimum entrainment rate

A common criticism from users of the old version of the KF scheme was that it sometimes produced widespread light precipitation in marginally unstable environments and, perhaps as a consequence, it tended to underpredict maximum rainfall amounts within the precipitation area (e.g., Warner and Hsu 2000; Colle et al. 2003). Furthermore, comparison with cloud-resolving model simulations suggested that updrafts were penetrating too far aloft (e.g., Liu et al. 2001). Early testing of the KF scheme in the Eta Model corroborated these observations. For example, the scheme often generated widespread “airmass thunderstorms” over the southeastern United States during the summer when observed convective activity was isolated or even nonexistent.

Diagnostic analysis of the scheme’s behavior revealed that one of the problems was related to the representation of entrainment–detrainment processes. As described in KF90; the rate that environmental air mixes with an updraft is specified in old versions of the system. The rate is specified in old versions of the scheme, but while some of that air typically mixes inward to dilute the mean properties of the updraft, some of it is allowed immediately to detrain back into the environment. If it detrains back into the environment, it does so within turbulent mixtures of updraft and the environment; that is, it extracts updraft air in the process. With this formulation, entrainment and detrainment rates are inversely proportional and they depend on the likelihood that negatively buoyant parcels can be generated when environmental air mixes with the updraft air, including its liquid water or ice. Entrainment of environmental air is favored when updrafts are much
Figure 1. Predicted updraft path (thick, dark solid line) in a marginally unstable, relatively dry environment using the (a) original and (b) updated version of the KF scheme.

Figure 2. UMF predicted by the original version of the KF scheme and DMF predicted by the old and new versions of the scheme for the sounding shown in Fig. 1. All profiles are normalized by the UMF at the top of the USL.

Figure 3. Normalized Mass Flux

In environments where updrafts are marginally buoyant and the environment is relatively dry, detrainment dominates because the evaporative cooling potential is relatively large and relatively little cooling is necessary to induce negative buoyancy. In contrast, when updrafts are marginally buoyant and the environment is relatively moist, entrainment dominates because the positive buoyancy is large before mixing and evaporative cooling potential is limited by the moist environment. In environments where updrafts are marginally buoyant and the environment is relatively moist, the updraft is only slightly warmer than the environment over a deep layer. In the first 100 hPa or so above the cloud base (i.e., approximately the 800–900-hPa layer), the environment is relatively moist and the updraft is just starting to accumulate liquid water, and so the entrainment process dominates, as reflected by a sharp increase in updraft mass flux (UMF; Fig. 2) and a decrease, or dilution, of updraft temperature (Fig. 3). However, in the relatively dry air of the next ~400 hPa, the detrainment processes dominate (Fig. 2) and the updraft parcels in the old KF scheme can ascend with very little dilution from the environment because environmental air that initially mixes with the updraft is left behind in negatively buoyant mixtures. Net entrainment can be minimal, and mean updraft thermodynamic properties may remain nearly undiluted over a deep layer of ascent. As a result, the unmodified KF scheme has a tendency to allow deep convection to activate too easily when the environmental lapse rate is neutral to slightly unstable, convective inhibition (CIN) is small, and the deep layer relative humidity (RH) is low. Furthermore, because the updraft shed most of its mass and moisture well below the equilibrium level in this circumstance, total condensation and production of precipitation can be very small.

This problem is largely responsible for the widespread light precipitation sometimes associated with the original KF scheme. It is mitigated in newer versions of the scheme by simply imposing a minimum entrainment rate for the updraft. In particular, the entrainment–detrainment calculations described in KF90 are performed initially, but the net environmental entrainment rate, $M_{ee}$ (kg s$^{-1}$, using the notation of KF90), is not allowed to fall below 50% of the total environmental air that mixes into the updraft:

$$M_{ee} \geq 0.5\delta M_e,$$

where $\delta M_e$ is the mixing rate (kg s$^{-1}$).

This change has a significant impact in some environments. For example, Fig. 1a shows the updraft path predicted by older versions of the scheme in one type of marginally unstable, relatively dry environment. The parameterized updraft is only slightly warmer than the environment over a deep layer. In the first 100 hPa or so above the cloud base (i.e., approximately the 800–900-hPa layer), the environment is relatively moist and the updraft is just starting to accumulate liquid water, and so the entrainment process dominates, as reflected by a sharp increase in updraft mass flux (UMF; Fig. 2) and a decrease, or dilution, of updraft temperature (Fig. 3). However, in the relatively dry air of the next ~400 hPa, the detrainment processes dominate (Fig. 2)
draft undergoes very little dilution (Fig. 3). In contrast, with the minimum entrainment rate imposed, relatively strong dilution of the updraft continues above 800 hPa, and so the updraft parcel loses its upward momentum by the time it reaches about 600 hPa (Figs. 1b, 3), yielding a cloud depth less than the minimum value required for deep convection. In older versions of the scheme, this would mean that no parameterized convection would occur, but in newer versions shallow (nonprecipitating) convection would be activated [see section 3a(4) below].

Real-time testing has shown that this modification has a favorable impact in reducing the areal coverage of widespread light precipitation and increasing maximum rainfall amounts within contiguous rainfall areas. At the same time, it has minimal impact in environments with higher instability and/or more humid cloud-layer conditions.

2) Variable Cloud Radius

In addition to ensuring some dilution of updraft parcels by imposing a minimum entrainment rate, it can be argued that the potential dilution should be a function of larger-scale forcing for convection (Frank and Cohen 1987). In other words, it seems reasonable to introduce additional factors that will promote convective initiation when larger-scale forcing is favorable and to suppress initiation when forcing is weak or negative. Indeed, the impact of large-scale destabilizing processes is included (either directly or indirectly) in the trigger functions of most other CPSs (e.g., Arakawa and Schubert 1974; Anthes 1977; Bougeault 1985; Tiedtke 1989; Grell 1993; Janjić 1994).

As a way of introducing this sensitivity in new versions of the KF scheme, the cloud radius is rendered as a function of larger-scale forcing. As indicated in KF90, cloud radius $R$ (m) controls the mixing rate (the maximum possible entrainment rate) according to

$$\delta M = M_{w_0} \left(\frac{-0.03\delta p}{R}\right),$$

where $M_{w_0}$ is the updraft mass flux (kg s$^{-1}$) at cloud base, $\delta p$ is the pressure depth of a model layer (Pa), and 0.03 is a constant of proportionality (m Pa$^{-1}$). In older versions of the KF scheme, $R$ is held constant, typically at a value of 1500 m. In the modified code, a conservative attempt to introduce some dependence of $R$ on larger-scale forcing has been included by making it dependent on the magnitude of vertical velocity at the LCL. To be specific, $R$ is defined as

$$R = \begin{cases} 1000, & W_{KL} < 0 \\ 2000, & W_{KL} > 10 \\ 1000(1 + W_{KL}/10), & 0 \leq W_{KL} \leq 10, \end{cases}$$

where $W_{KL}$ (cm s$^{-1}$) is the term inside brackets in (1): $W_{KL} = w_c - c(z)$. With this modification, the mixing rate increases as vertical velocity decreases near cloud base. Combined with the minimum entrainment rate (as a fraction of $\delta M$), discussed above, this modification typically results in higher dilution of cloud parcels when subcloud-layer forcing is weak or negative. It promotes weaker dilution when low-level forcing is stronger.

As alluded to above, this is a conservative (in that $R$ varies over a fairly limited range and never goes below 1000) approach to introducing the fundamental entrainment sensitivity advocated by Frank and Cohen (1987). Furthermore, it is another way of including a sensitivity to deep-layer relative humidity in the KF cloud model. It was motivated by evidence that midlevel moisture strongly modulates convective rainfall (e.g., Shepherd et al. 2001; Tompkins 2001) and by observations that the Betts–Miller–Janjić scheme, in which deep-layer moisture is effectively the trigger function (Baldwin et al. 2002), is very effective at capturing deep convective activity that is associated with organized mesoscale and larger-scale processes in day-to-day predictions from the Eta Model. However, the efficacy of this current formulation appears to be limited because it is dependent on vertical velocity at only one level. Alternative formulations, based on grid-resolved forcing over a deep layer, are being tested.

It is emphasized that, although cloud radius is the critical parameter in (5) and (6), we have little or no skill in actually predicting what the horizontal dimensions of convective clouds in the atmosphere will be. Furthermore, the basic entrainment relationships from which (5) is derived are based on idealized laboratory experiments involving fluids that are very different from latent-heat-driven convective clouds (e.g., Simpson 1983). The validity of these quantitative relationships for atmospheric convection is tenuous, at best (Emanuel 1994, p. 540). Thus, any specific value for cloud radius in (5) should not be taken too literally. Rather, application of (5) should be viewed simply as a mechanism to modulate the rate of dilution, and, thereby, cloud top,

![Fig. 3. Updraft $\theta_u$ as a function of height (pressure) for the sounding shown in Fig. 1 using the old and new (minimum entrainment) updraft algorithm. Note that the “Old E/D” curve corresponds to the UMF profile shown in Fig. 2.](image-url)
condensation rate, and so on, in parameterized clouds. This point of view is consistent with its use in one of the earliest and most enduring convective parameterizations, the Arakawa and Schubert (1974) scheme. In this scheme, $R$ (i.e., the entrainment rate) is manipulated systematically to generate an ensemble of entraining updrafts, such that each model computational level serves as the cloud top for at least one member of the ensemble (Lord 1982).

3) VARIABLE MINIMUM CLOUD-DEPTH THRESHOLD

Previous versions of the KF scheme used a specified minimum cloud-depth threshold, typically set at 3–4 km. The intention was to delineate between convective clouds that produce precipitation at the surface and/or a precipitation-induced downdraft, and those that do not. The specified value seemed to work effectively in most situations. However, in semioperational prediction with the ETA run, it was noted that this specified value can be inadequate. This fact was particularly evident in predictions of “lake-effect snow” (e.g., Niziol et al. 1995), in which observations indicate that significant snowfall rates can come from convective clouds that are only around 2 km deep.

To allow the KF scheme to parameterize this process, it was deemed necessary to decrease the minimum cloud depth to 2 km. In more general terms, it was reasoned that precipitation production is likely to be favored by active ice-phase processes, so that when cloud-base temperature is close to 0°C, precipitation is possible with relatively shallow convective clouds. In the absence of a known robust quantitative relationship, it was decided to make the minimum cloud depth a function of $T_{LCL}$ ($°C$). Minimum cloud depth $D_{\text{min}}$ (m) is now specified according to

\begin{align*}
D_{\text{min}} = \begin{cases} 
4000, & T_{LCL} > 20°C \\
2000, & T_{LCL} < 0°C \\
2000 + 100T_{LCL}, & 0° \leq T_{LCL} \leq 20°C.
\end{cases}
\end{align*}

4) SHALLOW (NONPRECIPITATING) CONVECTION

Parameterization of shallow convection has long been recognized as an important component of global climate models (e.g., Browning et al. 1993; Siebesma and Cuijpers 1995) and in recent years has become an important concern for mesoscale numerical weather prediction (NWP) models (Gregory and Rowntree 1990; Deng et al. 2003). Parameterized shallow convection transports moisture upward and heat downward within the shallow cloud layer. For NWP and operational forecasting, this process is particularly important because it affects vertical structures, including the boundary layer in some cases, and it often modulates the timing of deep convective initiation. At SPC, and elsewhere in the NWS, forecasters rely heavily on analysis of model-forecast soundings, and these soundings are strongly affected by parameterized shallow convection in the models (Kain et al. 2001; Baldwin et al. 2002). For example, fundamental derived sounding characteristics such as CAPE and CIN can be changed dramatically by parameterized shallow convection.

In the modified KF scheme, shallow convection is activated when all of the criteria for deep convection are satisfied except that the cloud model yields an updraft more shallow than the minimum cloud depth, similar to the convective schemes in the NCEP ETA model (Janjić 1994; Baldwin et al. 2002) and the European Centre for Medium-Range Weather Forecasts Integrated Forecast System (Gregory et al. 2000). As part of the shallow-convective modifications, $\delta T_{\text{env}}$ is set to zero if (1) yields a negative value. With this change, shallow convection is not suppressed by subsidence at the LCL, but parcels are assigned zero temperature perturbation in a subsidence regime. Without a positive perturbation, a parcel must be warmer than its environment at its LCL to satisfy the first test of the KF trigger function. This implies that the subcloud-layer lapse rate must be superadiabatic, as during strong daytime heating over land, for KF shallow convection to activate when air is sinking on resolved scales at the LCL.

Shallow convection is activated only after every potential USL in the lowest 300 hPa has been rejected as a candidate for deep convection. As the trigger function evaluates the potential for deep convection, the cloud depth associated with each USL is saved. If deep convection fails to activate, but one or more shallow clouds are found (i.e., cloud height > 0), the deepest “shallow” cloud is activated. For computational reasons, the value of $R$ is not changed for shallow clouds. The KF90 entrainment–detrainment algorithm is used initially to determine cloud properties and mass flux characteristics, but in the final updraft calculations total mass detrainment is specified to occur as a linear function of decreasing pressure between the LCL and cloud top. So, in effect, the rate of dilution of updraft air is determined by the KF algorithm, but the mass flux and detrainment profiles are specified to be at least qualitatively consistent with large-eddy simulation results (i.e., Siebesma and Cuijpers 1995).

Although shallow convective clouds are, by common definition, nonprecipitating, the KF algorithm typically generates precipitation in any cloud deeper than about 50 hPa. If an updraft has already been classified as shallow, any precipitation that is generated by the scheme is fed back to resolved scales as an additional moisture source. This is accomplished by setting the switch to activate elevated precipitation feedback (see section 2c) to 1.

Parameterized shallow clouds are also modulated by a different closure assumption. In particular, the cloud-base mass flux $M_{\text{e}}$ is assumed to be a function of turbulent kinetic energy (TKE) in the subcloud layer. This general relationship was initially deduced from physical
reasoning rather than quantitative measurements, but recent studies elsewhere have arrived at similar hypotheses regarding the likely relationship between TKE and \( M_{o0} \) (e.g., Grant 2001; Neggers et al. 2003, manuscript submitted to Mon. Wea. Rev., hereinafter NEG03). The quantitative relationship in the KF scheme is based on the concept of scaling \( M_{o0} \) by the maximum TKE in the subcloud layer. In sampling TKE values from the Eta Model’s turbulence parameterization, it was found that boundary layer (assuming most shallow clouds are driven by surface fluxes) TKE generally varies from 0 m\(^2\) s\(^{-2}\) in stable situations to about 10 m\(^2\) s\(^{-2}\) in unstable boundary layers with very strong heating from below. Thus, it was decided to assign the maximum value of \( M_{o0} \) when \( TKE_{\text{MAX}} \approx 10 \), ramping linearly down to zero when there was no TKE.

But what should the maximum value of \( M_{o0} \) be? A normalized updraft mass flux value, UMF\(^*\), has proven to be a useful diagnostic quantity in the KF scheme (Kain et al. 2003b). This quantity is based on the fraction of the USL that is processed by the convective scheme during each convective cycle,

\[
UMF^* = (M_{o0} \tau_c) / m_{\text{USL}}.
\]

where \( \tau_c \) is the convective time period, ranging from 1800 to 3600 s, and \( m_{\text{USL}} \) is the amount of mass in the USL (kg). Initial testing associated a value of UMF\(^*\) = 1 with \( TKE_{\text{MAX}} = 10 \), that is,

\[
M_{o0} = \begin{cases} \left( \frac{TKE_{\text{MAX}}}{k_0} \right) \left( \frac{m_{\text{USL}}}{\tau_c} \right), & \text{TKE}_{\text{MAX}} < 10 \\ \left( \frac{10}{k_0} \right) \left( \frac{m_{\text{USL}}}{\tau_c} \right), & \text{TKE}_{\text{MAX}} \approx 10, \end{cases}
\]

where \( k_0 = 10 \text{ m}^2 \text{ s}^{-2} \), implying that all of the mass in the USL would be processed during a convective cycle when subcloud layer \( TKE_{\text{MAX}} \approx 10 \text{ m}^2 \text{ s}^{-2} \). Although \( TKE_{\text{MAX}} \) values typically remain well below 10 m\(^2\) s\(^{-2}\) during warm-season diurnal cycles over land, this formulation seemed to produce tendencies that were too strong, and so \( k_0 \) was increased to 20. This formula seems to produce about the right magnitude of feedback tendencies, judging from the favorable impact in EtaKF forecasts (Kain et al. 2001; Baldwin et al. 2002). For example, Fig. 4a shows the updraft path predicted by the scheme in an environment in which deep convection is strongly inhibited. In this case, the predicted cloud is less than 1 km deep. The maximum subcloud-layer TKE is about 5 m\(^2\) s\(^{-2}\), so that \( M_{o0} \) is about 0.25\( M_{\text{USL}} / \tau_c \). Maximum temperature and moisture tendencies are on the order of 1 K h\(^{-1}\) and 1 g kg\(^{-1}\) h\(^{-1}\), respectively (Fig. 4b).

It is clear that this formulation could be better calibrated and, perhaps, adapted to more environmental parameters, such as cloud depth, boundary layer depth, and/or LCL position relative to the top of the boundary layer. As an alternative, an altogether different approach can be taken to implement the same fundamental closure assumption. For example, Grant (2001) argues that cloud-base mass flux for shallow clouds is proportional to the subcloud-layer vertical velocity scale, \( w_u \) (m s\(^{-1}\)), which is closely related to production of TKE and is readily available from most turbulence parameterizations. In particular, they reason that

\[
M_{o0} \approx k_i w_u,
\]

where \( k_i \) is a constant (kg m\(^{-1}\)). If this relationship is robust and \( k_i \) can be determined reliably, this would be a very simple closure. At this stage, the relationship between subcloud-layer TKE and cloud-base mass flux seems to be valid, but more work is needed to quantify, and perhaps qualify, better this relationship and its utility in shallow convective parameterization (NEG03).

### b. Downdraft

Convective downdrafts play an essential role in atmospheric convection. This is obvious in the lower troposphere, where downdrafts transport relatively low \( \theta_e \) air into the subcloud layer and strongly stabilize the local vertical structure. Downdrafts serve this same function in a parameterization scheme. Furthermore, as in the real atmosphere, they can enhance low-level convergence, favoring subsequent convective development at nearby points. However, parameterized downdrafts are also important for offsetting updraft mass flux in the lower troposphere. In particular, when downward mass flux is represented in moist, penetrative downdrafts, less environmental “compensating subsidence” is necessary and convective warming and drying tendencies in the lower part of the cloud layer tend to be more realistic (e.g., Johnson 1976; Cheng 1989).

The magnitude of these tendencies, which can have...
important implications for development of larger-scale precipitation processes and subsequent parameterized convection (Kain and Fritsch 1998), is modulated by the strength (mass flux) of the downdraft relative to the updraft. As a consequence, the performance of mass flux CPSs is very sensitive to parameters that control the ratio of these mass fluxes in the lower troposphere (Tiedtke 1989).

These parameters vary depending on the specific formulation for downdrafts. In general, parameterized downdrafts are driven by condensate from parameterized updrafts. They are typically conceived by an algorithm that determines their depth, (approximately) conserved thermodynamic properties (i.e., $\theta_c$), relative humidity, and the shape of their vertical mass flux profile. Once these properties are known, one can specify either (a) the downdraft mass flux value at some level (e.g., Tiedtke 1989; Frank and Cohen 1987) or (b) the amount of condensate available for evaporation in the downdrafts (e.g., Fritsch and Chappell 1980; Grell 1993) and then solve for the other.

The original KF scheme used the latter approach. It related an overall precipitation efficiency to vertical wind shear and cloud-base height (Zhang and Fritsch 1986). Its updraft model determined how much condensate was produced and how much of it detrained into the environment, whereas the empirical precipitation-efficiency relationship dictated the fraction of the condensate that reached the surface as precipitation. The remaining condensate was assumed to evaporate in the penetrative downdraft, providing a closure for the specification of downdraft mass flux.

This approach was used for a number of years, and the KF scheme has performed very well with it (e.g., Kuo et al. 1996; Wang and Seaman 1997; Gochis et al. 2002; Cohen 2002), but these precipitation-efficiency relationships are difficult to reconcile with observations and cloud-modeling studies. The relationship introduced by Fritsch and Chappell (1980), wherein precipitation efficiency is inversely proportional to vertical wind shear, does not seem to be valid over a wide range of conditions (e.g., Weisman and Klemp 1982; Fankhauser 1988; Ferrier et al. 1996). Furthermore, one can argue that the additional term included by Zhang and Fritsch (1986), which relates precipitation efficiency to cloud-base height, is not robust for general applications either. For example, a relatively high cloud base of 3 km can overlie a very dry, unstable convective boundary layer in which significant evaporation is likely to occur, but it can also lie at the top of a stable, saturated boundary layer, as is often the case in nocturnal heavy-rain events (e.g., Rochette and Moore 1996). The height of cloud base, by itself, is not a reliable indicator of the evaporation rate below cloud base. Last, even if these relationships were robust, the inverse of precipitation efficiency is not necessarily proportional to the local evaporation rate in the downdraft. For example, evaporation in the environment, but outside of lower-tropospheric downdrafts, is generally neglected but could be significant (Kreitzberg and Perkey 1976; Emanuel 1991).

Another problem with the KF downdraft formulation is the method for choosing the origination level of downdraft air. The starting level for the downdraft (level of minimum $\theta_e$; see section 2b) proves to be variable. It is not uncommon for this level to be as high as 300 hPa (e.g., Fig. 2) or as low as 850 hPa. Yet the amount of condensate available for maintaining the specified relative humidity does not change as a function of origination level. For a given amount of condensate and the same detrainment level (usually near the surface), a downdraft originating in the mid- to upper troposphere is relatively “tall and skinny” (i.e., has greater vertical depth but less mass flux at a given level), while one starting closer to the surface would be comparatively “short and fat” with the original KF formulation (e.g., compare the old and new downdraft mass flux profiles in Fig. 2). The corresponding ratios of downdraft to updraft mass flux near cloud base are relatively small (large) in the former (latter) case, implying larger (smaller) parameterized heating and drying rates. Furthermore, the mass of downdraft outflow in the subcloud layer is relatively small for taller updrafts and is larger for those originating lower. Thus, the original KF downdraft formulation leads to inconsistent predictions of lower-tropospheric heating and drying rates that are not justified by observational evidence or sound physical reasoning. The same is true for the Fritsch and Chappell (1980) approach, although it is somewhat different.

The new downdraft formulation in the KF scheme ameliorates some of these problems. It takes an approach in which key downdraft levels are linked specifically to the updraft. The downdraft is specified to start 150–200 hPa above the USL. This height is broadly consistent with most studies on this topic. For example, precipitation-driven downdrafts that penetrate into the subcloud layer appear to originate just above cloud base in relatively weak convective activity (e.g., Betts 1976; Zipser 1977; Knupp and Cotton 1985) and perhaps as much as a few kilometers above cloud base in more intense midlatitude convection over land (Knupp 1987).

The downdraft is formed entirely from environmental air, and it entrains equal amounts of air from all model layers within the downdraft source layer (DSL), which extends from the origination level to the top of the USL. Thus, at the top of the USL it is composed of a mass-weighted mixture of air from each model layer within the DSL. This approach is distinctly different from the old formulation that extracted most of the downdraft mass from a single origination level. Below the top of the USL, detrainment begins and entrainment ends. The downdraft is allowed to penetrate downward until it reaches the surface or it becomes warmer than its environment. Total detrainment is specified to occur as a linear function of pressure between the top of the USL and the base of the downdraft. Thus, the vertical profile of downdraft mass flux (DMF) shows a sharp peak at
the top of the USL and a linear decrease to zero above
and below (e.g., Fig. 2).

The downdraft is assumed to be completely saturated
above cloud base, with relative humidity decreasing by
20% km$^{-1}$ below this level, based loosely on the mod-
ing results of Srivastava (1985) and widespread ob-
servations of subsaturated downdrafts (e.g., Knupp and
Cotton 1985). The magnitude of the downdraft mass
flux at the top of the USL is specified currently as a
simple function of updraft mass flux and relative hu-
midity within the DSL.

$$\frac{-\Delta M_{U_{\text{USL}}}}{M_{U_{\text{USL}}}} = 2 \times (1 - RH), \quad \text{(11)}$$

where $RH$ is the mean (fractional) relative humidity in
the DSL. This formulation favors short, fat downdrafts
with maximum mass flux close to cloud base. Environ-
mental humidity is obviously important here, consistent
with its operative role in determining downdraft strength
(Knupp and Cotton 1985; Tompkins 2001) and precipi-
tation efficiency (Ferrier et al. 1996; Shepherd et al.
2001). Because $M_{U_{\text{USL}}}$ depends strongly on lower-tro-
pospheric lapse rates (Kain et al. 2003b), downdraft mass
flux is also very sensitive to environmental stability, an-
other factor emphasized by Knupp and Cotton (1985).

There is no longer a dependence on vertical wind
shear. Although wind shear undoubtedly plays a role,
especially with regard to precipitation efficiency, this
role appears to be complex and difficult to isolate from
other factors (Fankhauser 1988). In a conceptual sense,
windshear may induce a vertical tilt to updrafts, which
may reduce precipitation efficiency (Cheng 1989; Fer-
rier et al. 1996), but it is questionable whether this effect
can be quantified in a useful way for general application.
For example, Weisman and Klemp (1982) suggest that
precipitation efficiency may be inversely related to wind
shear for relatively low instability, low-shear environ-
ments in which pulse-type (i.e., single cell) convection
dominates, but they show that efficiency and shear ap-
pear to be positively correlated in environments with
higher shear and instability, in which mesoscale organ-
ization of convection is favored. In the face of this
uncertainty about quantitative relationships, it was de-
cided to exclude any dependence on wind shear in the
latest version of the KF scheme.

With this downdraft formulation, convective precipi-
tation is given by the residual condensate remaining after
updraft detrainment and downdraft evaporation. In some
cases, especially those with high cloud bases overlying
a deep convective boundary layer, the algorithm deter-
mines that no condensate remains after downdraft evap-
oration. In this circumstance, no convective precipitation
is produced but the scheme is still allowed to activate.
Fire-weather forecasters at SPC are investigating whether
predictions of deep convective UMF*, but no precipi-
tation, correspond to the occurrence of “dry lightning”
over the high terrain of the western United States (R.
Naden 2003, personal communication).

c. Closure assumption

As discussed in section 2c, the KF scheme uses a
CAPE closure. In specific terms, it increases mass fluxes
incrementally until CAPE is reduced by at least 90%,
where CAPE calculations are based on the mean char-
acteristics of air drawn from the USL before and after
the parameterized overturning.

In the original KF scheme, CAPE was computed on
the basis of undilute parcel ascent, as is typically done
for diagnostic calculations. However, it appears that the
scheme may overestimate convective rainfall and mass
flux [UMF*; see (8)] when it is programmed to eliminate
the relatively large positive area corresponding to un-
dilute ascent. For example, the positive area for an un-
dilute parcel can be much larger than the area associated
with an entraining parcel (cf. Figs. 5 and 1a). In this
case, when closure is based on undilute ascent the
scheme predicts much larger UMF* and precipitation
rate than when the dilute-parcel closure is used (Table
1). The scheme simply has more CAPE to eliminate
when calculations are based on undilute ascent.

In newer versions of the scheme, the closure is based
on the CAPE for an entraining parcel. This approach
provides reasonable rainfall rates for a broad range of
convective environments and it makes the scheme’s
UMF* field a better predictor of convective intensity. The
interested reader is referred to Kain et al. (2003b), in

\[\text{Rich Naden is a mesoscale assistant forecaster at SPC. His responsi-

\[\text{bilities include fire-weather forecasting.}\]
which detailed explanations of the UMF* field and the method for satisfying the closure assumption are given.

4. Summary

A number of modifications have been introduced in the Kain–Fritsch convective parameterization over the last decade or so. The purpose of this paper is to document these changes formally and provide some justification for their implementation. The changes were inspired by feedback from numerical modelers who use the scheme (e.g., Warner and Hsu 2000; Liu et al. 2001; Nagarajan et al. 2001; Cohen 2002; Colle et al. 2003) and from operational forecasters who utilize model output for daily forecasts (e.g., Kain et al. 2003a). The changes are briefly summarized here.

The updraft formulation was changed in several ways:

1) A minimum entrainment rate is imposed, primarily to suppress convective initiation in marginally buoyant, relatively dry environments. The minimum rate is 50% of the maximum possible entrainment rate defined by KF90.

2) The cloud radius, which controls the maximum possible entrainment rate, is specified to vary as a function of subcloud-layer convergence, similar to a formulation by Frank and Cohen (1987). This modification suppresses deep convective activation in weakly convergent or divergent environments and promotes activation in strongly convergent regimes.

3) A minimum cloud depth, required for activation of deep convection, is allowed to vary as a function of cloud-base temperature rather than remaining constant. This change is designed to allow the activation of deep convection for relatively shallow clouds when ice-phase processes are active.

4) Shallow (nonprecipitating) convective clouds are allowed. They are activated when the scheme’s cloud model determines that buoyant updrafts can form but cannot reach the imposed minimum cloud depth for deep convection. Cloud-base mass flux is based on TKE in the subcloud layer for shallow clouds, rather than CAPE.

The downdraft formulation was also changed:

1) A new downdraft algorithm is introduced. Downdrafts are formed from air in the layer at 150–200 hPa above cloud base, and they detrain over a fairly deep layer below cloud base. Downdraft mass flux is estimated as a function of the relative humidity and stability just above cloud base but is no longer related to vertical wind shear.

Last, changes to the closure assumption were made:

1) The scheme is still programmed to eliminate CAPE, but the calculation of CAPE is based on the path of an entraining (diluted) parcel rather than one that ascends without dilution.

This paper brings the formal documentation of the KF parameterization up to date with the latest working version of the scheme. As implied herein, traditional methods of convective parameterization, although rooted in scientific observations, become part engineering and part intuition when they are implemented. The implementation process necessarily involves a considerable amount of subjectivity, allowing room for continued calibration and improvement of these schemes.

Although convective parameterization for meso- and larger-scale models will be necessary for the foreseeable future, as computational power continues to increase, the greater challenge will be to develop parameterizations for higher-resolution models, particularly models with grid spacing on the order of 1–10 km. Over this range of scales, the processes and scales represented by traditional convective parameterizations become inconsistent with the features that are not well resolved by the model grid. Yet, the timing and evolution of explicitly simulated convective features degrades progressively as resolution is decreased over this range, implying that some parameterization of unresolved processes may be necessary and appropriate (Weisman et al. 1997). It is likely that convective parameterization on these scales will require something very different from traditional approaches.

Acknowledgments. Special thanks are given to Mike Baldwin of CIMMS/OU/NSSL, who provided valuable insight in the interpretation of model output, played a critical role in facilitating testing within the Eta Model, and provided helpful comments on this manuscript. I am very grateful to Peter Bechtold of ECMWF for his thoughtful suggestions on this manuscript and for his close collaboration over the last decade. I also appreciate Ted Mansell (CIMMS/OU/NSSL) and an anonymous reviewer who provided helpful comments on this manuscript. John Brown (NOAA/FSL) and Brian Mapes (NOAA/CDC/CIRES) volunteered many helpful suggestions for this work. I am grateful for their insight and encouragement. Many thanks are given to SPC forecasters and research scientists, especially Steve Weiss and Paul Janish, for their feedback on model performance and assistance with testing and evaluation of the convective scheme. This work was partially funded by NOAA-OU Cooperative Agreement NA17RJ1227 and COMET Cooperative Projects O99-15805 and S01-32796.
REFERENCES


