

# CAPE is predictable from large-scale environmental parameters

Funing LI<sup>1,1</sup> and Daniel Chavas<sup>2,2</sup>

<sup>1</sup>PURDUE UNIVERSITY

<sup>2</sup>Purdue University

November 30, 2022

## Abstract

A recent study by Agard and Emanuel (2017) proposed a simple equation for a quantity that scales with convective available potential energy (CAPE) that can be directly calculated from a limited number of environmental sounding parameters without lifting a hypothetical air parcel. This scaling CAPE was applied in a specific idealized framework, but the extent to which it can predict true CAPE in the real world has not been tested. This work uses reanalysis data over the U.S to demonstrate that this scaling CAPE does indeed scale very closely with CAPE, following a linear relationship with a scaling factor of 0.44. We then explain why they scale together via a step-by-step derivation of the theoretical assumptions linking scaling CAPE and real CAPE and their manifestation in the historical data. Overall, this work demonstrates that CAPE can be predicted from large-scale environmental parameters alone, which may be useful for a wide range of applications in weather and climate.

1           **Midlatitude continental CAPE is predictable from  
2           large-scale environmental parameters**

3           **Funing Li<sup>1</sup>, Daniel R. Chavas<sup>1</sup>**

4           <sup>1</sup>Purdue University, Department of Earth, Atmospheric, and Planetary Sciences, West Lafayette, IN

5           **Key Points:**

- 6           • CAPE can be predicted from environmental sounding parameters without lifting  
7           a hypothetical air parcel
- 8           • A step-by-step derivation demonstrates how CAPE scales with a recently-proposed  
9           CAPE-like quantity
- 10          • A simple predictive linear equation is presented based on 20 years of reanalysis  
11          data over the U.S.

---

Corresponding author: Funing Li, [li3226@purdue.edu](mailto:li3226@purdue.edu)

12 **Abstract**

13 A recent study by Agard and Emanuel (2017) proposed a simple equation for a quan-  
 14 tity that scales with convective available potential energy (CAPE) that can be directly  
 15 calculated from a limited number of environmental sounding parameters without lifting  
 16 a hypothetical air parcel. This scaling CAPE was applied in a specific idealized frame-  
 17 work, but the extent to which it can predict true CAPE in the real world has not been  
 18 tested. This work uses reanalysis data over the U.S to demonstrate that this scaling CAPE  
 19 does indeed scale very closely with CAPE, following a linear relationship with a scaling  
 20 factor of 0.44. We then explain why they scale together via a step-by-step derivation of  
 21 the theoretical assumptions linking scaling CAPE and real CAPE and their manifesta-  
 22 tion in the historical data. Overall, this work demonstrates that CAPE can be predicted  
 23 from large-scale environmental parameters alone, which may be useful for a wide range  
 24 of applications in weather and climate.

25 **Plain Language Summary**

26 Convective available potential energy (CAPE) is a key parameter commonly used  
 27 to measure the potential for thunderstorms. Its calculation requires lifting a hypothet-  
 28 ical air parcel through a column of atmosphere. This work combines theory and reanal-  
 29 ysis data to demonstrate that CAPE can be predicted using environmental data alone.  
 30 This can make it easier to quickly estimate CAPE in data and to understand the pro-  
 31 cesses that create CAPE in our atmosphere.

32 **1 Introduction**

Convective available potential energy (CAPE), a measure of conditional instability of the environment, is a key thermodynamic parameter in atmospheric research. It is proportional to the theoretical maximum vertical wind speed within the atmospheric column, and hence serves as an indicator of the potential intensity of deep convection if it is triggered (Holton, 1973). In practice, regular CAPE is estimated by the vertically-integrated buoyancy of a boundary-layer parcel ascending from the level of free convection (LFC) to the equilibrium level (EL) (Doswell III & Rasmussen, 1994), given by

$$CAPE = \int_{z_{LFC}}^{z_{EL}} g \frac{T_{vp} - T_{ve}}{T_{ve}} dz \quad (1)$$

where  $g$  is the acceleration due to gravity,  $z$  is height above ground level,  $T_{vp}$  is the virtual temperature of the rising air parcel and  $T_{ve}$  is that of the surrounding environment. Thus, calculating CAPE requires lifting a hypothetical parcel through a column of atmosphere defined by known vertical profiles of air temperature and moisture.

Recently, Agard and Emanuel (2017, hereafter AE17) proposed a simple equation for a quantity that scales with CAPE, here denoted  $CAPE_{AE17}$ , based on an idealized two-layer model for the atmospheric column. The AE17 model includes a dry adiabatic free troposphere overlying a cooler, moist, well-mixed boundary layer. Their proposed quantity scales with the difference between surface moist static energy ( $M_{ve}^{sfc}$ ) and free tropospheric dry static energy ( $\overline{D_{ve}^{FT}}$ ) multiplied by difference in the natural logarithm of virtual temperatures between boundary-layer top ( $T_{ve}^{BLT}$ ) and tropopause ( $T_{ve}^{trop}$ ):

$$CAPE_{AE17} = (M_{ve}^{sfc} - \overline{D_{ve}^{FT}}) \ln \frac{T_{ve}^{BLT}}{T_{ve}^{trop}} \quad (2)$$

The  $D_{ve}$  and  $M_{ve}$  are given by  $D_{ve} = c_p T_{ve} + gz$  and  $M_{ve} = c_p T_{ve} + gz + L_v r$ , respectively, where  $c_p$  and  $L_v$  are the specific heat of air and the latent heat of vaporization of water, and  $r$  is the water vapor mixing ratio. Note that Eq.2 is slightly different from the original formulation in AE17, as we use the free tropospheric mean dry static energy

( $\overline{D_{ve}^{FT}}$ ) rather than a constant  $D_{ve}$  of dry adiabatic free troposphere. In addition, we use virtual temperatures rather than temperatures for  $D_{ve}$  and  $M_{ve}$  to be consistent with definitions of CAPE in Eq.1 (detailed in Section 3). The CAPE<sub>AE17</sub> formula suggests that CAPE may to first order be determined by a limited number of environmental parameters within the boundary-layer and free troposphere. One significant benefit of this outcome is that this quantity may be calculated strictly from environmental sounding data without the need to lift a hypothetical air parcel.

Using this idealized framework, AE17 found that peak continental transient CAPE is expected to increase with global warming. Recent work used the AE17 framework to develop a simple physical model for a steady sounding for numerical simulations of severe convective storms (Chavas & Dawson, 2020). However, it remains unclear to what extent CAPE<sub>AE17</sub>, which represents CAPE in a highly idealized framework as we show below, directly predicts true CAPE in real soundings. Moreover, AE17 did not present a formal derivation of the relationship between CAPE<sub>AE17</sub> and CAPE.

To fill this gap, this work seeks to answer the following question: How closely does CAPE<sub>AE17</sub> scale with CAPE in real soundings, and why? To answer this question, we first directly compare CAPE<sub>AE17</sub> with CAPE over the U.S using reanalysis data and show that CAPE<sub>AE17</sub> does indeed scale closely with regular CAPE (Section 2). We then provide a step-by-step theoretical derivation and application to sounding data to explain why they scale together (Section 3). We end with a summary and discussion (Section 4).

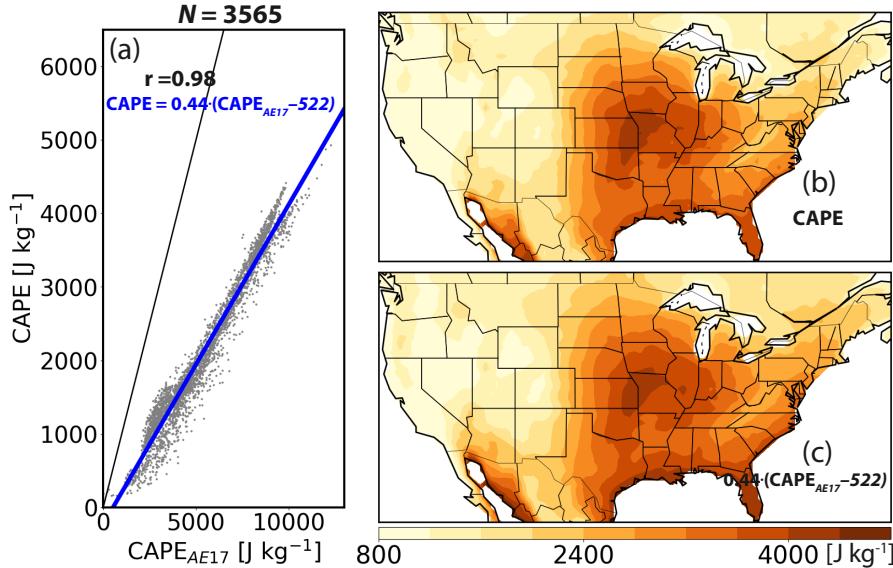
## 2 CAPE vs. CAPE<sub>AE17</sub>

We begin with an explicit comparison of CAPE and CAPE<sub>AE17</sub> in terms of 1) climatological extremes over the U.S, and 2) diurnal evolution during a significant tornado outbreak over the southern U.S.

### 2.1 Data

We use the 3-hourly surface and model-level (72 vertical levels) Modern-Era Retrospective analysis for Research and Applications version 2 (MERRA-2) reanalysis data for the period 2000–2019 in this work (Gelaro et al., 2017) (data accessed in March 2020 from [https://disc.gsfc.nasa.gov/datasets/M2I1NXASM\\_5.12.4/summary](https://disc.gsfc.nasa.gov/datasets/M2I1NXASM_5.12.4/summary) for the surface data and from [https://disc.gsfc.nasa.gov/datasets/M2I3NVASM\\_5.12.4/summary](https://disc.gsfc.nasa.gov/datasets/M2I3NVASM_5.12.4/summary) for the model-level data). The horizontal grid spacing of MERRA-2 is  $0.5^\circ \times 0.65^\circ$  in latitude and longitude. The model-level MERRA-2 data performs well in reproducing a reasonable magnitude and spatial distribution of CAPE over North America, though with a slight underestimation when compared against radiosonde data (Taszarek, Pilguy, et al., 2020). MERRA-2 also provides direct estimations of atmospheric properties at boundary-layer top and tropopause; this is especially useful for the calculation of CAPE<sub>AE17</sub>. Tegtmeier et al. (2020) found realistic representations of MERRA-2 derived boundary-layer top and tropopause temperatures as compared to radiosonde observations, with a small mean bias of less than 1 K; this may induce a bias percentage of less than  $\sim 1\%$  in CAPE<sub>AE17</sub>. Our domain of analysis focuses on the contiguous U.S, as it is a major hot spot for severe thunderstorm environments in the world (Brooks et al., 2003).

We generate a 20-year dataset of CAPE using Eq.1 and CAPE<sub>AE17</sub> using Eq.2 from the MERRA-2 reanalysis data over the U.S. Though CAPE estimation is sensitive to the origin of an air parcel, we select the near-surface parcel defined by 2-m temperature and moisture for simplicity, similar to past work (Riemann-Campe et al., 2009; Seeley & Romps, 2015; Li et al., 2020). Future work may seek to test alternate levels.



**Figure 1.** (a) Extreme values of CAPE (Eq. 1) vs. CAPE<sub>AE17</sub> (Eq. 2) over the contiguous U.S. Extreme values are defined as the 99th percentile of their respective full-period (2000–2019) time series from the MERRA-2 reanalysis data at each grid point (gray dots). Sample size is  $N=3565$ . Blue line denotes the linear least squares fit with linear correlation coefficient ( $r$ ). Black line denotes one-to-one fit. (b) Spatial distribution of extreme CAPE. (c) Predicted spatial distribution of extreme CAPE using the linear regression equation shown in (a).

88

## 2.2 Results

We first compare the representation of the climatological spatial distribution of extreme values of CAPE<sub>AE17</sub> against CAPE, as severe thunderstorms are typically associated with large values of CAPE (Brooks et al., 2003). We define extreme values by the 99th percentile of the full-period (2000–2019) time series of a given quantity at each grid point, in line with past work (Singh et al., 2017; Tippett et al., 2016; Li et al., 2020; Taszarek, Allen, et al., 2020). Results show that extreme CAPE<sub>AE17</sub> scales very closely with extreme CAPE (Figure 1a;  $r = 0.98$ ), with linear regression given by

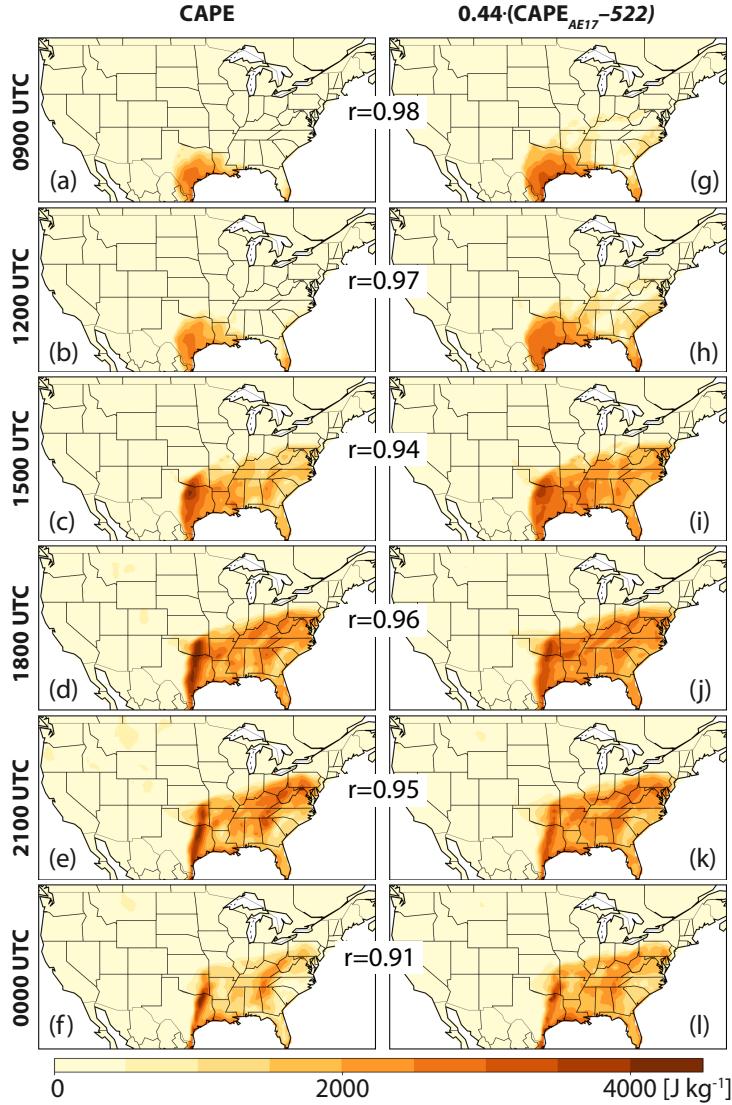
$$\text{CAPE} \approx 0.44 (\text{CAPE}_{\text{AE}17} - 522) \quad (3)$$

89  
90  
91

We then apply Eq. 3 to predicted extreme CAPE from extreme CAPE<sub>AE17</sub> (Figure 1c), which produces a spatial pattern that is quantitatively very similar to the observed extreme CAPE (Figure 1b).

92  
93  
94  
95  
96  
97  
98  
99  
100  
101  
102  
103

To further demonstrate how closely the two quantities scale, we present a case study comparison of their diurnal evolution during April 25, 2011, which is the first day of a three-day significant tornado outbreak event in the southeastern U.S (Knupp et al., 2014). The diurnal variation of CAPE indicates an initial generation of CAPE over southeastern Texas in the early morning (0900–1200 UTC; Figure 2a,b), followed by a strong enhancement at around 1500 UTC over eastern Texas (Figure 2c) and an eastward propagation of high CAPE in the afternoon (Figure 2d–f). The high CAPE values in the afternoon–evening over the southeastern U.S are associated with a swath of over 50 tornado reports extending from eastern Texas into the mid-Mississippi Valley (reference to the SPC Storm Reports: <https://www.spc.noaa.gov/exper/archive/event.php?date=20110425>). Compared to CAPE, CAPE<sub>AE17</sub> successfully reproduces the detailed spatial patterns and diurnal variations during the day (Figure 2g–l), with pattern correlation  $r \geq 0.90$



**Figure 2.** Spatial distributions of (a–f) CAPE vs. (g–l) predicted CAPE, using the equation in Fig 1(a), at (top–bottom) 0900, 1200, 1500, 1800, 2100, and 0000 UTC on April 25, 2011 from the MERRA-2 reanalysis data. The  $r$  denotes pattern correlation coefficient between CAPE and  $\text{CAPE}_{AE17}$  conditioned on gridpoints with  $\text{CAPE} \geq 100 \text{ J kg}^{-1}$ .

at each UTC time, though Eq. 3 slightly overestimates CAPE in the morning (Figure 2g,h vs. a,b) and slightly underestimates CAPE in the afternoon (Figure 2j,k vs. d,e).

Overall, our comparisons for both climatological extremes and the diurnal variation associated with a tornado outbreak case demonstrate a tight relationship between  $\text{CAPE}_{AE17}$  and CAPE distributions. This indicates that CAPE can be approximately predicted from  $\text{CAPE}_{AE17}$  via a simple linear equation. While this section focused on extreme values of CAPE to demonstrate its spatial variability, We show in Section 3 that such a close linear relation between CAPE and  $\text{CAPE}_{AE17}$  extends to the full distribution of CAPE.

113 **3 Theoretical foundation**

114 We next provide a theoretical derivation and explanation of the intermediate steps  
 115 and assumptions that link CAPE to CAPE<sub>AE17</sub>. We demonstrate each step both for a  
 116 single example radiosonde sounding (Figure 3) and statistically for all U.S gridpoints in  
 117 the full-period (2000–2019) MERRA-2 reanalysis database (Figure 4). Here the exam-  
 118 ple sounding was observed at 0000 UTC 07 June 2011 at the SGF (Springfield, MO) sta-  
 119 tion; we obtain it from the sounding database of the University of Wyoming (<http://weather.uwyo.edu/upperair/sounding.html>).  
 120

121 **3.1 A dry static energy view of CAPE**

122 As CAPE<sub>AE17</sub> is a function of an environmental static energy surplus between the  
 123 boundary layer and free troposphere, we first derive an alternative formula for estimat-  
 124 ing CAPE based on the parcel and environmental profiles of dry static energy rather than  
 125 temperature.

We begin from the environmental dry static energy relation ( $D_{ve}$ ),  $D_{ve} = c_p T_{ve} + gz$ . The environmental moist static energy ( $M_{ve}$ ) is given by  $M_{ve} = c_p T_{ve} + gz + L_v r$ . Heat capacities and latent heats are assumed to be constant. Counterparts for the parcel are given by  $D_{vp}$  and  $M_{vp}$ . Note that these static energies include the virtual temperature effect to be consistent with definitions of CAPE in Eq.1 as shown below. This virtual effect may add a small positive perturbation to regular static energies of approximately 0.9% and 0.8% of near-surface dry and moist static energy, respectively, given a surface temperature of 300 K and mixing ratio of 15 g kg<sup>-1</sup>, that will decrease with height. We may rewrite the  $D_{ve}$  equation for differential changes in height  $z$  as  $dz = -\frac{c_p}{g} dT_{ve} + \frac{1}{g} dD_{ve}$  and substitute into Eq.1. Doing so yields an alternative formulation of CAPE with limited approximations based on dry static energy profiles of the rising air parcel and the environment (derivation in Appendix A):

$$\text{CAPE} \approx \frac{\Gamma_d}{\Gamma} \mathcal{D} = -\frac{\Gamma_d}{\Gamma} \int_{T_{ve}^{LFC}}^{T_{ve}^{EL}} (D_{vp} - D_{ve}) d \ln T_{ve} \quad (4)$$

126 where  $\Gamma_d = g/c_p$  is the dry adiabatic lapse rate,  $\Gamma$  is the virtual temperature lapse rate  
 127 of the environment from LFC to EL, and  $T_{ve}^{LFC}$  and  $T_{ve}^{EL}$  are environmental virtual tem-  
 128 peratures at LFC and EL, respectively.

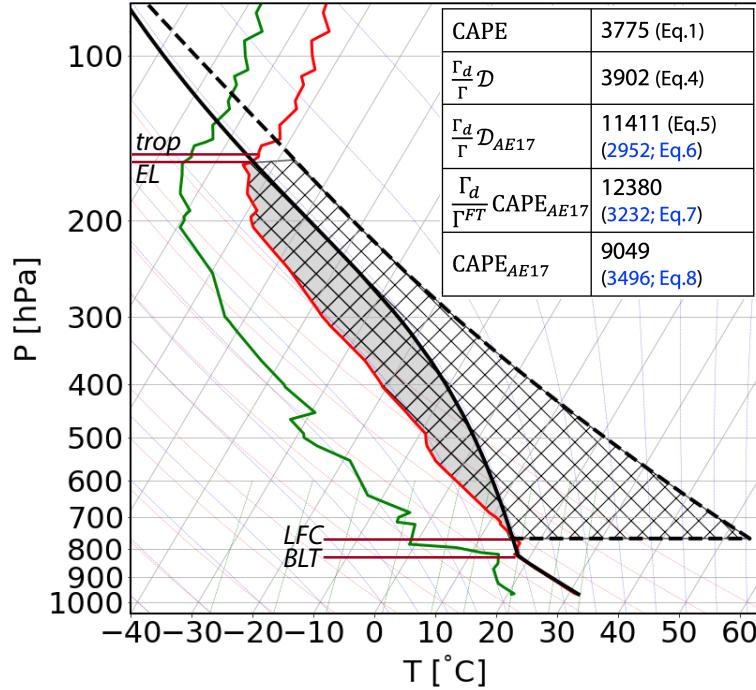
129 How well does  $\frac{\Gamma_d}{\Gamma} \mathcal{D}$  (Eq.4) compare to CAPE (Eq.1)? First, we compare  $\frac{\Gamma_d}{\Gamma} \mathcal{D}$  against  
 130 CAPE for our example sounding (Figure 3 inset). The two calculations yield similar val-  
 131 ues of CAPE (3775 vs. 3902 J kg<sup>-1</sup>). The slightly high bias in  $\frac{\Gamma_d}{\Gamma} \mathcal{D}$  relative to CAPE  
 132 (+3.4%) is due to the assumption of constant environmental virtual temperature lapse  
 133 rate ( $\Gamma$ ) from LFC to EL (Eq.A5). Second, we compare the two quantities for all grid-  
 134 points over the U.S in our MERRA-2 reanalysis dataset. The two quantities are indeed  
 135 nearly identical (Figure 4a;  $r > 0.99$ ) with linear regression given by CAPE = 0.98( $\frac{\Gamma_d}{\Gamma} \mathcal{D}$  +  
 136 18). The  $\frac{\Gamma_d}{\Gamma} \mathcal{D}$  formulation performs equally well in reproducing the detailed spatial dis-  
 137 tribution of extreme CAPE over the U.S (Figure S1b vs. S1a).

138 **3.2 Scaling of CAPE with CAPE<sub>AE17</sub>**

139 To obtain the CAPE<sub>AE17</sub> formula from Eq.4, we must assume that  $D_{vp} = M_{ve}^{sfc}$ ,  
 140 which yields

$$\frac{\Gamma_d}{\Gamma} \mathcal{D}_{AE17} = \frac{\Gamma_d}{\Gamma} (M_{ve}^{sfc} - \overline{D_{ve}}) \ln \frac{T_{ve}^{LFC}}{T_{ve}^{EL}} \quad (5)$$

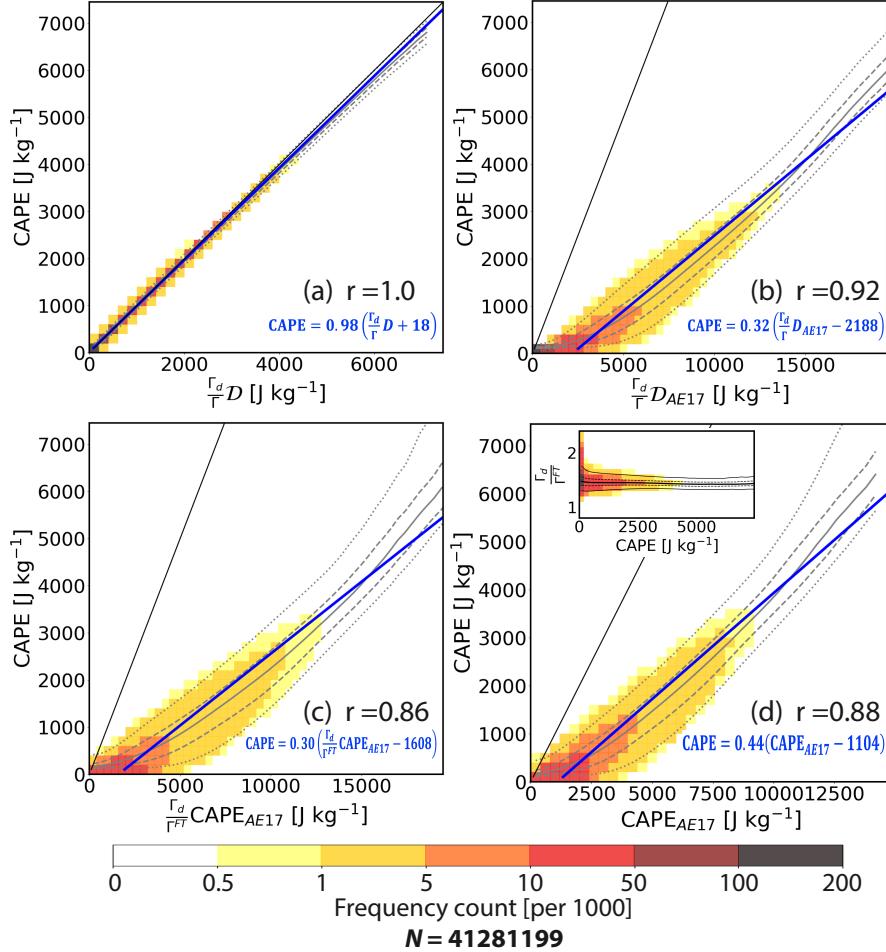
141 where  $\overline{D_{ve}} = \frac{\int_{T_{ve}^{LFC}}^{T_{ve}^{EL}} (D_{ve}) d \ln T_{ve}}{\int_{T_{ve}^{LFC}}^{T_{ve}^{EL}} d \ln T_{ve}}$  is the log-temperature-weighted average dry static en-  
 142 ergy of environment between LFC and EL. Though this assumption is not made explic-



**Figure 3.** The SGF (Springfield, MO) radiosonde observed virtual temperature (in red line) and dew-point temperature (in green line) profiles at 0000 UTC 07 June 2011 in a Skew-T diagram. Solid black line represents the virtual temperature profile of a surface air parcel ascending adiabatically. Dashed black line represents the virtual temperature profile of the idealized parcel ascending assuming that it converts all latent heat immediately to virtual dry static energy at LFC and perfectly conserves its virtual dry static energy thereafter. The *EL*, *LFC*, *trop*, and *BLT* are denoted by brown lines. Inset table lists values of CAPE (grey shading; Eq.1);  $\frac{\Gamma_d}{\Gamma} \mathcal{D}$  (Eq.4);  $\frac{\Gamma_d}{\Gamma} \mathcal{D}_{AE17}$  (Eq.5; approximates hatched region area = 10944 J kg<sup>-1</sup>);  $\frac{\Gamma_d}{\Gamma_{FT}} \text{CAPE}_{AE17}$  is the same as  $\frac{\Gamma_d}{\Gamma} \mathcal{D}_{AE17}$  but using virtual temperatures at *BLT* and *trop*, with CAPE<sub>AE17</sub> calculated from Eq.2. The inset table lists direct calculation of each quantity (black text) and prediction of true CAPE (blue text) using the relevant linear regression equation. The Python MetPy (May et al., 2008–2020) package is used to generate the parcel temperature profiles.

itly in AE17, it is an essential inference in order to derive CAPE<sub>AE17</sub> for a real atmosphere. Physically, this assumption implies that the lifted air parcel immediately convert all latent heat to sensible heat at LFC. Hence, the parcel will experience a sudden jump in dry static energy  $D_{vp}$  (to be equal to  $M_{vp}$ ) at the LFC, and above the LFC this quantity is conserved. Additionally, we must assume that the moist static energy of the surface parcel is conserved up to the LFC. Note that static energy is not perfectly conserved during adiabatic ascent because buoyancy acts as an enthalpy sink (Romps, 2015); because this static energy sink is not accounted for, the idealized parcel (Figure 3 black dashed) ends at a higher adiabat than the parcel following the regular moist adiabat (Figure 3 black solid). Taken together, the assumption results in  $D_{vp} = M_{vp} = M_{ve}^{sfc}$ .

We further use our example sounding (Figure 3) to help understand this assumption conceptually. As noted above, the above assumption implies that all latent heat within an air parcel is immediately converted to sensible heat at the LFC. Thus, the parcel is immediately warmed dramatically at the LFC and then subsequently rises dry adiabatically from the LFC to the EL. In this way, then,  $\frac{\Gamma_d}{\Gamma} \mathcal{D}_{AE17}$  is considered a “scaling” CAPE



**Figure 4.** Joint frequency fraction multiplied by 1000 (filled color) of (a) CAPE vs.  $\frac{r_d}{\Gamma} \mathcal{D}$ , (b) CAPE vs.  $\frac{r_d}{\Gamma} \mathcal{D}_{AE17}$ , (c) CAPE vs.  $\frac{r_d}{\Gamma^{FF}} \text{CAPE}_{AE17}$ , and (d) CAPE vs.  $\text{CAPE}_{AE17}$  (inset:  $\frac{r_d}{\Gamma^{FF}}$  vs. CAPE) for cases with  $\text{CAPE} \geq 100 \text{ J kg}^{-1}$  over all U.S gridpoints during 2000–2019 from the MERRA-2 reanalysis dataset (sample size  $N=41281199$ ). Black line denotes one-to-one line. Gray lines denote median (solid), interquartile range (dashed), and 5–95% range (dotted) of CAPE. Blue line denotes the linear regression with the correlation coefficient  $r$ .

because it represents a theoretical upper bound on how quickly a parcel can be warmed along its path (and hence on its integrated buoyancy). In the real atmosphere, latent heat is released gradually along the parcel path in accordance with the Clausius-Clapeyron relation that defines the moist adiabatic lapse rate. In a Skew-T diagram (Figure 3), this difference shows up as an expanded, angular region of positive buoyancy maximized above the LFC in  $\frac{r_d}{\Gamma} \mathcal{D}_{AE17}$ , which is larger than the true CAPE area. Thus,  $\frac{r_d}{\Gamma} \mathcal{D}_{AE17}$  is substantially larger than CAPE:  $\frac{r_d}{\Gamma} \mathcal{D}_{AE17} = 11411 \text{ J kg}^{-1}$  vs.  $\text{CAPE} = 3775 \text{ J kg}^{-1}$  (Figure 3 inset).  $\frac{r_d}{\Gamma} \mathcal{D}_{AE17}$  is slightly larger (+4.2 %) than the true value given by the hatched area ( $10944 \text{ J kg}^{-1}$ ), due to the assumption of constant  $\Gamma$  as noted earlier.

Though different in magnitude,  $\frac{r_d}{\Gamma} \mathcal{D}_{AE17}$  is still highly correlated with CAPE ( $r=0.92$ ) in the full reanalysis dataset over the U.S (Figure 4b), with linear regression given by

$$\text{CAPE} \approx 0.32 \left( \frac{r_d}{\Gamma} \mathcal{D}_{AE17} - 2188 \right) \quad (6)$$

For the example sounding, Eq.6 predicts a CAPE value ( $2952 \text{ J kg}^{-1}$ ) that is reasonably close to the true CAPE ( $3775 \text{ J kg}^{-1}$ ) (Figure 3 inset). Eq.6 also performs very well in reproducing the spatial distribution of extreme CAPE over the U.S (Figure S1c vs. S1a). Physically, the factor 0.32 is a manifestation of the large difference in the temperature profile of the parcel as it rises for the idealized parcel as compared to the normal parcel profile following the standard moist adiabatic lapse rate. The latter is a manifestation of the Clausius-Clapeyron relation governing the rate at which condensation occurs as the parcel cools adiabatically, and hence the rate at which latent heat is gradually converted to sensible heat (dry static energy) as the parcel rises through the troposphere. This contrasts with the idealized parcel where  $D_{vp}$  is set equal to  $M_{vp}$  immediately at the LFC, which equates to an instantaneous conversion of all latent heat to dry static energy. Geometrically, the factor 0.32 visually represents the ratio of the true CAPE area (grey shading in Figure 3) to the idealized parcel CAPE area (hatched in Figure 3). Indeed, for the case shown in Figure 3, that ratio is 0.33.

Finally, to produce a prediction with the original AE17 formulation ( $\text{CAPE}_{AE17}$ ), we must additionally assume that the temperatures of the EL and LFC may be replaced with that of the tropopause (*trop*) and boundary-layer top (*BLT*), respectively. This replaces  $\frac{\Gamma_d}{\Gamma} \mathcal{D}_{AE17}$  of Eq.5 with  $\frac{\Gamma_d}{\Gamma^{FT}} \text{CAPE}_{AE17}$ , where  $\Gamma^{FT}$  is defined by the lapse rate of virtual temperature of the free troposphere between the *BLT* and *trop*. These approximations are more quantitatively reasonable for higher-CAPE cases supportive of deep convection, as in the example sounding (Figure 3). This final approximation ( $\frac{\Gamma_d}{\Gamma^{FT}} \text{CAPE}_{AE17}$ ) is estimated solely by environmental parameters without lifting a hypothetical air parcel. We use the reanalysis dataset to examine its relationship to CAPE (Figure 4c), which indicates a close correlation ( $r=0.86$ ) with a linear regression given by:

$$\text{CAPE} \approx 0.30 \left( \frac{\Gamma_d}{\Gamma^{FT}} \text{CAPE}_{AE17} - 1608 \right) \quad (7)$$

Hence the scaling factor is similar to that for  $\frac{\Gamma_d}{\Gamma} \mathcal{D}_{AE17}$  above. For our example sounding, Eq.7 predicts a CAPE value ( $3232 \text{ J kg}^{-1}$ ) again reasonably close to the true CAPE ( $3775 \text{ J kg}^{-1}$ ) (Figure 3 insert). Eq.7 also quantitatively reproduces the spatial pattern of extreme CAPE over the U.S (Figure S1d vs. S1a).

Ultimately, then, Eq.7 offers a scaling of CAPE that depends only on a limited number of boundary-layer and free tropospheric variables. It differs from  $\text{CAPE}_{AE17}$  itself in the inclusion of the coefficient  $\frac{\Gamma_d}{\Gamma^{FT}}$ . This factor does not appear in the idealized model of AE17 because their model assumes a dry adiabatic free troposphere (i.e.,  $\Gamma^{FT} = \Gamma_d$ ), which yields  $\frac{\Gamma_d}{\Gamma^{FT}} = 1$ .

Given that CAPE was found to be predictable from  $\text{CAPE}_{AE17}$  alone in Section 2 (Eq.3), this result implies that the free tropospheric lapse rate ( $\Gamma^{FT}$ ) of the modern atmosphere does not vary too strongly and thus the factor  $\frac{\Gamma_d}{\Gamma^{FT}}$  remains relatively constant. We use our reanalysis dataset to calculate the statistics of  $\frac{\Gamma_d}{\Gamma^{FT}}$  as a function of CAPE (Figure 4d inset). The result is indeed a mean ( $\pm$  one standard deviation) value of  $1.47 \pm 0.06$ , with variance decreasing as CAPE increases. The resulting mean free tropospheric lapse rate ( $\Gamma^{FT}$ ) is roughly  $6.7 \text{ K km}^{-1}$ , which is close to that of the U.S Standard Atmosphere (COESA, 1976). These results indicate a relatively constant free tropospheric thermal structure at high values of CAPE, a result that is worthy of deeper investigation. As a result, we are able to directly scale CAPE with  $\text{CAPE}_{AE17}$  by assuming that  $\frac{\Gamma_d}{\Gamma^{FT}}$  is constant. We note that this behavior may differ in an alternate climate state. As a final test, we compare  $\text{CAPE}_{AE17}$  with CAPE for cases with  $\text{CAPE} \geq 100 \text{ J kg}^{-1}$  for the entire MERRA-2 database over the U.S and find a strong linear correlation between them as well ( $r = 0.88$ ; Figure 4d), with a linear regression of

$$\text{CAPE} \approx 0.44(\text{CAPE}_{AE17} - 1104). \quad (8)$$

This outcome is quite similar to the linear regression model we get from extreme cases alone in Eq.3. This is also close to the results of simply substituting  $\frac{\Gamma_d}{\Gamma^{FT}} = 1.47 \pm 0.06$  into

218 Eq.7, which yields a scaling factor of  $0.44 \pm 0.02$  and an offset of  $-1095 \pm 50$ . Using Eq.8  
 219 also successfully predicts the approximate CAPE for the example sounding (3496 vs. 3775  
 220 J kg<sup>-1</sup>; Figure 3 inset).

## 221 4 Conclusions

222 CAPE is a key thermodynamic parameter commonly calculated to evaluate the potential  
 223 for deep convection within a given environment. AE17 proposed a simple formula  
 224 for a quantity ( $\text{CAPE}_{AE17}$ ) that scales with CAPE that depends only on a limited number  
 225 of environmental variables and does not require lifting a hypothetical parcel.  $\text{CAPE}_{AE17}$   
 226 represents an expression of CAPE for a highly idealized column in which the EL and LFC  
 227 are exactly the tropopause and boundary-layer tops, the free tropospheric lapse rate is  
 228 dry adiabatic, and the rising parcel instantly converts all latent heat to sensible heat at  
 229 LFC; this requires idealizations of both the environmental and parcel thermal profiles.

230 This work used a 20-year reanalysis dataset over the U.S to examine the extent to  
 231 which this CAPE-like quantity can be used to predict true CAPE for real soundings, an-  
 232 alyzing both the spatial distribution of climatological extremes and the diurnal varia-  
 233 tion associated with a historical tornado outbreak case study. Results show a close scal-  
 234 ing relationship between  $\text{CAPE}_{AE17}$  and CAPE, yielding a simple linear equation for  
 235 predicting CAPE from environmental data. To understand the physics underlying this  
 236 relationship, we provided a step-by-step derivation linking the two quantities, which may  
 237 be summarized as:

$$\text{CAPE} \stackrel{\text{a1}}{\approx} \frac{\Gamma_d}{\Gamma} \mathcal{D} \stackrel{\text{a2}}{\approx} \frac{\Gamma_d}{\Gamma} \mathcal{D}_{AE17} \stackrel{\text{a3}}{\approx} \frac{\Gamma_d}{\Gamma_{FT}} \text{CAPE}_{AE17} \stackrel{\text{a4}}{\approx} \text{CAPE}_{AE17} \quad (9)$$

238 where (a1–a4) represent the assumptions: (a1) constant environmental virtual temper-  
 239 ature lapse rate from LFC to EL; (a2) the rising parcel immediately converts all latent  
 240 heat to sensible heat at the LFC; (a3) temperatures at the EL and LFC are equal to  
 241 the tropopause and boundary-layer top, respectively; (a4) free tropospheric lapse rate  
 242 of the present atmosphere does not vary strongly in space or time in environments with  
 243 non-negligible CAPE.

244 Though our assessment focused on the U.S. continent,  $\text{CAPE}_{AE17}$  also performs  
 245 well in predicting CAPE over the Gulf of Mexico and nearby tropical ocean (Figure S2a–  
 246 l). Additionally, we examined an existing analytical prediction for mean CAPE in the  
 247 tropics ( $\text{CAPE}_{R16}$ ; Eq. 17 in Romps (2016)), which also depends only on environmen-  
 248 tal parameters. We find that  $\text{CAPE}_{R16}$  does not reproduce the detailed spatial distri-  
 249 bution and temporal evolution of high CAPE values for the case study over the U.S. con-  
 250 tinent, though the performance is slightly improved over ocean (Figure S2m–r). The deriva-  
 251 tion of  $\text{CAPE}_{R16}$  assumes a zero-buoyancy plume under radiative–convective equilibrium.  
 252 This assumption applies very well for describing the tropical mean state, which is gov-  
 253 erned principally by the upward transfer of heat and moisture by persistent deep con-  
 254 vention (and its associated entrainment) that allows for an accurate prediction of the free  
 255 tropospheric thermodynamic structure from surface air properties alone. However, con-  
 256 tinental convective environments involve the time-dependent buildup and storage of CAPE  
 257 due to the presence of significant convective inhibition generated by the superposition  
 258 of distinct airmasses as well as variability in land surface-air interactions (Carlson et al.,  
 259 1983; Singh & O’Gorman, 2013; Romps, 2014, 2016; Agard & Emanuel, 2017). Hence,  
 260  $\text{CAPE}_{R16}$  would not be expected to perform well for such environments.

261 This work has significant practical benefits for the simple estimation of CAPE and  
 262 for understanding the processes that create CAPE in our atmosphere. The principal end  
 263 result of this work is a simple linear equation based on the 20-year reanalysis dataset over  
 264 the U.S (Eq.8) to predict CAPE from  $\text{CAPE}_{AE17}$ , which may be calculated strictly from  
 265 environmental data without the need to lift a hypothetical parcel. Meanwhile, the close  
 266 relationship between CAPE and  $\text{CAPE}_{AE17}$  indicates that there is significant potential

267 to use CAPE<sub>AE17</sub> to understand how CAPE is generated within the climate system. This  
 268 includes quantifying the roles of variability in surface moist static energy, free tropospheric  
 269 dry static energy, and temperatures at the top of boundary layer and tropopause and  
 270 the processes that govern each. This is a promising avenue of future research.

## 271 Appendix A Derivation of Eq.4

272 The equation for differential changes in environmental dry static energy may be  
 273 written as  $dz = -\frac{c_p}{g}dT_{ve} + \frac{1}{g}dD_{ve}$  and substituting into Eq.1 yields

$$274 \text{CAPE} = \int_{z_{LFC}}^{z_{EL}} g \frac{T_{vp} - T_{ve}}{T_{ve}} \left( -\frac{c_p}{g} dT_{ve} + \frac{1}{g} dD_{ve} \right) = \mathcal{D} + \mathcal{T} \quad (\text{A1})$$

This formulation decomposes CAPE into two terms. The first is given by

$$275 \mathcal{D} = - \int_{z_{LFC}}^{z_{EL}} \left( \frac{T_{vp} - T_{ve}}{T_{ve}} \right) d(c_p T_{ve}) = - \int_{z_{LFC}}^{z_{EL}} (D_{vp} - D_{ve}) d \ln T_{ve} \quad (\text{A2})$$

and represents differences in dry static energy integrated over changes in temperature.  
 The second is given by

$$276 \mathcal{T} = \int_{z_{LFC}}^{z_{EL}} \left( \frac{T_{vp} - T_{ve}}{T_{ve}} \right) dD_{ve} \quad (\text{A3})$$

274 and represents integrated differences in temperature over changes in dry static energy.  
 275 To further simplify Eq.A1, we can relate  $\mathcal{T}$  and  $\mathcal{D}$  by calculating their ratio. Using the  
 276 definition of buoyancy,  $b = \frac{T_{vp} - T_{ve}}{T_{ve}}$ , we may write this ratio as

$$\begin{aligned} 277 \frac{\mathcal{T}}{\mathcal{D}} &= \frac{\int_{z_{LFC}}^{z_{EL}} (b) dD_{ve}}{-\int_{z_{LFC}}^{z_{EL}} (b) d(c_p T_{ve})} \\ &= -(1 + \frac{g}{c_p} \frac{\int_{z_{LFC}}^{z_{EL}} (b) dz}{\int_{z_{LFC}}^{z_{EL}} (b) dT_{ve}}) \\ &= -(1 + \frac{g}{c_p} \frac{\bar{b}_1 \int_{z_{LFC}}^{z_{EL}} dz}{\bar{b}_2 \int_{z_{LFC}}^{z_{EL}} dT_{ve}}) \\ &= \frac{\bar{b}_1}{\bar{b}_2} \frac{\Gamma_d}{\Gamma} - 1 \end{aligned} \quad (\text{A4})$$

277 where  $\bar{b}_1 = \frac{\int_{z_{LFC}}^{z_{EL}} (b) dz}{\int_{z_{LFC}}^{z_{EL}} dz}$  and  $\bar{b}_2 = \frac{\int_{z_{LFC}}^{z_{EL}} (b) dT_{ve}}{\int_{z_{LFC}}^{z_{EL}} dT_{ve}}$  represent the mean value of  $b$  between  
 278 the LFC and EL weighted by height ( $z$ ) and environmental virtual temperature ( $T_{ve}$ ),  
 279 respectively.  $\Gamma_d = g/c_p$  is the dry adiabatic lapse rate and  $\Gamma = -\frac{\int_{z_{LFC}}^{z_{EL}} dT_{ve}}{\int_{z_{LFC}}^{z_{EL}} dz} = -\frac{T_{ve}^{EL} - T_{ve}^{LFC}}{z_{EL} - z_{LFC}}$   
 280 represents the average environmental virtual temperature lapse rate from LFC to EL.

If we take  $\Gamma$  to be constant between the LFC and EL, then  $\bar{b}_1 = \bar{b}_2$ , which yields

$$\frac{\mathcal{T}}{\mathcal{D}} = \frac{\Gamma_d}{\Gamma} - 1 \quad (\text{A5})$$

Substituting this result into Eq.A1 yields

$$281 \text{CAPE} \approx \frac{\Gamma_d}{\Gamma} \mathcal{D} = -\frac{\Gamma_d}{\Gamma} \int_{z_{LFC}}^{z_{EL}} (D_{vp} - D_{ve}) d \ln T_{ve} \quad (\text{A6})$$

This equation is shown to closely match the true CAPE in the main manuscript.

## 282 Acknowledgments

283 The authors acknowledge high-performance computing support from Cheyenne (doi:10.5065/  
 284 D6RX99HX) provided by NCAR's Computational and Information Systems Laboratory,

sponsored by the National Science Foundation, for the data analysis performed in this work. The authors gratefully acknowledge the open-source Python community, and particularly the authors and contributors to the Matplotlib (Hunter, 2007), NumPy (Oliphant, 2006), and MetPy (May et al., 2008–2020) packages that were used to generate many of the figures. Li and Chavas were supported by NASA FINESST grant 12000365 and NSF grant 1648681. The surface and model-level MERRA-2 reanalysis data during 2000–2019 were downloaded from [https://disc.gsfc.nasa.gov/datasets/M2I1NXASM\\_5.12.4/summary](https://disc.gsfc.nasa.gov/datasets/M2I1NXASM_5.12.4/summary) (DOI: 10.5067/3Z173KIE2TPD) and [https://disc.gsfc.nasa.gov/datasets/M2I3NVASM\\_5.12.4/summary](https://disc.gsfc.nasa.gov/datasets/M2I3NVASM_5.12.4/summary) (DOI: 10.5067/WWQSXQ8IVFW8), respectively. The example sounding was obtained from the sounding database of the University of Wyoming at <http://weather.uwyo.edu/upperair/sounding.html>.

## 296 References

- 297 Agard, V., & Emanuel, K. (2017). Clausius–Clapeyron scaling of peak CAPE in con-  
298 tinental convective storm environments. *Journal of the Atmospheric Sciences*,  
299 74(9), 3043–3054. doi: <https://doi.org/10.1175/JAS-D-16-0352.1>
- 300 Brooks, H. E., Lee, J. W., & Craven, J. P. (2003). The spatial distribution of se-  
301 vere thunderstorm and tornado environments from global reanalysis data. *At-  
302 mospheric Research*, 67, 79–94. doi: [https://doi.org/10.1016/S0169-8095\(03\)  
00045-0](https://doi.org/10.1016/S0169-8095(03)<br/>00045-0)
- 304 Carlson, T. N., Benjamin, S. G., Forbes, G. S., & Li, Y. F. (1983). Elevated  
305 mixed layers in the regional severe storm environment: Conceptual model  
306 and case studies. *Monthly Weather Review*, 111(7), 1453–1474. doi:  
307 [https://doi.org/10.1175/1520-0493\(1983\)111<1453:EMLITR>2.0.CO;2](https://doi.org/10.1175/1520-0493(1983)111<1453:EMLITR>2.0.CO;2)
- 308 Chavas, D. R., & Dawson, I., Daniel T. (2020). An idealized physical model for the  
309 severe convective storm environmental sounding. *Journal of the Atmospheric  
310 Sciences*, 78(2), 653–670. doi: <https://doi.org/10.1175/JAS-D-20-0120.1>
- 311 COESA. (1976). U.S. Standard Atmosphere. *National Oceanic and Atmospheric Ad-  
312 ministration*, 76(1562), 227pp.
- 313 Doswell III, C. A., & Rasmussen, E. N. (1994). The effect of neglecting the virtual  
314 temperature correction on CAPE calculations. *Weather and forecasting*, 9(4),  
315 625–629. doi: [https://doi.org/10.1175/1520-0434\(1994\)009<0625:TEONTV>2.0.  
.CO;2](https://doi.org/10.1175/1520-0434(1994)009<0625:TEONTV>2.0.<br/>.CO;2)
- 317 Gelaro, R., McCarty, W., Suárez, M. J., Todling, R., Molod, A., Takacs, L., ...  
318 others (2017). The modern-era retrospective analysis for research and appli-  
319 cations, version 2 (MERRA-2). *Journal of Climate*, 30(14), 5419–5454. doi:  
320 <https://doi.org/10.1175/JCLI-D-16-0758.1>
- 321 Holton, J. R. (1973). An introduction to dynamic meteorology. *American Journal of  
322 Physics*, 41(5), 752–754.
- 323 Hunter, J. D. (2007). Matplotlib: A 2d graphics environment. *Computing in Science  
324 & Engineering*, 9(3), 90–95. doi: 10.1109/MCSE.2007.55
- 325 Knupp, K. R., Murphy, T. A., Coleman, T. A., Wade, R. A., Mullins, S. A., Schultz,  
326 C. J., ... others (2014). Meteorological overview of the devastating 27 April  
327 2011 tornado outbreak. *Bulletin of the American Meteorological Society*, 95(7),  
328 1041–1062. doi: <https://doi.org/10.1175/BAMS-D-11-00229.1>
- 329 Li, F., Chavas, D. R., Reed, K. A., & Dawson II, D. T. (2020). Climatology of  
330 severe local storm environments and synoptic-scale features over North Amer-  
331 ica in ERA5 reanalysis and CAM6 simulation. *Journal of Climate*. doi:  
332 <https://doi.org/10.1175/JCLI-D-19-0986.1>
- 333 May, R. M., Arms, S. C., Marsh, P., Bruning, E., Leeman, J. R., Goebbert, K., ...  
334 Bruick, Z. S. (2008–2020). *Metpy: A Python package for meteorological data*.  
335 Boulder, Colorado: Unidata. doi: 10.5065/D6WW7G29
- 336 Oliphant, T. E. (2006). *A guide to NumPy* (Vol. 1). Trelgol Publishing USA.  
337 Riemann-Campe, K., Fraedrich, K., & Lunkeit, F. (2009). Global climatology

- 338 of convective available potential energy (CAPE) and convective inhibition  
339 (CIN) in ERA-40 reanalysis. *Atmospheric Research*, 93(1-3), 534–545. doi:  
340 <https://doi.org/10.1016/j.atmosres.2008.09.037>
- 341 Romps, D. M. (2014). An analytical model for tropical relative humidity. *Journal of*  
342 *Climate*, 27(19), 7432–7449.
- 343 Romps, D. M. (2015). MSE Minus CAPE is the True Conserved Variable for an  
344 Adiabatically Lifted Parcel. *Journal of the Atmospheric Sciences*, 72(9), 3639–  
345 3646. doi: <https://doi.org/10.1175/JAS-D-15-0054.1>
- 346 Romps, D. M. (2016). Clausius–clapeyron scaling of CAPE from analytical solutions  
347 to RCE. *Journal of the Atmospheric Sciences*, 73(9), 3719–3737.
- 348 Seeley, J. T., & Romps, D. M. (2015). The effect of global warming on severe thun-  
349 derstorms in the United States. *Journal of Climate*, 28(6), 2443–2458. doi:  
350 <https://doi.org/10.1175/JCLI-D-14-00382.1>
- 351 Singh, M. S., Kuang, Z., Maloney, E. D., Hannah, W. M., & Wolding, B. O. (2017).  
352 Increasing potential for intense tropical and subtropical thunderstorms under  
353 global warming. *Proceedings of the National Academy of Sciences*, 114(44),  
354 11657–11662. doi: <https://doi.org/10.1073/pnas.1707603114>
- 355 Singh, M. S., & O’Gorman, P. A. (2013). Influence of entrainment on the ther-  
356 mal stratification in simulations of radiative-convective equilibrium. *Geophys-  
357 ical Research Letters*, 40(16), 4398–4403.
- 358 Taszarek, M., Allen, J. T., Púćik, T., Hoogewind, K. A., & Brooks, H. E. (2020).  
359 Severe convective storms across Europe and the United States. Part 2: ERA5  
360 environments associated with lightning, large hail, severe wind and tornadoes.  
361 *Journal of Climate*, 1–53. doi: <https://doi.org/10.1175/JCLI-D-20-0346.1>
- 362 Taszarek, M., Pilguj, N., Allen, J. T., Gensini, V., Brooks, H. E., & Szuster, P.  
363 (2020). Comparison of convective parameters derived from era5 and merra2  
364 with rawinsonde data over europe and north america. *Journal of Climate*, 1 -  
365 55. doi: <https://doi.org/10.1175/JCLI-D-20-0484.1>
- 366 Tegtmeier, S., Anstey, J., Davis, S., Dragani, R., Harada, Y., Ivanciu, I., ... others  
367 (2020). Temperature and tropopause characteristics from reanalyses data in  
368 the tropical tropopause layer. *Atmospheric Chemistry and Physics*, 20(2),  
369 753–770.
- 370 Tippett, M. K., Lepore, C., & Cohen, J. E. (2016). More tornadoes in the most ex-  
371 treme US tornado outbreaks. *Science*, 354(6318), 1419–1423. doi: <https://doi.org/10.1126/science.aah7393>

# Supporting Information for "CAPE is predictable from large-scale environmental parameters"

Funing Li<sup>1</sup>, Daniel R. Chavas<sup>1</sup>

<sup>1</sup>Purdue University, Department of Earth, Atmospheric, and Planetary Sciences, West Lafayette, IN

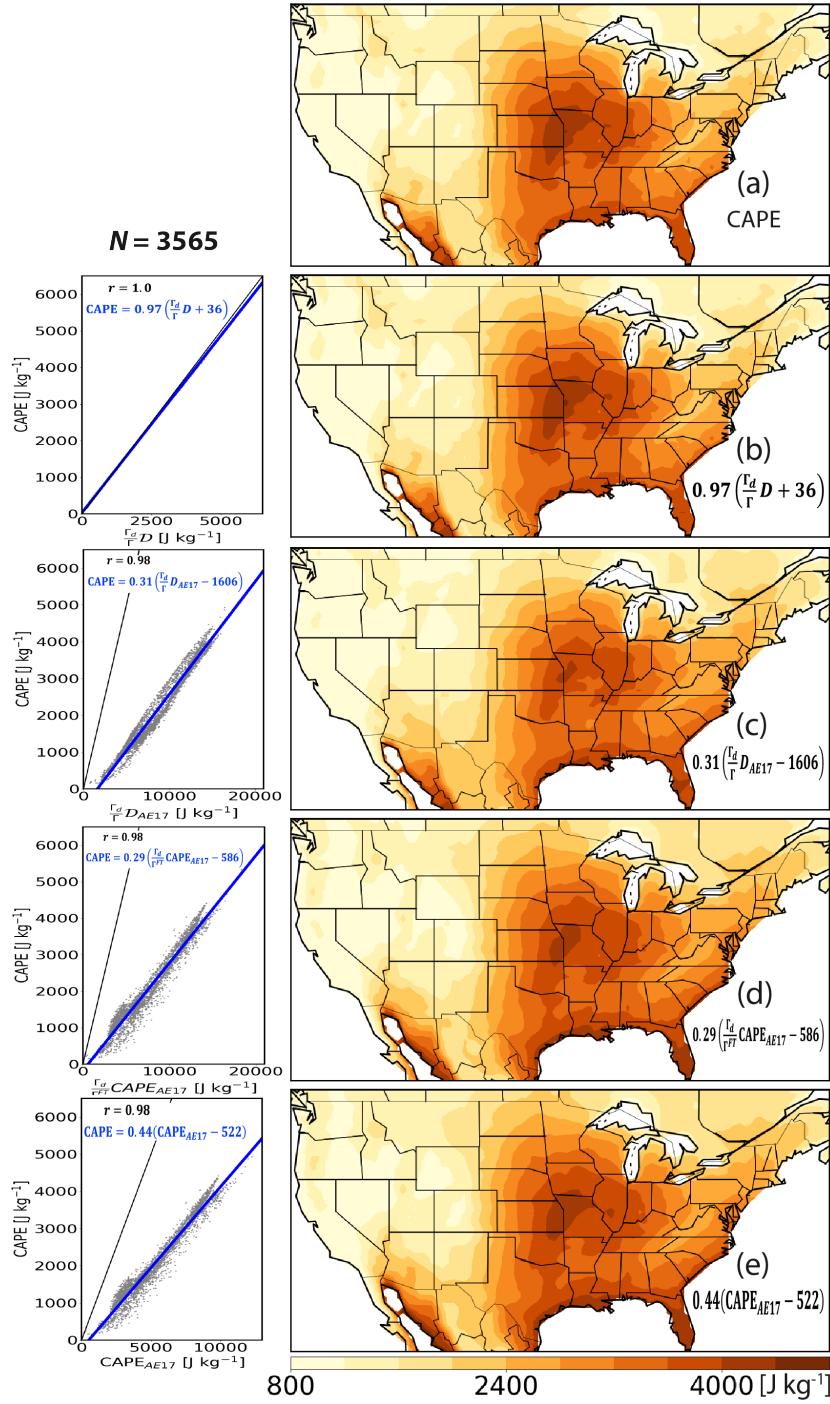
## Contents of this file

1. Figures S1

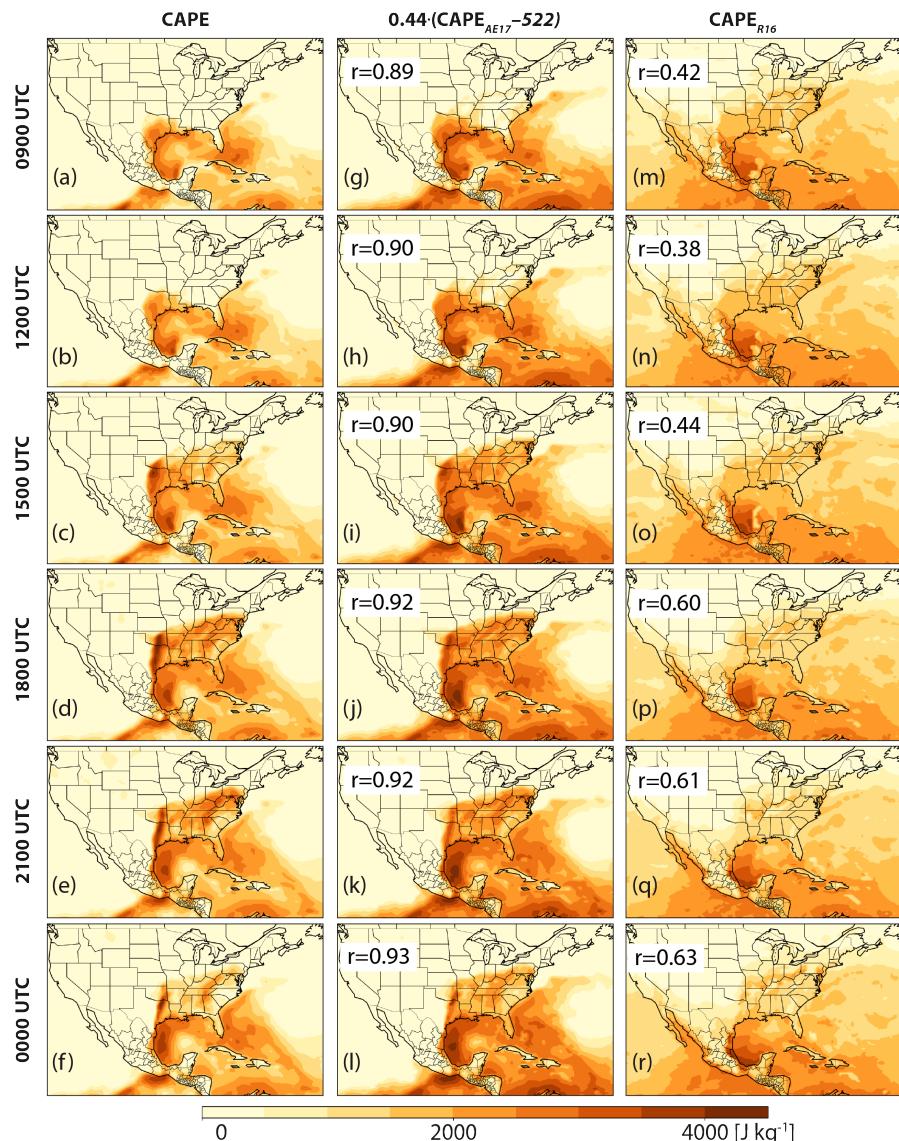
2. Figures S2

**Figure S1.**

**Figure S2.**



**Figure S1.** Extreme values of (a) CAPE compared to estimated CAPE from (b)  $\frac{r_d}{\Gamma} \mathcal{D}$ , (c)  $\frac{r_d}{\Gamma} \mathcal{D}_{AE17}$ , (d)  $\frac{r_d}{\Gamma F T} CAPE_{AE17}$ , and (e) CAPE<sub>AE17</sub> using their respective linear regression equation shown in the left (blue text). The (a) and (e) are the same as Figure 1.



**Figure S2.** As in Figure 2, but includes both the U.S. continent and tropical ocean area for spatial distributions of (a–f) CAPE vs. (g–l) predicted CAPE from  $\text{CAPE}_{AE17}$ , and (m–r)  $\text{CAPE}_{R16}$ , at (top–bottom) 0900, 1200, 1500, 1800, 2100, and 0000 UTC on April 25, 2011 from the MERRA-2 reanalysis data. The  $r$  indicates its pattern correlation coefficient with CAPE conditioned on gridpoints with  $\text{CAPE} \geq 100 \text{ J kg}^{-1}$ .