

# Application of Multivariate Sensitivity Analysis Techniques to AGCM-Simulated Tropical Cyclones

FEI HE

*University of California, Los Angeles, Los Angeles, California*

DEREK J. POSSELT

*Jet Propulsion Laboratory, California Institute of Technology, Pasadena, California*

NAVEEN N. NARISSETTY

*Department of Statistics, University of Illinois at Urbana–Champaign, Urbana, Illinois*

COLIN M. ZARZYCKI

*National Center for Atmospheric Research, Boulder, Colorado*

VIJAYAN N. NAIR

*Department of Statistics, University of Michigan, Ann Arbor, Michigan*

(Manuscript received 19 September 2017, in final form 23 April 2018)

## ABSTRACT

This work demonstrates the use of Sobol's sensitivity analysis framework to examine multivariate input–output relationships in dynamical systems. The methodology allows simultaneous exploration of the effect of changes in multiple inputs, and accommodates nonlinear interaction effects among parameters in a computationally affordable way. The concept is illustrated via computation of the sensitivities of atmospheric general circulation model (AGCM)-simulated tropical cyclones to changes in model initial conditions. Specifically, Sobol's variance-based sensitivity analysis is used to examine the response of cyclone intensity, cloud radiative forcing, cloud content, and precipitation rate to changes in initial conditions in an idealized AGCM-simulated tropical cyclone (TC). Control factors of interest include the following: initial vortex size and intensity, environmental sea surface temperature, vertical lapse rate, and midlevel relative humidity. The sensitivity analysis demonstrates systematic increases in TC intensity with increasing sea surface temperature and atmospheric temperature lapse rates, consistent with many previous studies. However, there are nonlinear interactions among control factors that affect the response of the precipitation rate, cloud content, and radiative forcing. In addition, sensitivities to control factors differ significantly when the model is run at different resolution, and coarse-resolution simulations are unable to produce a realistic TC. The results demonstrate the effectiveness of a quantitative sensitivity analysis framework for the exploration of dynamic system responses to perturbations, and have implications for the generation of ensembles.

## 1. Introduction

Understanding the response of weather extremes (e.g., tropical cyclone, extratropical cyclone, tornado, flood, etc.) to changes in multiple factors (e.g., atmospheric instability, sea surface temperature, atmospheric

moisture, etc.) is one of the most challenging problems in the analysis of Earth's weather and climate. Analysis of one factor at a time provides, at best, limited information on system sensitivity, as interactions among multiple factors may mitigate or amplify the response. Statistical quantitative sensitivity analysis algorithms have become powerful tools for the multivariate analysis of numerical models in uncertainty quantification (Hegstad and Henning 2001; O'Hagan et al. 1999;

---

*Corresponding author:* Derek J. Posselt, derek.posselt@jpl.nasa.gov

DOI: 10.1175/MWR-D-17-0265.1

© 2018 American Meteorological Society. For information regarding reuse of this content and general copyright information, consult the [AMS Copyright Policy](https://www.ametsoc.org/PUBSReuseLicenses) ([www.ametsoc.org/PUBSReuseLicenses](https://www.ametsoc.org/PUBSReuseLicenses)).

Murphy et al. 2004; Posselt and Vukicevic 2010; Posselt 2016; Zarzycki and Ullrich 2017), sensitivity analysis (Sobol 2005; Tushaus et al. 2015; Posselt 2016), data assimilation (Frolov et al. 2009), parameter estimation (Strounine et al. 2010; Posselt and Bishop 2012; Posselt et al. 2014), model calibration (Kennedy and O'Hagan 2001; Higdon et al. 2004; Habib et al. 2007; Sanso and Forest 2009), and prediction (Poole and Raftery 2000). In addition, the adjoint of a numerical model (Errico 1997), which computes linear sensitivity of the output of a numerical model to changes in one or more inputs, can be effectively used to assess sensitivity (e.g., Daescu and Todling 2010; Doyle et al. 2012, 2014; Reynolds et al. 2016). Ensemble forecasting is now the norm at most operational weather and climate prediction centers, and several studies have demonstrated the effectiveness of ensembles for sensitivity analysis as well (e.g., Ancell and Hakim 2007; Tong and Xue 2008; Torn and Hakim 2008; Boyle et al. 2015; McLay and Liu 2014; Hacker and Lei 2015; Hill et al. 2016).

Many modern sensitivity analysis studies combine random or quasi-random sampling of model inputs with a statistical approximation of the full model response (e.g., statistical emulators, response surface methodologies, or surrogate models). Emulation of the model is often necessary in cases for which 1) the number of variable parameters is large, 2) the model is computationally expensive, or both. With input settings selected using Monte Carlo random sampling or Latin hypercube sampling (LHS; McKay et al. 2000), statistical emulators can mimic the input–output relationships with a highly reduced computational burden, and thus enable deep exploration of model sensitivity (Živković et al. 1995; Helton and Davis 2003; Zhao et al. 2013; Iooss and Lemaitre 2015). Applications of these methods have spanned a wide range of Earth system analyses. For example, van der Merwe et al. (2007) used a neural network surrogate model to successfully emulate behaviors of the Columbia River and near-ocean regions. Habib et al. (2007) used a Gaussian process-based emulator for cosmic background radiation calibration. Sanso and Forest (2009) used statistical emulation to approximate climate models and calibrate climate system properties. Logemann et al. (2004) used simple regression models to represent the northeast Atlantic circulation. Lee et al. (2011) used a Gaussian process methodology to emulate the behavior of a complex global aerosol model, and to quantify sensitivity to uncertain parameters. Other studies that have used statistical emulators to approximate the behavior of complex numerical models include, but are not limited to the following: Bowman et al. (1993), Higdon et al. (2004), Williams et al. (2006), Sanso et al. (2008),

Tokmakian et al. (2012), and Lee et al. (2013). The choice of emulator depends on the complexity of the model parameter – response relationship. In cases for which responses are known to be approximately linear, a regression model may be sufficient. In cases with nonlinear relationships, more complex emulators are required.

The application of statistical emulators and quantitative sensitivity analysis methods to complex computational models has the following advantages: 1) the experiment design allows identification of the main effects of individual parameters, as well as the interaction effects among multiple parameters; 2) the statistical emulator allows a far more complete analysis of the model response than is possible with the numerical model itself, as emulators can produce thousands of surrogate simulations in a short period of time; and 3) the sensitivity analysis framework produces quantitative information on the relative importance of individual parameters and the interactions among them in a computationally affordable manner.

While statistical sensitivity analysis techniques have been applied to a wide range of topics, they have not yet been widely used to explore the behavior of simulated dynamic weather systems. Atmospheric dynamical systems, such as tropical cyclones, extratropical cyclones, mesoscale convective systems, etc., are the primary causes of weather-related economic damage and loss of life (Ranson et al. 2014). The growth and decay of these systems are strongly influenced by atmospheric dynamic and thermodynamic conditions, such as preexisting atmospheric disturbances, sea surface temperature, land surface vegetation, temperature, and soil moisture, atmospheric stability, relative humidity, etc. The sensitivity of dynamic systems to changes in their context, along with the interaction among scales and processes, has long been of interest.

While regression and other statistical techniques have been used to infer the sensitivity of tropical cyclones to changes in their environment (e.g., Saunders and Lea 2008; Villarini et al. 2010; Robbins et al. 2011), numerical models with varying levels of complexity are the primary tools used to understand the influence of environmental factors on weather systems. Sensitivity analysis is essentially a process of characterizing two types of relationships: 1) model input–output relationships (controls on model responses), and 2) model output–output relationships (correlations among responses). In this study, we use a response surface methodology to illustrate how proven statistical techniques may be used to examine dynamical systems. Specifically, we examine the sensitivity of tropical cyclone dynamics, precipitation, and cloud radiative effects to changes in cyclone

initial conditions and environment, using an idealized tropical storm in an atmospheric general circulation model (AGCM). This idealized tropical cyclone (TC) test case was designed specifically for the purpose of model evaluation (Reed and Jablonowski 2011a), and produces a single TC on an aquaplanet with constant sea surface temperature. It has been used to study uncertainty in AGCM-simulated TCs (Reed and Jablonowski 2011c), and the impact of modifications to the model physical parameterizations (Reed and Jablonowski 2011b; He and Posselt 2015). The idealized framework produces TC evolution consistent with TC representation in a full AGCM, but with highly reduced computational cost, freely tunable initial conditions, and easy-to-extract TC features. In addition, simulation of a single TC facilitates a robust evaluation of the perturbation–response relationship. The TC is initialized in gradient wind balance in a horizontally homogeneous environment, and the initial storm size and intensity, along with the complete description of the environment, can be specified using only five parameters. The relatively small dimensionality of the input parameter space allows us to conduct a thorough test of the efficacy of a sensitivity analysis framework for the analysis of weather systems. The well-known sensitivity of TCs to changes in their environment (e.g., Gray 1968; McCaul 1991; Emanuel et al. 2004; Emanuel 2007; Zeng et al. 2007; Davis and Ahijevych 2012; Manganello et al. 2012; Murakami et al. 2012; Camargo 2013; Emanuel 2013; Knutson et al. 2013) provides a rich context against which we may evaluate the performance of the sensitivity analysis algorithm.

In our experiments, perturbations to tropical cyclone initial conditions are created using an LHS method (Mckay et al. 2000; Iman and Conover 1980; Loh 1996; Isukapalli and Georgopoulos 2001; Hossain et al. 2006), then an ensemble of simulations are generated using these perturbed initial conditions. The input–output responses in the ensemble are used to train a Gaussian process emulator (Williams et al. 2006), which produces an approximation of the multivariate model response. The emulator is then used to examine the univariate and multivariate (“main” and “interaction”) effects of changes in TC initial conditions on TC characteristics using Sobol’s variance-based sensitivity analysis methodology. The goals of this study are twofold: first, we evaluate the effectiveness of a sensitivity analysis technique for the study of dynamical systems, and second, we explore the implications of the results for the generation of ensembles of TC simulations. We note that, while the idealized TC simulations produced by the model are considered to be realistic, the idealized nature of our model framework limits the applicability of our results

to real-world systems. That said, the model physical parameterizations and dynamical core are consistent with those used in the NCAR CAM model, and as such, our results are applicable to the interpretation of TC sensitivity in an AGCM. We also note that, while vertical wind shear is known to be an important factor in tropical cyclone development (Frank and Ritchie 1999, 2001; He et al. 2015), we do not include it as a perturbation factor because it is coupled to the temperature lapse rate (through thermal wind balance) and as such cannot be freely modified.

The remainder of this paper is organized as follows. Section 2 introduces a schematic summary of the proposed idea. Section 3 describes the hurricane simulation using CAM, the algorithm used to extract physical fields from each tropical cyclone, the LHS approach, the input parameters and their ranges, and the Gaussian process emulator. Section 4 diagnoses the interrelationship among the simulated TC characteristics, and analyzes the sensitivities of the final simulated characteristics of TCs to initial conditions. Section 5 compares the sensitivity of simulated TC intensity to the five initial factors in three typically used model resolutions. Section 6 provides a discussion of the results and their limitations. Section 7 summarizes the main conclusions.

## 2. Schematic summary

Essentially, a multivariate sensitivity analysis framework can be considered to be a design for conducting numerical experiments, so as to retrieve detailed information on a system of interest. For a given dynamical system, the user first selects a numerical model that can be used to realistically simulate it. Then the user chooses the model input factors of interest and the key characteristics of the dynamical system as model output. Once the inputs and outputs are defined, the user then specifies the range of input factors. Figure 1 provides a schematic summary of the methodology. The first step is to apply the Latin hypercube sampling method to model input to obtain  $N$  samples. In our previous experiments, we found that  $N$  should be set larger than  $50m$ , where  $m$  is the number of input control variables. Setting  $N \geq 50$  provides sufficient number of samples to construct a robust emulator. The numerical model is then run  $N$  times with the  $N$  sets of model input variables, producing  $N$  model output vectors.

Let  $\mathbf{x} = (x_1, x_2, \dots, x_m)$  denote the  $m$  model inputs. These could be sea surface temperature, relative humidity, wind speed, translational speed, atmospheric stability, etc.; any output factor of interest in the targeted dynamical system. Then, let  $\mathbf{y} = (y_1, y_2, \dots, y_k)$  denote the length  $k$  vector of characteristics of the

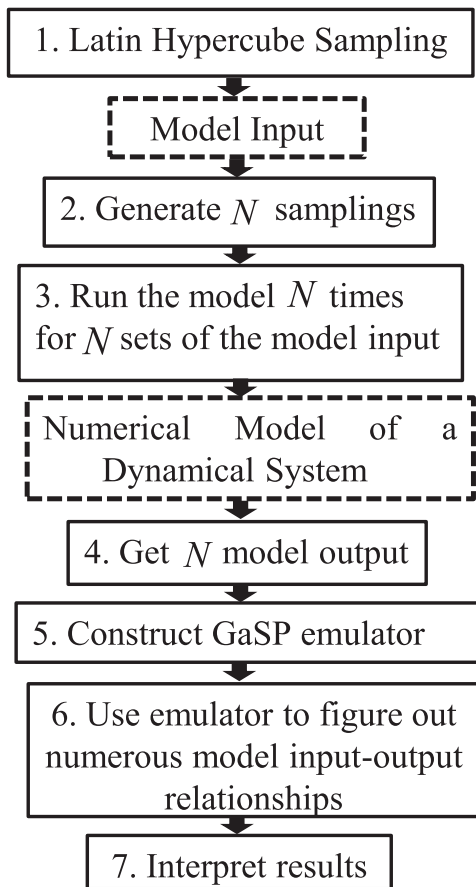


FIG. 1. The schematic flowchart of implementing the multivariate sensitivity analysis techniques on dynamical system in a step-by-step way.

dynamical system as model output. These could be storm size and/or intensity, precipitation, location, and so on. To illustrate this idea, we imagine a case with multiple input parameters and one model output variable ( $k = 1$ ). Recall that, implementing these steps in practice requires us to first identify the range of  $\mathbf{x}$ , then use LHS to generate  $N$  samples of  $\mathbf{x}$ , which means  $\mathbf{x}$  becomes a matrix

$$\mathbf{X} = \begin{pmatrix} x_{11} & x_{21} & x_{31} & \cdots & x_{m1} \\ x_{12} & x_{22} & x_{32} & \cdots & x_{m2} \\ x_{13} & x_{23} & x_{33} & \cdots & x_{m3} \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ x_{1N} & x_{2N} & x_{3N} & \cdots & x_{mN} \end{pmatrix}.$$

If we set the model input to the vector  $(x_{11}, x_{21}, x_{31}, \dots, x_{m1})$ , and run the model, we will obtain the corresponding model output  $y_1$ . Each row of matrix  $\mathbf{X}$  comprises a set of inputs to the model, and hence the  $N$  rows of matrix  $\mathbf{X}$  correspond to  $N$  runs of the numerical

model, yielding a model output vector  $\mathbf{Y}$ , which is an  $N \times 1$  column vector

$$\mathbf{Y} = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ \vdots \\ \vdots \\ y_N \end{pmatrix}.$$

The next step is to construct a statistical emulator using the matrix  $\mathbf{X}$  and (row) vector  $\mathbf{Y}$ . The statistical emulator is also referred to as a surrogate model, which in this study is a Gaussian spatial process model (GaSP; Sacks et al. 1989). The emulator is designed to fill the void between discrete values of model output and input via curve fitting or interpolation, and is used to establish a functional relationship between the output variable  $\mathbf{y}$  and input vector  $\mathbf{x}$ . That is, we use samples in matrix  $\mathbf{X}$  and the output vector  $\mathbf{Y}$  to construct an approximate model:  $\hat{y} = \hat{f}(x_1, x_2, \dots, x_m)$ . While the ensemble produced by Latin hypercube sampling can be used to obtain an initial estimate of the model sensitivity, analysis of the multivariate response, including the interactions among various input parameters, requires orders of magnitude more simulations than are feasible for most computational models. The statistical emulator is used alongside the more limited ensemble of model simulations to produce a large number of input–response pairs. These can then be used in variance-based sensitivity analysis to examine univariate and multivariate sensitivities in detail.

Having defined a functional relationship that emulates the model input–output response, one can input other values of  $(x_1^{\text{new}}, x_2^{\text{new}}, \dots, x_m^{\text{new}})$  to the emulator and obtain corresponding model output variables  $(\hat{y}^{\text{new}})$ . As the statistical emulator is much more efficient than the full numerical model, it allows generation of numerous input–output pairs in a computationally affordable fashion. The emulator facilitates generation of the thousands of simulations necessary for the computation of Sobol' sensitivity indices, which allow us to attribute the sensitivity of the output to the inputs, both in terms of the individual (main) effects and interactions between any given pair of inputs. The following section provides information on how this general framework is applied to the analysis of sensitivity in an idealized TC simulation. For a case involving multiple model output variables  $\mathbf{y}$ , one simply builds the emulator multiple times, once for each output variable. In our TC example, there are six model output variables [see

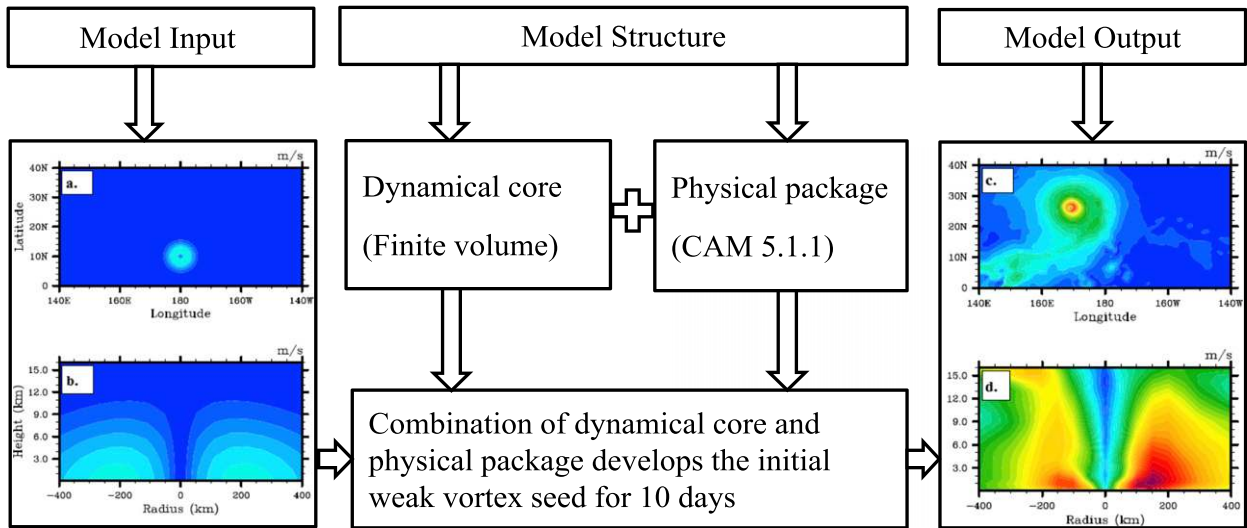


FIG. 2. The flowchart of idealized tropical cyclone simulation in the NCAR Community Atmosphere Model. (a)–(d) The snapshots shows the evolution of tropical cyclone–like vortex at days (a),(b) 0 and (c),(d) 10 at the resolution  $0.5^\circ$ , L30. (a),(c) Horizontal cross section of the wind speed at a height of 100 m. (b),(d) Longitude–height cross section of the wind speed through the center of the vortex. [For details, the reader is referred to Fig. 7 in Reed and Jablonowski (2011a).]

section 3a(1) for details], and hence we need to build six functional relationships:  $\hat{y}_1 = \hat{f}_1(x_1, x_2, \dots, x_m)$ ,  $\hat{y}_2 = \hat{f}_2(x_1, x_2, \dots, x_m)$ ,  $\hat{y}_3 = \hat{f}_3(x_1, x_2, \dots, x_m)$ ,  $\hat{y}_4 = \hat{f}_4(x_1, x_2, \dots, x_m)$ ,  $\hat{y}_5 = \hat{f}_5(x_1, x_2, \dots, x_m)$ , and  $\hat{y}_6 = \hat{f}_6(x_1, x_2, \dots, x_m)$ .

### 3. Data and methodology

#### a. Idealized tropical cyclone simulation

##### 1) MODEL CONFIGURATION

In this study, we use the National Center for Atmospheric Research (NCAR)–Department of Energy (DOE) Community Atmosphere Model (CAM), version 5.1.1 (Neale et al. 2010), with horizontal resolutions of  $1.0^\circ \times 1.0^\circ$ ,  $0.5^\circ \times 0.5^\circ$ , and  $0.25^\circ \times 0.25^\circ$ . A  $1.0^\circ$ ,  $0.5^\circ$ , and  $0.25^\circ$  latitudinal and longitudinal horizontal grid corresponds to a grid spacing of approximately 110, 55, and 28 km, respectively, in the equatorial region. The coarser 1.0 resolution is chosen so as to match the horizontal resolutions commonly used in phase 5 of the Coupled Model Intercomparison Project (CMIP5; Taylor et al. 2012) archive. Many high-resolution studies specifically investigating tropical cyclones in climate models use  $0.5^\circ$  or  $0.25^\circ$  resolution (Zhao et al. 2009; Bacmeister et al. 2014; Zarzycki and Jablonowski 2014; Wehner et al. 2014).

Figure 2 shows the flowchart of the tropical cyclone simulation. The model is provided an idealized initial condition, which consists of a three-dimensional (3D)

weak vortex seed overlaid on horizontally homogeneous large-scale tropical environmental conditions at  $10^\circ\text{N}$  latitude. The background wind is set equal to zero, the surface pressure is set to 1015 hPa, and the surface relative humidity is set to 83.6%. The surface air temperature is set equal to the sea surface temperature, and the relative humidity linearly decreases from 80% to a user-specified value at 500 hPa, above which it remains constant. Based on the specified surface pressure, air temperature, relative humidity profile, and lapse rate, the environmental temperature and pressure are set according to hydrostatic balance. The specific humidity profile is determined from the relative humidity profile and the temperature. Given a radius of maximum wind and maximum wind speed at the surface, the temperature and wind field of the initial vortex seed are derived from hydrostatic and gradient wind balance. The model surface is covered with a global ocean, and SST is set constant everywhere (aquaplanet setup). After initialization, the model is integrated for 10 days, and produces a TC that ranges from weak and disorganized to a strong and well-developed storm (Reed and Jablonowski 2011b,c, 2012). For a detailed description of the idealized tropical cyclone simulation set up in CAM, the reader is referred to Reed and Jablonowski (2011a). This model configuration is constructed deliberately to favor TC development. All of the necessary ingredients are in place: organized vortex seed, high SST, no shear, sufficient planetary vorticity, and favorable lapse rate. As we will show, this leads to the production of, on average, stronger TCs than are typically



TABLE 1. List of names for the TC-relevant output variables used in this paper (vortex center is defined by the minimum surface pressure).

Variables	Description	Units
Intensity (MWS/Vmax)	Max wind speed of 100-m height above surface	$\text{m s}^{-1}$
PRECT	Total precipitation rate (large-scale precipitation rate + convective precipitation rate)	$\text{mm h}^{-1}$
SWCF	Shortwave cloud radiative forcing	$\text{W m}^{-2}$
LWCF	Longwave cloud radiative forcing	$\text{W m}^{-2}$
LWP	Cloud liquid water path	$\text{g m}^{-2}$
IWP	Cloud ice water path	$\text{g m}^{-2}$

observed in the real world (particularly at higher resolutions); however, the idealized environment is useful in that it allows a detailed examination of TC development that is unclouded by the influence of external sources of variability.

## 2) ALGORITHM USED TO EXTRACT MEAN OUTPUT FIELDS FROM SIMULATED TROPICAL CYCLONES

The sensitivities of six TC-relevant output variables to the initial conditions are investigated (Table 1). They are the following: maximum wind speed (MWS; intensity; Vmax), total precipitation rate (PRECT), shortwave cloud radiative forcing (SWCF), longwave cloud radiative forcing (LWCF), cloud ice water path (IWP), and cloud liquid water path (LWP). The five physical variables are spatially averaged over the region surrounding the center of the tropical cyclone using the following procedure: 1) find the minimum surface pressure, which is defined to be the storm center; 2) search the tropical cyclone in a latitude/longitude region within  $\pm 20^\circ$  of the storm center; and 3) select the points that satisfy each of the following criteria: (i) wind speed at 100-m height above surface greater than  $6 \text{ m s}^{-1}$ , and (ii) outgoing longwave radiation (OLR) less than  $230 \text{ W m}^{-2}$ . Figure 3 shows an example of the shortwave cloud radiative forcing before and after applying the TC identification algorithm. The purpose of the algorithm is to filter out features that are not directly associated with tropical cyclone development, while preserving the TC-relevant fields. The first two thresholds locate the TC and its surrounding region, circumscribed in a square subdomain. The third threshold refines the specific shape of the TC, and ensures that the resulting fields accurately represent the processes directly involved in TC development. Tests of other slightly higher or lower threshold values of wind speed and OLR to isolate the TC did not alter our conclusions. After selecting those points that compose the TC, an area-weighted average of each variable calculation is computed.

## b. Multivariate sensitivity analysis framework

Multivariate sensitivity analysis methods comprise a set of techniques now widely used to examine uncertainties in numerical models associated with initial conditions, model physics parameters, and model construction (Stensrud et al. 2000; Murphy et al. 2004; van den Heever and Cotton 2004; Posselt and Vukicevic 2010; Collins et al. 2011; Berner et al. 2011; Posselt et al. 2016; Posselt 2016). In this paper, it specifically includes a combination of LHS with a surface response emulator and a variance-based sensitivity analysis technique. LHS is used to sample the points for the initial parameters, a Gaussian process emulator is used to approximate the model response surface, and Sobol's variance-based sensitivity analysis algorithm is used to quantify main and interaction effects. In this context, the "main effect" denotes the influence of variations in a single parameter on each model response variable (holding the other parameters fixed), while the "interaction effect" describes the response to changes in a variable when it is changed in concert with other control parameters.

### 1) LATIN HYPERCUBE SAMPLING

LHS methods generate a space-filling ensemble of samples of the model input vector  $\mathbf{x}$ . The input parameter range is divided into equiprobable nonoverlapping intervals, and LHS requires that every row and column in the hypercube of partitions has exactly one sample. This restriction to one sample per row-column combination in the hypercube yields more uniform coverage of the range of input variables, relative to traditional Monte Carlo sampling. By filling the input space in this manner, LHS seeks to accurately span the range of the various input-output relationships using a limited number of sample points. In our experiments, we used the R package "lhs" and function "randomLHS" to generate samples in the hypercube. For more information on the method, we refer to the work of Mckay et al. (2000) and Helton and Davis (2003). We choose five input parameters for study that are both tunable, and have known relevance to the development of tropical cyclones, and we divide their ranges into 300 segments (Table 2). Ranges for each parameter are set based on a survey of the TC literature, and some are adjusted to ensure the model does not crash. While we perturb the 500-hPa relative humidity, the surface relative humidity is set to the same constant 83.6% for each individual simulation. The LHS sampling is typically followed by the use of a response surface emulator to fill the space between quasi-randomly selected initial conditions and their corresponding output vectors, for example,

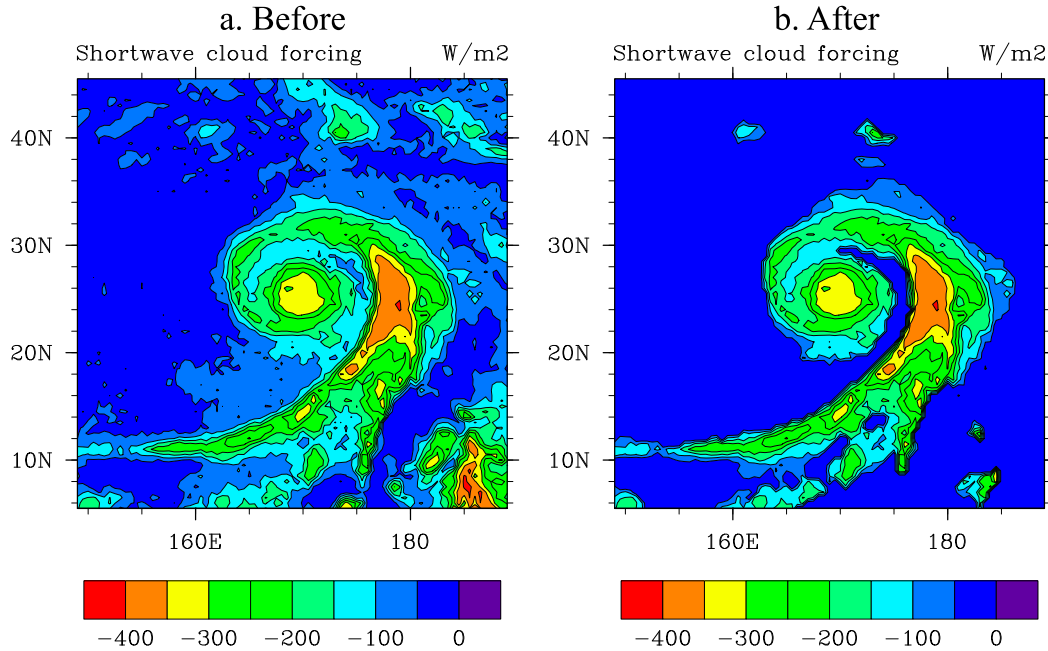


FIG. 3. Example of the shortwave cloud radiative forcing (a) before and (b) after applying the TC identification algorithm.

multivariate adaptive regression splines (Friedman 1991; Friedman and Roosen 1995), GaSP (Sacks et al. 1989), polynomial chaos, etc. In our paper, we use the GaSP emulator for fitting the input–output relationship.

### 2) GAUSSIAN SPATIAL PROCESS EMULATOR

Numerical models (e.g., CAM in this study) can be very computationally demanding to run and generation of hundreds of thousands of simulations is generally not feasible. Statistical emulators approximate the behavior of a numerical model by replacing the full model with a simpler interpolation or curve-fitting algorithm. In GaSP, the output is a realization of a Gaussian stochastic process (Sacks et al. 1989). Let  $\mathbf{X}$  be the  $N \times m$  model input matrix,  $\mathbf{Y}$  be the  $N \times 1$  model output (column) vector,  $N$  be the sample size, and  $m$  be the number of input parameters. The  $i$ th input to a numerical simulation can be denoted as an  $m$ -dimensional vector,  $\mathbf{x}_i = (x_{i1}, \dots, x_{im})$ , while the corresponding output is denoted by  $y_i$ , respectively. The emulator built with a Gaussian process (GP) algorithm is expressed as

$$y_i = f(\mathbf{x}_i) = \mu + z(\mathbf{x}_i); \quad i = 1, \dots, N,$$

where  $\mu$  is the overall mean, and  $z(\mathbf{x}_i)$  is a GP with  $E[z(\mathbf{x}_i)] = 0$ ,  $\text{VAR}[z(\mathbf{x}_i)] = \sigma^2$ , and  $\text{Cov}[z(\mathbf{x}_i), z(\mathbf{x}_j)] = \sigma^2 R_{ij}$ .

Generally speaking,  $y(\mathbf{X})$  follows a multivariate normal distribution,  $N_N(\mathbf{1}_N, \mu, \Sigma)$  with  $\Sigma = \sigma^2 \mathbf{R}$ , where  $\mathbf{R}$  is a correlation matrix having elements  $R_{ij}$ , and  $\mathbf{1}_N$  is a

$N \times 1$  vector of all ones. The GP algorithm used in our paper follows Ranjan et al. (2011) and MacDonald et al. (2015), and the correlation matrix  $\mathbf{R}$  is given by

$$R_{ij} = \prod_{k=1}^m \exp\{-\theta_k |x_{ik} - x_{jk}|^2\}, \quad \text{for all } i, j,$$

where  $\boldsymbol{\theta} = (\theta_1, \dots, \theta_m) \in [0, \infty)^m$  is a vector of hyperparameters. With  $\mathbf{R}$  defined, we can estimate  $\mu$  and  $\sigma^2$  to be

$$\hat{\mu}(\boldsymbol{\theta}) = \left[ (\mathbf{1}_N)^T \mathbf{R}^{-1} \mathbf{1}_N \right]^{-1} \left[ (\mathbf{1}_N)^T \mathbf{R}^{-1} \mathbf{Y} \right] \quad \text{and} \quad \hat{\sigma}^2(\boldsymbol{\theta}) = \frac{[\mathbf{Y} - \mathbf{1}_N \hat{\mu}(\boldsymbol{\theta})]^T \mathbf{R}^{-1} [\mathbf{Y} - \mathbf{1}_N \hat{\mu}(\boldsymbol{\theta})]}{N}.$$

Using the correlation matrix  $\mathbf{R}$ , along with  $\hat{\mu}(\boldsymbol{\theta})$  and  $\hat{\sigma}^2(\boldsymbol{\theta})$ , the negative profile log-likelihood (also referred to as deviance) can be obtained as

TABLE 2. Short names, minimum and maximum values, and descriptions of the five input parameters in the ensemble simulation with  $0.5^\circ \times 0.5^\circ$  grid spacing.

Parameter	Min	Max	Description (units)
RMW	175	300	Radius of max wind speed (km)
MWS	12.5	25	Max wind speed ( $\text{m s}^{-1}$ )
GAMMA	5.5	7.5	Environmental temperature lapse rate ( $\text{K km}^{-1}$ )
SST	22.5	34.0	Sea surface temperature ( $^\circ\text{C}$ )
500-hPa RH	0.4	0.7	500-hPa relative humidity (unitless)

$$-2\log(L_\theta) \propto \log(|\mathbf{R}|) + N \log \{ [\mathbf{Y} - \mathbf{1}_N \hat{\boldsymbol{\mu}}(\boldsymbol{\theta})]^\top \mathbf{R}^{-1} [\mathbf{Y} - \mathbf{1}_N \hat{\boldsymbol{\mu}}(\boldsymbol{\theta})] \},$$

where  $|\mathbf{R}|$  denotes the determinant of  $\mathbf{R}$ , and  $L_\theta$  is the likelihood function of  $\boldsymbol{\theta}$  given observations  $\mathbf{Y}$ . The deviance is the metric for estimating the hyperparameters  $\boldsymbol{\theta}$  via the maximum likelihood approach.

In our study, we used the GaSP developed by MacDonald et al. (2015), and implemented using the R package: “GPfit.” GPfit is distributed under the General Public License and available from the Comprehensive R Archive Network. The parameters used in building the GPfit model are set to the recommended default values, which are further explained in the GPfit documentation. Tests with other parameter values yielded only slight variation in our results, as we explain below. Note that the choice of specific emulator is often made based on personal preference and familiarity with an algorithm. We chose GaSP because it produced a robust fit to the model output, has strong support in the literature, and is capable of representing nonlinear interactions among parameters without having to know in advance which nonlinear terms are important. We did conduct a comparison of several different emulation techniques (not shown). In addition to GaSP, we tested multiple linear regression, multivariate adaptive regression splines (MARS), random forest, and Bayesian additive regression trees (BART). All produced comparable results, but GaSP provided the smallest fitting errors.

### 3) VARIANCE-BASED SENSITIVITY ANALYSIS

Sobol’s (1993) variance-based sensitivity analysis has recently been used in atmospheric and climate science (Lee et al. 2013; Zhao et al. 2013; Marzban et al. 2014). It is capable of assessing the relative effect of each of a number of model input parameters, along with their interactions, on one or more model output variables. In it, the variance of  $\mathbf{y}$  is decomposed into several variance terms that are contributed by each input parameter  $x$  and their interactions (Sobol 1993; Homma and Saltelli 1996; Bolado-lavin and Badea 2008; Marzban 2013):

$$V(\mathbf{y}) = \sum_{i=1}^n V_i + \sum_{i < j} V_{ij} + \text{higher order.}$$

Here,  $n$  is the number of input parameters.

Traditionally, each variance term is proportional to the total variance and converted into an index referred as Sobol’s sensitivity index:

$$S_i = \frac{V_i}{V(\mathbf{y})}$$

$$S_{ij} = \frac{V_{ij}}{V(\mathbf{y})}$$

$$S_{T_i} = \sum_{k \in \text{all the sets of indices containing the index } i} S_k.$$

The terms  $S_i$ ,  $S_{ij}$ , and  $S_{T_i}$  are referred to as the first-order effect index, second-order effect index, and total effect index, respectively. For estimating these indices, we first compute  $V_i$  and  $V_{ij}$  using the definitions:  $V_i = \text{Var}(E[\mathbf{y}|\mathbf{x}_i])$  and  $V_{ij} = \text{Var}(E[\mathbf{y}|\mathbf{x}_i, \mathbf{x}_j]) - V_i - V_j$ . Then, we compute  $V(\mathbf{y})$  with  $V(\mathbf{y}) = \sum_{i=1}^n V_i + \sum_{i < j} V_{ij}$ . By doing so, we neglect the third- and higher-order effects that contribute to  $V(\mathbf{y})$ , as we are only interested in the first-order effect  $V_i$  and second-order effect  $V_{ij}$ . Finally, Sobol’s sensitivity indices are computed. The GaSP emulator described above is used to generate the inputs for computing  $V_i$  and  $V_{ij}$ .

## 4. Ensemble simulation results

### a. Output–output relationships among the TC-relevant characteristics

In this study, our main focus is the examination of input–output relationships, and the use of quantitative sensitivity analysis techniques to evaluate them. Prior to doing so, we exploit the opportunity afforded by our ensemble of 300  $0.5^\circ \times 0.5^\circ$  model simulations to explore the relationships among the output variables themselves. Specifically, we assess the degree to which our simulations reflect known interrelationships among intensity and other outcomes, and obtain an initial view of the response of the model to perturbations in the initial conditions. This analysis yields information on the correlation among the model output variables, and also provides useful context for the input–output sensitivity analysis. Briefly, we hypothesize that the simulations should adhere to known relationships among TC intensity, precipitation rate, and water path. For example, increased intensity should be associated with larger precipitation rate and larger values of liquid and ice water path as the storm processes greater amounts of water vapor. However, the relationships among TC intensity and the LWCF and SWCF are less clear, as the LWCF is fundamentally related to the relative amount of cirrus compared with deep clouds and low clouds. The SWCF is a function of cloud brightness, which is primarily related to cloud water content; however, the integrated SWCF will depend on the



distribution of highly reflective clouds versus less reflective (more diffuse) clouds.

1) INTERRELATIONSHIP BETWEEN INTENSITY AND THE OTHER FIVE TC-RELEVANT CHARACTERISTICS

Figure 4 displays the interrelationship between TC intensity and each of the other five physical variables considered in our study: precipitation rate (Fig. 4a), LWCF (Fig. 4b), SWCF (Fig. 4c), LWP (Fig. 4d), and IWP (Fig. 4e). It is clear that larger TC mean precipitation rate is associated with stronger TC intensity in the model, consistent with northwest Pacific observation data. LWCF exhibits a positive linear relationship with TC intensity (Fig. 4b), while TC SWCF bears a weak and nonmonotonic relationship with intensity (Fig. 4c). The LWP has a nonmonotonic relationship with intensity: as TC intensity increases from 30 to 50  $\text{ms}^{-1}$ , LWP decreases. Above TC intensity of 50  $\text{ms}^{-1}$ , LWP exhibits a slight increase with increasing intensity (Fig. 4d), while IWP has a positive linear relationship with TC intensity (Fig. 4e) throughout. The response in IWP is consistent with the notion that increased TC intensity, and by extension, larger vertical transport, should be associated with larger ice water path.

2) RELATIONSHIPS TO TC PRECIPITATION RATE

Figure 5 depicts the interrelationships among precipitation rate and the four other output variables, as well as the relationship between LWCF and SWCF. Larger mean precipitation rate tends to be correlated with larger mean TC LWCF, up to a mean precipitation rate of approximately 1.7  $\text{mm h}^{-1}$  (Fig. 5a). The decrease in LWCF above 1.7  $\text{mm h}^{-1}$  appears to be related to a decrease in IWP (Fig. 5d), and we hypothesize that the smaller LWCF may be due to a reduction in the amount of cirrus cloud in the model. SWCF exhibits a nonmonotonic relationship with precipitation rate, reaching a relative maximum when the mean TC precipitation rate is approximately 1.4  $\text{mm h}^{-1}$ . Above this value, larger mean precipitation rates are associated with smaller magnitudes in SWCF. The range of values of SWCF is large at intermediate precipitation rates, and it is not clear how to interpret the physics behind this apparent signal. The functional form appears to mirror the change in IWP with precipitation rate (Fig. 5d). The precipitation rate is strongly (and positively) correlated with LWP (Fig. 5c), indicating that the initiation mechanism of rain from droplet growth by collision and coalescence works efficiently in the AGCM physical parameterizations. TC LWCF tends to bear a good linear relationship with TC SWCF (Fig. 5e). Larger LWCF comes with larger magnitude of SWCF.

3) RELATIONSHIPS AMONG LWCF, SWCF, LWP, AND IWP

Figure 6 depicts the interrelationships among the cloud radiative forcing and the liquid and ice water path, as well as the relationship between LWP and IWP. LWCF appears to be largely independent of LWP (Fig. 6a), but has a positive linear relationship with IWP (Fig. 6b). This makes sense, given the control high clouds place on the longwave radiative budget. SWCF is linearly correlated with LWP (Fig. 6c), which is consistent with an increase in cloud brightness with increasing LWP. SWCF has a nonmonotonic relationship with IWP (Fig. 6d), but we note that the ranges of IWP for most values of SWCF are large, and this relationship is likely not robust, or perhaps is indicative of interactions with other parameters. While IWP is clearly related to LWCF, it has a somewhat ambiguous relationship with all other variables, including LWP (Fig. 6e). This leads to complexity in the response, which we will explore in more detail in the next section.

b. Input–output relationships

1) VALIDATION OF GASP

The normalized root-mean-square error (RMSE) is used to measure how well GaSP fits the CAM model in simulating the relationships between tropical cyclone characteristics and model initial conditions. Let  $\mathbf{y}$  denote the output from simulated CAM model and let  $\hat{\mathbf{y}}$  be the output calculated from the GaSP model. The normalized RMSE can then be defined as

$$\text{RMSE} = \sqrt{\frac{1}{n} \sum_i (\hat{y}_i - y_i)^2} / \left( \frac{1}{n} \sum_i y_i \right).$$

RMSE in this case is the proportion of mean square error to the mean value over the ensemble of CAM simulations. Specifically, we use cross-validation analysis to validate the model. Consider TC intensity (MWS) as an example. In each of several trials, we use 299 sampling points to construct the GaSP fit, and then predict the remaining one point. This procedure is commonly referred to as “leave-one-out cross validation” in statistics. We conducted 300 trials, in which we compared each of the 300 predicted values with their corresponding simulated values. The cross-validation analysis is conducted for all the six variables. Figure 7 shows the cross-validated mean square prediction error results. GaSP performs well for MWS, PRECT, LWCF, SWCF, and IWP (as indicated by the fact that the fitting errors are on the order of 10% or less), but returns a poorer fit for LWP. The reason that GaSP does not perform well

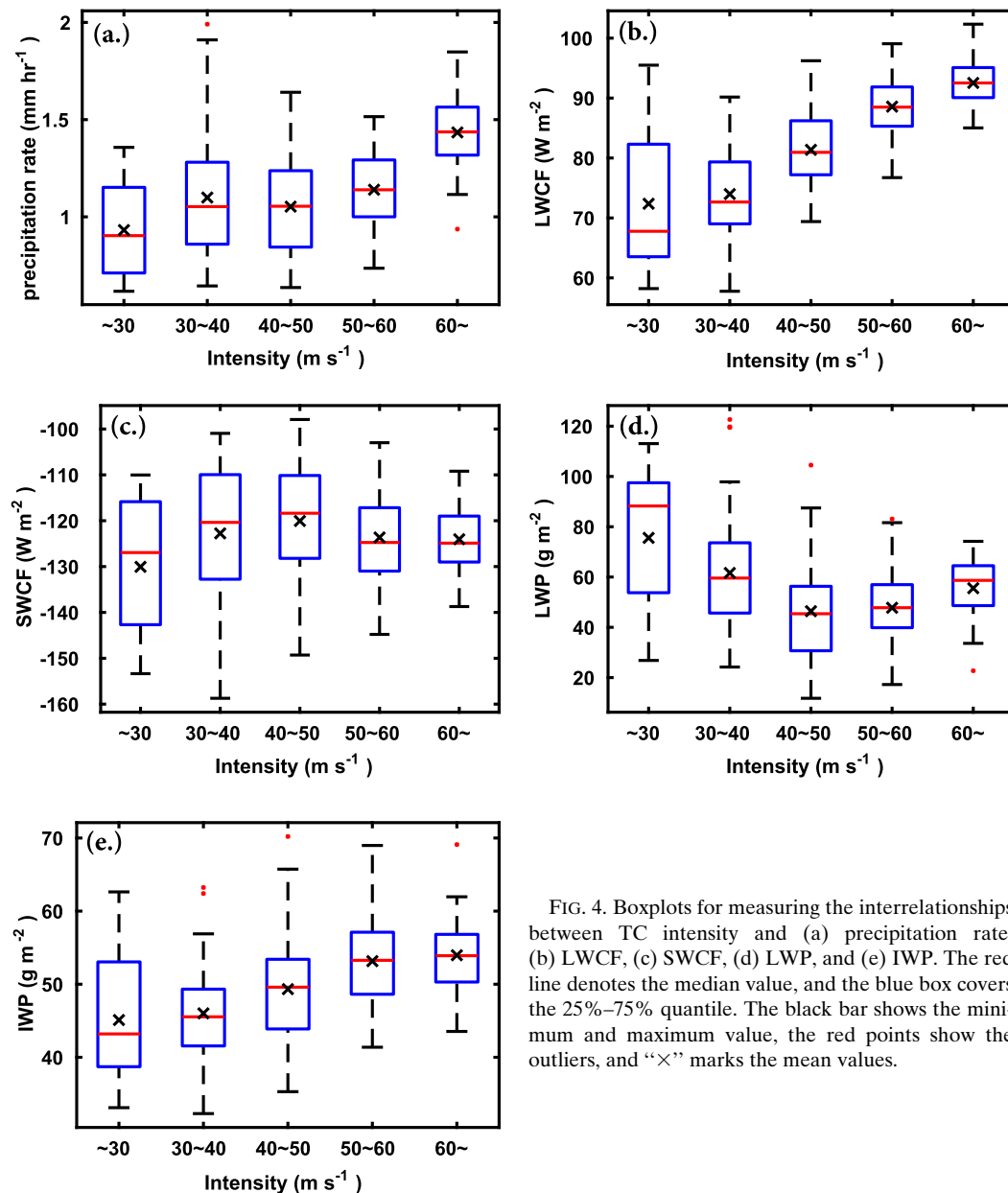


FIG. 4. Boxplots for measuring the interrelationships between TC intensity and (a) precipitation rate, (b) LWCf, (c) SWCF, (d) LWP, and (e) IWP. The red line denotes the median value, and the blue box covers the 25%–75% quartile. The black bar shows the minimum and maximum value, the red points show the outliers, and “×” marks the mean values.

for LWP may be that the relationship between LWP and the five initial conditions is much more complicated, or that there are outliers produced by extreme behavior in the model.

## 2) VARIANCE-BASED SENSITIVITY ANALYSIS

The Sobol total sensitivity index is used to quantify the relative importance of the five initial parameters to the six TC characteristics. The larger the sensitivity index is, the more important the parameter is to the relevant TC characteristic. From each subplot in Fig. 8, it is

easy to distinguish the dominant factors. Figure 8a shows that simulated TC intensity is most sensitive to the lapse rate, followed by the SST. The initial size (RMW) and intensity (MWS) of the vortex seed and RH 500 have only a minor influence on TC intensity. Precipitation rate (Fig. 8b) is most sensitive to SST while LWCf (Fig. 8c) is mainly affected by lapse rate and SST. The three dominant factors for SWCF are SST, MWS, and lapse rate (Fig. 8d). LWP (Fig. 8e) is most sensitive to lapse rate, while IWP (Fig. 8f) is most sensitive to SST. Overall, the vertical lapse rate and SST are the two

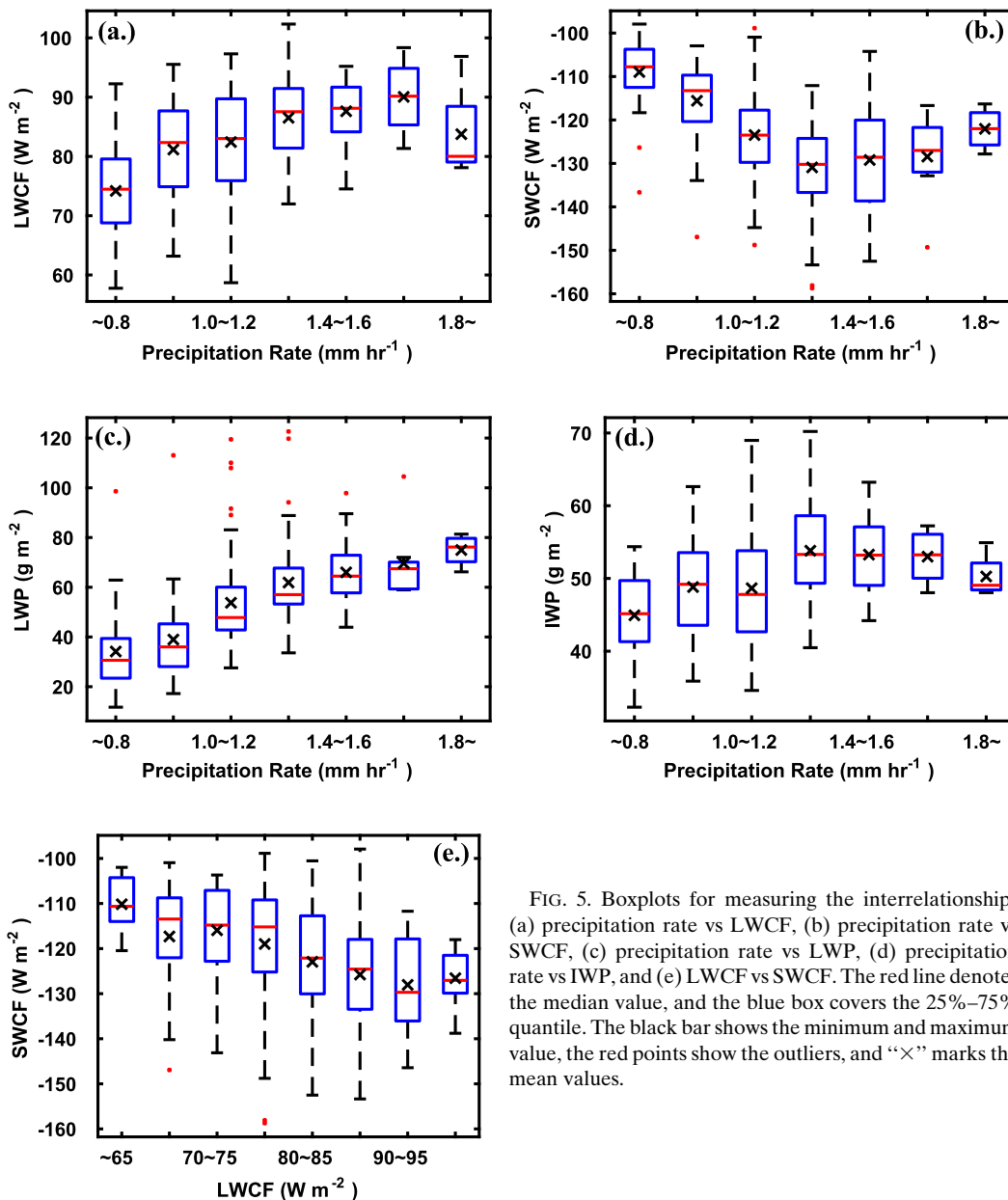


FIG. 5. Boxplots for measuring the interrelationships (a) precipitation rate vs LWCF, (b) precipitation rate vs SWCF, (c) precipitation rate vs LWP, (d) precipitation rate vs IWP, and (e) LWCF vs SWCF. The red line denotes the median value, and the blue box covers the 25%–75% quartile. The black bar shows the minimum and maximum value, the red points show the outliers, and “x” marks the mean values.

dominant controlling factors for all six of the TC characteristics. Sources of sensitivity may be analyzed in greater detail by examining the first-order and second-order sensitivity indices (Table 3). Recall that these describe the individual effect of a change in a parameter (first order), and the effect of interactions between two different parameters (second order) on a particular output variable. The TC intensity, LWCF, and SWCF are primarily influenced by linear (independent) changes in each of the five input factors, as second-order effects are very small. For example, the linear effect of

the five input factors ( $S1 + S2 + S3 + S4 + S5$ ) contributes 94.5%, 96.8%, and 99.9% of the total variance for intensity, LWCF, and SWCF, respectively. It is worthwhile to note that the midlevel RH shows almost no effect on TC intensity change. This may be explained, in part, by the lack of vertical wind shear (Ge et al. 2013). It is also possible that parameter settings that control the entrainment of air into convective clouds are preventing dry midlevel air from having an effect on simulated storms [see He and Posselt (2015) for additional details]. The size and intensity of initial vortex seed (RMW and

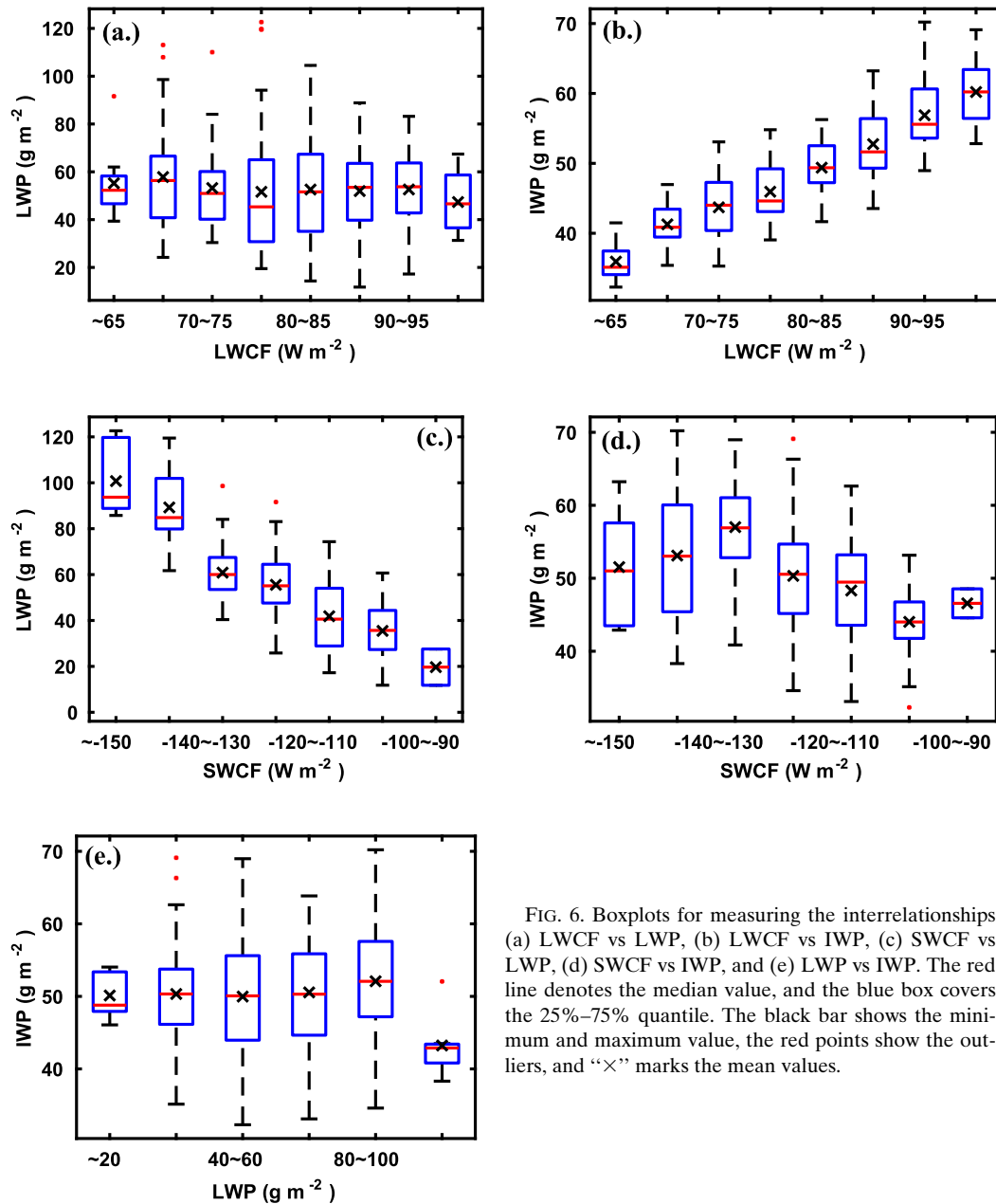


FIG. 6. Boxplots for measuring the interrelationships (a) LWCF vs LWP, (b) LWCF vs IWP, (c) SWCF vs LWP, (d) SWCF vs IWP, and (e) LWP vs IWP. The red line denotes the median value, and the blue box covers the 25%–75% quantile. The black bar shows the minimum and maximum value, the red points show the outliers, and “x” marks the mean values.

MWS) cause weak changes in TC intensity throughout their predetermined ranges. It is widely assumed that factors favoring TC genesis can be considered separately and are additive or multiplicative (Nolan and Rappin 2008). This assumption is fundamental to the genesis parameters of Gray (1968), and DeMaria et al. (2001), as well as statistical hurricane intensity forecast models such as SHIPS (DeMaria and Kaplan 1999). These genesis parameters, and statistical prediction models, do not include the initial vortex characteristics. Our results

show that the combined effect on AGCM-simulated TCs is indeed negligible.

Second-order indices are nonnegligible for precipitation rate, LWP, and IWP. For precipitation rate, the linear effect covers 76% of the total variance, and the main contribution is from SST with 66.6%. The interaction effect between GAMMA and SST covers 15.33%, which indicates that the interaction between these two factors is the second most important effect on the final simulated precipitation rate. Interaction

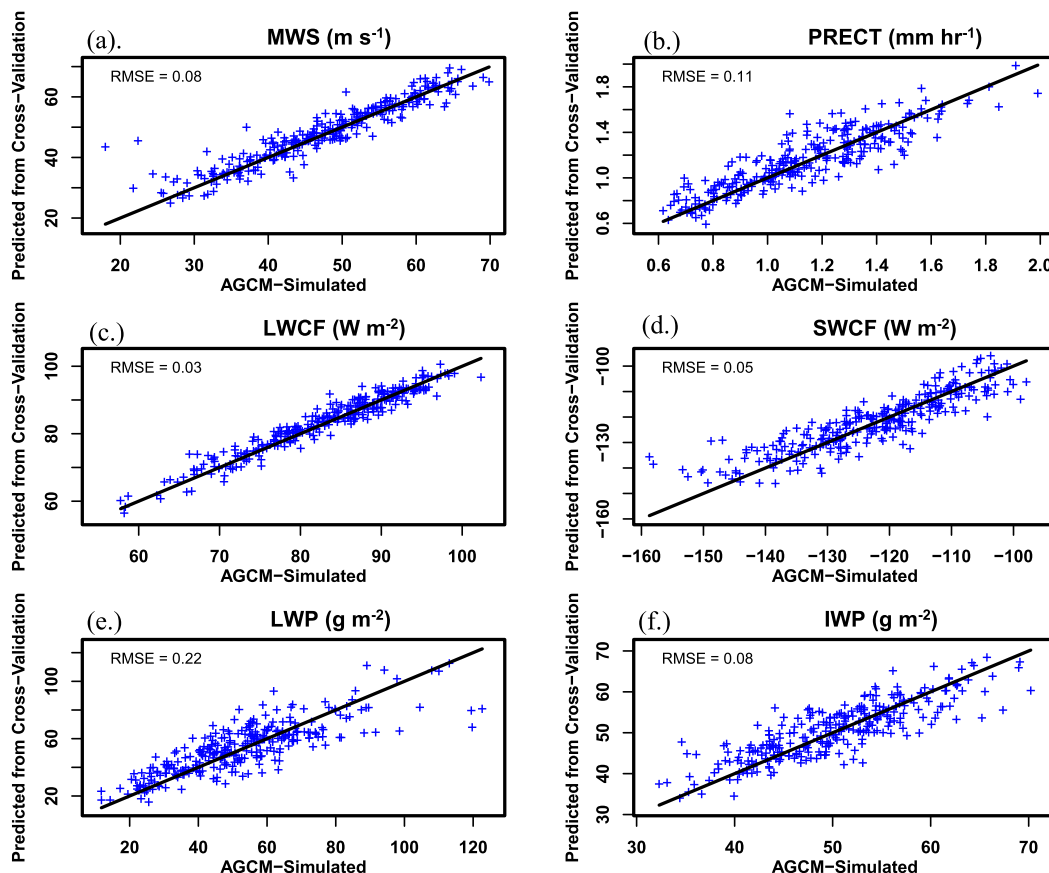


FIG. 7. Predicted values from GaSP using cross-validation analysis vs the actual values of (a) MWS, (b) total precipitation rate, (c) LWCF, (d) SWCF, (e) LWP, and (f) IWP from the numerical model for the initial 300 runs.

between GAMMA and SST indicates the two are not linearly independent, and that a joint increase in these two parameters causes a larger change in storm strength than the combined change due to each individual factor alone. This makes sense from a buoyancy perspective, as the lapse rate controls the stability of the profile, but the SST sets the properties of the boundary layer air that is ingested into (and lifted in) the TC. As such, an increase in SST serves to amplify the effect of an increase in lapse rate (and vice versa).

For LWP, the linear effect contributes 71.1% to the total variance, with the largest contribution from changes in the lapse rate. It is worthwhile to note that there are two second-order sensitivity indices greater than 10%: S23 and S34. The large value of S23 indicates that interactions between MWS and GAMMA are very important to the final simulated LWP. In essence, this means that an increase in MWS and an increase in GAMMA both produce an increase in LWP, but the combination of the two produces an amplified response. We speculate that, in this case, a given change in initial maximum wind speed leads to an increase in LWP via an

increase in total storm intensity, and that this increase is amplified by an increase in lapse rate. Relatively large values of S34 mean that the interaction between GAMMA and SST is also very important to the final simulated LWP. The mechanism is likely the same as in the case of increasing storm intensity described above. In addition, we note that both interaction effects include GAMMA. The relatively large interaction effects show that the contributions from the input factors to LWP is complicated, and may help to explain why LWP is much harder for GaSP to predict than the other five output variables. For IWP, the linear effect covers 75.6% of the total variance, while the interaction effect between GAMMA and SST covers 16%. The first-order, second-order, and total sensitivity index provide information as to which input factor is more important, and whether the variance stems mainly from its linear effect or interaction effect. In the following section, we take a detailed look at the marginal relationships and covarying relationships for the important factors.



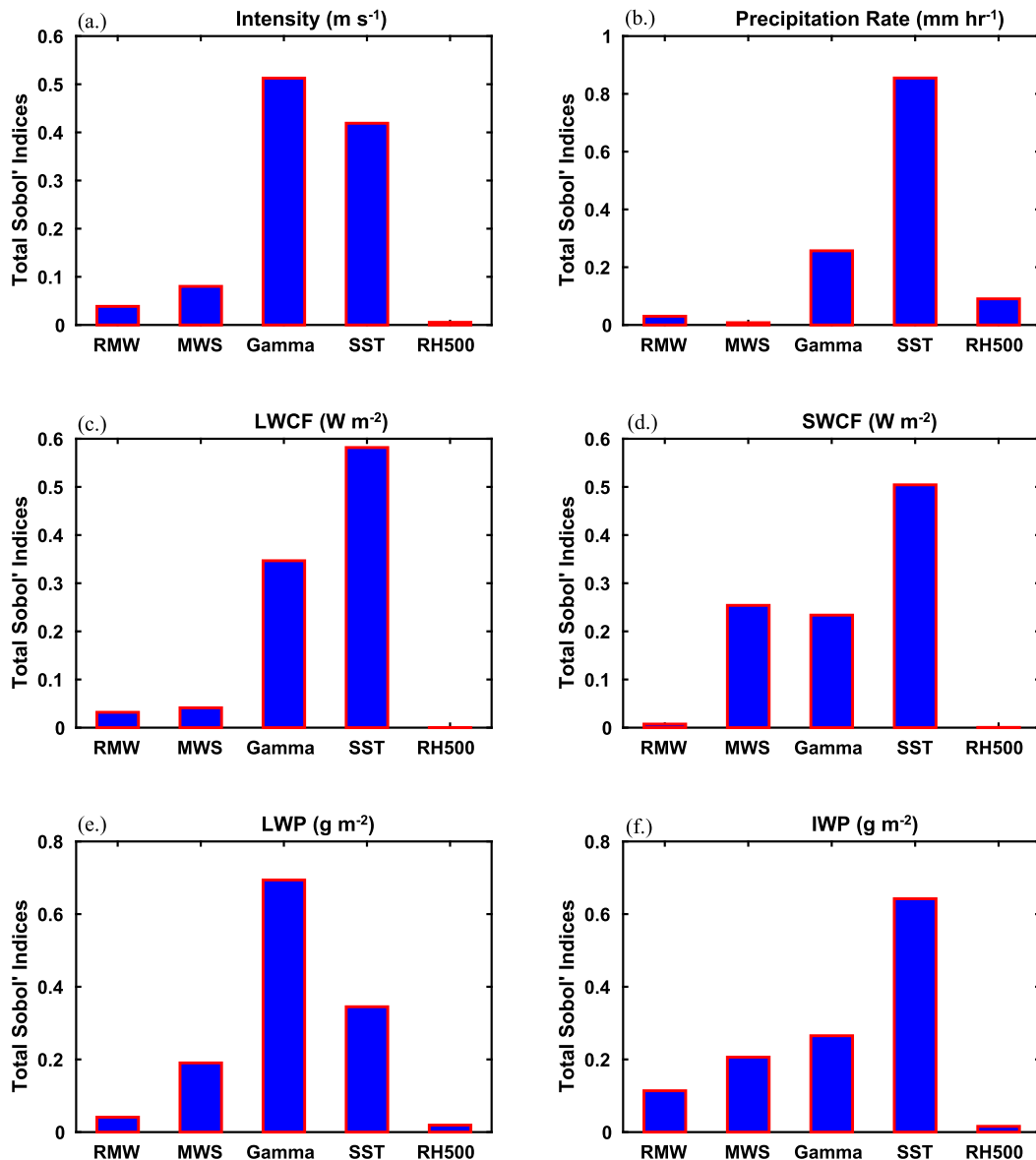


FIG. 8. The total sensitivity index of the six output variables to the five initial parameters: (a) intensity, (b) precipitation rate, (c) LWCF, (d) SWCF, (e) LWP, and (f) IWP.

### 3) INPUT–OUTPUT RELATIONSHIPS

#### (i) Intensity

The effect of atmospheric conditions on tropical cyclone intensity change has been extensively studied using observational data and high-resolution regional models. Even so, few studies have attempted to quantify the effect of an isolated factor on TC intensity with statistical robustness. This is in large part because it is difficult to isolate a single factor from observational analysis, as Earth's atmosphere changes its various conditions simultaneously. In addition, production of

large numbers of high-resolution regional model simulations is limited by computational constraints. Having generated 300 simulations using quasi-random LHS sampling, and with the aid of a statistical emulator, we are able to assess the relationship between TC intensity and a single factor, quantify its associated uncertainty, and assess the changed TC intensity because of two simultaneously changing factors. Figure 9 illustrates this point using TC intensity. Figure 9a assesses the relationship between TC intensity and atmospheric lapse rate (GAMMA), while the other four perturbed parameters (RMW, MWS, SST, 500-hPa RH) are set equal

TABLE 3. The first-order and second-order Sobol sensitivity indices for the six output variables. For the second-order indices, only the values greater than 1% are listed. The five input factors and their order are as follows: 1) RMW, 2) MWS, 3) GAMMA, 4) SST, and 5) 500-hPa RH. S1 denotes the first-order sensitivity index from the linear effect of the first input factor: RMW. S24 denotes the second-order sensitivity index from the interaction effect between the second input factor (MWS) and the fourth input factor (SST).

	Intensity	Precipitation	LWCF	SWCF	LWP	IWP
S1	0.0277	0.0030	0.0314	0.0074	0.0116	0.0831
S2	0.0418	0.0018	0.0400	0.2540	0.0586	0.1613
S3	0.4917	0.0540	0.3450	0.2337	0.4415	0.0873
S4	0.3812	0.6661	0.5513	0.5043	0.1974	0.4229
S5	0.0015	0.0346	0.0001	0.0005	0.0018	0.0018
S13	—	0.0163	—	—	0.0210	—
S14	—	—	—	—	—	0.0232
S23	—	—	—	—	0.1071	0.0126
S24	0.0252	—	—	—	0.0197	0.0267
S34	—	0.1533	0.0305	—	0.1167	0.1596
S35	—	0.0317	—	—	—	—
S45	—	0.0223	—	—	—	—

to three different values: 25% (low, red line), 50% (median, green line), and 75% (high, blue line) of their respective ranges (Table 2). Figure 9b is the same as Fig. 9a, but shows the relationship between TC intensity and sea surface temperature. While it is well known that large atmospheric lapse rates (more unstable atmosphere) and high sea surface temperatures produce larger TC intensity, the relationships have not been quantified. Our results show that the TC intensity is linearly affected by atmospheric instability and sea surface temperature. Not only are the relationships between both parameters and the TC intensity monotonic, but there is a systematic change in intensity as the remaining parameters are increased from 25% to 50% and to 75% of their ranges. When GAMMA and SST are allowed to vary simultaneously (Fig. 9c), they exhibit near-linear influence on TC intensity change, with little evidence of intervariable interactions.

(ii) Longwave cloud radiative forcing

TC LWCF is linearly dependent on GAMMA (Fig. 10a); however, the slope of this relationship is different when the other four input factors are set to different values. Specifically, the largest response per unit change in GAMMA occurs when all other parameters are set to relatively low values, while the smallest response occurs when other parameters are set to relatively high values. LWCF increases with increasing SST (Fig. 10b), and this relationship tends to flatten when SST is higher than 30°C. Unlike GAMMA, the SST–LWCF response does not appear to change with changes in the values of other

parameters. Higher SST produces stronger storms and also larger vapor content in the atmospheric, both of which could increase high cloud fraction, and thus lead to higher LWCF. High LWCF occurs in high GAMMA and high SST conditions, while low LWCF occurs at low values of GAMMA, and in low SST regions (Fig. 10c). The flattening of the response in LWCF to changes in SST appears to be a function of the joint relationship between GAMMA and SST and the LWCF. At low values of GAMMA, increasing SST produces a monotonic increase in SWCF; however, as GAMMA increases, the response of LWCF to increases in SST saturates. It is possible that, at large GAMMA (and higher storm intensity; Fig. 9b), the storm processes water more efficiently, leading to a saturation in the LWCF.

(iii) Shortwave cloud radiative forcing

Figure 11 quantifies the same types of relationship for SWCF, and only the factors that have a significant effect on the TC-relevant SWCF are shown. As shown in Figs. 11a, 11b, and 11c, the magnitude of SWCF is linearly affected by three factors: the intensity of the initial vortex seed (MWS), atmospheric lapse rate (GAMMA), and sea surface temperature (SST). Larger magnitudes of SWCF (absolute value of SWCF) are associated with smaller initial MWS, a more stable atmosphere, and larger SST. Recall that the SWCF is essentially unrelated to storm intensity (Fig. 4), but is strongly related to LWP (Fig. 6c). It is possible that a lower MWS and lower GAMMA lead to larger liquid (vs ice) production in the storm and as such a stronger SWCF. Increasing SST, for a given low-level RH, implies an increase in the low-level water vapor content, which may also be related to an increase in LWP and a consequent increase in SWCF. It is worthwhile to note that changes in GAMMA and SST have the opposite effect on the SWCF change. Recall that SST and GAMMA were shown to interact to produce stronger storms. The inverse sensitivity of SWCF to GAMMA and SST indicates that storms of equal intensity may have very different SWCF.

(iv) Cloud ice water path

The TC cloud ice water path is most sensitive to changes in SST (Fig. 8f), but it also responds to changes in the initial vortex characteristics and atmospheric stability. In fact, IWP is the only TC-relevant variable that shows strong sensitivity to four of the five initial parameters (RMW, MWS, GAMMA, and SST). Examination of the input–response relationships indicates IWP exhibits a small, but discernible, linear decrease with increasing initial vortex seed intensity (Fig. 12a); however, the response is a strong function of the values

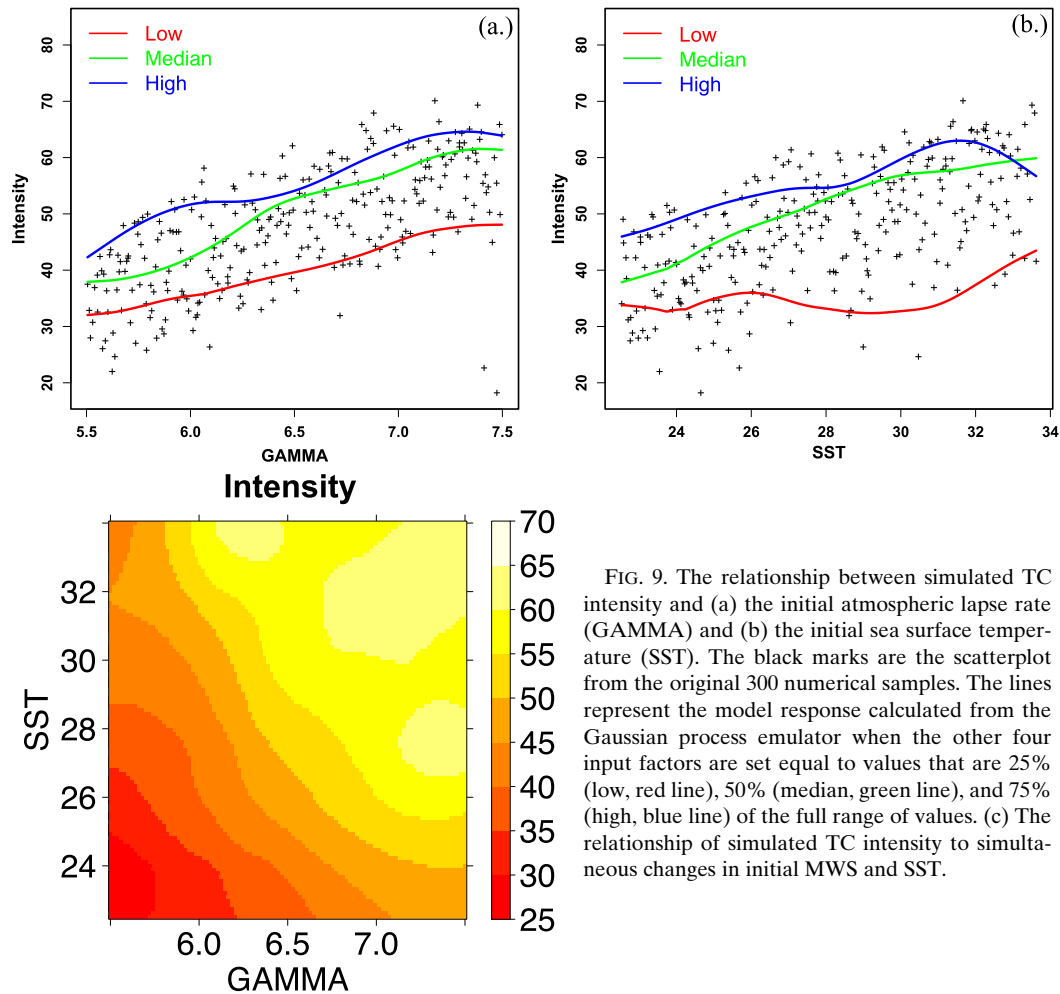


FIG. 9. The relationship between simulated TC intensity and (a) the initial atmospheric lapse rate (GAMMA) and (b) the initial sea surface temperature (SST). The black marks are the scatterplot from the original 300 numerical samples. The lines represent the model response calculated from the Gaussian process emulator when the other four input factors are set equal to values that are 25% (low, red line), 50% (median, green line), and 75% (high, blue line) of the full range of values. (c) The relationship of simulated TC intensity to simultaneous changes in initial MWS and SST.

of the other control parameters. There is a somewhat stronger increase in IWP with increasing sea surface temperature (Figs. 12b,d); however, the overall trend is masked by significant variability. We have regenerated the GaSP model using subsets of the data and various parameter settings in the GaSP training, but the high-frequency variability in the relationship between SST and IWP remains (Fig. 12c). In essence, this means that, while the univariate response plots suggest an increase in IWP with respect to increases in SST, the actual response is a complex function of the input parameters. This is reflected in the Sobol sensitivity indices (Table 3), which show nonnegligible second-order effects for several input variable combinations, with interactions between SST and lapse rate particularly strong. The sensitivity in the IWP to changes in multiple variables may simply reflect the fact that the amount of ice in the simulated TC is a function of multiple different processes that affect the vertical transport of water, the conversion from liquid to ice, or both.

In each of the sets of results presented in Figs. 9–12, the sensitivity analysis derived from the GaSP emulator exhibits an overarching structure along with small-scale variations. For example, in the relationships among SST, GAMMA, and LWCF, there is a general trend toward increasing LWCF with increasing GAMMA and SST, with evidence of nonmonotonicity (decreasing LWCF with increasing SST). There are also very finescale variations in the relationship (Fig. 10c). We have conducted tests in which we retrained the GaSP model using different tuning parameter values, and have found that the larger-scale sensitivity is robust, while the small-scale variation is not.

##### 5. Marginal sensitivity of intensity to five initial conditions at three horizontal resolutions

An examination of the sensitivity of simulated TC characteristics to changes in atmospheric environmental factors in CAM provides insight into the key controls on

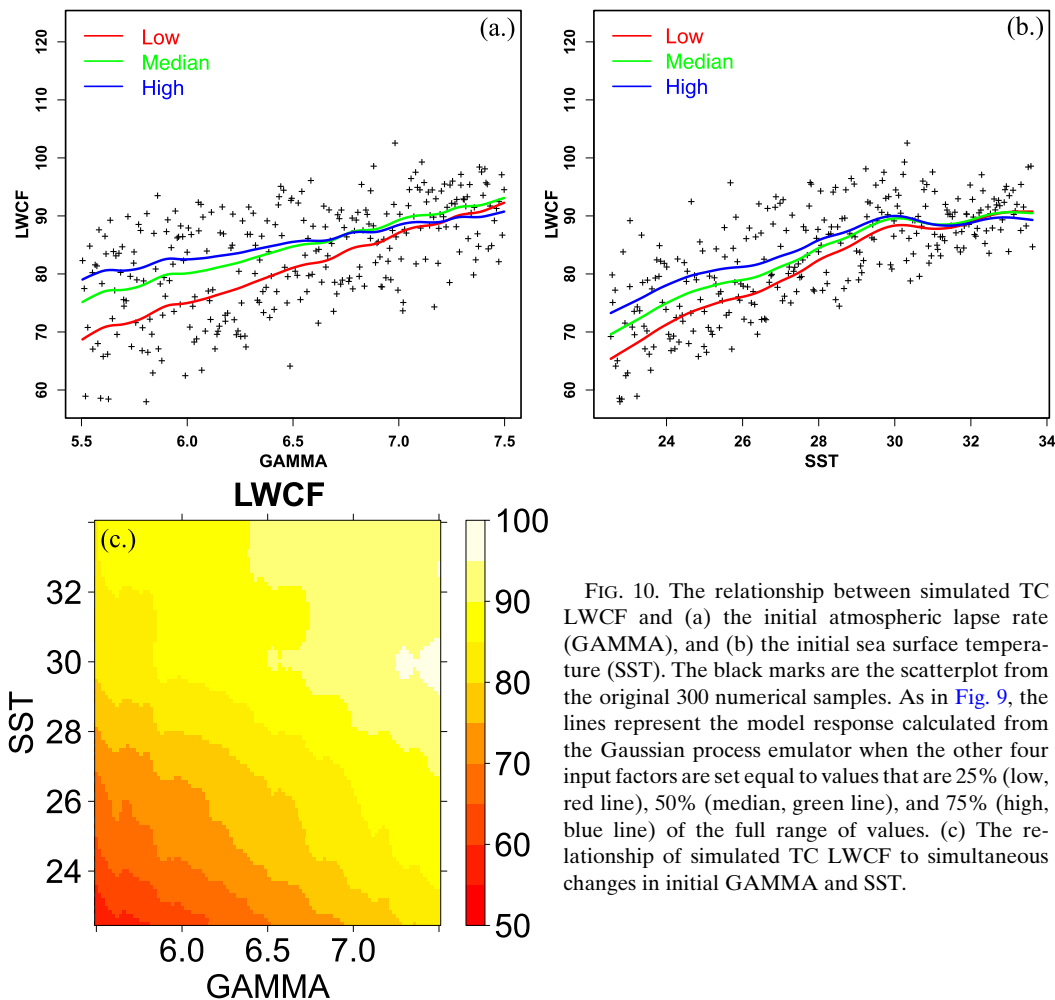


FIG. 10. The relationship between simulated TC LWCF and (a) the initial atmospheric lapse rate (GAMMA), and (b) the initial sea surface temperature (SST). The black marks are the scatterplot from the original 300 numerical samples. As in Fig. 9, the lines represent the model response calculated from the Gaussian process emulator when the other four input factors are set equal to values that are 25% (low, red line), 50% (median, green line), and 75% (high, blue line) of the full range of values. (c) The relationship of simulated TC LWCF to simultaneous changes in initial GAMMA and SST.

TC evolution in an AGCM. However, all of our results are valid for the specific horizontal grid resolution of  $0.5^\circ \times 0.5^\circ$ . Here, we compare the sensitivity of simulated TC intensity to changes in the five initial conditions for three different horizontal grid resolutions commonly used in full AGCMs:  $1.0^\circ \times 1.0^\circ$ ,  $0.5^\circ \times 0.5^\circ$ , and  $0.25^\circ \times 0.25^\circ$ . The  $1.0^\circ \times 1.0^\circ$  and  $0.25^\circ \times 0.25^\circ$  experiments utilize the same LHS-generated control parameter sets, yielding a sample of 300 input–output vector pairs each.

Figure 13 shows the marginal relationship between simulated TC intensity and changes in each of the five control factors, and suggests that these relationships are resolution dependent. First, note that, in simulations run on a  $1.0^\circ \times 1.0^\circ$  grid, the TC intensity never exceeds  $30 \text{ ms}^{-1}$ . This is generally consistent with a previous study that showed that the NCAR CAM 5.0 model with a resolution of  $1.0^\circ \times 1.0^\circ$  was only able to capture nine tropical cyclones per year globally (Walsh et al. 2015) and was not able to produce strong tropical

cyclones (e.g., category 4 or 5 on the Saffir–Simpson hurricane wind scale). While coarse-resolution simulations exhibit moderate sensitivity to the initial size and strength of the vortex seed (Figs. 13a and 13b), there is little discernible sensitivity to changes in any of the environmental factors (Figs. 13c–e). As was demonstrated above, there is moderate sensitivity of simulated TC intensity to changes in the vortex seed characteristics for simulations run at  $0.5^\circ \times 0.5^\circ$  resolution, but the response is quite different from the one obtained at  $1.0^\circ \times 1.0^\circ$  resolution. At  $0.25^\circ \times 0.25^\circ$ , there is essentially no response in intensity to changes in the vortex seed. In contrast to the  $1.0^\circ \times 1.0^\circ$  simulations, sensitivities to changes in the environmental factors (Figs. 13c–e) are quite similar for the two finer resolutions. While we cannot comment on convergence, these results indicate the response of TCs to changes in environment at  $0.5^\circ \times 0.5^\circ$  and  $0.25^\circ \times 0.25^\circ$  resolution in CAM is similar. Consistency between the two sets of simulations is

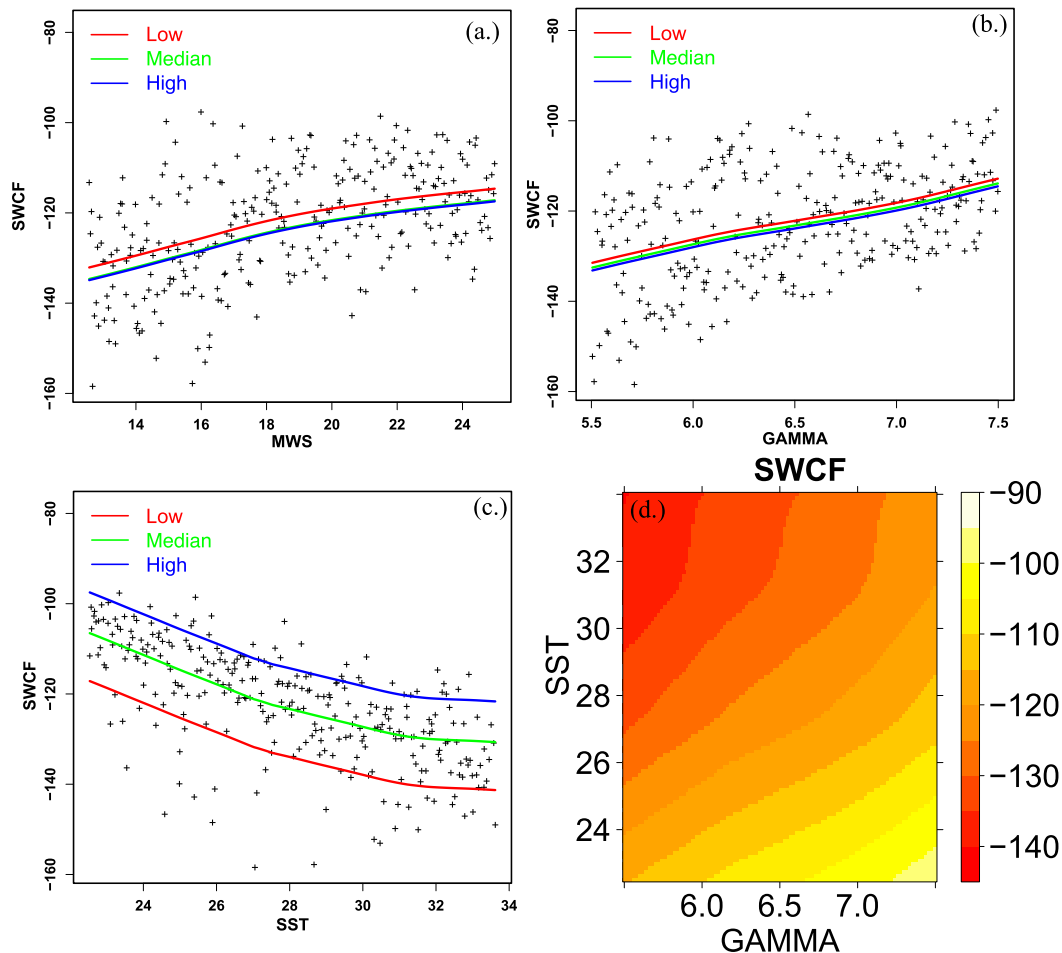


FIG. 11. The relationship between simulated TC SWCF and (a) the initial MWS, (b) the initial atmospheric lapse rate (GAMMA), and (c) the initial sea surface temperature (SST). The black marks are the scatterplot from the original 300 numerical samples. As in Figs. 9 and 10, the lines represent the model response calculated from the Gaussian process emulator when the other four input factors are set equal to values that are 25% (low, red line), 50% (median, green line), and 75% (high, blue line) of the full range of values. (d) The relationship between simulated TC SWCF and simultaneous changes in initial GAMMA and SST.

likely due to representation of mesoscale circulations, which are largely absent at  $1.0^\circ \times 1.0^\circ$  resolution. TC intensity at  $0.25^\circ \times 0.25^\circ$  resolution is uniformly stronger than it is at  $0.5^\circ \times 0.5^\circ$ , consistent with findings in earlier studies that higher-resolution models tend to generate stronger TCs. Wehner et al. (2014) showed that the high-resolution ( $0.25^\circ \times 0.25^\circ$ ) configuration of NCAR CAM, version 5.1, simulates a much more realistic global number of tropical cyclones and can produce tropical cyclones up to category 5. Bacmeister et al. (2014) also showed that the high-resolution ( $0.23^\circ \times 0.31^\circ$ ) CAM 5 was capable of simulating realistic tropical cyclone distributions, and reproducing the interannual variability in TC statistics. Our results suggest that, when run with relatively coarse resolution of  $1.0^\circ \times 1.0^\circ$ , the NCAR CAM cannot simulate TCs well, and that this

may be because it could not simulate the response to the SST and GAMMA in a realistic manner. It also shows that, while storms are still somewhat underresolved at  $0.5^\circ \times 0.5^\circ$ , this resolution more closely approximates the response exhibited by simulations run at  $0.25^\circ \times 0.25^\circ$ , a resolution increasingly used for tropical cyclone studies in climate models.

## 6. Discussion and limitations of this study

This work uses a multivariate sensitivity analysis technique to study the response of individual dynamical systems to changes in multivariate parameters. It provides a specific example of the use of LHS, a GaSP emulator, and Sobol's variance-based sensitivity analysis for sensitivity assessment in AGCM-simulated TCs.



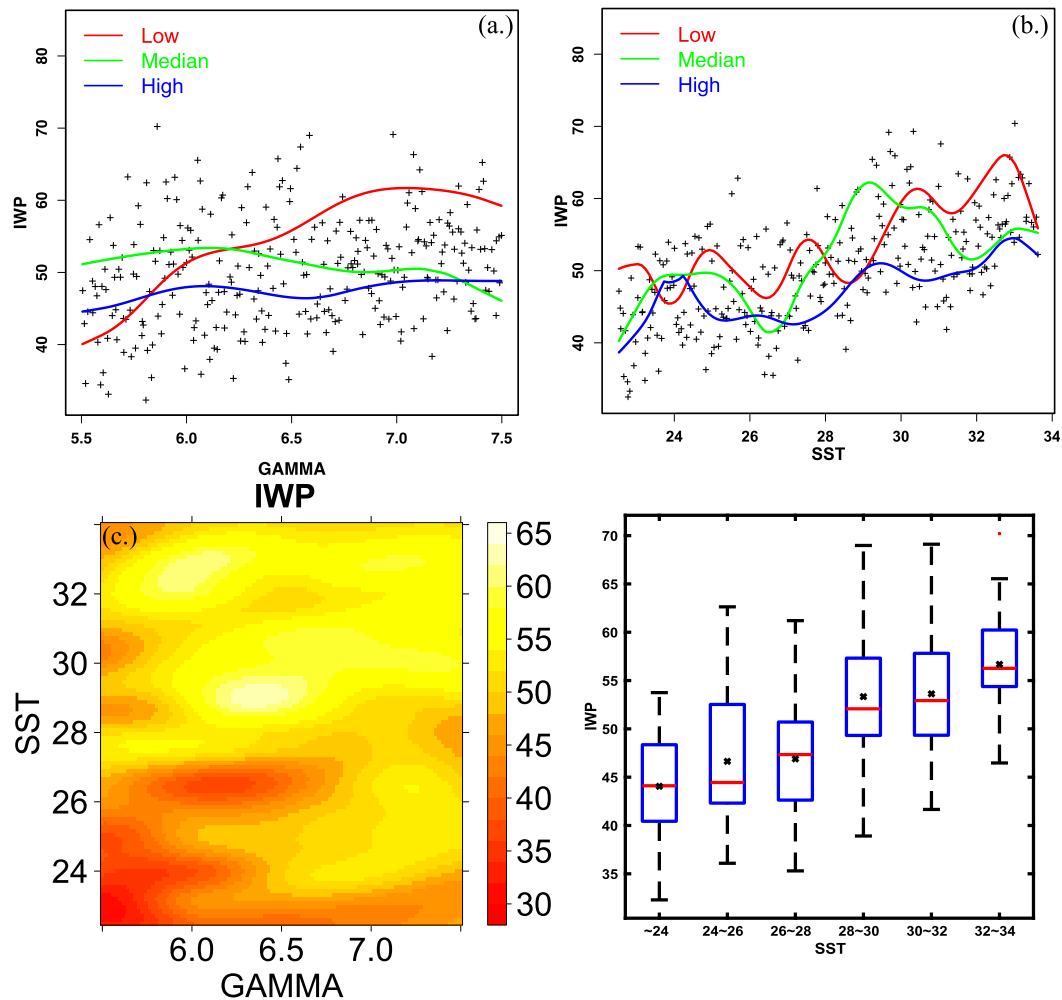


FIG. 12. The relationship between simulated TC IWP and (a) the initial atmospheric lapse rate (GAMMA) and (b) the initial sea surface temperature (SST). The black marks are the scatterplot from the original 300 numerical samples. As in Figs. 9–11, the lines represent the model response calculated from the Gaussian process emulator when the other four input factors are set equal to values that are 25% (low, red line), 50% (median, green line), and 75% (high, blue line) of the full range of values. (c) The relationship between simulated TC IWP and simultaneous changes in initial GAMMA and SST. (d) The boxplot measures the relationship between IWP and SST using the original 300 numerical simulated points.

The sensitivity analysis technique could easily be extended to other dynamical systems, such as extratropical cyclones, non-TC tropical convective systems, and so on. It also has the potential to quantify the uncertainties associated with variability in model internal parameters in physical parameterization schemes, such as cloud, precipitation, and turbulence, which are key components in climate model representation of small-scale processes.

It is worthwhile to note that the sensitivities of AGCM-simulated TC characteristics to the five initial conditions found in our study utilized a highly idealized framework. While our results are consistent with the known sensitivity of TCs to changes in SST and lapse rate, our goal is not primarily to understand how TCs

function in the real world. Such a study requires a simulation framework that accounts for additional known factors that influence the development of TCs (e.g., vertical wind shear, characteristics of individual convective updrafts, ocean mixing, proximity to land, etc.). Rather, we use an idealized model setup to obtain insight into sensitivities of TCs simulated in AGCMs, which include a multitude of interacting and complex factors. The lack of vertical wind shear in our idealized configuration enables us to focus on the response of the dynamical system to environmental factors like SST, GAMMA, and so on. Incorporation of vertical wind shear into the framework will require the adjustment of vertical temperature and thus lapse rate based on

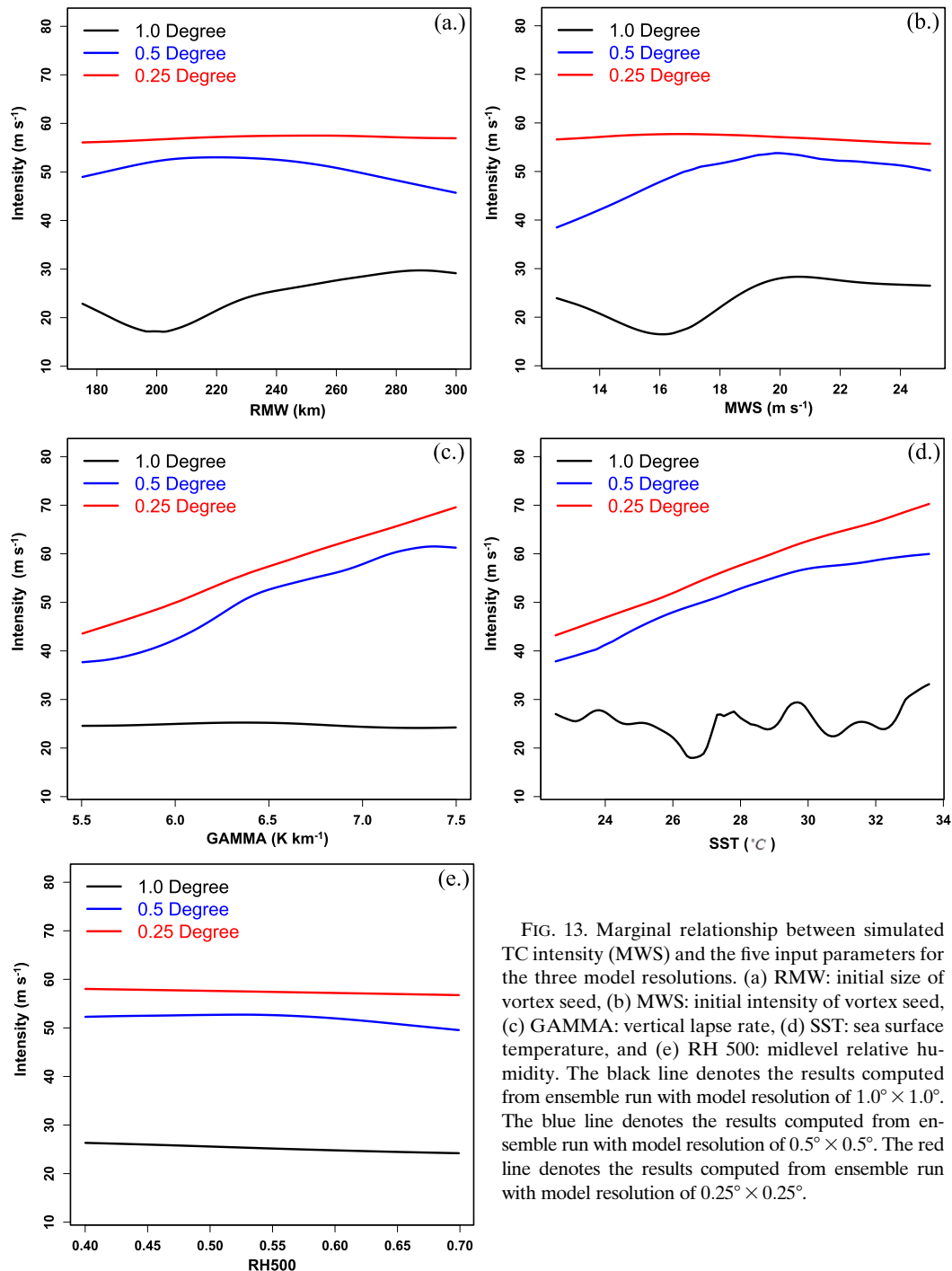


FIG. 13. Marginal relationship between simulated TC intensity (MWS) and the five input parameters for the three model resolutions. (a) RMW: initial size of vortex seed, (b) MWS: initial intensity of vortex seed, (c) GAMMA: vertical lapse rate, (d) SST: sea surface temperature, and (e) RH 500: midlevel relative humidity. The black line denotes the results computed from ensemble run with model resolution of  $1.0^{\circ} \times 1.0^{\circ}$ . The blue line denotes the results computed from ensemble run with model resolution of  $0.5^{\circ} \times 0.5^{\circ}$ . The red line denotes the results computed from ensemble run with model resolution of  $0.25^{\circ} \times 0.25^{\circ}$ .

thermal wind balance. This is not an insurmountable challenge, but is one that we leave for future work.

## 7. Summary

This case study–based idealized sensitivity study utilized a quantitative sensitivity analysis technique to

map changes in TC initial conditions to responses of TC intensity, precipitation, water path, and radiative forcing. The sensitivity analysis framework allowed us to explore not only the univariate responses, but also the interactions among input factors and output variables. Examination of the response functions indicates the

following behavior in the idealized TC framework: 1) Environmental lapse rate (instability) and sea surface temperature linearly affect the growth of TCs. 2) The TC total precipitation rate, cloud radiative forcing, and cloud water content are all sensitive to the size and intensity of the initial vortex seed, the environmental lapse rate, and the sea surface temperature, but have weak or no sensitivity to the initial 500-hPa relative humidity. 3) Changes in TC initial conditions interact nonlinearly to affect the final simulated characteristics of the TCs, in addition to the linear relationships mentioned in behavior 1. 4) Interrelationships among TC relevant characteristics exist: specifically, stronger TCs tend to be associated with larger LWCF and larger total precipitation rate tends to be associated with larger values of LWCF and larger cloud LWP. 5) The sensitivities of simulated TC intensity to the five initial conditions change for different model resolutions, with  $1^\circ$  resolution producing notably different sensitivity response, relative to simulations run at  $0.5^\circ$  and  $0.25^\circ$ . This highlights the reduced utility in using  $1.0^\circ \times 1.0^\circ$  for tropical cyclone studies, since the resolution is unable to correctly simulate storm sensitivities found at higher resolutions.

We have shown that the multivariate sensitivity analysis technique can be applied to the study of a particular dynamical system, and yields valuable information on the sensitivity of the system and its interaction with environmental factors. Specifically, we found the combination of Latin hypercube sampling for generation of an ensemble of model simulations, surrogate models for computationally efficient approximation of numerical simulation outcomes, and variance-based sensitivity analysis for computation of individual and interaction effects constitutes a powerful set of tools for the diagnosis of dynamic systems. It allows simultaneous variation of multiple model input factors, and thus can be used to examine not only univariate sensitivity, but also the effects of interactions among different parameters on simulated model output. The results have implications not only for understanding how AGCMs simulate TCs, but also potentially for the design of ensemble simulations. Quantification of the primary individual control factors, along with their interaction with other parameters, allows for a more well-informed ensemble experiment that more completely spans the range of variability in the system. To the extent that the model construction (e.g., dynamical core, physical parameterizations, etc.) are consistent between the idealized framework and the fully realistic AGCM, an ensemble based on perturbations to the factors we identified may be expected to span a broader range of TC evolution than random perturbations alone.

In closing, we note that the sensitivity analysis we conducted can easily be extended to other dynamic systems, given a suitable model configuration and identification of key control factors. We are currently in the process of using these tools for the analysis of extratropical cyclones, orographic precipitation, and mesoscale convective systems, and expect to report on these results in subsequent manuscripts.

*Acknowledgments.* The authors thank Christiane Jablonowski, Kevin Reed, and Weiye Yao for their advice and help on tropical cyclone simulations in the NCAR CAM. The authors also thank Erika Roesler and Zach Zhanyang Zhang for their advice and help with the Latin hypercube sampling method. The authors thank the Advanced Research Computing Technology Services at the University of Michigan. Dr. Posselt was supported by the Cyclone Global Navigation Satellite System (CYGNSS) mission under NASA Science Mission Directorate Contract NNL13AQ00C. A portion of this research was carried out at the Jet Propulsion Laboratory, California Institute of Technology, under a contract with the National Aeronautics and Space Administration.

#### REFERENCES

- Ancell, B., and G. J. Hakim, 2007: Comparing adjoint- and ensemble-sensitivity analysis with applications to observation targeting. *Mon. Wea. Rev.*, **135**, 4117–4134, <https://doi.org/10.1175/2007MWR1904.1>.
- Bacmeister, J. T., M. F. Wehner, R. B. Neale, A. Gettelman, C. Hannay, P. H. Lauritzen, J. M. Caron, and J. E. Truesdale, 2014: Exploratory high-resolution climate simulations using the Community Atmosphere Model (CAM). *J. Climate*, **27**, 3073–3099, <https://doi.org/10.1175/JCLI-D-13-00387.1>.
- Berner, J., S.-Y. Ha, J. P. Hacker, A. Fournier, and C. Snyder, 2011: Model uncertainty in a mesoscale ensemble prediction system: Stochastic versus multiphysics representations. *Mon. Wea. Rev.*, **139**, 1972–1995, <https://doi.org/10.1175/2010MWR3595.1>.
- Bolado-Lavin, R., and A. C. Badaea, 2008: Review of sensitivity analysis methods and experience for geological disposal of radioactive waste and spent nuclear fuel. European Commission Joint Research Centre Scientific and Technical Rep. EUR 23712 EN-2008, 84 pp., [http://publications.jrc.ec.europa.eu/repository/bitstream/JRC49536/reqno\\_jrc49536\\_eur\\_23712en1.pdf](http://publications.jrc.ec.europa.eu/repository/bitstream/JRC49536/reqno_jrc49536_eur_23712en1.pdf).
- Bowman, K., J. Sacks, and Y. Chang, 1993: Design and analysis of numerical experiments. *J. Atmos. Sci.*, **50**, 1267–1278, [https://doi.org/10.1175/1520-0469\(1993\)050<1267:DAAONE>2.0.CO;2](https://doi.org/10.1175/1520-0469(1993)050<1267:DAAONE>2.0.CO;2).
- Boyle, J. S., S. A. Klein, D. D. Lucas, H.-Y. Ma, J. Tannahill, and S. Xie, 2015: The parametric sensitivity of CAM5's MJO. *J. Geophys. Res. Atmos.*, **120**, 1424–1444, <https://doi.org/10.1002/2014JD022507>.
- Camargo, S. J., 2013: Global and regional aspects of tropical cyclone activity in the CMIP5 models. *J. Climate*, **26**, 9880–9902, <https://doi.org/10.1175/JCLI-D-12-00549.1>.

- Collins, M., B. B. Booth, B. Bhaskaran, G. R. Harris, J. M. Murphy, D. M. H. Sexton, and M. J. Webb, 2011: Climate model errors, feedbacks and forcings: A comparison of perturbed physics and multi-model ensembles. *Climate Dyn.*, **36**, 1737–1766, <https://doi.org/10.1007/s00382-010-0808-0>.
- Daescu, D. N., and R. Todling, 2010: Adjoint sensitivity of the model forecast to data assimilation system error covariance parameters. *Quart. J. Roy. Meteor. Soc.*, **136**, 2000–2012, <https://doi.org/10.1002/qj.693>.
- Davis, C. A., and D. A. Ahijevych, 2012: Mesoscale structural evolution of three tropical weather systems observed during PREDICT. *J. Atmos. Sci.*, **69**, 1284–1305, <https://doi.org/10.1175/JAS-D-11-0225.1>.
- DeMaria, M., and J. Kaplan, 1999: An updated Statistical Hurricane Intensity Prediction Scheme (SHIPS) for the Atlantic and eastern North Pacific basins. *Wea. Forecasting*, **14**, 326–337, [https://doi.org/10.1175/1520-0434\(1999\)014<0326:AUSHIP>2.0.CO;2](https://doi.org/10.1175/1520-0434(1999)014<0326:AUSHIP>2.0.CO;2).
- , J. A. Knaff, and B. H. Connell, 2001: A tropical cyclone genesis parameter for the tropical Atlantic. *Wea. Forecasting*, **16**, 219–233, [https://doi.org/10.1175/1520-0434\(2001\)016<0219:ATCGPF>2.0.CO;2](https://doi.org/10.1175/1520-0434(2001)016<0219:ATCGPF>2.0.CO;2).
- Doyle, J. D., C. A. Reynolds, C. Amerault, and J. Moskaitis, 2012: Adjoint sensitivity and predictability of tropical cyclogenesis. *J. Atmos. Sci.*, **69**, 3535–3557, <https://doi.org/10.1175/JAS-D-12-0110.1>.
- , C. Amerault, C. A. Reynolds, and P. A. Reinecke, 2014: Initial condition sensitivity and predictability of a severe extratropical cyclone using a moist adjoint. *Mon. Wea. Rev.*, **142**, 320–342, <https://doi.org/10.1175/MWR-D-13-00201.1>.
- Emanuel, K., 2007: Environmental factors affecting tropical cyclone power dissipation. *J. Climate*, **20**, 5497–5509, <https://doi.org/10.1175/2007JCLI1571.1>.
- , 2013: Downscaling CMIP5 climate models shows increased tropical cyclone activity over the 21st century. *Proc. Natl. Acad. Sci. USA*, **110**, 12 219–12 224, <https://doi.org/10.1073/pnas.1301293110>.
- , C. DesAutels, C. Holloway, and R. Korty, 2004: Environmental control of tropical cyclone intensity. *J. Atmos. Sci.*, **61**, 843–858, [https://doi.org/10.1175/1520-0469\(2004\)061<0843:ECOTCI>2.0.CO;2](https://doi.org/10.1175/1520-0469(2004)061<0843:ECOTCI>2.0.CO;2).
- Errico, R. M., 1997: What is an adjoint model? *Bull. Amer. Meteor. Soc.*, **78**, 2577–2591, [https://doi.org/10.1175/1520-0477\(1997\)078<2577:WIAAM>2.0.CO;2](https://doi.org/10.1175/1520-0477(1997)078<2577:WIAAM>2.0.CO;2).
- Frank, W. M., and E. A. Ritchie, 1999: Effects of environmental flow upon tropical cyclone structure. *Mon. Wea. Rev.*, **127**, 2044–2061, [https://doi.org/10.1175/1520-0493\(1999\)127<2044:EOEFUT>2.0.CO;2](https://doi.org/10.1175/1520-0493(1999)127<2044:EOEFUT>2.0.CO;2).
- , and —, 2001: Effects of vertical wind shear on the intensity and structure of numerically simulated hurricanes. *Mon. Wea. Rev.*, **129**, 2249–2269, [https://doi.org/10.1175/1520-0493\(2001\)129<2249:EOVWSO>2.0.CO;2](https://doi.org/10.1175/1520-0493(2001)129<2249:EOVWSO>2.0.CO;2).
- Friedman, J., 1991: Multivariate adaptive regression splines. *Ann. Stat.*, **19**, 1–67, <https://doi.org/10.1214/aos/1176347963>.
- , and C. B. Roosen, 1995: An introduction to multivariate adaptive regression splines. *Stat. Methods Med. Res.*, **4**, 197–217, <https://doi.org/10.1177/096228029500400303>.
- Frolov, S., A. M. Baptista, T. K. Leen, Z. Lu, and R. van der Merwe, 2009: Fast data assimilation using a nonlinear Kalman filter and a model surrogate: An application to the Columbia River estuary. *Dyn. Atmos. Oceans*, **48**, 16–45, <https://doi.org/10.1016/j.dynatmoce.2008.10.004>.
- Ge, X., T. Li, and M. Peng, 2013: Effects of vertical shears and midlevel dry air on tropical cyclone developments. *J. Atmos. Sci.*, **70**, 3859–3875, <https://doi.org/10.1175/JAS-D-13-066.1>.
- Gray, W. M., 1968: Global view of the origin of tropical disturbances and storms. *Mon. Wea. Rev.*, **96**, 669–700, [https://doi.org/10.1175/1520-0493\(1968\)096<0669:GVOTOO>2.0.CO;2](https://doi.org/10.1175/1520-0493(1968)096<0669:GVOTOO>2.0.CO;2).
- Habib, S., K. Heitmann, D. Higdson, C. Nakhleh, and B. Williams, 2007: Cosmic calibration: Constraints from the matter power spectrum and the cosmic microwave background. *Phys. Rev. D*, **76**, 083503, <https://doi.org/10.1103/PhysRevD.76.083503>.
- Hacker, J. P., and L. Lei, 2015: Multivariate ensemble sensitivity with localization. *Mon. Wea. Rev.*, **143**, 2013–2027, <https://doi.org/10.1175/MWR-D-14-00309.1>.
- He, F., and D. J. Posselt, 2015: Impact of parameterized physical processes on simulated tropical cyclone characteristics in the Community Atmosphere Model. *J. Climate*, **28**, 9857–9872, <https://doi.org/10.1175/JCLI-D-15-0255.1>.
- , —, C. M. Zarzycki, and C. Jablonowski, 2015: A balanced tropical cyclone test case for AGCMs with background vertical wind shear. *Mon. Wea. Rev.*, **143**, 1762–1781, <https://doi.org/10.1175/MWR-D-14-00366.1>.
- Hegstad, B., and O. Henning, 2001: Uncertainty in production forecasts based on well observations, seismic data, and production history. *SPE J.*, **6**, 409–424, <https://doi.org/10.2118/74699-PA>.
- Helton, J. C., and F. J. Davis, 2003: Latin hypercube sampling and the propagation of uncertainty in analyses of complex systems. *Reliab. Eng. Syst. Saf.*, **81**, 23–69, [https://doi.org/10.1016/S0951-8320\(03\)00058-9](https://doi.org/10.1016/S0951-8320(03)00058-9).
- Higdson, D., M. Kennedy, J. C. Cavendish, J. A. Cafeo, and R. D. Ryne, 2004: Combining field data and computer simulations for calibration and prediction. *SIAM J. Sci. Comput.*, **26**, 448–466, <https://doi.org/10.1137/S1064827503426693>.
- Hill, A. J., C. C. Weiss, and B. C. Ancell, 2016: Ensemble sensitivity analysis for mesoscale forecasts of dryline convection initiation. *Mon. Wea. Rev.*, **144**, 4161–4182, <https://doi.org/10.1175/MWR-D-15-0338.1>.
- Homma, T., and A. Saltelli, 1996: Importance measures in global sensitivity analysis of nonlinear models. *Reliab. Eng. Syst. Saf.*, **52**, 1–17, [https://doi.org/10.1016/0951-8320\(96\)00002-6](https://doi.org/10.1016/0951-8320(96)00002-6).
- Hossain, F., E. N. Anagnostou, and A. C. Bagtzoglou, 2006: On Latin Hypercube sampling for efficient uncertainty estimation of satellite rainfall observations in flood prediction. *Comput. Geosci.*, **32**, 776–792, <https://doi.org/10.1016/j.cageo.2005.10.006>.
- Iman, R. L., and W. J. Conover, 1980: Small sample sensitivity analysis techniques for computer models with an application to risk assessment. *Commun. Stat. Theory Methods*, **9**, 1749–1842, <https://doi.org/10.1080/03610928008827996>.
- Iooss, B., and P. Lemaître, 2015: A review on global sensitivity analysis methods. *Uncertainty Management in Simulation-Optimization of Complex Systems*, G. Dellino and C. Meloni, Eds., Operations Research/Computer Science Interfaces Series, Vol. 59, Springer, 101–122, [https://doi.org/10.1007/978-1-4899-7547-8\\_5](https://doi.org/10.1007/978-1-4899-7547-8_5).
- Isukapalli, S. S., and P. G. Georgopoulos, 2001: Computational methods for sensitivity and uncertainty analysis for environmental and biological models. Rep. EPA/600/R-01-068, U.S. Environmental Protection Agency, Washington, DC, 158 pp.
- Kennedy, M. C., and A. O'Hagan, 2001: Bayesian calibration of computer models. *Roy. Stat. Soc.*, **63**, 425–464, <https://doi.org/10.1111/1467-9868.00294>.

- Knutson, T. R., and Coauthors, 2013: Dynamical downscaling projections of twenty-first-century Atlantic hurricane activity: CMIP3 and CMIP5 model-based scenarios. *J. Climate*, **26**, 6591–6617, <https://doi.org/10.1175/JCLI-D-12-00539.1>.
- Lee, L. A., K. S. Carslaw, K. J. Pringle, G. W. Mann, and D. V. Spracklen, 2011: Emulation of a complex global aerosol model to quantify sensitivity to uncertain parameters. *Atmos. Chem. Phys.*, **11**, 12 253–12 273, <https://doi.org/10.5194/acp-11-12253-2011>.
- , K. J. Pringle, C. L. Reddington, G. W. Mann, P. Stier, D. V. Spracklen, J. R. Pierce, and K. S. Carslaw, 2013: The magnitude and causes of uncertainty in global model simulations of cloud condensation nuclei. *Atmos. Chem. Phys.*, **13**, 8879–8914, <https://doi.org/10.5194/acp-13-8879-2013>.
- Logemann, K., J. O. Backhaus, and I. H. Harms, 2004: SNAC: A statistical emulator of the north-east Atlantic circulation. *Ocean Modell.*, **7**, 97–110, [https://doi.org/10.1016/S1463-5003\(03\)00039-8](https://doi.org/10.1016/S1463-5003(03)00039-8).
- Loh, W. L., 1996: On latin hypercube sampling. *Ann. Stat.*, **24**, 2058–2080, <https://doi.org/10.1214/aos/1069362310>.
- MacDonald, B., P. Ranjan, and H. Chipman, 2015: GPfit: An R Package for fitting a Gaussian process model to deterministic simulator outputs. *J. Stat. Software*, **64** (12), <https://doi.org/10.18637/jss.v064.i12>.
- Manganello, J. V., and Coauthors, 2012: Tropical cyclone climatology in a 10-km global atmospheric GCM: Toward weather-resolving climate modeling. *J. Climate*, **25**, 3867–3893, <https://doi.org/10.1175/JCLI-D-11-00346.1>.
- Marzban, C., 2013: Variance-based sensitivity analysis: An illustration on the Lorenz'63 model. *Mon. Wea. Rev.*, **141**, 4069–4079, <https://doi.org/10.1175/MWR-D-13-00032.1>.
- , S. Sandgathe, and J. D. Doyle, 2014: Model tuning with canonical correlation analysis. *Mon. Wea. Rev.*, **142**, 2018–2027, <https://doi.org/10.1175/MWR-D-13-00245.1>.
- McCaul, E. W., 1991: Buoyancy and shear characteristics of hurricane-tornado environments. *Mon. Wea. Rev.*, **119**, 1954–1978, [https://doi.org/10.1175/1520-0493\(1991\)119<1954:BASCOH>2.0.CO;2](https://doi.org/10.1175/1520-0493(1991)119<1954:BASCOH>2.0.CO;2).
- McKay, M. D., R. J. Beckman, and W. J. Conover, 2000: A comparison of three methods for selecting values of input variables in the analysis of output from a computer code. *Technometrics*, **42**, 55–61, <https://doi.org/10.1080/00401706.2000.10485979>.
- McLay, J. G., and M. Liu, 2014: Detecting dependence in the sensitive parameter space of a model using statistical inference and large forecast ensembles. *Mon. Wea. Rev.*, **142**, 3734–3755, <https://doi.org/10.1175/MWR-D-13-00340.1>.
- Murakami, H., R. Mizuta, and E. Shindo, 2012: Future changes in tropical cyclone activity projected by multi-physics and multi-SST ensemble experiments using the 60-km-mesh MRI-AGCM. *Climate Dyn.*, **39**, 2569–2584, <https://doi.org/10.1007/s00382-011-1223-x>.
- Murphy, J. M., D. M. H. Sexton, D. N. Barnett, G. S. Jones, M. J. Webb, M. Collins, and D. A. Stainforth, 2004: Quantification of modelling uncertainties in a large ensemble of climate change simulations. *Nature*, **430**, 768–772, <https://doi.org/10.1038/nature02771>.
- Neale, R. B., and Coauthors, 2010: Description of the NCAR Community Atmosphere Model (CAM5.0). NCAR Tech. Note NCAR/TN-486+STR, 268 pp., [www.cesm.ucar.edu/models/cesm1.0/cam/docs/description/cam5\\_desc.pdf](http://www.cesm.ucar.edu/models/cesm1.0/cam/docs/description/cam5_desc.pdf).
- Nolan, D. S., and E. D. Rappin, 2008: Increased sensitivity of tropical cyclogenesis to wind shear in higher SST environments. *Geophys. Res. Lett.*, **35**, L14805, <https://doi.org/10.1029/2008GL034147>.
- O'Hagan, A., M. C. Kennedy, and J. E. Oakley, 1999: Uncertainty analysis and other inference tools for complex computer codes. *Bayesian Stat.*, **6**, 503–524.
- Poole, D., and A. E. Raftery, 2000: Inference for deterministic simulation models: The Bayesian melding approach. *J. Amer. Stat. Assoc.*, **95**, 1244–1255, <https://doi.org/10.1080/01621459.2000.10474324>.
- Posselt, D. J., 2016: A Bayesian examination of deep convective squall-line sensitivity to changes in cloud microphysical parameters. *J. Atmos. Sci.*, **73**, 637–665, <https://doi.org/10.1175/JAS-D-15-0159.1>.
- , and T. Vukicevic, 2010: Robust characterization of model physics uncertainty for simulations of deep moist convection. *Mon. Wea. Rev.*, **138**, 1513–1535, <https://doi.org/10.1175/2009MWR3094.1>.
- , and C. H. Bishop, 2012: Nonlinear parameter estimation: Comparison of an ensemble Kalman smoother with a Markov chain Monte Carlo algorithm. *Mon. Wea. Rev.*, **140**, 1957–1974, <https://doi.org/10.1175/MWR-D-11-00242.1>.
- , D. Hodyss, and C. H. Bishop, 2014: Errors in ensemble Kalman smoother estimates of cloud microphysical parameters. *Mon. Wea. Rev.*, **142**, 1631–1654, <https://doi.org/10.1175/MWR-D-13-00290.1>.
- , B. Fryxell, A. Molod, and B. Williams, 2016: Quantitative sensitivity analysis of physical parameterizations for cases of deep convection in the NASA GEOS-5. *J. Climate*, **29**, 455–479, <https://doi.org/10.1175/JCLI-D-15-0250.1>.
- Ranjan, P., R. Haynes, and R. Karsten, 2011: A computationally stable approach to Gaussian process interpolation of deterministic computer simulation data. *Technometrics*, **53**, 366–378, <https://doi.org/10.1198/TECH.2011.09141>.
- Ranson, M., C. Kousky, M. Ruth, L. Jantarasami, A. Crimmins, and L. Tarquinio, 2014: Tropical and extratropical cyclone damages under climate change. *Climatic Change*, **127**, 227–241, <https://doi.org/10.1007/s10584-014-1255-4>.
- Reed, K. A., and C. Jablonowski, 2011a: An analytic vortex initialization technique for idealized tropical cyclone studies in AGCMs. *Mon. Wea. Rev.*, **139**, 689–710, <https://doi.org/10.1175/2010MWR3488.1>.
- , and —, 2011b: Impact of physical parameterizations on idealized tropical cyclones in the Community Atmosphere Model. *Geophys. Res. Lett.*, **38**, L04805, <https://doi.org/10.1029/2010GL046297>.
- , and —, 2011c: Assessing the uncertainty in tropical cyclone simulations in NCAR's Community Atmosphere Model. *J. Adv. Model. Earth Syst.*, **3**, M08002, <https://doi.org/10.1029/2011MS000076>.
- , and —, 2012: Idealized tropical cyclone simulations of intermediate complexity: A test case for AGCMs. *J. Adv. Model. Earth Syst.*, **4**, M04001, <https://doi.org/10.1029/2011MS000099>.
- Reynolds, C. A., J. D. Doyle, and X. Hong, 2016: Examining tropical cyclone–Kelvin wave interactions using adjoint diagnostics. *Mon. Wea. Rev.*, **144**, 4421–4439, <https://doi.org/10.1175/MWR-D-16-0174.1>.
- Robbins, M. W., R. B. Lund, C. M. Gallagher, and Q. Lu, 2011: Change points in the North Atlantic tropical cyclone record. *J. Amer. Stat. Assoc.*, **106**, 89–99, <https://doi.org/10.1198/jasa.2011.ap10023>.
- Sacks, J., W. J. Welch, T. J. Mitchell, and H. P. Wynn, 1989: Design and analysis of computer experiments. *Stat. Sci.*, **4**, 409–423, <https://doi.org/10.1214/ss/1177012413>.



- Sanso, B., and C. Forest, 2009: Statistical calibration of climate system properties. *J. Roy. Stat. Soc.*, **58**, 485–503, <https://doi.org/10.1111/j.1467-9876.2009.00669.x>.
- , —, and D. Zantedeschi, 2008: Inferring climate system properties using a computer model. *Bayesian Anal.*, **3**, 1–62.
- Saunders, M. A., and A. S. Lea, 2008: Large contribution of sea surface warming to recent increase in Atlantic hurricane activity. *Nature*, **451**, 557–560, <https://doi.org/10.1038/nature06422>.
- Sobol, I. M., 1993: Sensitivity estimates for nonlinear mathematical models. *Math. Model. Comput. Exp.*, **1**, 407–414.
- , 2005: Global sensitivity indices for nonlinear mathematical models and their Monte Carlo estimates. *Math. Comput. Simul.*, **55**, 271–280, [https://doi.org/10.1016/S0378-4754\(00\)00270-6](https://doi.org/10.1016/S0378-4754(00)00270-6).
- Stensrud, D. J., J.-W. Bao, and T. T. Warner, 2000: Using initial condition and model physics perturbations in short-range ensemble simulations of mesoscale convective systems. *Mon. Wea. Rev.*, **128**, 2077–2107, [https://doi.org/10.1175/1520-0493\(2000\)128<2077:UICAMP>2.0.CO;2](https://doi.org/10.1175/1520-0493(2000)128<2077:UICAMP>2.0.CO;2).
- Strouine, K., S. Kravtsov, D. Kondrashov, and M. Ghil, 2010: Reduced models of atmospheric low-frequency variability: Parameter estimation and comparative performance. *Physica D*, **239**, 145–166, <https://doi.org/10.1016/j.physd.2009.10.013>.
- Taylor, K. E., R. J. Stouffer, and G. A. Meehl, 2012: An overview of CMIP5 and the experiment design. *Bull. Amer. Meteor. Soc.*, **93**, 485–498, <https://doi.org/10.1175/BAMS-D-11-00094.1>.
- Tokmakian, R., P. Challenor, and Y. Andrianakis, 2012: On the use of emulators with extreme and highly nonlinear geophysical simulators. *J. Atmos. Oceanic Technol.*, **29**, 1704–1715, <https://doi.org/10.1175/JTECH-D-11-00110.1>.
- Tong, M., and M. Xue, 2008: Simultaneous estimation of microphysical parameters and atmospheric state with simulated radar data and ensemble square root Kalman filter. Part I: Sensitivity analysis and parameter identifiability. *Mon. Wea. Rev.*, **136**, 1630–1648, <https://doi.org/10.1175/2007MWR2070.1>.
- Torn, R. D., and G. J. Hakim, 2008: Ensemble-based sensitivity analysis. *Mon. Wea. Rev.*, **136**, 663–677, <https://doi.org/10.1175/2007MWR2132.1>.
- Tushaus, S. A., D. J. Posselt, M. M. Miglietta, R. Rotunno, and L. Delle Monache, 2015: Bayesian exploration of multivariate orographic precipitation sensitivity for moist stable and neutral flows. *Mon. Wea. Rev.*, **143**, 4459–4475, <https://doi.org/10.1175/MWR-D-15-0036.1>.
- van den Heever, S. C., and W. R. Cotton, 2004: The impact of hail size on simulated supercell storms. *J. Atmos. Sci.*, **61**, 1596–1609, [https://doi.org/10.1175/1520-0469\(2004\)061<1596:TIOHSO>2.0.CO;2](https://doi.org/10.1175/1520-0469(2004)061<1596:TIOHSO>2.0.CO;2).
- van der Merwe, R., T. K. Leen, Z. Lu, S. Frolov, and A. M. Baptista, 2007: Fast neural network surrogates for very high dimensional physics-based models in computational oceanography. *Neural Networks*, **20**, 462–478, <https://doi.org/10.1016/j.neunet.2007.04.023>.
- Villarini, G., G. A. Vecchi, and J. A. Smith, 2010: Modeling the dependence of tropical storm counts in the North Atlantic basin on climate indices. *Mon. Wea. Rev.*, **138**, 2681–2705, <https://doi.org/10.1175/2010MWR3315.1>.
- Walsh, K. J., and Coauthors, 2015: Hurricanes and climate: The U.S. CLIVAR working group on hurricanes. *Bull. Amer. Meteor. Soc.*, **96**, 997–1017, <https://doi.org/10.1175/BAMS-D-13-00242.1>.
- Wehner, M. F., and Coauthors, 2014: The effect of horizontal resolution on simulation quality in the Community Atmospheric Model, CAM5.1. *J. Adv. Model. Earth Syst.*, **6**, 980–997, <https://doi.org/10.1002/2013MS000276>.
- Williams, B., D. Higdon, J. Gattiker, L. Moore, M. McKay, and S. Keller-McNulty, 2006: Combining experimental data and computer simulations, with an application to flyer plate experiments. *Bayesian Anal.*, **1**, 765–792, <https://doi.org/10.1214/06-BA125>.
- Zarzycki, C. M., and C. Jablonowski, 2014: A multidecadal simulation of Atlantic tropical cyclones using a variable-resolution global atmospheric general circulation model. *J. Adv. Model. Earth Syst.*, **6**, 805–828, <https://doi.org/10.1002/2014MS000352>.
- , and P. A. Ullrich, 2017: Assessing sensitivities in algorithmic detection of tropical cyclones in climate data. *Geophys. Res. Lett.*, **44**, 1141–1149, <https://doi.org/10.1002/2016GL071606>.
- Zeng, Z., Y. Wang, and C.-C. Wu, 2007: Environmental dynamical control of tropical cyclone intensity—An observational study. *Mon. Wea. Rev.*, **135**, 38–59, <https://doi.org/10.1175/MWR3278.1>.
- Zhao, C., and Coauthors, 2013: A sensitivity study of radiative fluxes at the top of atmosphere to cloud-microphysics and aerosol parameters in the Community Atmosphere Model CAM5. *Atmos. Chem. Phys.*, **13**, 10 969–10 987, <https://doi.org/10.5194/acp-13-10969-2013>.
- Zhao, M., I. M. Held, S. J. Lin, and G. A. Vecchi, 2009: Simulations of global hurricane climatology, interannual variability, and response to global warming using a 50-km resolution GCM. *J. Climate*, **22**, 6653–6678, <https://doi.org/10.1175/2009JCLI3049.1>.
- Živković, M., J. Louis, and J. Moncet, 1995: Sensitivity analysis of a radiation parameterization. *J. Geophys. Res.*, **100**, 13 827–13 840, <https://doi.org/10.1029/95JD00983>.