Estimating Convection's Moisture Sensitivity: An Observation–Model Synthesis Using AMIE-DYNAMO Field Data

BRIAN MAPES AND ARUNCHANDRA S. CHANDRA

University of Miami, Miami, Florida

ZHIMING KUANG

Harvard University, Cambridge, Massachusetts

SIWON SONG

University of New South Wales, Sydney, New South Wales, Australia

PAQUITA ZUIDEMA

University of Miami, Miami, Florida

(Manuscript received 3 May 2018, in final form 23 January 2019)

ABSTRACT

We seek to use ARM MJO Investigation Experiment (AMIE)-DYNAMO field campaign observations to significantly constrain height-resolved estimates of the parameterization-relevant, causal sensitivity of convective heating Q to water vapor q . In field data, Q profiles are detected via Doppler radar wind divergence D while balloon soundings give q . Univariate regressions of D on q summarize the information from a 10-layer time–pressure series from Gan Island (0° , 90°E) as a 10×10 matrix. Despite the right shape and units, this is not the desired causal quantity because observations reflect confounding effects of additional q -correlated casual mechanisms. We seek to use this matrix to adjudicate among candidate estimates of the desired causal quantity: Kuang's matrix $\mathbf{M}$ of the linear responses of a cyclic convection-permitting model (CCPM) at equilibrium. Transforming $\mathbf{M}$ to more observation-comparable forms by accounting for observed autocorrelations, the comparisons are still poor, because (we hypothesize) larger-scale vertical velocity, forbidden by CCPM methodology, is another confounding cause that must be permitted to covary with q . By embedding $\mathbf{M}$ and modified candidates in an idealized GCM, and treating its outputs as virtual field campaign data, we find that observations favor a factor of 2 (rather than 0 or 1) to small-domain $\mathbf{M}$'s free-tropospheric causal q sensitivity of about 25% rain-rate increment over 3 subsequent hours per $+1 \text{ g kg}^{-1}$ q impulse in a 100-hPa layer. Doubling this sensitivity lies partway toward Kuang's $\mathbf{M}$ for a long domain that organizes convection into squall lines, a weak but sign-consistent hint of a detectable parameterization-relevant (causal) role for convective organization in nature. Caveats and implications for field campaign proposers are discussed.

1. Introduction

Precipitating deep convection is a central process in the dynamics of the tropical troposphere. It is tautological that moist convection depends on moisture [see Sherwood et al. (2010) for a lucid broad review], and the vertical dependence of rainfall's sensitivity is arguably the most important of the “importance” profiles

for water vapor surveyed in Mapes et al. (2017). Here we wish to further refine and quantify that sensitivity by estimating the *profile* of convective heating responses to vertically resolved humidity variations. Another driving impetus for this work is to do so in a way that makes some meaningful, statistically significant use of special field campaign data, in the age of computer modeling. Specifically, we attempt to utilize observations from the several-month 2011/12 ARM MJO Investigation Experiment (AMIE)-Dynamics of the Madden–Julian Oscillation (DYNAMO)

Corresponding author: Brian Mapes, mapes@miami.edu

DOI: 10.1175/JAS-D-18-0127.1

© 2019 American Meteorological Society. For information regarding reuse of this content and general copyright information, consult the AMS Copyright Policy (www.ametsoc.org/PUBSReuseLicenses).

field campaign (Yoneyama et al. 2013) in the equatorial Indian Ocean to estimate or infer convection's moisture sensitivity, a major justification for conducting that campaign.

Work in this area is ultimately motivated by the desire to parameterize the convective process for models of larger-scale phenomena, where “models” is meant very broadly (as in Arakawa 2004) to include the scientific problem of understanding large-scale moist convectively coupled dynamics, as well as the related technological problem of building process-emulation algorithms for scale-truncated numerical models. Both problems require a foundational understanding of the whole phenomenology of latent-heated convection in a stratified environment, including the broad clear-air subsiding branch as well as the narrow condensate-containing mostly ascending branch. Only then can we understand how scale truncation mangles the phenomenon, the better to design a strategy for treating that truncation wound.

The humidity sensitivity of the narrow cloudy branch of convection is fortunately very local. The chain of causality mainly involves actual mixing of environmental air into convective updrafts, affecting their bulk buoyancy, with secondary effects such as reevaporation of precipitation also potentially playing a role. This locality makes balloon and radar field campaigns salient and indeed even crucial because they resolve the vertical dimension, compensating for the biggest weakness of satellite observations. Surely observed associations of humidity and deep convective activity profiles contain some echoes of the causal dependency information we seek, even though correlation and regression coefficients cannot be directly interpreted as indicative of causation.

Unfortunately, inferring the causal core of the parameterization problem (bulk sensitivities at a given filter or truncation scale) from field observations of the total flow is far from straightforward. At this stage of tropical meteorology, qualitative descriptions of detailed field situations and scenarios—the major strength of field campaigns in an earlier exploratory era—may no longer be the frontier or even a limiting resource. Instead, progress seems to require bringing statistically significant sample sizes of geophysical variables into some well-conceived quantitative mathematical framework or model, again in the broad sense of that term.

Assimilation into state-of-the-art numerical weather prediction (NWP) systems is one obvious choice of framework. As stated by Kim et al. (2018), “Ultimately, our knowledge and understanding of the physical processes of the DYNAMO MJO events must be built upon accumulated information from observations, data assimilation products, and numerical model simulations.” But accumulation is perhaps a telling word there, and that

triad is far from equal in their relative volume of “information” contributions. The value of temporary local special observations is tricky to quantify, generalize, or even utilize in NWP: for instance, research radar data have no real way in, except perhaps for statistical verifications of high-resolution simulations driven by NWP analyses (e.g., Takemi 2015; Hagos et al. 2014) or by sounding array-derived forcing sets. Given the sheer abundance of global operational data entering cutting-edge NWP, and the technical hurdles of research with such complex systems, special field data may not find their highest and best use there. Still, analysis needs some framework. Here we use only basic statistics (regression coefficients).

The first step in any statistical analysis is a decision of how much to pool the data versus splitting them into separate “regimes,” within which intraregime associations will be separately characterized. Such choices are uncomfortably subjective, except (arguably) for the agnostic choice: to pool all the trusted data from a uniformly sampled set. The next choice is whether standard linear statistics suffice (correlations and regressions of deviations from the mean), or if a more elaborate statistical model must be entertained to capture the available information content.

Quantitative field data are inaccurate and noisy estimates of the desired geophysical quantities at filter scales. For instance, rawinsonde arrays need a lot of averaging to combat the representativeness errors of such sparse point samples (Ooyama 1987; Mapes et al. 2003). Radar data also have a host of challenges, despite their better coverage: raw measurements do not give our important geophysical quantities, and missing values in Doppler velocity are both numerous and preferentially not sampled in a weather-dependent way (i.e., in conditions of undetectably small reflectivity). Even if measurements were flawless, geophysical noise limits us fundamentally: convection is governed by many unmeasured factors, of both small (subfilter) and large scales. All these limitations (reflected in the scatter of scatterplots) militate strongly for simplicity in analysis, so in this paper we use only standard linear statistics. While some justifiable screenings or elaborations might modestly change the values of some of our correlation or regression coefficients, structural issues of inference and interpretation loom so much larger that such efforts are not, in our view, on point.

For these reasons, our strategy begins with characterizing field data on deep convection statistically, in simple ways. Our larger efforts (as narrated below) are expended instead on the project of bringing much more abundant and accurate data from models (which embody sophisticated and detailed hypotheses) into a state of comparability with the statistically summarized field data. In this approach, the role of problematic but

epistemically unique field data is to adjudicate prior hypotheses from models. We consider this the strongest use of field data's quantitative information content, especially for the parameterization-relevant (scale specific but situation generalizable) questions that ultimately drive research in this area.

While the success of the present effort is only partial, as described below, the methodological aspects may be useful to help clarify justifications and strategies for future field campaigns. For this reason, results from a seemingly disparate line of activities (our titular “synthesis”) are kept together in this paper, and narrated in a logical order from the simplest statistics toward more causality-relevant inference methodologies.

Section 2 describes more details of the mathematical and modeling frameworks used later. Section 3 describes the observational data we used and shows the observational results: a statistical characterization against which to judge model-derived hypotheses. Section 4 describes the parameterization-relevant (causal) convective sensitivity matrix $\mathbf{M}$ that we seek to constrain, and discusses why it is not directly comparable to the observational regressions, despite having the same units and shape. Section 5 narrates our journey toward more observation-adjudicable prior predictions, first merely accounting for autocorrelation, and then using an approach to compensate for a structural limitation of CCPMs. Section 6 summarizes the findings and their caveats, and returns to this discussion of prospects for field campaign research in our computation-rich age.

2. Framework: Definitions of convection, humidity, and sensitivity

We seek to estimate or constrain the sensitivity of vertically resolved deep convective heating $Q(p)$ to vertically resolved specific humidity perturbations $q(p)$. For clarity, let us establish a notation of partial derivatives of Q in discrete pressure layer j with respect to q or T values in layer i . These *sensitivity* derivatives can be expressed as coefficients $\partial Q_j/\partial q_i$ and $\partial Q_j/\partial T_i$ in a series expansion over NP pressure layers:

$$Q_j = Q_{j0} + \sum_{i=1}^{\text{NP}} \frac{\partial Q_j}{\partial q_i} \delta q_i + \sum_{i=1}^{\text{NP}} \frac{\partial Q_j}{\partial T_i} \delta T_i + \dots \quad (1)$$

Unfortunately, partial-derivative notation is ambiguous without saying what all is being held constant, including philosophical status flags (like nature vs model) as well as other continuous or discrete quantities, as elaborated for instance in Mapes et al. (2017). In nature, sensitivity may in general depend (strongly or weakly) on the weather “state” at the scale implied

by an analyst's estimation procedures, while in models, sensitivity depends on all the aspects and parameters of the experimental configuration, again perhaps strongly but perhaps weakly. In this synthesis paper where observations meet models, a mathematical description would require so many indicator symbols, themselves requiring careful words of symbol definition, that we decided to instead narrate the logic verbally with proper care. Mathematical formulas are used only where it helps clarify operations underlying the figures.

This paper is an attempt to test the hypothesis that nature's effective $(\partial Q_j/\partial q_i)|_{\text{nat}}$ implied by field observations is robustly different from Kuang's linear response function $(\partial Q_j/\partial q_i)|_{\text{CCPM}}$ derived from interrogation of a cyclic convection-permitting model (CCPM) in a state of radiative–convective equilibrium. Observations are inadequate (for various reasons narrated below) to permit direct inference of $(\partial Q_j/\partial q_i)|_{\text{nat}}$, so the hypothesis test must be conducted in some common informational domain, with uncertainty considerations in that domain sufficient to give meaning to the “robustly” part of this hypothesis.

Prior work with CCPMs makes plausible that linearizability holds to a remarkable degree, even for fairly large perturbations (say, up to $\pm 50\%$ in precipitation rate) around a steadily convecting state of an atmospheric column of global atmospheric circulation model (GCM) grid relevant size (Tulich and Mapes 2010; appendix of Kuang 2010). Furthermore, the state dependence of Kuang's linearized response function matrix $\mathbf{M}$ for different domains (2D and 3D, different sizes and shapes), and for a couple of different background forcings (i.e., different steadily convecting basic states) is also modest, when normalized for the mean strength of the convective heating (Kuang 2010, 2012; Mapes et al. 2017; Kuang 2018), certainly when compared to the gross fundamental errors of simple lifted-parcel CAPE calculations sometimes used in GCM parameterization schemes. Based on those considerations, we chose an isotropic 3D CCPM-derived $\mathbf{M}$ (discussed in section 4) as a sufficiently interesting baseline measure of sensitivity, and one that matches in complexity the statistic (a matrix of linear regression coefficients) to be derived from the observations.

Physically, the T tendency or heating rate $Q(p)$ as used here is the sum of latent heat release and the convergence of vertical eddy temperature flux, where “eddy” refers to deviations from a horizontal average over a filter scale relevant to the parameterization problem, and/or to the effective spatial coarseness of the observational estimation, usually considered to be of order 10^2 km. Radiation is excluded: $Q = Q_1 - Q_R$ in the classical terminology of Yanai et al. (1973). The linearized partial sensitivity of Q to q can then be expressed as $\mathbf{M}^{Qq} = \partial Q_j/\partial q_j$, where i and j are centered indices of discrete pressure layers

(slabs of mass) in the atmosphere. Matrix $\mathbf{M}^{Qq}$ is a quadrant of the full matrix $\mathbf{M}$ of sensitivities of both temperature and moisture tendencies to both temperature T and q perturbations, in a CCPM framework that embodies a set of conditions (things held constant) on those partial derivatives. More discussion of how to interpret Kuang's CCPM-estimated $\mathbf{M}$ will be taken up again in [section 4](#), after we present the same-shaped but differently interpreted regression coefficient matrix from field observations.

3. Field campaign data analysis

As explained in the introduction, we begin with simple linear regression of pooled data, using Q and q time series drawn one by one from a set of altitude layers. Data appear too few and too noisy to allow robust estimation of a meaningful multiple-predictor regression, even with principal component preparatory data reductions (appendix C of [Song 2015](#)). Specific humidity q was computed from the hygrometer and thermometer of balloon soundings launched over the Gan Island DOE ARM site (0° , 90°E) every 4 h during AMIE-DYNAMO, and was further averaged in pressure and interpolated in time to yield the 50-hPa-layered hourly data product utilized here. Hourly microwave sounding retrievals ([Zhang et al. 2018](#)) were also studied, but with only 2–3 vertical degrees of freedom they were deemed less useful for present purposes. It must be noted that the spatial scale of q measurements is essentially a point, as compared to around 100 km for the radar-derived divergence with which we will correlate it.

Unfortunately, convective heating Q is not measured directly. Because gravity is strong and thus efficient at flattening density surfaces, on parameterization-relevant scales in the tropics (much smaller than the Rossby radius of deformation), the only field-measurable quantity with information content about the vertical profile of hourly time scale Q above observational noise levels is horizontal wind divergence. The next subsection describes our effort to derive estimates of divergence on the $\sim 100\text{-km}$ scale, using the Doppler radar VAD techniques of [Mapes and Lin \(2005\)](#).

a. Divergence D and diabatic divergence D_d (a measure of Q)

Two Doppler radar datasets with full azimuthal coverage are available in the AMIE-DYNAMO datasets: one from the Texas A&M Shared Mobile Atmospheric Research and Teaching Radar (SMART-R) truck-mounted C-band radar on Gan Island ([DePasquale et al. 2014](#)), and one from the NASA TOGA radar on the R/V *Revelle* ([Xu et al. 2015](#)). Raw data were kindly binned by

researchers in the Schumacher group at Texas A&M University (the SMART-R data, which we utilized for 6 October–31 December 2011) into hourly histograms in a cylindrical coordinate system centered on the antenna as described in [Mapes and Lin \(2005\)](#). Similar processing was kindly performed by Dr. Paul Hein of the Rutledge research group at Colorado State University (the *Revelle* data), but since results from that dataset were similar but noisier they were not in the end shown in this paper. Altitude layers of 500 m were further pooled into 50-hPa pressure layers. Doppler dealiasing (unfolding) was done in this cylindrical space by a two-step process: (i) the hourly histogram of raw radial velocities was used to dealias the radial velocities in each spatial bin into a common Nyquist interval, and then (ii) first-guess wind speed and direction were used to absolutely unfold the spatial bin mean radial velocities. More details, and the resulting dataset of VAD-derived quantities, can be accessed at [Mapes and Chandra \(2017\)](#). For the present paper, divergence was estimated from line integrals of pooled data falling within the 8-km-wide range annulus centered at 76-km range, which by Stokes's theorem gives an area average over that circle of approximately 76-km radius.

On dynamically small scales, where density gradients and time changes are small ([Charney 1963](#); [Sobel and Bretherton 2000](#)), wind divergence can be interpreted to within observational error as “diabatic divergence,” denoted herein as D_d ([Raymond 1983](#); [Mapes and Houze 1995](#)), which is the pressure vertical derivative of the quotient of total heating rate $Q_1 = Q + Q_R$ divided by a static stability profile. In this work, static stability is taken to be time independent, from a time-mean tropical profile of virtual temperature, and we neglect the contribution of Q_R .

One check that our radar-derived divergence D is a valid measure of D_d is to regress it against precipitation, which is a measure of column-integrated Q . The result in [Fig. 1](#) compares well to [Fig. 13](#) of [Mapes and Lin \(2005\)](#) showing the identical regression for nine other radar deployments in tropical field campaigns. Another indicator of the physical realism of D is that the vertical integral should vanish to high precision, since tropical surface pressure variations are very small. Despite the fact that divergence at each altitude is estimated from independently measured Doppler shifts, the regression in [Fig. 2](#) evidently satisfies this physical constraint rather well, except at large negative lags (indicative of sparse echo and the associated Doppler data shortcomings at upper levels in advance of deep rainstorms).

At peak rainfall time (lag 0), convergence near the surface up to 600 hPa is overlain by divergence, indicating deep upward motion. At positive lags (after peak rain),

FIG. 1. Lagged regression of the Doppler divergence time series at different pressure levels onto the surface rain-rate time series. Units are $10^{-6} \text{ s}^{-1} (\text{mm h}^{-1})^{-1}$, and negative (positive) lags indicate times before (after) the rainfall signal.

midlevel convergence is seen, indicative of upper-level ascent in precipitating stratiform clouds and/or descent in the lower troposphere where that precipitation evaporates (Houze 1997). In the negative lags leading up to peak rainfall, a deepening layer of lower-tropospheric ascent is implied, sometimes expressed in the shorthand of (inferred) “cumulus congestus” convection, the middle-topped mode of tropical trimodal convection (Johnson et al. 1999). These features confirm again the results of similar regression results in Mapes and Lin (2005) and Mapes et al. (2006).

b. Regression coefficients of Q (as indicated by D) on q

This section describes the heart of our observational results, the outdoor deep convective data version of associational studies like Bellenger et al. (2015) for shallow convection and Takemi (2015) using model data. Figure 2 displays simple univariate regression coefficients of $D(p, t)$ on a base time series of temporal anomalies $q'_{500-600}(t)$ averaged over the 500–600-hPa layer, as a function of time lag. In broad terms, Fig. 2 may be related back to Fig. 1 in light of the fact that midlevel humidity is positively correlated with rainfall (e.g., Holloway and Neelin 2009, 2010). That is, Fig. 2 resembles Fig. 1 in some ways, weighted especially to the structure at positive lags with midlevel convergence sandwiched between layers of divergence. Although the midlevel convergence in Fig. 2 rises somewhat with lag, its vertical tripole pattern remains stratiform-like at all lags, indicating that midlevel humidity is especially strongly correlated with stratiform-rich mesoscale convective systems.

FIG. 2. Lagged regression coefficients of divergence time series from the SMART-R Doppler radar at different altitudes onto the time series $q_{500-600}$ of specific humidity averaged over the 500–600-hPa layer. Coefficients with magnitudes less than their uncertainty values (because of either weak signal or large noise, assuming each hour of data is 1 degree of freedom) are masked in black. Negative (positive) lags indicate times before (after) the moisture signal.

Since the results of Fig. 2 are only a weak function of lag, it is informative to extract the lag = 0 column alone. Figure 3 juxtaposes that with the similar columns from equivalent univariate regressions of D on q in other layers. Again, the tendency for cancellation between positive and negative values in the columns of Fig. 3 is indicative of physical validity of the divergence measurement. Recalling that regression coefficient

FIG. 3. Regression coefficients of divergence time series in different altitude layers onto base time series of moisture from different layers, with no time lag. The column centered on 550 hPa is identical to the lag = 0 column of Fig. 2. White areas are coefficients with magnitudes less than their uncertainty values (statistically indistinguishable from 0 because of either weak signal or large noise).

FIG. 4. (a) The heating-from-moisture (Qq) quadrant of response function $\mathbf{G}_{3h}$, remapped appropriately to our 10 equally spaced pressure levels [cf. Fig. 8c (middle row) of Kuang (2012) or Fig. 4 of Kelly et al. (2017)]. The color-scale values in (a) are nearly equal to $\text{K day}^{-1} (\text{g kg}^{-1})^{-1}$. (b) As in (a), but for the QT quadrant.

$R(D, q) = c(D, q) \sigma_D / \sigma_q$, where c the correlation coefficient is bounded in $[-1, 1]$, we can understand that the larger values of regression coefficients for time series $q(t)$ from higher altitudes reflects the small standard deviation σ_q there. The results could be recast to be more uniform by converting to relative humidity as a measure, as discussed in Mapes et al. (2017).

The great difference of interpretation between the regression associations in Fig. 3 and the true causal response of convection to moisture is stark in the case of the 900–1000-hPa layer. An increment of q in the low levels would cause a substantial increment of deep convection, *all else being equal* (as shown Fig. 4a and section 4 below). But all else is not equal in natural variability. Instead, low-level moisture during a typical episode of enhanced deep convection over the ocean is not observed to be especially elevated, because cool and often absolutely dry low- q outflow (postconvective) air expands to cover large areas and therefore to strongly influence spatially random balloon samples, whether or not that is representative of the inflow to convective updrafts. At a statistical association level, then, balloon-measured $q_{900-1000}$ is only weakly correlated to Q excursions, as indicated by the weak divergence signature in the 950-hPa column in Fig. 3.

Despite this impossibility of causal interpretation, the spatial coherence of the pattern in Fig. 3 does suggest that statistically significant, systematic information content is present. Can this information be utilized to constrain, evaluate, or improve a first-guess causality estimate?

4. Unobservable prior causality estimates: The sensitivity Jacobian

The information content of observations is precious but feeble: There are never enough or precise-enough

data, and it is difficult or impossible to trace the propagation of the many sources of errors (accuracy, coverage, sampling representativeness) through complex data processing steps. In addition, natural variability involves correlated excursions of many variables, many of them unobserved, so that associational relationships like Fig. 4 cannot be interpreted directly in parameterization-relevant (causal) terms.

Instead, to exploit observations effectively, we will transform a few candidate prior estimates of convection's causal moisture sensitivity into a format where Fig. 3 can be used to assess their relative veracity. As mentioned above, we have from the work in Kuang (2010, 2012) a linearized response function $\mathbf{M}$ derived from interrogating a CCPM. To approach observations, the heating Q in the quadrant $\mathbf{M}^{Qq} = \partial Q / \partial q_j$ must at a minimum be converted into diabatic divergence units, so that it can be even nominally comparable to Fig. 4 in its shape and values (including units c). Fortunately, the required vertical differentiation is a linear operation and commutes with the procedures used to derive $\mathbf{M}$.

Matrix $\mathbf{M}$ is a fitted linear surrogate model of a CCPM's sensitivities around its base state of statistical equilibrium between an imposed destabilization forcing and the resulting explicit convective processes. Here we take Kuang's highest-resolution 3D CCPM results on the parameterization-relevant and radar-comparable $128 \text{ km} \times 128 \text{ km}$ domain scale. Matrix $\mathbf{M}$ estimates the anomalous convective tendencies of T and q that would occur in the CCPM in response to instantaneous, horizontally uniform but vertically local perturbations of the mean T or q profiles. In other words, $\mathbf{M}$ expresses genuine causality, the thing we want for parameterization, not merely an observed association. Ideally, as advocated by Mapes et al. (2017), the conditions for these partial derivatives should be explicitly notated, because

the value of the derivative depends on their values, although the list may be long:

$$\mathbf{M}^{Qq} = \frac{\partial Q_i}{\partial q_i} \Big|_{T_{\text{all}}, q_{\text{all}} \neq i; \text{RCE}(\{w=0, \text{geometry}, \text{SST}, Q_{\text{rad}}, \dots\})} \quad (2)$$

This formulation clarifies that T at all levels, and q at all levels but i , are held constant in assessing this local q sensitivity. Direct attempts to estimate $\mathbf{M}$ -comparable quantities observationally are not straightforward, because in nature T and q covary with each other and across altitudes. While a multipredictor regression could in principle tease apart the dependences on T and q at all levels, *if* sufficiently many linearly independent variations were observed accurately enough, other unmeasured variables would still confound the effort. An excellent new general-science discussion of modern causal inference is offered in [Pearl and Mackenzie \(2018\)](#), and some attempts at atmospheric applications may be seen for instance in [Ebert-Uphoff and Deng \(2017\)](#), but such methods are beyond the present scope.

The term “convective” heating and moistening includes both turbulence and thunderstorms, which can be separated by time scale in $\mathbf{M}$. The largest values in $\mathbf{M}$ are vertical tripole (+/-/+) tendencies straddling the diagonal, expressing the process of vertical diffusion, which produces +/-/+ tendencies straddling a positive vertically local T or q impulse. However, that strong diffusion response does not last long once a sharp spike is smoothed out. In mathematical terms, that aspect of $\mathbf{M}$ is characterized by large negative eigenvalues. To better see the longer-lasting responses of the deep cloud ensemble, it is clearer to display time-averaged responses over a few hours following an impulse.

Time-averaged tendencies are typically expressed in the form of a finite-time propagator matrix, $\mathbf{G}_\tau = [\exp(\mathbf{M} \times \tau) - \exp(\mathbf{M} \times 0\text{h})]/\tau = [\exp(\mathbf{M}\tau) - \mathbf{I}]/\tau$. Its two terms predict the CCPM’s state evolution over a time interval $[0, \tau]$ by evaluating the state at those two times in the exponential solution to the matrix differential equation $d\mathbf{x}/dt = \mathbf{M}\mathbf{x}$. The lack of a vertical advection term in this equation embodies the CCPM’s lack of vertical velocity response to its predicted Q . The 4-h-averaged tendencies $\mathbf{G}_{4\text{h}}$ are shown in Fig. 8 of [Kuang \(2012\)](#); here we show 3-h versions arbitrarily, to avoid redundancy. The 1-h values are broadly similar ([Kelly et al. 2017](#)), and somewhere in the 1–4-h range is arguably the right choice for our goal of eventually being comparable to observations: Doppler radar D is hourly averaged (pooled), while sounding q is 4 hourly and balloon ascents take about an hour. This explicit and forward-in-time specification of the several-hour τ in $\mathbf{G}_\tau$ is a convenient label for emphasizing that the response is

averaged over all stages of the life cycle of convective clouds (although really it is averaged over an all-life-stages *ensemble* in a state of equilibrium, so the lifetime of individual clouds is not actually involved mathematically). That average response occurs at all altitudes throughout the vertical domain, not just locally. In mathematical terms, the matrix exponentiation in $\mathbf{G}$ diminishes the strong but transient diffusion process, because of its large negative eigenvalues.

Figure 4 shows the quadrants $\mathbf{G}_{3\text{h}}^{Qq}$ and $\mathbf{G}_{3\text{h}}^{QT}$, comparable to parts of Fig. 8 of [Kuang \(2012\)](#) but on our 10-level vertical pressure coordinate. The vertical red columns at the left edges of both panels indicate that a positive perturbation to either humidity or temperature in the 1000–900-hPa layer *causes* enhanced convective heating throughout the troposphere in the subsequent 3 h. This can easily be understood as a simple buoyancy enhancement effect for low-level parcels rising in the deep convection. More subtly, the mixed red and blue column centered on 350 hPa in the Qq panel indicates that a positive q perturbation in the 300–400-hPa layer enhances heating (presumably condensational) at and above that layer, but also boosts cooling (presumably evaporation of that enhanced precipitation) below it by a similar amount, yielding a vertical dipole in the heating profile Q . The mass-weighted vertical integral of Q is total convective heating, nearly proportional to rain rate. Those column integrals of $\mathbf{G}_{3\text{h}}$ (3-h rainfall response functions) are plotted as profiles in Fig. 5 of [Mapes et al. \(2017\)](#).

The cells of Figs. 3 and 4a are still not in the same *units*. To bring them closer to comparability, we need to convert $\mathbf{G}_\tau$ into its equivalent for diabatic divergence D_d instead of Q , a matrix we will call $\mathbf{G}'$. This involves division of matrix columns $Q(p)$ by a time-mean tropical static stability profile, then differentiation vertically (in pressure coordinates). Fortunately, these are again linear operations that commute with the operations used to develop $\mathbf{M}$ and $\mathbf{G}$. The centered evaluation of this derivative puts the D_d columns of $\mathbf{G}'$ on levels that are offset by 50 hPa from the Q levels in the columns of $\mathbf{G}$. For this reason, we left Figs. 3 and 4 on a 50-hPa grid.

Converting Q to diabatic divergence D_d and summing correlated responses

As discussed in [section 3a](#), area-averaged convective heating Q is not a measurable quantity. However, the linear transformation from $\mathbf{G}$ to $\mathbf{G}'$ yields *diabatic divergence* D_d , which we have argued would be indistinguishable from the actual divergence D in Fig. 3 to within observational uncertainty. Figure 5 shows $\mathbf{G}'_{3\text{h}}$. The tall red columns at left in Figs. 4a and 4b translate into near-surface diabatic convergence and near-tropopause diabatic divergence in Figs. 5a and 5b. These are the

FIG. 5. Quadrants of the response matrix $\mathbf{G}'_{3h}$, transformed from the submatrices shown in Fig. 4. The columns of Fig. 4 were transformed from Q to diabatic divergence D_d by taking the vertical pressure derivative of $Q/(ds/dp)$, where s is dry static energy. Notice that the vertical differencing puts these cells on a staggered grid relative to Fig. 4.

divergent flows that would occur very rapidly if the deep heating profile Q occurred in a dynamically small local area in a weakly rotating stratified fluid while gravity kept density surfaces horizontal and constant (Sobel and Bretherton 2000).

The column in Fig. 5a centered on 550 hPa is the diabatic divergence response which would occur in a convecting CCPM, as a *causal reaction* to imposed 500–600-hPa humidity perturbations, with q at all other levels and T at all levels held constant, as in the list of conditions in Eq. (2). Values in this column are replotted as a dashed-blue curve in Fig. 6a. Comparing to the red curve, which is the corresponding observational regression (a column of Fig. 3), we see that the CCPM-derived causal response curve is an order of magnitude smaller, as well as different in shape. The red and dashed-blue curves for other base levels (Figs. 6b,c,d) are likewise incommensurate in value. The solid blue curves in Fig. 6 are more similar to the red in magnitude, and will be described after the necessary viewing of Fig. 7.

In nature, humidity varies in coherent layers deeper than 100 hPa, so when one layer is anomalously humid and enhancing convection, the adjacent ones tend to amplify the effect. In addition, q variations are correlated with T , and those associated T perturbations cause additional impacts on convective heating. Estimating the impact of these correlated influences allows us to bring model-predicted sensitivity quantities one step closer to comparability with observations.

Autoregression profiles of q from Gan sounding data in our 100-hPa layers are shown in Fig. 7a. Regressions of T on q are further shown as columns in Fig. 7b. The diagonal is unity in Fig. 7a, and values elsewhere are all positive: when any level is more humid, all levels tend to be more humid. Temperature regressions exhibit both signs, and tend to indicate a cooler lower troposphere but

warmer upper troposphere during moist times, a known aspect of variability such as the Madden–Julian oscillation that was prominent during AMIE-DYNAMO (Johnson and Ciesielski 2013).

Accounting for the impact of *observed* autoregression associations on our CCPM-predicted causal impacts brings the latter closer to observational comparability, yielding the solid blue curves in Fig. 6. This accounting may be appreciated by writing the CCPM-derived dashed-blue curve in Fig. 6a (a column of $\mathbf{G}'$) as a partial derivative $\partial(D_d)_j/\partial q_{500-600}$, as in Eq. (1). Observational regression coefficients (columns of Fig. 7) are the slope of lines fit to scatterplots, so they are more appropriately represented as ratios of finite-amplitude (indicated by symbol δ) observed excursions, $\delta q_i/\delta q_{500-600}$ and $\delta T_i/\delta q_{500-600}$. Using the “total derivative” symbol d for the result, we can then express the solid blue curves in Figs. 6a and 6e as the two summation terms in

$$\frac{d(D_d)_j}{dq_{550}} = \sum_{i=1}^{10} \left. \frac{\partial(D_d)_j}{\partial q_i} \right|_{\text{CCPM}} \frac{\delta q_i}{\delta q_{550}} + \sum_{i=1}^{10} \left. \frac{\partial(D_d)_j}{\partial T_i} \right|_{\text{CCPM}} \frac{\delta T_i}{\delta q_{550}} + \dots, \quad (3)$$

which is a variant of Eq. (1). Solid blue curves in the center and right panels in Fig. 6 display the same results with 550 (representing the 500–600-hPa layer) replaced by 650-, 750-, 850-hPa-layer means. The sensitivity of CCPM convection to q is larger than to q -correlated T fluctuations (evidenced by larger magnitudes of the solid blue curves on the top panels of Fig. 6 than those in the bottom panels). The similarity of all the solid blue curves in Fig. 6 stems from the propensity for the entire troposphere to be moist when any part is moist (all values are positive in Fig. 7a).

FIG. 6. (a)–(d) Comparison of SMART-R radar divergence regressed on q anomalies in 100-hPa layers (red; columns from Fig. 3) with causal quantities drawn from matrix $\mathbf{G}_{\text{sh}}$ summarizing the CCPM’s behavior (blue). Broken blue profiles indicate single columns from Fig. 5a, responses to isolated 100 hPa fluctuations. Solid blue profiles are sums of the columns of Fig. 5a, weighted according to the observed autoregression of humidity from Fig. 7a [first summation in Eq. (2)], for 100-hPa-layer base time series centered on each “Pert-layer” value as indicated. (e)–(h) As in (a)–(d), respectively, but blue curves are sums of the columns of Fig. 5b, weighted according to the observed regression of T on q from Fig. 7b [second summation in Eq. (2)].

This accounting using Eq. (3) for the deep vertical coherence of natural variations goes some way toward making the CCPM-predicted causal sensitivity (solid blue curves) at least commensurate in magnitude with the observed regressions (red curves). However, magnitudes in the red curves are still substantially greater in terms of Q (although the differentiation to yield D_a somewhat obscures this fact). The red curves from linear regression represent that portion of the observed D variance that can be linearly explained by all factors in nature that are linearly correlated with observed q . By this reasoning, if we take $\mathbf{M}$ to be a realistic estimate of convection’s causal sensitivity (the null hypothesis to this paper’s motivating hypothesis), then the gross *magnitude shortfall* of Q implied by the solid blue curves relative to the red suggests the existence of additional q -correlated factors in nature that make broadly *positive* contributions to natural variability in Q . Beyond this magnitude shortfall, the incorrect shape of the blue curves suggest that those other q -correlated factors may

also contribute importantly to the vertical structure of convective heating variability. But of course we cannot rule out the possibility that $\mathbf{M}$ is extremely different from nature’s effective sensitivity matrix.

One clear lesson in modern causality inference (Pearl and Mackenzie 2018) is that causal hypotheses need to be stated explicitly if observational studies are to have any chance of estimating their terms. In the next section, we speculate on one confounding (important but unmeasured) governing variable of natural convective variability, and attempt to address its absence with a sophisticated but idealized tool. This exercise is nonunique and a bit complicated, as discussed in the conclusions, but deserves the reader’s consideration in our view.

5. Coupling with $[w]$: Toward an $\mathbf{M}$ -based quantity more comparable to observations

We speculate that one major confounder (an important but unmeasured variable) shaping filter-scale

FIG. 7. Columns are regressions of (a) q and (b) T on q in the indicated base layer (i.e., the layer whose mean is the base time series for the temporal regression). Note the matrix convention (with the diagonal from upper left to lower right) in the axis conventions.

Q variability in nature is filter-scale vertical velocity $[w]$, which CCPMs (and therefore their causal response matrices $\mathbf{M}$ and $\mathbf{G}$) lack for structural reasons. Advection of basic-state vertical gradients by a Q -proportional $[w]$ under the weak temperature gradient approximation could be incorporated as a linear effect in a modified matrix (Kuang 2012), or captured in other parameterized large-scale dynamics approaches (e.g., Daleu et al. 2016, 2017), but here we shall instead utilize a more explicit mechanism for coupling CCPM sensitivities to larger-scale dynamics, in order to bring those sensitivities into an observation-comparable domain.

The sign of the effect can be anticipated logically: the trapping of heating-induced subsidence by the $[w] = 0$ condition in “closed” CCPMs (i.e., the noninclusion of dynamical adjustments to heating) acts as an artificial *negative* feedback, making CCPM convection *less responsive* to, say, a boost from anomalous moisture than a comparable-sized patch of atmosphere would be in a dynamically responsive or “open” atmosphere. Thus $[w]$ could fill the speculated role of an unmeasured q -correlated variable with positive effect [i.e., a positive contribution by the “...” terms in Eq. (3)].

To achieve closer comparability to observations by coupling to dynamics, we immersed matrix $\mathbf{G}_{1h}$ in a GCM, specifically the 10-level model described in Kelly et al. (2017). The coupling strategy is a tangent linear version of “superparameterization” (e.g., Randall 2013), an approach that the reader should broadly understand in order to fully appreciate the description below.

In the present experiments, we used a 10-layer global primitive-equation solver (dynamical core), with an advected tracer q but no condensation scheme other than the time-independent 3D forcing field and the convective tendencies contained in $\mathbf{G}_{1h}$. This dynamical core was calibrated to operate in a realistic boreal autumn basic

climatology by means of time-independent forcings (3D source terms) on all the state variables (u, v, T, q). Those forcings were created empirically, following Hall (2000), by negating the ensemble average of the first-time-step tendencies after an ensemble of initializations of the dynamical core only with a representative set of atmospheric reanalysis states in that season. In addition to that time-independent forcing, which allows model “weather” to develop only through hydrodynamic (momentum) instabilities of the base state (Sardeshmukh and Sura 2007), we also permit the model to develop moist tropical variability by permitting anomaly coupling to matrix $\mathbf{G}_{1h}$.

The anomaly convection coupling works as follows. Convective tendencies of T and q are computed as the matrix product $\mathbf{G}\mathbf{x}$, where $\mathbf{x}$ is the 20-member state vector $[T'(p), q'(p)]$ of anomalies (deviations from a climatological time mean of model output from the uncoupled model). To confine the resulting anomalous convection tendencies to convecting regions in the tropics, they are scaled by a time-invariant geographical pattern that is zero outside convecting tropical regions, and inside the tropics is a smooth version of the climatological tropical precipitation pattern in boreal autumn. This approach embodies the assumption that the linearized convective tendencies $\mathbf{G}\mathbf{x}$ depend on the background or base state via a constant factor, proportional to an assumed background convective rain rate. A conceptual model of convection consistent with this assumption is that the linear response $\mathbf{M}$ is proportional to the *areal number density* of convective cloud systems, a reasonable description of deep convective variability (Robe and Emanuel 1996; Parodi and Emanuel 2009), with each of these cloud systems having the same vertical structure of sensitivities and impacts.

FIG. 8. Regressions of divergence D on $q_{500-600}$. The observation-derived curve (red) is repeated in all panels as a reference. Black curves show regression coefficients computed identically to the red curve, but using as input data the 1800 days of time–height outputs of Q and q from the 10-layer GCM anomaly coupled to the $\mathbf{G}$ matrix with its free-tropospheric q sensitivity scaled by $\alpha =$ (a) 0, (b) 1, and (c) 2.

With this anomaly coupling to the simple GCM, convectively coupled tropical waves develop and interact with the advected moisture field and the hydrodynamic weather in the model’s fully nonlinear midlatitudes, even as the forcing that sets the model climate remains constant in time. While the simulations (Kelly et al. 2017) are not as realistic as we had hoped, there is enough variability to allow time-dependent “virtual field campaign” datasets (Mapes et al. 2009) to be extracted and compared statistically to AMIE-DYNAMO field data.

This virtual field campaign approach is arguably the most observation-comparable quantity we can derive from experimental CCPM sensitivity matrices $\mathbf{M}$ or $\mathbf{G}$. On this basis, we attempt here to test our main hypothesis (from section 2) that nature’s effective $(\partial Q_j / \partial q_i)_{\text{obs}}$ implied by field observations is robustly different from Kuang’s linear response function $(\partial Q_j / \partial q_i)_{\text{CCPM}}$. Three experimental candidate matrices are considered, with an experiment factor $\alpha = \{0, 1, 2\}$ multiplying the columns representing the CCPM’s convective response to q in $p < 900$ hPa. This doubling or zeroing out of the relevant columns of $\mathbf{G}_{\text{th}}$ by factor α strongly affects the model’s tropical variability, as shown in Kelly et al. (2017). However, that simulated free variability was not considered realistic enough to permit the candidate values of α to be declared better or worse in that study. The results here are an additional approach to evaluate this model against *local* observations in terms of temporal correlations. A third approach (hindcast skill) is relegated to the appendix, since its findings are weak but it slightly affects the regressions here.

Figure 8 shows actual and virtual field data regressions, our cleanest comparison involving experimental CCPM-derived sensitivity matrices. Regressions of D on q_{550} from AMIE-DYNAMO (red, from Fig. 6) are mimicked

from time–pressure sections at an equatorial Indian Ocean gridpoint in the $\mathbf{G}$ -coupled GCM (black). With no free-tropospheric moisture sensitivity ($\alpha = 0$; Fig. 8a) the regression values are clearly too small, indicating too little *association* of q_{550} with modeled convective heating Q when its *causal impacts* on Q are forcibly suppressed. In contrast, the case with $\alpha = 2$ (doubled sensitivity; Fig. 8c) is the best match to the observational regression, in shape as well as magnitude. Interpretation follows in section 6.

6. Summary and conclusions

At face value, Fig. 8 indicates that AMIE-DYNAMO observations suggest that natural convection’s sensitivity to free-tropospheric moisture may be roughly twice that computed from a small-domain cloud-resolving CCPM. Before asserting this finding too strongly, let us review this paper’s tall stack of logical inference steps and assumptions.

Doppler radar data on horizontal wind divergence D contain our most informative vertically resolved information about Q (or more strictly, about the dynamically small-scale component of total heating Q_1) on the filter scale desired for parameterization. Rawinsonde humidity is significantly correlated with D variability (Figs. 2, 3), indicative of real physical relationships but with unknown chains of causality. Scatterplots (not shown) give little indication that valuable refinements could be gained by subsetting or by fancier curve fits. Unfortunately, natural variability is also governed by additional unobserved and observed variables that are linearly correlated with q , so these regressions are not direct indicators of causality, although they are informative about it (Pearl and Mackenzie 2018).

The matrix of regression coefficients (ratios of finite excursions notated $\delta Q_i/\delta q_j$ related to the red curves in Figs. 6 and 8 through the Q – D relationship) is wildly larger in its values than the parameterization-relevant (locally causal) CCPM-derived matrix quadrant $\mathbf{M}^{Qq} = \partial Q_i/\partial q_j$ (dashed-blue curves in Fig. 6). One reason for the discrepancy is that natural q variations are coherent over deep layers and correlated with T , while partial derivatives apply with all other dependencies held constant. A closer to salient comparison is thus given by the rhs sums (solid blue curves in Fig. 6) in Eq. (3) for the “total” signal dQ_i/dq_j that would be expected if $\mathbf{M}$ were the true sensitivity of natural convection, but if it acted on coherent (autocorrelated) q and T fluctuations like the observed. However, even that total remains deficient, probably in part because of the effects of additional q -correlated positive influences on convection [the ellipsis in Eq. (3)].

We speculate that a big problem involves the CCPM’s cyclic boundary conditions, which enforce $[w] = 0$, distorting convectively driven environmental motions that feed back quickly and importantly on convection changes. The *causal sensitivities* of closed (CCPM) columns are fundamentally different quantities from the *statistics* of fluctuations in dynamically open air columns (in nature or the GCM). Unfortunately, the former are what we need for subfilter-scale parameterizations whose putative improvement is an overarching motivation for field campaigns.

Virtual field campaigns bridge this fundamental difference, as shown here from a GCM coupled to a family ($\alpha = \{0, 1, 2\}$) of prescribed experimental sensitivity specifications, made possible by the GCM’s linearization assumptions and the “superparameterization” coupling approach (Randall 2013) underlying the title of Kelly et al. (2017).

It is this real versus GCM time series statistics comparison which suggests that the $\mathbf{M}$ summarizing the responses of a small (128-km square) isotropic 3D CCPM has inadequate (by roughly a factor of 2) free-tropospheric moisture sensitivity. One immediate concern is that the model’s variability is different in character from the AMIE-DYNAMO variability, since the hindcast strategy was not very skillful (see appendix). However, correlations in filtered data (e.g., Mapes et al. 2006, Inoue and Back 2015, Yasunaga et al. 2019) show that lower-frequency variability exhibits bigger excursions of q per unit rainfall (smaller $\delta Q/\delta q$, in the notation of this paper). Since the AMIE-DYNAMO intraseasonal variability is lower frequency than our GCM’s Kelvin waves, this effect would strengthen not weaken the case for concluding $\alpha > 1$.

One plausible reason for $\alpha > 1$ could be that “organized” convection in nature differs systematically from

the intermittent cumulonimbi in the small CCPM. We can quantify this notion by comparing the control ($\alpha = 1$) matrix $\mathbf{M}$ used above to an equivalent matrix $\mathbf{M}_{\text{org}}$ for Kuang’s (2012) long narrow CCPM domain, which enforces squall-like structure: spatially confined, steadier convection summarized there as *layer overturning*. Matrix $\mathbf{M}_{\text{org}}$ has a linear rainfall sensitivity to q that is statistically noisier [yellow band in Fig. 5 of Mapes et al. (2017)], and varies a lot with height, but is clearly greater than that of $\mathbf{M}$, which is seen in that figure to have a fairly vertically uniform value of about 25% rain-rate increment averaged over 3 subsequent hours per 1 g kg^{-1} q impulse in a 100-hPa layer. Eyeball averaging over the 900–400-hPa layer (ignoring the puzzlingly large sensitivity of $\mathbf{M}_{\text{org}}$ to upper-level q), we may say that $\alpha = 2$ would lie about halfway between $\mathbf{M}$ and $\mathbf{M}_{\text{org}}$. While intriguingly consistent in sign, this evidence that nature has roughly half-organized convection is frail. For one thing, $\mathbf{M}_{\text{org}}$ is a linearization about a much drier mean state reflecting the large coherent subsidence area in its domain. Probing that whole system with horizontally uniform q increments does not correspond cleanly to the parameterization-relevant sense of causality we are trying to constrain. Indeed, survival of convection in those dry states might actually be an indication that it is less sensitive to q , as discussed below Fig. 5 of Mapes et al. (2017). However, the domain mean q is arguably irrelevant to the *local* sensitivity of the cloudy ascending branch, which is really the process we need to parameterize. In short, interpretation hinges, as always, on subtle scale separation issues requiring careful thought in future research.

Other idealized ways of allowing $[w]$ to respond to Q (parameterized large-scale dynamics) should be explored, although those like our GCM have long litany of assumptions. In addition, other neglected causal confounders [excluded but important variables; ellipsis terms in Eq. (3)] must be acknowledged as possible drivers of the AMIE-DYNAMO observed variability. Raymond and Flores (2016) combine these two suggestions in their model study of the importance of surface flux in “predicting” convection (in a statistical sense in observational work; and causally in the model). Quantitative analysis is needed to reconcile these senses, in light of the correlations among causative factors, which differ for convective variability of different types (that is, filtered correlations for different space and time scales). Internal feedbacks from subfilter-scale structures (Colin et al. 2019) may also affect such estimates and require an accounting. Again, the analyst’s whole causal model must be made explicit and scale specific for individual causality inferences to be meaningfully assessed and compared (Pearl and Mackenzie 2018). In light of these methodological uncertainties,

the paucity and errors of observational data begin to seem like secondary concerns.

Where does all this leave a student of moist (convectively coupled) tropical dynamics who seeks a trustworthy new conclusion about the causes of natural (or at least realistic) convection variability? Or a parameterization developer seeking truly empirically supported improvements? Or an observationalist seeking motivations for more and better outdoor campaigns in the tropics?

For the dynamicist, the precision and controllability of CCPMs remains extremely valuable for inferring coupling mechanisms, despite some systematic shortcomings of the models (e.g., Varble et al. 2014). Dynamically “open” protocols allow CCPM-simulated convection to be interrogated cleanly for responses to certain forcings while coupled to parameterized or simplified large-scale dynamics (e.g., Sobel and Bretherton 2000; Raymond and Zeng 2005; Herman and Raymond 2014; Raymond and Flores 2016; Daleu et al. 2016) that are less chaotic than our tangent linear GCM’s dynamics. Direct if rather method-dependent contacts could then be made with observations, at least “to some degree” (Wang et al. 2013, 2016) or statistically (Sentić et al. 2015), using idealized probing signals for sensitivity characterizations (Sessions et al. 2015). Perhaps with the right feedbacks and stochastic forcings, output variability suitable for comparison to observational relationships could allow stronger causal inferences than ours here, a strong use for field observations.

For the parameterization developer, it would be straightforward to build data objects comparable to Fig. 8 from initialized or free-running model outputs at virtual campaign sites [e.g., revisiting datasets from Mapes et al. (2009)]. However, local field campaigns are a weak evidentiary basis for globally applied parameterizations, and the routine observing network filtered through well-designed data assimilation activities probably offers more resolving power. Observations aside, comparisons of parameterizations with CCPMs (e.g., appendix of Herman and Kuang 2013) could be a direct path to “improvement,” if performing more like superparameterization (Randall 2013) is viewed as good. Deeper still, there are always framing issues in algorithmic treatments of convection that should be best devised by minds steeped in rich observations and the questions they evoke. Thus, even if parameterization improvement is becoming a bit threadbare as a narrow justification for special observing campaigns, participation experience could still be transformative.

For the campaign advocate, then, exploratory work has the strongest data justifications: the $N^{-1/2}$ decay of sampling noise reduction with number N flattens very quickly as a driver in cost–benefit analyses. The question

then arises about what composes a *new* phenomenon or regime, as opposed to more samples of an existing one. Is tropical convection one process? Or is it robustly and importantly different by location or season or setting, in ways we can realistically hope to discern with logistically feasible sample sizes? In this paper, all of AMIE-DYNAMO’s duration at a site was pooled. Subdividing the data by subregimes or adding data from another site (the *Revelle*) seemed to offer little clear signal above the noise toward addressing our driving questions. While intercampaign comparisons reveal some differences [e.g., comparing Fig. 13 of Mapes and Lin (2005) to Fig. 1 here], the strongest of those concerned merely radar calibration.

Beyond the data benefits, field campaigns may exert their most positive influence through the experience of participants. Detail-rich albeit qualitative impressions, including semiquantitative displays of data but also sense impressions and the thoughts and discussions they provoke among diverse assembled participants, profoundly shape our science over time. These social virtues militate strongly for a continuing role of observational campaigns in our equation-rooted and increasingly simulation-led science. But in this perspective, inclusive participation (by theoreticians and parameterization developers, not just specialized observationalists) must be viewed as essential, not merely as a labor cost to be phased out if measurements can be automated or delegated to field technicians. Unfortunately, such considerations do not translate readily into ways of prioritizing competing campaign proposals or sampling strategies, except to help breadth hold its own against narrow specificity. As grateful beneficiaries of such field campaign experiences, we will continue to try to extract gold from them, respecting alchemy’s lesson about quality that “one must start with gold” (Ooyama 1987, p. 2501). Science’s eternal challenge of synthesizing observations and models into knowledge also requires new thinking at the level of inference methodologies (e.g., Pearl and Mackenzie 2018), not merely additional samples. Closer contact with data assimilation and modeling will be crucial to keeping field observations relevant to the scientific project of inferring the interplay of causes and effects, from data about a complicated world we can only imperfectly simulate and predict.

Acknowledgments. This article is based on work supported by U.S. NOAA Grant NA13OAR4310156 and NASA Grant NNX15AD11G. We are grateful for the efforts of countless people behind the success of the AMIE-DYNAMO campaign. We especially thank Dr. Paul Hein at the Rutledge research group at Colorado State University (the *Revelle* data) and Fiaz Ahmed and Aaron Funk in the Schumacher group at Texas A&M

University (the SMART-R data) for large initial data reduction efforts on the Doppler radar data used here. The manuscript was greatly improved by comments from David Raymond, Wojtek Grabowski, and three anonymous reviewers.

APPENDIX

Hindcasts with the Kelly et al. (2017) Model

As another step toward evaluating CCPM-based sensitivity estimates against observations, we performed hindcasts with the Kelly et al. (2017) model by initializing its state (u , v , T , q fields) with reanalysis states. Large-scale hindcast skill as a function of α is reported here for completeness. The linear regressions of Fig. 8 used these hindcasts.

Hindcasts of 30-day duration were launched from initial conditions once per day during November and December 2011. Total model integration time was thus 61×30 days. All initializations and skill evaluations utilize ERA-Interim (Dee et al. 2011), interpolated to the centers of the model's 10 equal pressure layers of 100 hPa each. Skill was assessed for total column precipitable water (PW) field in the time–longitude domain averaged over 15°S – 15°N , verified against ERA-Interim itself. Other fields such as 850-hPa zonal wind give a qualitatively similar picture (not shown).

Initial tests showed that model skill was limited by initialization and spinup shock after the cold start, owing to the model's mean-state bias relative to the reanalyses. While we have not diagnosed it closely, this bias seems to reflect a shortcoming of our calibration strategy (following Hall 2000) in devising the model's time-independent forcing, rather than a scientifically interesting nonlinear dynamics rectification of the anomaly coupled linearized convection matrix. At any rate, to combat this bias, we turned to *anomaly initialization*. Specifically, we initialized each hindcast by adding the daily *anomaly* fields from reanalysis to a mean state derived from a previous free integration of the matrix-coupled GCM for each of the three candidate matrices with $\alpha = \{0, 1, 2\}$.

Visual examination of longitude–time sections indicates that the model produces Kelvin waves, but a disappointing MJO, as noted in Kelly et al. (2017). Figure A1 shows the growth of root-mean-square errors (RMSEs) and decay of correlation coefficients r between model and verification series in the longitude domain, averaged over all 60 hindcast initialization times. Because the climatological moisture structure (humid over warm longitudes, dry over cold longitudes) is retained in these longitude-domain statistics, this averaged spatial correlation r asymptotically

FIG. A1. Hindcast skill, as measured by (top) the root-mean-square error and (bottom) the correlation coefficient, of precipitable water vapor for various values of sensitivity to free-tropospheric moisture: $\alpha = 0, 1, 2$.

approaches a finite value (about 0.8). The curves in Fig. A1 indicate that the model loses tropical weather skill in about 5 days. Values of the RMS and correlation measures asymptotically approach different values for different α because the model simulates greater variance with increasing α , but all of that variability is uncorrelated with nature after a few days. These results add depth to Kelly et al. (2017). The important point for section 5b is that, given merely 5-day skill in 30-day integrations sampled for the statistics, the model's free variability—mainly convectively coupled Kelvin waves, affected slightly by storage and advection of moisture (Kelly et al. 2017)—predominates in the statistics of the virtual field campaign evaluations of Fig. A1. While such variability does exist in nature (Kiladis et al. 2009), it is not entirely comparable to the broad spectrum of variability underlying the AMIE-DYNAMO field campaign time series.

REFERENCES

- Arakawa, A., 2004: The cumulus parameterization problem: Past, present, and future. *J. Climate*, **17**, 2493–2525, [https://doi.org/10.1175/1520-0442\(2004\)017<2493:RATCPP>2.0.CO;2](https://doi.org/10.1175/1520-0442(2004)017<2493:RATCPP>2.0.CO;2).
- Bellenger, H., K. Yoneyama, M. Katsumata, T. Nishizawa, K. Yasunaga, and R. Shirooka, 2015: Observation of moisture tendencies related to shallow convection. *J. Atmos. Sci.*, **72**, 641–659, <https://doi.org/10.1175/JAS-D-14-0042.1>.
- Charney, J. G., 1963: A note on large-scale motions in the tropics. *J. Atmos. Sci.*, **20**, 607–609, [https://doi.org/10.1175/1520-0469\(1963\)020<0607:ANOLSM>2.0.CO;2](https://doi.org/10.1175/1520-0469(1963)020<0607:ANOLSM>2.0.CO;2).
- Colin, M., S. Sherwood, O. Geoffroy, S. Bony, and D. Fuchs, 2019: Identifying the sources of convective memory in cloud-resolving simulations. *J. Atmos. Sci.*, **76**, 947–962, <https://doi.org/10.1175/JAS-D-18-0036.1>.

- Daleu, C. L., and Coauthors, 2016: Intercomparison of methods of coupling between convection and large-scale circulation: 2. Comparison over nonuniform surface conditions. *J. Adv. Model. Earth Syst.*, **8**, 387–405, <https://doi.org/10.1002/2015MS000570>.
- , R. S. Plant, and S. J. Woolnough, 2017: Using the weak-temperature gradient approximation to evaluate parameterizations: An example of the transition from suppressed to active convection. *J. Adv. Model. Earth Syst.*, **9**, 2350–2367, <https://doi.org/10.1002/2017MS000940>.
- Dee, D. P., and Coauthors, 2011: The ERA-Interim reanalysis: Configuration and performance of the data assimilation system. *Quart. J. Roy. Meteor. Soc.*, **137**, 553–597, <https://doi.org/10.1002/qj.828>.
- DePasquale, A., C. Schumacher, and A. Rapp, 2014: Radar observations of MJO and Kelvin wave interactions during DYNAMO/CINDY2011/AMIE. *J. Geophys. Res. Atmos.*, **119**, 6347–6367, <https://doi.org/10.1002/2013JD021031>.
- Ebert-Uphoff, I., and Y. Deng, 2017: Causal discovery in the geosciences—Using synthetic data to learn how to interpret results. *Comput. Geosci.*, **99**, 50–60, <https://doi.org/10.1016/j.cageo.2016.10.008>.
- Hagos, S., Z. Feng, C. D. Burleyson, K.-S. S. Lim, C. N. Long, D. Wu, and G. Thompson, 2014: Evaluation of convection-permitting model simulations of cloud populations associated with the Madden-Julian oscillation using data collected during the AMIE/DYNAMO field campaign. *J. Geophys. Res. Atmos.*, **119**, 12 052–12 068, <https://doi.org/10.1002/2014JD022143>.
- Hall, N. M. J., 2000: A simple GCM based on dry dynamics and constant forcing. *J. Atmos. Sci.*, **57**, 1557–1572, [https://doi.org/10.1175/1520-0469\(2000\)057<1557:ASGBOD>2.0.CO;2](https://doi.org/10.1175/1520-0469(2000)057<1557:ASGBOD>2.0.CO;2).
- Herman, M. J., and Z. Kuang, 2013: Linear response functions of two convective parameterization schemes. *J. Adv. Model. Earth Syst.*, **5**, 510–541, <https://doi.org/10.1002/jame.20037>.
- , and D. J. Raymond, 2014: WTG cloud modeling with spectral decomposition of heating. *J. Adv. Model. Earth Syst.*, **6**, 1121–1140, <https://doi.org/10.1002/2014MS000359>.
- Holloway, C. E., and J. D. Neelin, 2009: Moisture vertical structure, column water vapor, and tropical deep convection. *J. Atmos. Sci.*, **66**, 1665–1683, <https://doi.org/10.1175/2008JAS2806.1>.
- , and —, 2010: Temporal relations of column water vapor and tropical precipitation. *J. Atmos. Sci.*, **67**, 1091–1105, <https://doi.org/10.1175/2009JAS3284.1>.
- Houze, R. A., Jr., 1997: Stratiform precipitation in regions of convection: A meteorological paradox? *Bull. Amer. Meteor. Soc.*, **78**, 2179–2196, [https://doi.org/10.1175/1520-0477\(1997\)078<2179:SPIROC>2.0.CO;2](https://doi.org/10.1175/1520-0477(1997)078<2179:SPIROC>2.0.CO;2).
- Inoue, K., and L. E. Back, 2015: Column-integrated moist static energy analysis on various time scales during TOGA COARE. *J. Atmos. Sci.*, **72**, 1856–1871, <https://doi.org/10.1175/JAS-D-14-0249.1>.
- Johnson, R. H., and P. E. Ciesielski, 2013: Structure and properties of Madden-Julian oscillations deduced from DYNAMO sounding arrays. *J. Atmos. Sci.*, **70**, 3157–3179, <https://doi.org/10.1175/JAS-D-13-065.1>.
- , T. M. Rickenbach, S. A. Rutledge, P. E. Ciesielski, and W. H. Schubert, 1999: Trimodal characteristics of tropical convection. *J. Climate*, **12**, 2397–2418, [https://doi.org/10.1175/1520-0442\(1999\)012<2397:TCOTC>2.0.CO;2](https://doi.org/10.1175/1520-0442(1999)012<2397:TCOTC>2.0.CO;2).
- Kelly, P., B. Mapes, I.-K. Hu, S. Song, and Z. Kuang, 2017: Tangent linear superparameterization of convection in a 10 layer global atmosphere with calibrated climatology. *J. Adv. Model. Earth Syst.*, **9**, 932–948, <https://doi.org/10.1002/2016MS000871>.
- Kiladis, G. N., M. C. Wheeler, P. T. Haertel, K. H. Straub, and P. E. Roundy, 2009: Convectively coupled equatorial waves. *Rev. Geophys.*, **47**, RG2003, <https://doi.org/10.1029/2008RG000266>.
- Kim, J., C. Zhang, G. N. Kiladis, and P. Bechtold, 2018: Heating and moistening of the MJO during DYNAMO in ECMWF reforecasts. *J. Atmos. Sci.*, **75**, 1429–1452, <https://doi.org/10.1175/JAS-D-17-0170.1>.
- Kuang, Z., 2010: Linear response functions of a cumulus ensemble to temperature and moisture perturbations and implications for the dynamics of convectively coupled waves. *J. Atmos. Sci.*, **67**, 941–962, <https://doi.org/10.1175/2009JAS3260.1>.
- , 2012: Weakly forced mock-Walker cells. *J. Atmos. Sci.*, **69**, 2759–2786, <https://doi.org/10.1175/JAS-D-11-0307.1>.
- , 2018: Linear stability of moist convecting atmospheres. Part I: From linear response functions to a simple model and applications to convectively coupled waves. *J. Atmos. Sci.*, **75**, 2889–2907, <https://doi.org/10.1175/JAS-D-18-0092.1>.
- Mapes, B. E., and R. A. Houze Jr., 1995: Diabatic divergence profiles in western Pacific mesoscale convective systems. *J. Atmos. Sci.*, **52**, 1807–1828, [https://doi.org/10.1175/1520-0469\(1995\)052<1807:DDPIWP>2.0.CO;2](https://doi.org/10.1175/1520-0469(1995)052<1807:DDPIWP>2.0.CO;2).
- , and J. Lin, 2005: Doppler radar observations of mesoscale wind divergence in regions of tropical convection. *Mon. Wea. Rev.*, **133**, 1808–1824, <https://doi.org/10.1175/MWR2941.1>.
- , and A. S. Chandra, 2017: Retrievals of Doppler wind divergence profiles from SMART-R and TOGA radars during DYNAMO, version 1.0. UCAR/NCAR Earth Observing Laboratory, accessed 9 June 2017, <https://doi.org/10.5065/D6Q52N2P>.
- , P. E. Ciesielski, and R. H. Johnson, 2003: Sampling errors in rawinsonde-array budgets. *J. Atmos. Sci.*, **60**, 2697–2714, [https://doi.org/10.1175/1520-0469\(2003\)060<2697:SEIRB>2.0.CO;2](https://doi.org/10.1175/1520-0469(2003)060<2697:SEIRB>2.0.CO;2).
- , S. Tulich, J. Lin, and P. Zuidema, 2006: The mesoscale convection life cycle: Building block or prototype for large-scale tropical waves? *Dyn. Atmos. Oceans*, **42**, 3–29, <https://doi.org/10.1016/j.dynatmoce.2006.03.003>.
- , J. Bacmeister, M. Khairoutdinov, C. Hannay, and M. Zhao, 2009: Virtual field campaigns on deep tropical convection in climate models. *J. Climate*, **22**, 244–257, <https://doi.org/10.1175/2008JCLI2203.1>.
- , A. S. Chandra, Z. Kuang, and P. Zuidema, 2017: Importance profiles for water vapor. *Surv. Geophys.*, **38**, 1355–1369, <https://doi.org/10.1007/s10712-017-9427-1>.
- Ooyama, K., 1987: Scale-controlled objective analysis. *Mon. Wea. Rev.*, **115**, 2479–2506, [https://doi.org/10.1175/1520-0493\(1987\)115<2479:SCOA>2.0.CO;2](https://doi.org/10.1175/1520-0493(1987)115<2479:SCOA>2.0.CO;2).
- Parodi, A., and K. Emanuel, 2009: A theory for buoyancy and velocity scales in deep moist convection. *J. Atmos. Sci.*, **66**, 3449–3463, <https://doi.org/10.1175/2009JAS3103.1>.
- Pearl, J., and D. Mackenzie, 2018: *The Book of Why: The New Science of Cause and Effect*. Basic Books, 432 pp.
- Randall, D. A., 2013: Beyond deadlock. *Geophys. Res. Lett.*, **40**, 5970–5976, <https://doi.org/10.1002/2013GL057998>.
- Raymond, D. J., 1983: Wave-CISK in mass flux form. *J. Atmos. Sci.*, **40**, 2561–2574, [https://doi.org/10.1175/1520-0469\(1983\)040<2561:WCIMFF>2.0.CO;2](https://doi.org/10.1175/1520-0469(1983)040<2561:WCIMFF>2.0.CO;2).
- , and X. Zeng, 2005: Modelling tropical atmospheric convection in the context of the weak temperature gradient approximation. *Quart. J. Roy. Meteor. Soc.*, **131**, 1301–1320, <https://doi.org/10.1256/qj.03.97>.
- , and M. M. Flores, 2016: Predicting convective rainfall over tropical oceans from environmental conditions. *J. Adv. Model. Earth Syst.*, **8**, 703–718, <https://doi.org/10.1002/2015MS000595>.

- Robe, F. R., and K. Emanuel, 1996: Moist convective scaling: Some inferences from three-dimensional cloud ensemble simulations. *J. Atmos. Sci.*, **53**, 3265–3275, [https://doi.org/10.1175/1520-0469\(1996\)053<3265:MCSSIF>2.0.CO;2](https://doi.org/10.1175/1520-0469(1996)053<3265:MCSSIF>2.0.CO;2).
- Sardeshmukh, P. D., and P. Sura, 2007: Multiscale impacts of variable heating in climate. *J. Climate*, **20**, 5677–5695, <https://doi.org/10.1175/2007JCLI1411.1>.
- Sentić, S., S. L. Sessions, and Ž. Fuchs, 2015: Diagnosing DYNAMO convection with weak temperature gradient simulations. *J. Adv. Model. Earth Syst.*, **7**, 1849–1871, <https://doi.org/10.1002/2015MS000531>.
- Sessions, S. L., M. J. Herman, and S. Sentić, 2015: Convective response to changes in the thermodynamic environment in idealized weak temperature gradient simulations. *J. Adv. Model. Earth Syst.*, **7**, 712–738, <https://doi.org/10.1002/2015MS000446>.
- Sherwood, S. C., R. Roca, T. M. Weckwerth, and N. G. Andronova, 2010: Tropospheric water vapor, convection, and climate. *Rev. Geophys.*, **48**, RG2001, <https://doi.org/10.1029/2009RG000301>.
- Sobel, A. H., and C. S. Bretherton, 2000: Modeling tropical precipitation in a single column. *J. Climate*, **13**, 4378–4392, [https://doi.org/10.1175/1520-0442\(2000\)013<4378:MTPIAS>2.0.CO;2](https://doi.org/10.1175/1520-0442(2000)013<4378:MTPIAS>2.0.CO;2).
- Song, S. W., 2015: Analyzing characteristics of convection and the relationship with its environment. Ph.D. dissertation, University of Miami, 152 pp., https://scholarlyrepository.miami.edu/oa_dissertations/1548/.
- Takemi, T., 2015: Relationship between cumulus activity and environmental moisture during the CINDY2011/DYNAMO field experiment as revealed from convection-resolving simulations. *J. Meteor. Soc. Japan*, **93A**, 41–58, <https://doi.org/10.2151/jmsj.2015-035>.
- Tulich, S. N., and B. E. Mapes, 2010: Transient environmental sensitivities of explicitly simulated tropical convection. *J. Atmos. Sci.*, **67**, 923–940, <https://doi.org/10.1175/2009JAS3277.1>.
- Varble, A., and Coauthors, 2014: Evaluation of cloud-resolving and limited area model intercomparison simulations using TWP-ICE observations: 1. Deep convective updraft properties. *J. Geophys. Res. Atmos.*, **119**, 13 891–13 918, <https://doi.org/10.1002/2013JD021371>.
- Wang, S., A. H. Sobel, and Z. Kuang, 2013: Cloud-resolving simulation of TOGA-COARE using parameterized large-scale dynamics. *J. Geophys. Res. Atmos.*, **118**, 6290–6301, <https://doi.org/10.1002/jgrd.50510>.
- , —, and J. Nie, 2016: Modeling the MJO in a cloud-resolving model with parameterized large-scale dynamics: Vertical structure, radiation, and horizontal advection of dry air. *J. Adv. Model. Earth Syst.*, **8**, 121–139, <https://doi.org/10.1002/2015MS000529>.
- Xu, W., S. A. Rutledge, C. Schumacher, and M. Katsumata, 2015: Evolution, properties, and spatial variability of MJO convection near and off the equator during DYNAMO. *J. Atmos. Sci.*, **72**, 4126–4147, <https://doi.org/10.1175/JAS-D-15-0032.1>.
- Yanai, M., S. Esbensen, and J. Chu, 1973: Determination of bulk properties of tropical cloud clusters from large-scale heat and moisture budgets. *J. Atmos. Sci.*, **30**, 611–627, [https://doi.org/10.1175/1520-0469\(1973\)030<0611:DOBPOT>2.0.CO;2](https://doi.org/10.1175/1520-0469(1973)030<0611:DOBPOT>2.0.CO;2).
- Yasunaga, K., S. Yokoi, K. Inoue, and B. E. Mapes, 2019: Space–time spectral analysis of the moist static energy budget equation. *J. Climate*, **32**, 501–529, <https://doi.org/10.1175/JCLI-D-18-0334.1>.
- Yoneyama, K., C. Zhang, and C. N. Long, 2013: Tracking pulses of the Madden–Julian oscillation. *Bull. Amer. Meteor. Soc.*, **94**, 1871–1891, <https://doi.org/10.1175/BAMS-D-12-00157.1>.
- Zhang, J., P. Zuidema, D. D. Turner, and M. P. Cadetdu, 2018: Surface-based microwave humidity retrievals over the equatorial Indian Ocean: Applications and challenges. *J. Appl. Meteor. Climatol.*, **57**, 1765–1782, <https://doi.org/10.1175/JAMC-D-17-0301.1>.