Abstracts for Posters, 2014 STATMOS Annual Meeting

Presenter: Won Chang, University of Chicago and Argonne National Laboratory

Title: A conditional simulation approach to future precipitation scenario generation

Abstract: Future precipitation scenarios are important input for the development of adaptation strategy for climate change. Climate models allow a physically consistent simulation of future precipitation changes, but their application for future scenario generation is complicated due to their inability to reproduce the current precipitation patterns. Our goal is to propose a conditional simulation approach, which modifies the observed precipitation based on the changes suggested by climate models. In order to capture the highly-localized behavior of precipitation, we use high-resolution ensemble from the WRF regional climate model to derive the changes in precipitation patterns. This problem poses nontrivial computational and inferential challenges due to the non-Gaussian and high-dimensional nature of the data.

Presenter: Zhen Zhang, University of Chicago and Argonne National Laboratory

Title: Modeling Large-scale Nonstationary Gaussian Processes with Dependent Random Partitions

Abstract: For large-scale spatial and spatio-temporal data, it is particularly useful to randomly partition the large study domain into a group of smaller regions using Bayesian techniques such as treed partition or Voronoi Tessellations. With considerable efforts on modeling the within-subregion spatial dependence, the dependence between subregions under a partition is conventionally ignored, partially for fully employing the parallel computing techniques for fast computation. In this study, we consider modeling the nonstationary Gaussian processes with local means and smooth, anisotropic covariance functions. We introducing the dependence structure between subregions by defining the boundaries, and compare the performance to the partitioning models with block-independence assumption via a series of examples. Simulation results suggest that under certain circumstances, the model fitting can be improved by allowing dependence between subregions, while maintaining the computational efficiency. We also discuss the model implementation with partition/boundary search, and the prior specifications.
The proposed methodologies can readily accommodate extremely large-scale data by recursively defining such partitions.

Presenter: Andrew Poppick, University of Chicago

Title: Time-Varying Spectral Methods for Modeling Projected Future Changes in Temperature Variability and Producing Data-Driven Simulations of Future Transient Climates

Abstract: Impacts assessments for future climate change depend on changes in temporal variability in addition to changes in mean climate. General circulation models (GCMs) do predict changes in variability, but do not reproduce the means or variability of observed climate. Our work addresses the ensuing need for climate simulations that combine observational data with GCM projections of changes in both mean and variability. Leeds et al. (2014) introduced a methodology for modifying existing temperature observations based on GCM projections of changes in means and covariances between two equilibrium (stationary) climates under different forcing conditions. Here, we extend that methodology to account for GCM projections of transient (nonstationary) climates. The method relies on the time-varying spectral representation of nonstationary processes that are locally approximately stationary. Other existing methods can account for changes in mean climate (the “Delta method”) or model-observation discrepancy in marginal probability distributions (quantile-based methods), but a key advantage of the methodology proposed here is that it accounts for changes in temporal dependence, and therefore in different changes in variability at distinct timescales. Our methodology can be combined with climate model emulation to study changes in unmodeled forcing scenarios and impacts under these scenarios, assuming that the changes in covariance structure can be described by a simple function of the forcing scenario. We illustrate our methodology with runs from the NCAR CCSM3 model, comparing daily temperatures under continuously increasing CO2 forcing scenarios with temperatures from a run forced at preindustrial conditions. In most regions, temperature’s overall temporal variability decreases as temperature increases, and furthermore, its temporal dependence structure changes in time.

Presenter: Michael Horrell, University of Chicago

Title: Half-spectral space-time covariance models and a natural condition

We develop two new classes of space-time Gaussian process models by specifying covariance structure, $K$, using what we call a half-spectral representation
of $K$. Half-spectral representations are a special case of spectral representations and have been studied previously by Cressie and Huang (1999), Gneiting (2002) and Stein (2005). Here, we focus on developing desirable theoretical properties of certain half-spectral forms. In particular, for a model, $K$, we determine mean-square differentiability of a Gaussian process governed by $K$, and we determine whether or not $K$ meets a natural condition posed by Stein (2011) which in some cases guarantees nice interpolation properties in $K$. We fit models we develop in this paper to the Irish wind dataset first analyzed in Haslett and Raftery (1989), and we show our models fit these data better than other separable and non-separable space-time models developed in Cressie and Huang (1999) and Gneiting (2002).

**Presenter:** Charlotte L. Haley, Argonne National Laboratory  
**Title:** A simple method for the estimation and detection of individual solar oscillations  
**Abstract:** The study of the systematic oscillations of the Sun, or solar modes, has lead to better understanding of the Sun’s inner structure and dynamics. Estimates of individual mode frequencies are desirable as inputs for solar inversion models and have been obtained in the literature using either least-squares models fit to spectrum estimates or directly to Fourier-transformed helioseismic data. Recent studies have concluded that solar modal structure remains coherent past turbulence in the convection zone and imprints its signatures on the solar wind and the interplanetary magnetic field fluctuations, so data from these sources can be expected to contain several hundreds of modes in very short bands in frequency. Estimation of modal parameters from these data pushes the limits of conventional spectrum-based approaches. An alternative parametric approach described here uses maximum likelihood estimation on a simple Lorentzian model with four parameters: $A$, the integrated power of the mode, $a$ the damping constant, $f_0$ the center frequency and $\sigma^2$ the noise power. Assuming a Gaussian process model, the likelihood can be simplified using a Karhunen-Loève expansion where the associated eigenvalues and eigencoefficients have convenient analytic expressions. This method may also be of interest for modeling terrestrial free oscillations.  

**Presenter:** Jiali Wang, Argonne National Laboratory  
**Title:** Model performance in spatio-temporal patterns of precipitation: New methods for identifying value added by a regional climate model  
**Abstract:** Regional climate models (RCMs) are a standard tool for down-
scaling climate forecasts to finer spatial scales. The evaluation of RCMs against observational data is an important step in building confidence in the use of RCMs for future prediction. In addition to model performance in climatological means and marginal distributions, a model’s ability to capture spatio-temporal relationships is important. This study develops two approaches: (1) spatial correlation/variogram for a range of spatial lags, with total monthly precipitation and non-seasonal precipitation components used to assess the spatial variations of precipitation; and (2) spatio-temporal correlation for a wide range of distances, directions, and time lags, with daily precipitation occurrence used to detect the dynamic features of precipitation. These measures of spatial and spatio-temporal dependence are applied to a high-resolution RCM run and to the National Center for Environmental Prediction (NCEP)-U.S. Department of Energy (DOE) AMIP II reanalysis data (NCEP-R2), which provides initial and lateral boundary conditions for the RCM. The RCM performs better than NCEP-R2 in capturing both the spatial variations of total and non-seasonal precipitation components and the spatio-temporal correlations of daily precipitation occurrences. The improvements are apparent not just at resolutions finer than that of NCEP-R2, but also when the RCM and observational data are aggregated to the resolution of NCEP-R2.

Presenter: Whitney Huang, Purdue University
Title: Climate extremes in CCSM3 under different CO2 concentrations
Abstract: In a context of global warming, it is of great interest to determine how the tail behavior of temperature and precipitation will affect by the increasing CO2 concentrations. However, this inquiry is complicated by the fact that 1) the observational data records are not long enough to reliably detect the changes 2) the current climate is under a transient state. In this study, we focus on the 1000-year model output from the Community Climate System Model version 3 (CCSM3) under three different CO2 concentrations in the equilibrium climate. The Generalized Extreme Value (GEV) distribution is fitted using the block maxima approach. A comparison of quantile functions of the fitted extreme value distributions among different CO2 scenarios at US was performed. The results suggest that the tail behavior of both temperature and precipitation may change especially in the in-land areas.

Presenter: Mark Risser, Ohio State University
Title: Regression-Based Covariance Functions for Nonstationary Spatial Modeling

Abstract: Much of the recent literature on Gaussian process models for spatial data focuses on nonstationary methods, but many of these more appropriate models suffer from difficulties in model fitting, parameter identification, and interpretability. To overcome these issues, we build on the growing literature of covariate-driven nonstationary spatial modeling. Using process convolution techniques, we propose a Bayesian model for continuously-indexed spatial data based on a flexible covariance regression structure for a convolution-kernel covariance matrix. Unlike current approaches, our approach constrains the components of the convolution-kernel matrix to vary smoothly over space according to spatially-varying covariate information. We explore the properties of the implied model, including a description of the implied nonstationary covariance function and the interpretational benefits in the kernel parameters, and demonstrate that our parsimonious model provides a compromise between stationary and overly parameterized nonstationary models that do not perform well in practice. We illustrate our approach through simulation and an analysis of precipitation data.

Presenter: Shivi Vaidyanathan, Ohio State University

Title: Hierarchical Bayesian Prediction of Coastal Sea Levels

Abstract: We present a Bayesian hierarchical model of coastal sea levels using data from 1107 sites distributed across the globe. The data are monthly averages obtained from tide gauge measurements at each site. The length of the time series is different for each site, with the earliest records starting from January 1850 and the latest going through December 2013. The model assumes that the data at each site are distributed according to a statistical time series model with location specific model parameters. The sites are linked by assuming that the site specific parameters are samples from a regional prior distribution which in turn have parameters that are samples from a global prior. The model uses monthly hemispheric temperatures as a covariate. Hence, it can be used to make forecasts of future sea levels using forecasts of future temperature. We test the performance of the model by holding out data from the last 10 years and then comparing predicted sea levels from the model with observed sea levels.

Presenter: Oksana Chkrebtii, Ohio State University

Title: Quantifying Numerical Uncertainty in Dynamical Climate Models
Abstract: Climate models can be very sensitive to small perturbations, including those arising from truncation error in the numerical solution of systems of ordinary or partial differential equations defining the system states. We present a new formalism for characterizing and propagating this source of uncertainty through the statistical inverse problem of inference. Our Bayesian approach is illustrated on classical chaotic systems, and allows the trade-off between accuracy and computational load arising from the choice of discretization grid in a principled way.

Presenter: Dorit Hamerling, NCAR
Title: A Hierarchical Bayesian Model for Climate Change Detection and Attribution

Abstract: Climate change detection and attribution methodology has evolved substantially within the climate community over the last decades and has started to attract the attention of statisticians. We propose to add to this development by presenting a Bayesian hierarchical statistical model for regression-based climate-change detection and attribution. The proposed model addresses some of the challenges of current approaches, such as high-dimensional covariance estimation and how to optimally choose the truncation of basis function expansions. Another feature of the proposed model is the comprehensive treatment and propagation of uncertainties from all components of the model. The methodology is applied to global temperature-change data in the lower troposphere from CMIP5 climate models and a large ensemble of satellite observations.

Presenter: Robert Yuen, University of Michigan
Title: A Gauss-Pareto process model for spatial prediction of extreme precipitation

Abstract: In order to develop adaptive strategies for dealing with consequences of extreme precipitation such as insufficient drainage and various aspects of flooding, it is necessary to be able to estimate extremes at unobserved sites. We introduce a hierarchical Gauss-Pareto model for spatial prediction of precipitation given nearby observations that are extreme. The model belongs to the max-domain of attraction of popular Brown-Resnick max-stable processes (Brown and Resnick, 1977; Kabluchko et al., 2009) and retains the essential dependence structure of their corresponding generalized Pareto processes (Ferreira and DeHaan, 2012). An MCMC algorithm is developed for inference. The algorithm allows for left censored data from pre-
cipitation that accumulates below instrument precision, which often happens despite nearby observations that are extreme. The model and methodology is applied to summer 24 hour cumulative precipitation over south central Sweden. We discuss some extensions and challenges for future work.

Presenter: Frederick Bingham, UNC Wilmington
Title: Comparison of Aquarius and Argo sea surface salinity across the statistical distribution
Abstract: Using Argo in situ and Aquarius version 3.0 sea surface salinity, we compare the measured salinity across a range of distributions. In addition to comparison between mean or median values as is traditionally done, we compare the datasets at the 1%, 10%, 25%, 50%, 75%, 90% and 99% quantiles of the statistical distributions. The datasets compare well at the median, which is to be expected since the Aquarius retrieval algorithm is based on this. The datasets are less similar at the ends of the distributions, especially in the lower tail. In general, Argo data are much peakier, with more of the observations concentrated in the middle of the distribution. This is true across seasons, ocean basins and hemispheres. Aquarius is better able to capture the low and high outliers which are most prevalent during periods when the surface salinity is rapidly changing.

Presenter: Fei He, University of Michigan
Title: Assessing the Tropical Cyclone Characteristics Simulated in AGCMs
Abstract: This work explores the utility of proven quantitative sensitivity analysis techniques for analysis of dynamical systems in atmospheric general circulation models (AGCMs). Specifically, it uses an Extended Multivariate Adaptive Regression Splines (EMARS) algorithm to examine the response of cloud radiative forcing, cloud content, and precipitation rate in idealized AGCM-simulated tropical cyclones (TCs) to changes in TC initial conditions and environment. The use of a multivariate sensitivity analysis algorithm allows the simultaneous examination of multiple controlling factors and their interactions. Use of a well-established idealized TC test case framework (1) allows for a thorough evaluation of the sensitivity analysis algorithm, and (2) extends the test case methodology to a quantitative and multivariate examination of dynamical system sensitivity. Consistent with previous coarse and fine resolution idealized and case study analyses of TCs, the EMARS algorithm yields increases in TC intensity for higher sea surface temperature and larger atmospheric lase rates. However, multivariate perturbation of initial
conditions and environment allow a more nuanced exploration of the sensitivities in the system. It is found that: (1) atmospheric lapse rate and sea surface temperature behave strong but different impact on the TC characteristics. Larger SST is beneficial for all the TC characteristics, such as TC intensity, shortwave cloud radiative forcing (SWCF), and precipitation rate. In contrast, larger atmospheric lapse rate is beneficial for TC intensity, longwave cloud radiative forcing, and cloud ice water path (IWP), but harmful for SWCF, cloud liquid water path (LWP) and precipitation; (2) Most TC characteristics show very weak sensitivity to the size and strength of initial disturbance except SWCF and IWP; (3) TC intensity shows linear relationship with TC LWCF, but no dependence on SWCF. More relationships are quantified and the implication is discussed.

Presenter: Xian He, Purdue University
Title: Calibration of an Ecosystem Model
Abstract: An ecosystem model is a representation of the real complex ecological system, and is usually described by sophisticated mathematical models. Terrestrial Ecosystem Model (TEM) is one of the ecosystem models, that describes the dynamics of carbon, nitrogen, soil, water and other vegetation related variables. Dynamic computer model with unknown parameters is developed from the differential equations of TEM. The process of estimating those unknown parameters is called calibration which is essential for following application and forecasting using the model. Besides data assimilation methods which are usually used by environmental scientists, statisticians have already developed Bayesian calibration method to calibrate complex computer models. However, the limit to the existing calibration framework is the application on dynamic computer model when the output is time series. In this work, we develop an efficient calibration method for dynamic computer model with application on Terrestrial Ecosystem model.

Presenter: Piyas Chakraborty
Title: Modeling the Frequencey and Magnitude of Extreme Events
Abstract: In environmental data analysis, the occurrence of extreme values and their intensities can be modeled by two-dimensional Poisson processes. This model is particularly useful in formulating non-stationarity in threshold excesses. To tackle non-uniformity of exceedance times, a time-varying threshold is usually used which is often subjective. A general model to separately handle the frequency and magnitude of extreme events will be dis-
cussed in this poster. This new approach is capable of modeling events that become more extreme in magnitude but less frequent in time or vice versa. Some simulated and real-life data are analyzed using the proposed model to illustrate its usefulness.

Presenter: Yen-Ning Huang, Purdue University
Title: Empirical Bayesian Modeling of Bivariate Intensity-marked Log Gaussian Cox processes
Abstract: Spatial marked point processes describe the locations of randomly distributed events in a region, with a measured quantity associated with each observed point. The simplest assumption is that marks and points are independent. However, most of the time this is not true in practice. In this work we focus on intensity-dependent spatial marked point processes. We consider both univariate and bivariate intensity marked Log Gaussian Cox processes and apply an empirical Bayesian estimation procedure with implementation of Markov Chain Monte Carlo methodology for statistical inference. We apply this approach to a limited range white oak dataset.

Presenter: Paul D. Sampson, University of Washington
Title: A partial warp parameterization for the spatial deformation model for nonstationary covariance
Abstract: The spatial deformation model introduced by Sampson and Guttorp has been around for a long time, but it (and most of the other modeling strategies for nonstationary covariance) have suffered from a lack of flexible and convenient software. A simple reparameterization in terms of partial warps of the thin-plate spline approach to spatial deformations provides a convenient framework for more efficient numerical estimation, strategies for dimension reduction, and application to larger problems than has been practical. We explain the nature of the partial warp decomposition of splines and present preliminary illustrations of the fitting of deformations for simulated spatio-temporal and spatial datasets. We note a number of research directions and extensions for this work in progress.

Presenters: Joanne Lo, Elizabeth Mannshardt, North Carolina State University
Title: Spatial and Temporal Modeling of Extreme Precipitation Events
Abstract: Quantification of heavy precipitation is crucial for prevention and relief of disaster events such as flooding which could have economic impacts
such as crop production disruptions and infrastructure damages. Although
extreme value theory has been well developed for modeling univariate ex-
tremes, capturing spatial dependency and characterizing the joint distribu-
tion of extreme values is essential. This is demonstrated using National
Climatic Data Center’s COOP data on historical daily observations of pre-
cipitation between 1949 and 2008. Initial results from fitting the univariate
Generalized Extreme Value (GEV) model marginally suggest that the uni-
variate GEV model captures the general trend of extreme precipitation but
misses some of the more extreme events, which are central to prevention
planning. In addressing the scientific question of whether extreme precipi-
tation events are increasing over time, the models also signify the need to
incorporate spatial dependency. Recent methods address the issue of spatial
dependency, but limit the use of all available data due to restrictions in high
dimensions. To capitalize on the spatial information in surrounding sites, we
are developing methods to incorporate neighboring data in a Conditional Au-
toregressive (CAR) model framework while preserving the GEV marginals at
each location. Using this approach the model gains strength from additional
data, sites with missing time series values can utilize neighboring data, and
computational cost can be reduced.

Presenter: Marcela Alfaro-Córdoba, North Carolina State University
Title: Hierarchical Modeling of Spatial-Temporal Tropical Cyclone Occur-
rences with Application to Seasonal Cyclone Forecasting
Abstract: Tropical cyclones can be classified as tropical storms, mild hurri-
canes, and strong hurricanes. Because of the threat they represent to pop-
ulated areas, it is crucial to understand the relationship between different
variables and the number of cyclones according to their intensity and lo-
cation. We introduce a statistical framework to forecast the probability of
occurrence and the intensity of a tropical cyclone at any spatial location
in the Atlantic Basin. The different intensities are modeled simultaneously
across space and time as a function of sea surface temperatures and latent
heat fluxes. We work with massive remote sensing data and computer model
output to characterize sea surface temperatures and heat fluxes. A hierarchi-
cal Bayesian spatial model is introduced to model dependence between the
proportion of tropical cyclones for each response, time and location and the
correspondent covariates.

Presenter: Kimberly Kaufeld, SAMSI/North Carolina State University
Title: Estimating spatial impacts on mortality from climate change
Studies on mortality from exposure to air pollution and extreme weather events under changing climatic conditions use deterministic model output for future projections of the environmental stressors. Prior modeling frameworks have integrated data from climate model outputs, historical meteorology and air pollution observations using spatial quantile functions. A quantile regression approach characterizes the tails of future climate projections. Calibrated climate model data on future projections can be used to look at the impact on risk assessment such as mortality from air pollution. We explain how calibrated climate model output using Bayesian spatial quantile regression can be used to provide an approach to studying the spatio-temporal associations of risk assessment from air pollution, while characterizing potential sources of risks associated with climate change. In addition, we mention how spatial modeling and prediction of different climate and mortality can incorporate chemistry models, monitoring data, meteorological and other spatio-temporal variables and confounding influences, to take into account different sources of uncertainty in the data.