Successful Preliminary Exam Prospectus Samples

Colorado State University Department of Atmospheric Science

1. Measuring Net Ammonia Fluxes in the U.S. Midwest Corn Belt
2. GREMLIN: GOES Radar Estimation via Machine Learning to Inform NWP
3. Environmental Controls on Tropical Land-Aerosol-Cloud Interactions

Research Prospectus

Measuring Net Ammonia Fluxes in the U.S. Midwest Corn Belt

Jakob Lindaas

Department of Atmospheric Science
Colorado State University
Fort Collins, CO

Advisor
Emily V. Fischer

Committee
Jeffrey L. Collett Jr.
A.R. Ravishankara
Shantanu H. Jathar

submitted October 16th, 2018

1. Introduction and Research Questions

Within the last century, humans have significantly perturbed the global nitrogen cycle (Galloway et al., 2008). Sparked by the discovery and refinement of the Haber-Bosch process for artificially fixing atmospheric N_2 and driven by the need to feed a growing world population, humans have more than doubled the quantity of reactive nitrogen (defined as nitrogen-containing compounds that are chemically, biologically or radiatively active: N_r) cascading through the earth system (Fowler et al., 2013). Understanding and managing the impact of these anthropogenic perturbations to the global nitrogen cycle has been described as a "grand challenge" for humanity in the next few decades (Reis et al. 2016). As additional reactive nitrogen has been added to the earth system, fluxes of nitrogen across ecosystem boundaries have likewise increased (Vitousek et al., 1997). One of the primary conduits for these fluxes is through the atmosphere. Reactive nitrogen is most often emitted into the atmosphere as ammonia (NH_3), nitric oxide (NO), or nitrous oxide (N2O) with NH_3 emissions accounting for roughly half the total (Flechard et al., 2013). Of these NH_3 emissions, agricultural emissions comprise around two thirds, with the rest of global NH_3 emissions coming from natural soils, oceans, biomass burning, and transportation. The increasing contribution from agriculture to NH_3 emissions is driven largely by agricultural intensification of livestock and crop farming, involving the use of machinery, fertilizers, pesticides and other means to maximize yields from available land and animals.

Atmospheric NH_3 has a number of important impacts on human and earth environments. NH_3 is the most abundant gas-phase base in the atmosphere, and neutralizes acidic gases such as nitric acid or sulfuric acid. Compounds such as ammonium sulfate and ammonium nitrate exist primarily in the solid and liquid phases, and contribute to particle nucleation and growth, thus positioning NH_3 as an important precursor for fine particulate matter ($PM_{2.5}$). High concentrations of $PM_{2.5}$ have harmful effects on human health (*e.g.,* Pope et al., 2002), and also impact regional and global climate by altering radiation budgets (*e.g.,* Pinder et al., 2013). Gas-phase ammonia and particulate ammonium may also be transported through the atmosphere and deposited via either wet or dry deposition. This deposition of reduced nitrogen may have detrimental effects on nitrogen-limited ecosystems, causing eutrophication, species assemblage shifts, and biodiversity loss (*e.g.,* Krupa et al., 2003).

Despite the important role NH_3 plays in the earth system, there are relatively sparse measurements of atmospheric NH_3 compared to many other atmospheric constituents. This contributes to large uncertainties in the fluxes of NH_3 into or out of many systems. Sparse measurements are due in large part to the hydrophilic and polar nature of NH_3, making it highly reactive and prone to adsorbing, or "sticking" to sampling surfaces. Thus, many measurements of atmospheric NH_3 have been restricted to denuder, impinger, or filter sampling techniques, which generally provide concentrations on minute- to week- long time scales. Within the last 15-20 years, however, techniques such as tunable infrared laser absorption and chemical ionization mass spectrometry have been developed for NH_3 that are capable of making measurements with much higher precision and temporal resolution than in the past (Nowak et al., 2007, von Bobrutski et al., 2010). Coupled with recent innovation in inlet design to further increase measurement time response, these instruments are now being use to make measurements of gas-phase NH_3 at rates greater than 1Hz. Leveraging this capability, I propose to sample gas-phase NH_3 from a small aircraft platform in order to calculate net fluxes of NH_3 out of or into areas of intensive agriculture. The overarching question I aim to answer is:

What is the magnitude and spatial distribution of net fluxes of ammonia from intensive crop farming in the U.S. Midwest Corn Belt across the growing season?

In addressing this broader question, the proposed research will also tackle several other sub questions:

How do inventories and models of NH_3 sources/sinks compare to the empirical estimates of net NH_3 fluxes from croplands?

What is the magnitude and seasonality of NH_3 deposition to areas adjacent to intensive crop farming?

I propose to employ the University of Wyoming King Air aircraft with a specific instrument payload to make measurements of NH_3 and atmospheric state variables over Iowa and Illinois. These measurements will be made over the course of a growing season and will be analyzed using the eddy covariance flux technique to calculate empirical net fluxes of NH_3 over regional spatial scales. In the following sections, I will describe the motivation for this work, hypotheses that will be tested, and the research approach, including analysis techniques and sampling strategies. Lastly, I present a discussion of the merits of this research along with potential related work and obstacles that may need to be overcome.

2. Background

As discussed in the introduction, agricultural NH_3 emissions (from both livestock and crop farming) are the dominant source of NH_3 to the atmosphere. On aggregate in the U.S., NH_3 emissions from livestock farming operations (55% of the total) are larger than emissions from chemical fertilizer use in croplands (26%), but in regions with a high density of intensive crop farming, such as the U.S. Midwest, emissions from both may be comparable (EPA 2018). Furthermore, the volatilization of NH_3 from fertilizer use is a function of many factors, which contributes to large uncertainties around the spatial and temporal distribution of fertilizer emissions. Fertilizer type and application, soil properties, meteorological conditions, and plant canopy characteristics may all influence rates of volatilization, with positive relationships between NH_3 volatilization at air temperature, soil PH, and urea content of the fertilizer for example (*e.g.,* Terman 1979, Sharpe and Harper 1995, Sommer et al., 2004). A meta-analysis by Pan et al., (2016) suggests that between 0.6 – 64% of total nitrogen applied as fertilizer may be volatilized. NH_3 emissions and volatilization rates have been measured in a small number of crop systems and locations in the U.S. Examples include multiple crops in California (Potter et al., 2001), corn in North Carolina (Walker et al., 2013) and Illinois (Nelson et al., 2018), and soybeans (Myles et al., 2011), and short grass (Phillips et al., 2004) in North Carolina.

The most common method to estimate NH_3 volatilization across broad spatial scales has been the use of emission factors for different fertilizers, crop species, and soil types, combined with disaggregated economic data on fertilizer sales and agricultural land use (*e.g.,* Goebes et al., 2003; Balasubramanian et al., 2015). These emission factors are assumed to be constant, and do not account for the biological and chemical processes inherent to the volatilization of NH_3 that strongly depend on environmental conditions and properties. In the National Emissions Inventory 2011, the Sparse Matrix Operator Kernel Emissions model is used to break emission down further into spatial and temporal resolutions suitable for input into chemical transport models (EPA 2013).

Other research has used chemical transport models and inverse modeling to test and apportion spatial and temporal distributions of NH_3 emissions in the U.S. Heald et al. (2012) compared GEOS-Chem NH_3 and ammonium nitrate concentrations to observations from several networks, field campaigns, and IASI satellite measurements. Gilliland et al. (2006) and Paulot et al. (2014) both used measurements of wet deposition of ammonium to constrain national emissions, employing the Community Multiscale Air Quality model and the GEOS-Chem adjoint model respectively. Zhu et al., (2013) used TES satellite observations in coordination with GEOS-Chem to constrain NH_3 inventories. One main takeaway from these studies and others is that while overall NH_3 emissions in the NEI may be reasonable, the spatial and temporal distribution of those emissions is consistently incorrect.

Efforts to improve inventories and model estimates of NH_3 emissions have taken the form of including various types of additional data or models. Balasubramanian et al. (2015) merged emission factors with a high resolution map of cropland and a biogeochemical process model to estimate NH_3 emissions in Illinois and surrounding states (model name: Improved Spatial Surrogate with DenitrificationDecomposition, ISS DNDC). Other process-based models such as the Environmental Policy Integrated Climate (EPIC) model (Williams et al., 2008) have been used to simulate fertilizer application driven by plant demand. EPIC is a part of a bidirectional flux scheme within CMAQ (Cooter et al., 2012; Bash et al., 2013), and has recently been included in the NEI2014 framework (EPA 2018). And Paulot et al. (2014) developed semi-empirical relationships between environmental parameters and NH_3 emissions into the Magnitude and Seasonality of Agricultural Emission model (MASAGE_NH3) as part of their inverse modeling study. Each of these models or inventories claims to result in improved spatial and temporal representation of NH_3 fluxes, but the availability of NH_3 flux measurements to test them limits the further development of NH_3 emission estimates from fertilizer use (Flechard et al., 2013; Balasubramaniamn et al., 2015).

While site-scale NH_3 fluxes yield detailed temporal resolution and important insights in processes (*e.g.,* Walker et al., 2013), flux measurements on a larger spatial scale would afford the opportunity to observe spatial variability and test the ability of various models and inventories to capture this variability. Airborne eddy covariance flux measurements promise precisely this opportunity to sample fluxes on regional spatial scales, and repeated measurements over time offer the opportunity to sample temporal variability and probe process-level understanding as well (*e.g.,* Wolfe et al., 2018).

3. Region of Study

While many regions in the U.S. have high NH_3 emissions from agriculture, the region chosen for this study is the U.S. Midwest, specifically focusing on the states of Iowa and Illinois. These states have a very high density of intensive crop farming (Figure 1), and comparisons of models with observations (Heald et al., 2012) and models with inventories (Balasubramanian et al., 2015) suggest that both the spatial and temporal distribution of NH_3 fluxes is uncertain. This region is both a source of NH_3 and a sink for NH_3 via dry deposition. Li et al., (2016) find that reduced N_r now accounts for over 50% of total N_r deposition. Furthermore, recent studies by Steiner et al. (2012) and Evanoski-Cole et al. (2017) have highlighted the importance of NH_3 emissions to $PM_{2.5}$ formation during late winter/early spring (February-April). Lastly, the high density of agricultural activities and relatively homogeneous terrain are well suited to airborne

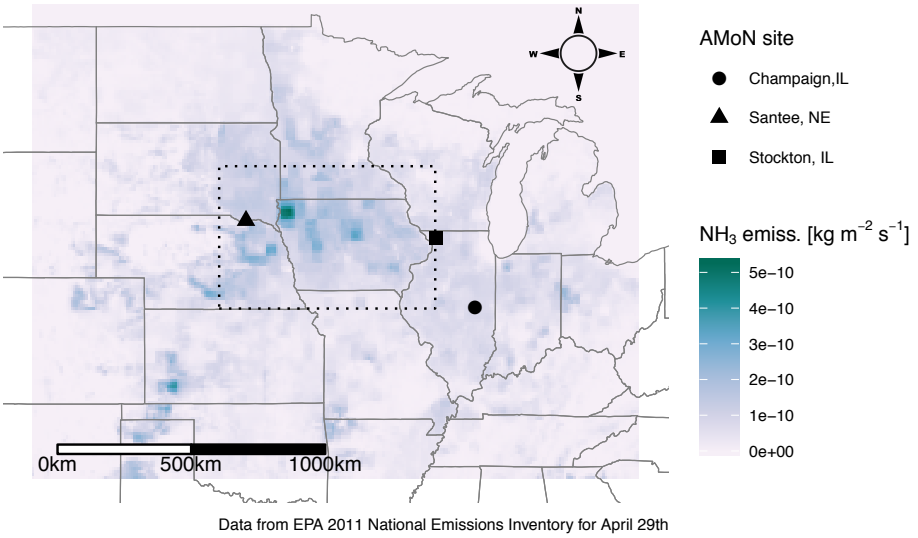


Figure 1. Spatial distribution of NH_3 emissions over the U.S. Midwest on April 29th from the EPA 2011 National Emissions Inventory. The three Ambient Ammonia Monitoring Network locations closest to Iowa are displayed with black symbols. The specific region of interest is outlined in a black dotted line.

eddy covariance flux measurements. These factors all combine to make the U.S. Midwest a relevant and exciting location for this research.

4. Hypotheses

Median seasonal cycles for ambient NH_3 concentrations at three sites within the study region are shown in Figure 2. These data are 2-week composite measurements from denuder filters as part of the Ambient Ammonia Monitoring Network. They show clear peaks in the spring and fall in ambient NH_3 concentrations, matching up with the primary fertilizer application periods for this region, with significant variability in individual measurements (plotted as points behind the lines). This general seasonal cycle generally agrees well with sparse site flux observations suggesting ~80% of NH_3 emissions occurred after fertilizer application in a corn field in Illinois (Nelson et al., 2018)

Following these observations, I hypothesize that the greatest NH_3 emissions will occur in the early to mid spring, as fertilizer is applied to fields. Biogeochemical models suggest that corn cultivation may account for up to 95% of the emissions of NH_3 in this region (Balasubramanian et al., 2015). Corn and other crops such as soybeans are intermixed widely throughout these states, and thus I hypothesize that flux measurements at fine scales will be variable, with strong emission sources in close proximity to small sources or even areas of net deposition. During summer, emissions from fertilizer will likely be much smaller, even with warmer air temperatures, and strong livestock sources in the region may lead to widespread measurements of net deposition to croplands. In the fall, harvesting and additional applications of fertilizer will likely lead to net emissions once again.

While the rationale for these hypotheses follows insights from small-scale measurements, widely dispersed routine measurement locations, and process based models, empirical NH_3 flux estimates of regional scale fluxes from intensive crop farming have not been made. Therefore, measurements of regional NH_3 fluxes across a growing season will be a valuable test of these sources of knowledge.

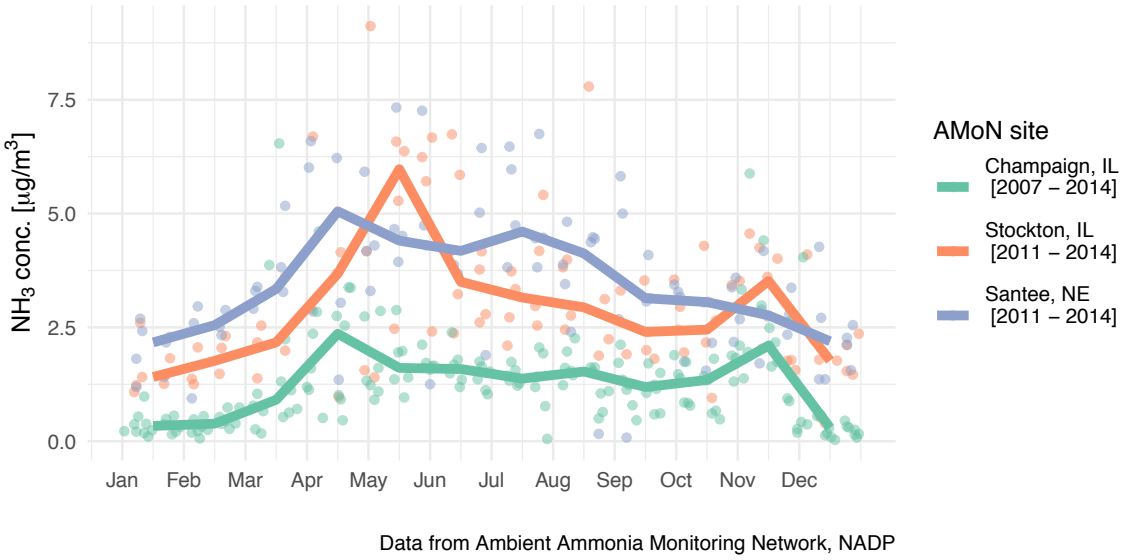


Figure 2. Median seasonal cycles for three Ambient Ammonia Monitoring Network locations from Figure 1. Two-week average concentrations are plotted as points, and the median over 3-7 years for each month is plotted as solid lines.

5. Approach

5.1 Proposed Platform

The aircraft research platform I propose to deploy is the Univeristy of Wyoming King Air (UWKA). The UWKA is a small twin-propeller aircraft that has a ~4 hour flight time with an airspeed of around 85 m/s, and has participated in airborne flux measurement campaigns in the past (*e.g.,* Stull et al., 1996; Gerbush et al., 2008). A small research crew flies on board to operate instrumentation and guide the sampling strategy. While the payload (~1650 lbs) is smaller than other aircraft that have been used to measure airborne fluxes (*e.g.,* NASA C-23 Sherpa: Wolfe et al., 2018; NSF C-130: Yuan et al., 2015), as a smaller aircraft the UWKA has the capability to fly within several hundred feet of the ground, which decreases errors in the calculated fluxes (more in sections 5.3 and 7).

5.2 Proposed Measurements

Standard state parameters are always measured on the UWKA. These include altitude, position, air temperature, pressure, heading, ground velocity, vertical velocity, pitch/roll angle, and 3-dimensional winds. Requestable instruments in service of fluxes of energy, and carbon dioxide (CO_2) include a Friehe type air temperature measurement with UW modifications, and a LI-COR 6262 infrared gas analyzer for fast time response measurements of *in situ* water vapor and carbon dioxide. Specifications of the water vapor measurements include a range of 0-75 mb, accuracy of 1%, and precision of 0.1 mb while the CO2 measurements have a range of 0-3000 ppmv, accuracy of 1 ppmv at 350 ppmv and a precision of 1 ppmv. Air temperature and water vapor measurements allow the calculation of latent and sensible heat fluxes to use along with

CO_2 fluxes. Energy and CO_2 fluxes have been widely studied in agricultural systems. Therefore, these fluxes can used to test the credibility of the NH_3 fluxes.

In situ NH_3 mixing ratios will be measured using an Aerodyne Quantum Cascade Tunable Infrared Laser Direct Absorption Spectrometer (QC-TILDAS) (Ellis et al., 2010), owned and operated by my research group. This instrument was previously used to sample *in situ* NH_3 in wildfire smoke aboard the NSF/NCAR C-130 aircraft during the Western Wildfire Experiment on Cloud Chemistry, Aerosol Absorption, and Nitrogen (WECAN) field project in summer 2018. Inlet modifications including a quartz glass inertial inlet for the separation of aerosol particles greater than 300 nm, heated PFA Teflon tubing, and the use of a strong base, 1H,1H perfluorooctylamine, to coat inlet surfaces serve to increase time response such that operation in flight at 10Hz has a precision of 1 ppbv and an accuracy of 12%, with a range of at least 0-400 ppbv. QC-TILDAS instruments have been used to make ground-based measurements of eddy covariance fluxes (*e.g.,* von Bobrutski et al., 2010; Zoll et al., 2016). As flown during WECAN, the instrument rack weighed 356 lbs, with the pump adding an additional 80 lbs.

Space, weight, and power permitting, two other instruments could be requested from the National Center for Atmospheric Research (NCAR) to fly on the UWKA. The NCAR NO/NO_2 2-channel chemiluminesence instrument measures *in situ* nitrogen oxides with a 1Hz uncertainty of 10%, and precision of 10 pptv (Ridley and Grahek, 1990). And the NCAR N_2O/CO QC-TILDAS (Lebegue et al., 2016) uses the same direct infrared absorption technique as the NH_3 instrument to measure nitrous oxide and carbon monoxide with precisions of 30 pptv and 100 pptv respectively at 1Hz (and is also capable of operating at 10Hz). These instruments also flew on the C-130 during WECAN, and while they would not be critical to this research, they would provide important tracers for urban, industrial, and traffic emissions. Additionally, NO_x and N_2O have significant sources from agricultural soils, and extensions to this research could investigate fluxes of these oxidized reactive nitrogen species from intensive crop farming within the study region.

5.3 Proposed Analysis Techniques

Several methods have been developed to estimate surface fluxes from airborne *in situ* data, including mass balance approaches, column integration, and eddy covariance (Wolfe et al., 2018). Each method requires different flight patterns and is best suited to different flux processes. Airborne eddy covariance flux directly quantifies the vertical turbulent fluxes of mass and energy within the boundary layer by calculating the covariance between fluctuations in vertical wind speed and perturbations in a scalar quantity (such as temperature or gas concentration). It has the main advantage of being able to resolve spatial gradients in fluxes over broad regions at relatively fine scale (~1 km), and is well suited to disperse fluxes such as those from widespread vegetation or intensive crop farming (Misztal et al., 2014). Recent examples of aircraft measurements used in eddy covariance flux calculations include estimates of water and CO_2 fluxes over agriculture and forests in Europe (Gioli et al., 2004), coastal mangroves in Mexico (Zulueta et al., 2013), and the eastern U.S. (Wolfe et al., 2018), estimates of methane fluxes from Arctic tundra (Sayres et al., 2017) and shale gas production regions in the U.S. (Yuan et al., 2015), nitrogen oxide fluxes over London, U.K. (Vaughan et al., 2016), and isoprene fluxes from the California Central Valley (Misztal et al., 2014).

Requirements for successful airborne eddy covariance flux measurements include high frequency and high precision measurements, to resolve the range of timescales on which

turbulent eddies may transport the quantity of interest, level flight legs on the order of tens of kilometers to facilitate reduction in error, and relatively homogeneous terrain to limit the variability in turbulent properties along the flight path (Wolfe et al., 2018). Traditionally, the average flux along a flight track has been calculated by applying a fast fourier transform (FFT) algorithm to the normalized covariance of the wind and gas concentration fluctuations. This method generates a single average flux over the flight track and further requires stationarity, or that statistical properties of the turbulence do not change along the flight path. Most recent applications, including all the examples cited earlier, employ a different method that does not require stationarity and delivers spatial resolved fluxes. This method is called continuous wavelet transform or CWT. Here, both timeseries of pertubations in vertical wind speed and gas concentration are convolved with a wavelet function (Torrence and Campos, 1998), the cospectrum of which calculates contributions to the total flux from eddies of all sizes/frequencies along the flight track. Different types of wavelets can be chosen, with the Morlet wavelet the most commonly used (Wolfe et al., 2018). By preserving energy from all distances and eddy frequencies along the flight track, CWT integrates information from the entire sampling leg to derive fluxes. Further information on the implementation of CWT and calculation of uncertainties is given in section 7.

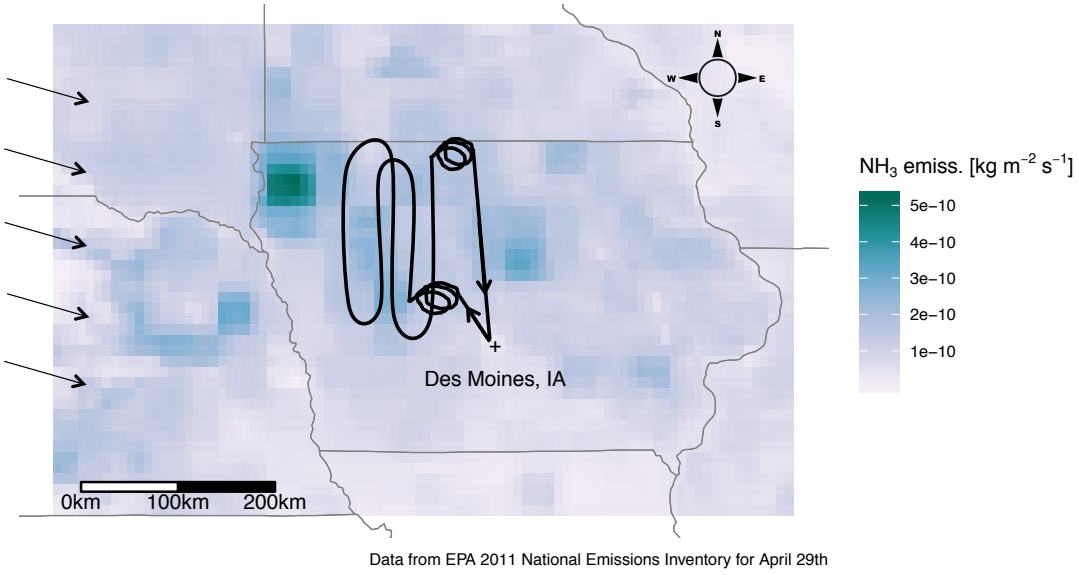


Figure 3. A sample flight plan in black overlaid on the spatial distribution of NH_3 emissions over the U.S. Midwest on April 29th from the EPA 2011 National Emissions Inventory. Arrows on the left indicate the prevailing wind direction.

6. Sampling Strategy

As stated in section 4, NH_3 fluxes from intensive crop farming in the U.S. Midwest are likely to vary considerably across the growing season due to the timing of fertilizer application along with the seasonal cycle in meteorological conditions. In order to sample across this temporal variability, I propose to use ~180 flight hours, split up between three one month deployments in the month of April, mid July through mid August, and the month of October. Roughly fifteen 4-hour flights will be completed during each deployment, offering dense temporal sampling rates and the opportunity to make multiple flights in a single day to sample diurnal variability as well. The primary flight times will be late morning to early afternoon, when the boundary layer is well defined and well mixed. Flights will target synoptically simple days, with consistent wind speeds and directions so as to simplify the analysis and interpretation of measurements.

For each deployment, I propose to base the UWKA out of Des Moines, IA. Des Moines is centrally located in Iowa. The UWKA has an average airspeed of 85 m/s which translates to a roughly 1200 km range over a flight duration of 4 hours. Thus, the UWKA will be able to sample anywhere in Iowa and over much of the area of surrounding states during any given flight. Each flight will have 6-12 level altitude sampling legs, with each eddy covariance flux leg being between 50 and 150 km long. This range offers the opportunity to sample multiple areas or to stack legs at multiple altitudes. The most common flight altitude will be within 100-200 m above ground level. Vertical spirals through the boundary layer will be performed at least once per flight to estimate the boundary layer height. Figure 3 shows a sample flight path, departing and returning to Des Moines. A "racetrack" pattern enables sampling of long flight tracks through multiple areas, and the spirals denote vertical soundings throughout the boundary layer. Lastly, flights from Des Moines also offer the opportunity to overfly two Ameriflux sites: Brooks Field, IA (http://dx.doi.org/10.17190/AMF/1246038), and Bondville, IL (http://dx.doi.org/10.17190/AMF/1246036). These sites maintain eddy covariance flux towers measuring energy and CO_2 fluxes, which can be used to compare and validate the aircraft fluxes.

7. Data Analysis

Data analysis will have three main components. First will be the production of final data and the calculation of net NH_3 fluxes and associated errors. Next will be the analysis of the spatiotemporal variability of surface fluxes with crop surface information and meteorological conditions. And last will be the comparison of fluxes with several models and inventories.

The calculation of eddy covariance fluxes and errors will follow the detailed methods laid out in Wolfe et al. (2018). Briefly, these first involve preprocessing the data to align the time base of each measured quantity. Next, the Continuous Wavelet Transform analysis is applied, and the cospectra may be compared to a Fast Fourier Transform. Systematic and random errors must be calculated, along with the vertical flux divergence, which is the change in the flux with height. The vertical flux divergence is calculated by comparing stacked flight legs, and provides a scaling factor by which the airborne flux may be extrapolated to a surface flux. This correction factor is applied everywhere along the flight track and may account for a 10-50% increase in the surface flux as compared to the lowest level airborne fluxes (Wolfe et al., 2018). Lastly, the flux footprint is calculated. To calculate the total uncertainty, the systematic error, random error, and flux divergence correction error are summed in quadrature. Errors may be averaged over the

timeseries as at short timescales errors may be well is excess of 100%. On spatial scales of 2 km errors will often average down to the range of 40-90%. This error range is sufficient to distinguish differences in the magnitudes of fluxes along different parts of the flight track, and to clearly distinguish net sinks from net sources.

The analysis of fluxes with crop surface information and meteorological conditions will utilize the flux footprints to compare spatially resolved NH_3 fluxes from different crop and land cover types (NASS, 2014). The influence of meteorological variability on fluxes from different crop and land cover types will be analyzed within each deployment, and the overall seasonal variability will be investigated. This analysis should be able to separate out the respective influence from crop type, environmental conditions, and spatial and temporal variability on net NH_3 fluxes in the U.S. Midwest.

The third analysis will be a comparison with three models and inventories. Net NH_3 fluxes will be compared to three of the models described earlier:, the Improved Spatial Surrogate with DenitrificationDecomposition (Balasubramanian et al., 2015), NEI2014 (EPA 2018), and Magnitude and Seasonality of Agricultural Emission model (Paulot et al., 2014).

8. Work Plan

Recognizing that proposal timelines, budget requirements, and facility scheduling constraints would require a longer timeframe than a PhD, here is an example work plan for this research, along with a timeline for meeting PhD requirements:

	2019				2020				2021	
	Q1	Q2	Q3	Q4	Q1	Q2	Q3	Q4	Q1	Q2
Upload to UW King Air	■									
Research Flight Period		■	■							
Final Data Production				■	■					
Calculation of Fluxes					■	■	■			
Comparison to Models/Inventories							■	■	■	
Preparation of Publications					■	■	■	■	■	
Preliminary Exam X										
Committee Meeting						X				
PhD defense										X

9. Discussion and Summary

In this prospectus, I have proposed a set of airborne measurements of *in situ* NH_3 and atmospheric state variables in order to calculate the magnitude and spatial distribution of net fluxes of NH_3 from intensive crop farming in the U.S. Midwest Corn Belt across the growing season. Based out of Des Moines, IA, the UW King Air will sample areas of chemical fertilizer use in crop-based agriculture within Iowa and surrounding states. This region is a major source of NH_3 in the U.S. The U.S. Midwest is expected to continue to be a major agricultural production region in the future and while NH_3 is currently unregulated, it may become so in the next few decades. As it stands now, with current technology and practices, Davidson et al. (2012), estimate that nitrogen loss from farms and livestock operations could be reduced by 30-

50%. Interdisciplinary work aimed at estimating the economic cost to human and environmental health due to "nitrogen leakage" (Sobota, 2015) may strengthen the economic and policy case for innovative policies and incentives for agricultural producers to reduce reactive nitrogen loss and usage. Thus, regional estimates of NH_3 net fluxes in the U.S. Midwest from this work will be a baseline for comparison with future studies. Furthermore, comparison of empirical fluxes with models and the national inventory will help improve inputs into chemical transport models, thereby resulting in better predictions of nitrogen deposition and particulate matter distributions within the midwestern and eastern U.S. Possible extensions to this research could include validation or comparison to satellite column measurements of NH_3 from CrIS, IASI or AIRS (*e.g.,* Van Damme et al., 2015).

One possible weakness in the sampling strategy presented here is the partitioning of gas-phase NH_3 to aerosol. NH_3 has a lifetime of around 1 day (Galperin and Sofiev, 1998) and can be appreciably lost to particulates on the timescale of hours (Wagstrom and Pandis, 2011). While the timescale of boundary layer mixing means that a low level sampling strategy should be measuring air within minutes to an hour of interaction with the surface, this chemical loss of NH_3 could bias our flux calculations. However, this is a case where calculating the vertical flux divergence and performing many stacked legs will help address this issue. The vertical flux divergence is the sum of entrainment/detrainment, horizontal advection, or production/loss processes. Thus, the empirical calculation of the vertical flux divergence wraps chemical production or loss into the surface flux correction factor (Wolfe et al., 2018). In theory, analyzing differences in the vertical flux divergence correction factor across different flights may produce insight into the relative importance of gas-particle partitioning during the sampling period. While direct measurements of the partitioning rate could be made by adding an Aerosol Mass Spectrometer to the King Air payload, space, weight, and power constraints would make this additional research activity difficult.

Other possible constraints to this research may be logistical in nature. Deployments of aircraft platforms across entire seasons are often infeasible due to scheduling conflicts. Were the use of the King Air to be restricted to less than three month-long deployments, this research could still produce valuable estimates of the magnitude and spatial variability of net NH_3 fluxes across one spring planting season, or a spring and summer timeframe. Therefore, the deployment schedule could be adjusted to yield valuable scientific insights while acknowledging wider community needs.

References

Balasubramanian, S., S. Koloutsou-Vakakis, D. M. McFarland, and M. J. Rood (2015), Reconsidering emissions of ammonia from chemical fertilizer usage in Midwest USA, J. Geophys. Res. Atmos., 120, 6232–6246, doi:10.1002/2015JD023219

Bash JO, Cooter EJ, Dennis RL, Walker JT, Pleim JE. Evaluation of a regional air-quality model with bidirectional NH3 exchange coupled to an agroecosystem model. Biogeosciences. 2013;10:1635–45. doi: 10.5194/Bg-10-1635-2013.

Cooter EJ, Bash JO, Benson V, Ran L. Linking agricultural crop management and air quality models for regional to national-scale nitrogen assessments. Biogeosciences. 2012;9:4023–35. doi: 10.5194/Bg-9-4023-2012.

Davidson E.A., David, M.B., J.N. Galloway, C. L. Goodale, R. Haeuber, J.A. Harrison, R.W. Howarth, D. B. Jaynes, R. Lowrance, B. T. Nolan, J. L. Peel, R. W. Pinder, E. Porter, C. S. Snyder, A. R. Townsend, and M. H. Ward, Excess Nitrogen in the U.S. Environment: Trends, Risks, and Solutions, Issues in Ecology Ecology, ESA, Winter 2012, report number 15.

Ellis RA, Murphy JG, Pattey E, van Haarlem R, Obrien JM, Herndon SC. Characterizing a Quantum Cascade Tunable Infrared Laser Differential Absorption Spectrometer (QC-TILDAS) for measurements of atmospheric ammonia. Atmos Meas Tech. 2010;3:397–406.

Evanoski-Cole, A. R., Gebhart, K. A., Sive, B. C., Zhou, Y., Capps, S. L., Day, D. E., Prenni, A. J., Schurman, M. I., Sullivan, A. P., Li, Y., Hand, J. L., and Collett, Jr., J.L. (2017) Composition and sources of winter haze in the Bakken oil and gas extraction region. Atmos. Environ., 156, 77-87, doi: 10.1016/j.atmosenv.2017.02.019.

Flechard, C. R., R.-S. Massad, B. Loubet, E. Personne, D. Simpson, J. O. Bash, E. J. Cooter, E. Nemitz, and M. A. Sutton (2013), Advances in understanding, models and parameterizations of biosphere-atmosphere ammonia exchange, Biogeosciences, 10, 5183–5225, doi:10.5194/bg-10-5183-2013.

Fowler D, Coyle M, Skiba U, Sutton MA, Cape JN, Reis S, et al. The global nitrogen cycle in the twenty-first century. Philos T R Soc B. 2013;368:20130164.

Galloway JN, Townsend AR, Erisman JW, Bekunda M, Cai ZC, Freney JR, et al. Transformation of the nitrogen cycle: recent trends, questions, and potential solutions. Science. 2008;320:889–92. doi: 10.1126/Science.1136674.

Galperin MV, Sofiev MA. The long-range transport of ammonia and ammonium in the Northern Hemisphere. Atmos Environ. 1998;32:373–80. doi: 10.1016/S1352-2310(97)00045-9.

Gilliland AB, Appel KW, Pinder RW, Dennis RL. Seasonal NH_3 emissions for the continental United States: inverse model estimation and evaluation. Atmos Environ. 2006; 40:4986–98.

Goebes, M. D., R. Strader, and C. Davidson (2003), An ammonia emission inventory for fertilizer application in the United States, Atmos. Environ., 37, 2539–2550, doi:10.1016/S1352-2310(03)00129-8.

Gioli, B., Miglietta, F., De Martino, B., Hutjes, R. W. A., Dolman, H. A. J., Lindroth, A., Schumacher, M., Sanz, M. J., Manca, G., Peressotti, A., and Dumas, E. J.: Comparison between tower and aircraft-based eddy covariance fluxes in five European regions, Agr. Forest Meteorol., 127, 1–16, 2004

Gerbush, M.R., D.A. Kristovich, and N.F. Laird, 2008: Mesoscale Boundary Layer and Heat Flux Variations over Pack Ice–Covered Lake Erie. J. Appl. Meteor. Climatol., 47,668-682, https://doi.org/10.1175/2007JAMC1479.1

Heald CL, Collet JL, Lee T, Benedict KB, Schwandner FM, Li Y, et al. Atmospheric ammonia and particulate inorganic nitrogen over the United States. Atmos. Chem. Phys. 2012;12: 10295–312. doi: 10.5194/acp-12-10295-2012.

Krupa SV. Effects of atmospheric ammonia (NH3) on terrestrial vegetation: a review. Environ Pollut. 2003; 124:179–221. doi: 10.1016/S0269-7491(02)00434-7.

Li, Y., Schichtel, B. A., Walker, J. T., Schwede, D. B., Chen, X., Lehmann, C. M. B., Puchalski, M. A., Gay, D. A., and Collett, Jr., J.L. (2016) Increasing importance of deposition of reduced nitrogen in the United States. PNAS, 113 (21), 5874-5879, doi:10.1073/pnas.1525736113.

Lebegue, Benjamin, et al. "Comparison of nitrous oxide (N2O) analyzers for high-precision measurements of atmospheric mole fractions." Atmospheric Measurement Techniques 9.3 (2016): 1221-1238.

Misztal, P. K., Karl, T., Weber, R., Jonsson, H. H., Guenther, A. B., and Goldstein, A. H.: Airborne flux measurements of biogenic isoprene over California, Atmos. Chem. Phys., 14, 10631– 10647, https://doi.org/10.5194/acp-14-10631-2014, 2014.

Myles, L., J. Kochendorfer, M. W. Heuer, and T. P. Meyers (2011), Measurement of trace gas fluxes over an unfertilized agricultural field using the flux-gradient technique, J. Environ. Qual., 40, 1359–1365, doi:10.2134/jeq2009.0386.

NASS (2014), Crop. data layer, US Department of Agriculture. [Available at http://nassgeodata.gmu.edu/CropScape/

Nelson, A.J., Koloutsou-Vakakis, S., Rood, M.J., Myles, L., Lehmann, C., Bernacchi, C., Balasubramanian, S., Joo, E., Heuer, M., Vieira-Filho, M., Lin, J., 2017. Season-long ammonia flux measurements above fertilized corn in central Illinois, USA, using relaxed eddy accumulation. Agric. For. Meteorol. 239:202–212. https://doi.org/10.1016/j. agrformet.2017.03.010.

Nowak JB, Neuman JA, Kozai K, Huey LG, Tanner DJ, Holloway JS, et al. A chemical ionization mass spectrometry technique for airborne measurements of ammonia. J Geophys Res-Atmos 2007:112.

Paulot F, Jacob DJ, Pinder RW, Bash JO, Travis K, Henze DK. Ammonia emissions in the United States, European Union, and China derived by high-resolution inversion of ammonium wet deposition data: interpretation with a new agricultural emissions inventory (MASAGE_NH3). J Geophys Res-Atmos. 2014;119:4343–64. doi: 10.1002/2013jd021130.

Phillips, S., S. Pal Arya, and V. P. Aneja (2004), Ammonia flux and dry deposition velocity from near-surface concentration gradient measurements over a grass surface in North Carolina, Atmos. Environ., 38, 3469–3480, doi:10.1016/j.atmosenv.2004.02.054

Pinder RW, Bettez ND, Bonan GB, Greaver TL, Wieder WR, Schlesinger WH, et al. Impacts of human alteration of the nitrogen cycle in the US on radiative forcing. Biogeochemistry. 2013;114:25–40. doi: 10.1007/S10533-012-9787-Z.

Pope CA, Burnett RT, Thun MJ, Calle EE, Krewski D, Ito K, et al. Lung cancer, cardiopulmonary mortality, and long-term exposure to fine particulate air pollution. J Am Med Assoc. 2002;287:1132–41.

Potter, C., C. Krauter, and S. Klooster (2001), Statewide inventory estimates of ammonia emissions from native soils and chemical fertilizers in California, ARB Contract Number ID 98-716, California State Univ., Fresno, Calif.

Reis, S., Bekunda, M., Howard, C., Karanja, C, Winiwarter, W., Yan, X., Bleeker, A., Sutton, M.A., Synthesis and review: Tackling the nitrogen management challenge: from global to local scales**,** Environmental Research Letters, Volume 11, Number 12 https://doi.org/10.1088/1748-9326/11/12/120205

Ridley, B.A.; Grahek, F.E. (1990), A small, low flow, high-sensitivity reaction vessel for NO chemiluminescence detectors, J. Atmos. Oceanic Technol, 7, 307-311.

Sayres, D. S., Dobosy, R., Healy, C., Dumas, E., Kochendorfer, J., Munster, J., Wilkerson, J., Baker, B., and Anderson, J. G.: Arctic regional methane fluxes by ecotope as derived using eddy covariance from a low-flying aircraft, Atmos. Chem. Phys., 17, 8619– 8633, https://doi.org/10.5194/acp-17-8619-2017, 2017.

Sharpe, R. R., and L. A. Harper (1995), Soil, plant and atmospheric conditions as they relate to ammonia volatilization, Fertil. Res., 42, 149–158, doi:10.1007/BF00750509

Sobota D J, Compton J E, McCrackin M L and Singh S 2015 Cost of reactive nitrogen release from human activities to the environment in the United States Environ. Res. Lett. 10 025006

Sommer, S. G., J. K. Schjoerring, and O. T. Denmead (2004), Ammonia emission from mineral fertilizers and fertilized crops, Adv. Agron., 82, doi:10.1016/S0065-2113(03)82008-4.

Stanier, C., et al. (2012), Overview of the LADCO winter nitrate study: Hourly ammonia, nitric acid and PM2.5 composition at an urban and rural site pair during PM2.5 episodes in the US Great Lakes region, Atmos. Chem. Phys., 12, 11,037–11,056, doi:10.5194/acp-12-11037-2012

Stull, R., Santoso, E., Berg, L. and Hacker J. Boundary Layer Experiment 1996 (BLX96), Bulletin of the American Meteorological Society Vol. 78, No. 6 (June 1997), pp. 1149-1158

Torrence, C. and Compo, G.: A practical guide to wavelet analysis, B. Am. Meteorol. Soc., 79, 61–78, 1998 (code available at: http://atoc.colorado.edu/research/wavelets/

U.S. EPA (2014), 2011 National Emissions Inventory, version 1, Technical Support Document, Office of Air Quality Planning and Standards, Air Quality Assessment Division, North Carolina

U.S. EPA (2018), 2014 National Emissions Inventory, Technical Support Document, Office of Air Quality Planning and Standards, Air Quality Assessment Division, North Carolina

Van Damme M, Clarisse L, Dammers E, Liu X, Nowak JB, Clerbaux C, et al. Towards validation of ammonia (NH3) measurements from the IASI satellite. Atmos. Meas. Tech. 2015;8:1575–91. doi: 10.5194/amt-8-1575-2015.

Vaughan, A. R., Lee, J. D., Misztal, P. K., Metzger, S., Shaw, M. D., Lewis, A. C., Purvis, R. M., Carslaw, D. C., Goldstein, A. H., Hewitt, C. N., Davison, B., Beevers, S. D., and Karl, T. G.: Spatially resolved flux measurements of NOx from London suggest significantly higher emissions than predicted by inventories, Faraday Discuss., 189, 455–472, 2016

Vitousek PM, Aber JD, Howarth RW, Likens GE, Matson PA, Schindler DW, et al. Human alteration of the global nitrogen cycle: sources and consequences. Ecol Appl. 1997;7:737–50. doi: 10.2307/2269431.

von Bobrutzki K, Braban CF, Famulari D, Jones SK, Blackall T, Smith TEL, et al. Field inter-comparison of eleven atmospheric ammonia measurement techniques. Atmos Meas Tech. 2010;3:91–112. doi: 10.5194/Amt-3-91-2010.

Wagstrom KM, Pandis SN. Source-receptor relationships for fine particulate matter concentrations in the Eastern United States. Atmos Environ. 2011;45:347–56. doi: 10.1016/J.Atmosenv.2010.10.019.

Walker, J. T., M. R. Jones, J. O. Bash, L. Myles, T. Meyers, D. Schwede, J. Herrick, E. Nemitz, and W. Robarge (2013), Processes of ammonia air-surface exchange in a fertilized Zea mays canopy, Biogeosciences, 10, 981–998, doi:10.5194/bg-10-981-2013.

Williams, J. W., R. C. Izaurralde, and E. M. Steglich (2008), Agricultural Policy/Environmental eXtender Model: Theoretical documentation version 0604, Texas AgriLIFE Research, Texas A & M Univ., Temple, Tex. [Available at http://epicapex.brc.tamus.edu].

Wolfe, G. M., Kawa, S. R., Hanisco, T. F., Hannun, R. A., Newman, P. A., Swanson, A., Bailey, S., Barrick, J., Thornhill, K. L., Diskin, G., DiGangi, J., Nowak, J. B., Sorenson, C., Bland, G., Yungel, J. K., and Swenson, C. A.: The NASA Carbon Airborne Flux Experiment (CARAFE): instrumentation and methodology, Atmos. Meas. Tech., 11, 1757-1776, https://doi.org/10.5194/amt-11-1757-2018, 2018.

Yuan, B., Kaser, L., Karl, T., Graus, M., Peischl, J., Campos, T. L., Shertz, S., Apel, E. C., Hornbrook, R. S., Hills, A., Gilman, J. B., Lerner, B. M., Warneke, C., Flocke, F. M., Ryerson, T. B., Guenther, A. B., and de Gouw, J. A.: Airborne flux measurements of methane and volatile organic compounds over the Haynesville and Marcellus shale gas production regions, J. Geophys. Res.- Atmos., 120, 6271–6289, 2015

Zhu L, Henze DK, Cady-Pereira KE, Shephard MW, Luo M, Pinder RW, et al. Constraining U.S. ammonia emissions using TES remote sensing observations and the GEOS-Chem adjoint model. J Geophys Res-Atmos. 2013; 118:3355–68. doi: 10.1002/Jgrd.50166.

Zöll, U., Brümmer, C., Schrader, F., Ammann, C., Ibrom, A., Flechard, C. R., Nelson, D. D., Zahniser, M., and Kutsch, W. L.: Surface–atmosphere exchange of ammonia over peatland using QCL-based eddy-covariance measurements and inferential modeling, Atmos. Chem. Phys., 16, 11283-11299, https://doi.org/10.5194/acp-16-11283-2016, 2016.

Zulueta, R. C., Oechel, W. C., Verfaillie, J. G., Hastings, S. J., Gioli, B., Lawrence, W. T., and Paw U, K. T. P.: Aircraft Regional-Scale Flux Measurements over Complex Landscapes of Mangroves, Desert, and Marine Ecosystems of Magdalena Bay, Mexico, J. Atmos. Ocean. Tech., 30, 1266–1294, 2013.

PROSPECTUS

GREMLIN: GOES Radar Estimation via Machine Learning to Inform NWP

Submitted by

Kyle A. Hilburn

Department of Atmospheric Science

In partial fulfillment of the requirements

For the Degree of Doctor of Philosophy

Colorado State University

Fort Collins, Colorado

April 8, 2022

Committee:

Advisor: Steven D. Miller (CSU ATS)

Co-Advisors: Christian D. Kummerow (CSU ATS), Elizabeth A. Barnes (CSU ATS)

External Advisors: Imme Ebert-Uphoff (CSU ECE), Curtis R. Alexander (NOAA GSL)

1. Introduction

Imagery from the Geostationary Operational Environmental Satellite (GOES) has been a key element of U.S. operational weather forecasting since 1975, supporting the need for high-resolution rapidly refreshing imagery for situational awareness (Line et al. 2016). Despite the well demonstrated value to human forecasters, usage of GOES imagery in data assimilation (DA) for initializing numerical weather prediction (NWP) has been limited. The obstacles impeding the use of GOES for DA include: (a) gaps in scientific knowledge that limit the accuracy of forward operators, (b) computational limitations on simulating all relevant physical processes within the scope of the NWP operational production cycle, and (c) unmeasured and uncertain quantities that complicate inversion of the forward operator. This dissertation is specifically concerned with the use of GOES radiances and lightning observations in cloudy and precipitating scenes, where all three of these obstacles are present. However, the broader significance of this problem is that these same obstacles are encountered in other areas of Earth System Modeling, particularly when coupling different Earth System components, such as the atmosphere and land surface. The **central question** of this dissertation is: *how can artificial intelligence/machine learning (AI/ML) be used in combination with DA to better utilize remotely sensed observations for improving high-resolution NWP forecasts of high-impact weather hazards?* By providing a rich and powerful library of nonlinear statistical tools, ML can address these scientific gaps by providing statistically based forward and backward operators. While it can take considerable computational resources to train ML models, once trained, the computational cost at inference time is generally a tiny fraction of that required by physical models, thus addressing the limitations during the operational production cycle. Unmeasured and uncertain quantities remain an issue with ML, which is a motivating factor behind the need for trustworthy models that are explainable or interpretable. This dissertation will explore the usage of ML to enhance DA for the purpose of connecting models and observations.

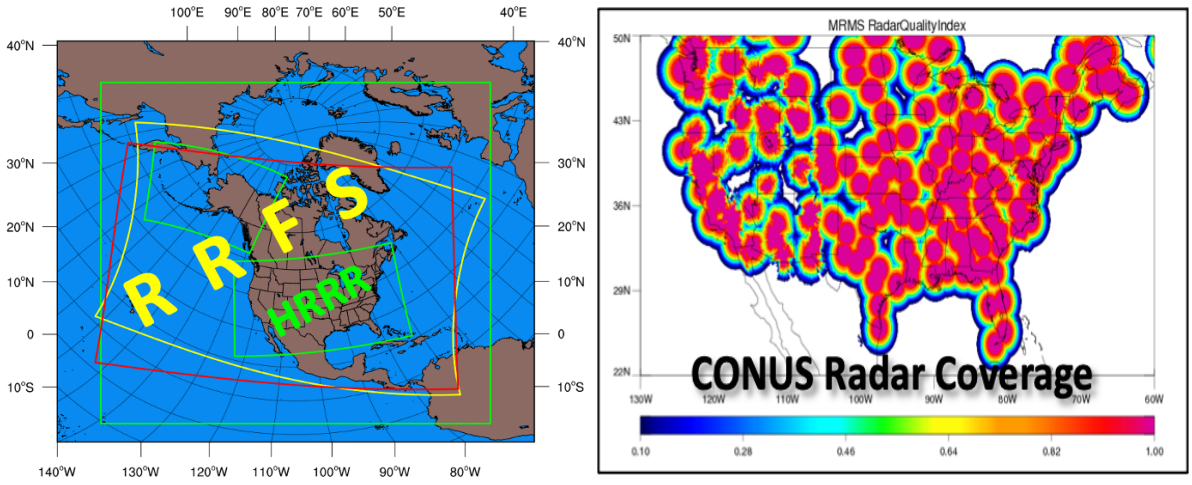


Figure 1. (Left) Rapid Refresh Forecast System (RRFS) domain in yellow and High-Resolution Rapid Refresh (HRRR) domain in green. (Right) Multi-Radar Multi-Sensor (MRMS) radar coverage over the contiguous United States (CONUS).

The broader significance and timeliness of this problem is illustrated in **Figure 1**. The High-Resolution Rapid Refresh (HRRR) model is NOAA's current operational NWP system, and it

uses ground-based radar to initialize convection over the CONUS domain (Benjamin et al. 2016). However, NOAA is in the process of updating its NWP models to use the Unified Forecast System (UFS) framework (Alexander et al. 2020, Back et al. 2022). When completed, the HRRR model will be replaced by the Rapid Refresh Forecast System (RRFS), which covers a much larger area, most of which lacks ground-based radar coverage. Therefore, geostationary (GEO) satellite products will be key to initializing and assessing the RRFS system. This is made possible by the high spatial resolution (2 km) and rapid temporal refresh (10 minutes) of the GOES-R Series. In addition, the GOES-R Series provides observations at more wavelengths, particularly in the near-infrared, which have become popular with National Weather Service (NWS) forecasters for severe weather applications in the form of the Day Cloud Phase RGB product (Connell et al., 2020). The GOES-R Series also provides lightning mapping capabilities, which I will show are crucial for this specific application, using a combination of Explainable AI and Interpretable AI techniques. The **hypothesis** of this dissertation is that *by leveraging the power of ML, GOES-R capabilities can be used to provide "radar everywhere" for initializing convection in high-resolution NWP models*. Note that even over CONUS, in regions of complex terrain and sparsely populated areas, there are gaps in radar coverage where GOES-R-based estimates of synthetic radar reflectivity can provide value to NWP and help alleviate social injustices resulting from uneven radar coverage.

2. Background

Currently, DA makes greater usage of microwave and infrared sounder observations from low-Earth-orbiting (LEO) satellites (Lin et al. 2017) than GEO. Sounders have the advantage of providing more vertically resolved information than imagers but have the disadvantage that LEO satellites provide coarse temporal resolution (e.g., twice per day) and longer latency that can reach 1.5 h or more. Thus, there is an opportunity for operational DA to benefit from the high volume of low-latency complementary data coming from the constellation of GEO imagers. While DA for convective-scale NWP has made steady scientific advances (Gustafsson et al. 2018), all-sky assimilation of infrared radiances has yet to be demonstrated operationally (Geer et al. 2018). That means the most dynamically interesting areas where precipitation is occurring, which have significant impacts on human activities, are also areas with the least amount of data to constrain estimates of the current atmospheric state. The current state-of-the-art for utilizing satellite observations in NWP is radiance assimilation (RA). Zhang et al. (2018, 2019) and Jones et al. (2020) have demonstrated RA using GOES-16 Advanced Baseline Imager (ABI). RA has the advantage of being physically based, which aids interpretation, although the use of background error covariances (BECs) estimated from ensembles means that GOES observations do not always clearly map into DA increments. The usage of RA is complicated by different assumptions that need to be made regarding how to inflate observation and background errors and how to weight information in the vertical. Moreover, errors in model microphysical and radiative transfer (RT) parameterizations will be inherited by RA, and the land surface is relevant to the interpretation of window channels. A limitation of RA is that it does not provide a means for assimilating lightning observations, so an observation operator is required to convert lightning observables into control variable increments (Kong et al. 2020, Apodaca and Zupanski 2018). Finally, a fundamental limitation of RA for infrared radiances is that information content saturates in cloudy and precipitating pixels through production of abundant cloud ice particles in convective cores and their horizontal spreading by wind shear (Grasso and Greenwald 2004). For

example, the maximum optical depths retrieved by GOES ABI are 160 and 8 during day and night (Walther et al. 2013), which roughly correspond to composite reflectivity (REFC) values of 20-25 dBZ for daytime and 0-5 dBZ at nighttime (Rutledge et al. 2020). This truncated sensitivity is also an issue for pixelwise physically based cloud retrievals (Jones et al. 2015).

ML provides three significant benefits with the potential to "disrupt" RA. The first benefit is that Convolutional Neural Networks (CNNs) can capture the information content in multi-scale image gradients (Hilburn et al. 2021a), thereby improving on the radiance saturation problem present in pixelwise RA. I will demonstrate that gradient information can provide reliable information up to REFC of about 50 dBZ. Note that this represents an entirely new source of information, which is currently going unused in DA systems. A key aspect of this dissertation will be understanding the nature of spatial information content, which gives CNNs their seemingly "magical" ability to mimic human interpretation of satellite imagery. The second benefit of ML is that it provides a convenient framework for performing data fusion, which makes it possible to utilize radiance and lightning observations simultaneously. In fact, I will demonstrate that the interaction between information in radiances and lightning combine to provide more skill than the sum of its parts. Finally, the third benefit is that ML provides a computationally efficient implementation for utilizing remotely sensed data. This is crucially important because of limited time during the operational production cycle and the cost of complex scattering calculations in RT parameterizations is prohibitive for operational NWP.

Traditional (non-ML) retrieval techniques tend to either treat individual satellite pixels as being independent or they make use of simple spatial information, such as the standard deviation in a neighborhood (Grecu and Anagnostou 2001, Olson et al. 2001). Even simple spatial information can improve predictions, for example in classifying raining pixels over cold surfaces (Kummerow et al. 2001). Convolutional ML approaches capture multi-scale spatial information, which is an important factor in their ability to outperform traditional methods (Guilloteau and Foufoula-Georgiou 2020). This is typically accomplished using pooling layers, which subsample the input images by a factor of two along each dimension. In doing so, the ML forms an image pyramid (Burt and Adelson 1983, Adelson et al. 1984), which provides a multi-scale representation. Section 3 will elaborate on the role of image pyramids in ML.

Criticisms of ML are that (a) it provides "black box" models that are difficult to understand, (b) does not generalize well to data outside of the training set, and (c) provides temporally inconsistent predictions. These are all valid criticisms, and an important aspect of the dissertation research is in addressing these issues. The first issue regarding understanding ML model predictions can be addressed in two ways, both of which are used in this research. One way is to use Explainable AI (XAI) techniques (Ebert-Uphoff and Hilburn 2020, McGovern et al. 2019), which are methods that are applied to trained ML models to provide information about which parts of the input images are most influential in the prediction. There has been an explosion of research in XAI in recent years, and it has become apparent that not all XAI methods are equally trustworthy (Mamalakis et al. 2022). The other way is to use Interpretable AI (IAI) techniques, which focus on building the explainability into the model right from the start, before the model has been trained. Chapter 2 of the dissertation will develop an IAI model, which is crucial for gaining a thorough understanding of exactly what information is present in the "spatial context" utilized by CNNs. The second issue is regarding the abilities of ML models to generalize to

unseen data, or what has been referred to as "Brittle AI" (Heaven 2019). In a classical linear regression framework, this can be seen as a case of covariance shift between the training data and the data used for predictions (Simon Pfreundschuh 2022). The dissertation will address this issue using a hierarchy of datasets, described in Section 3, and by evaluating ML predictions using datasets that were unseen during training. The third issue regrading temporal consistency of predictions has been referred to as "Twitchy AI" (Flora 2021). Defining what is meant by "temporal consistency" is challenging, but we have developed some approaches (Section 3). It will be addressed through assessing scales of spatial and temporal variability for a hierarchy of models, ranging from linear regression to CNNs. This is a difficult but important issue to address in order to use ML predictions for applications such as nowcasting, where the temporal trends of the predictions are important to a human forecaster's decision making.

3. Data and Methodology

Table 1 provides a list of the datasets, where GOES data serve as the inputs and MRMS data serve as the outputs. GOES-16 is used for training, but GREMLIN[1] predictions from GOES-18 will be evaluated once it has moved into West position in June 2022. The GPM DPR data will be used for validation over the ocean, and the code has already been developed to process DPR reflectivity profiles into composite reflectivity maps. The use of HRRR and GFS fields is described in Section 4. All fields are resampled to a common grid, as described in Hilburn et al. (2021a). The justification for including these ABI channels is given in Hilburn et al. (2018). Lightning data must also be accumulated in time, and a 15-minute period was found to give the best results by reducing the flickering in stratiform regions produced using shorter periods. Lightning "groups" are more spatially extensive than "flashes" and have higher spatial correlation with radar reflectivity. The value of lighting area for better distinguishing convective from stratiform echoes was described in Hilburn et al. (2021b).

Table 1. List of datasets used in this research.

Data Source	Parameter
GOES ABI	C02 0.64 μm Red Visible
	C05 2.2 μm Cloud Particle Phase/Size
	C07 3.9 μm Shortwave IR Window
	C09 6.9 μm Mid-level Water Vapor
	C13 10.3 μm Clean Longwave IR Window
GOES GLM	Lightning Group Extent Density
	Lightning Group Area
MRMS	Composite Reflectivity
	Convective/Stratiform/Frozen Precipitation Flag
GPM DPR	Composite Reflectivity
HRRR	Temperature and Relative Humidity at 850, 750, 500, and 300 hPa
GFS	Temperature and Relative Humidity at 850, 750, 500, and 300 hPa

[1] GOES Radar Estimation via Machine Learning to Inform NWP

After resampling, the fields must be matched in time and collected into a form suitable for ML training. This includes quality control, handling missing data areas, scaling the variables, and splitting training, testing, and validation datasets. The resulting ML datasets are listed in **Table 2**. The first dataset, CONUS1, was constructed by hand (with help from Alex Libardoni) to provide a sampling of severe storms maximizing diversity in location, time-of-day, and convective mode. A subset of this dataset, along with a "toy" version of GREMLIN was shared with Jamin Rader for use in his ML class project. The second dataset, CONUS2, provides a sampling of warm season convective storms that were selected in an automated fashion based on storm reports. The first two datasets are storm-centered, while the latter two datasets cover all of CONUS. A subset of the CONUS3 dataset was already used by Eric Goldenstern (2022) for his master's thesis. Finally, CONUS4 was prepared to transition GREMLIN from a 3 km grid to a 2 km grid. GREMLIN predictions will be made over the portion of the Full Disk (FD) that has GLM coverage (about 2/3^{rds}), but the training will be restricted to the CONUS portion of the FD.

Table 2. List of ML datasets produced by this research.

Dataset	Number of Images	Image Size	Grid	Construction	Reference
CONUS1	225	256 x 256	HRRR 3 km	Manual	Hilburn et al. (2019)
CONUS2	1,800	256 x 256	HRRR 3 km	Storm Reports	Hilburn et al. (2021a)
CONUS3	35,040 per year	1799 x 1059	HRRR 3 km	Every 15 minutes	Hilburn et al. (2020)
CONUS4	35,040 to 52,560 per year	4660 x 2678 (3072 x 1536)	ABI FD 2 km	Every 10-15 minutes	Hilburn et al. (2021b)

This ML application takes images as inputs and returns images as outputs, making this an image-to-image transition problem. A model architecture in the U-Net family (Ronneberger et al. 2015) is an ideal choice for this type of problem, and the details of Version-1 GREMLIN architecture are depicted in **Figure 2** left panel. The model was fit using `TensorFlow` with a weighted mean squared error (MSE) loss function. The weighting was found to be necessary to compensate for imbalance in the output quantity, since radar reflectivity follows an exponentially decreasing probability distribution function, using an unweighted MSE produces strong underprediction at high radar reflectivity values. The weights were tuned to minimize the categorical bias (Wilks 2006) across radar reflectivity bins, which produces a performance diagram (Roebber 2009) that falls along the 1:1 line. This means that the model is balancing underprediction with overprediction across the range of radar reflectivity values, which I found subjectively produces the most realistic looking output images. Baseline performance statistics for the warm season are RMSD = 5.53 dBZ, R^2 = 0.740, and 35 dBZ CSI = 0.33.

After training GREMLIN, the question becomes how does the model make these skillful predictions, and where/when can we trust the predictions? To gain insight, I'm using the XAI technique called Layerwise Relevance Propagation (LRP, Bach et al. 2015, Montavon et al. 2018, Lapuschkin et al. 2019, Toms et al. 2019), which reveals where in the inputs (which pixels in which channels) the neural network (NN) was primarily looking to derive the output pixel's estimate. LRP was found to be particularly powerful when combined with ablation studies, that

is where certain capabilities of the NN are removed (and the model is retrained) to see their impact on the predictions. For example, one can "turn off" lightning by setting it to zero or turn off spatial context by replacing 3x3 filters with 1x1 filters. Using this methodology, it was found that GREMLIN predictions depend on three factors, in order: presence of lightning, strong brightness temperature gradients, and cold brightness temperatures. These are all physically reasonable strategies for the NN to employ in predicting radar reflectivity. However, the disadvantage of XAI is that it only provides insights, not guarantees. To get such guarantees about how the model will perform, I have developed an Interpretable GREMLIN. An interpretable model is one where the explainability is built in from the start, and this has the benefit of also clarifying exactly what is meant by "spatial context" and separating out the role of nonlinearity in producing skillful predictions. The key insight enabling Interpretable GREMLIN is that the U-Net architecture shown in **Figure 2** left panel is equivalent to an image pyramid (**Figure 2** right panel) and filter bank, which is the mechanism by which CNNs can capture multiresolution information content. Thus, this dissertation will demonstrate that the power of CNNs for remote sensing problems comes from three factors: multiresolution spatial context, gradients and spatial patterns, and nonlinearity. The development of an interpretable model consists of taking the work performed by the CNN "under the hood" and bringing it out into the open, as part of a data preprocessing stage. The number of actual pieces of information being considered in making predictions at each pixel are the product of the number of input channels, number of pyramid levels, and number of image kernels. GREMLIN Version-1 has four input channels and is four levels deep. While the GREMLIN CNN uses 32 kernels (filters) per layer, such a large filter bank of kernels is not necessary. The dissertation will demonstrate that the Interpretable model can reproduce the skill of the CNN using only four 3x3 kernels: identity, Sobel dx, Sobel dy, and Laplacian. Other problems, such as super-resolution, appear to require more kernels. Thus, the number of inputs to the interpretable model is 4*4*4 = 64.

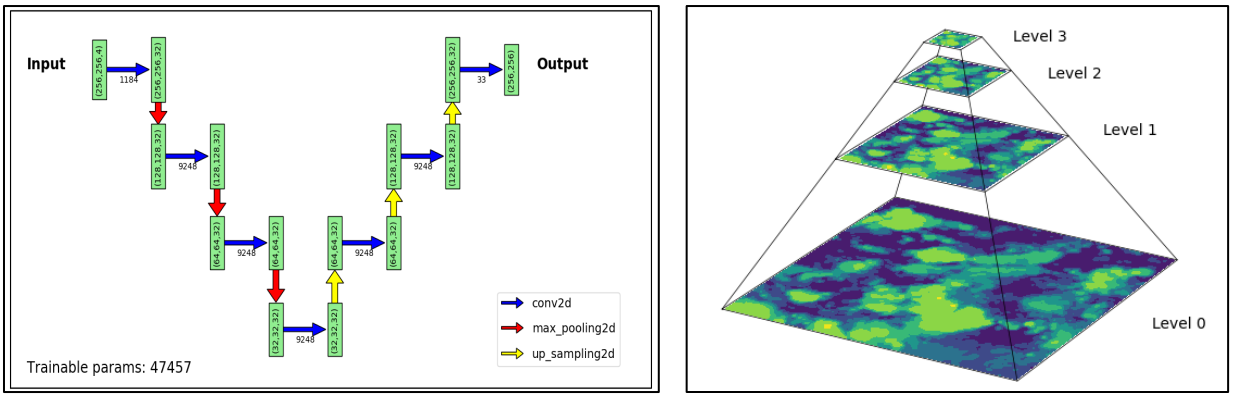


Figure 2. (Left) GREMLIN model architecture where number of parameters are given under the blue arrows and image sizes are shown in green boxes. (Right) Image pyramid corresponding to GREMLIN.

Given those inputs to the interpretable model, we still need a regression framework for making the predictions. Two approaches are used. First is to use a fully connected dense NN, which serves as a nonlinear function approximator. Using this approach is a quick way to confirm that indeed, the interpretable model can reproduce the skill of the CNN, and in some cases even performs better. However, a dense NN is not very interpretable, and so the second approach is to replace that with a linear model. Linear models represent the gold standard of interpretability

because given a set of inputs, the weights of the model tell you exactly how much each input contributes to the output, and that is why human forecasters still make use of linear models (e.g., the Statistical Hurricane Intensity Prediction Scheme, SHIPS, DeMaria 2022). Given that the ability to represent nonlinearity is an important contributor to ML skill, a linear regression must include terms to represent nonlinear functions. It was found that GREMLIN skill could be reproduced with a model of the form

$$y = \sum_{i=1}^{n} w_i x_i + \sum_{i=1}^{n} \sum_{j=1}^{j \le i} w_{i,j} x_i x_j = \sum_{i=1}^{N} w_i x_i \quad (1)$$

where n is the number of inputs. This model includes linear functions of each input, two-way interaction terms, and quadratic functions of each input. The number of terms in this model is given by $N = n + n(n+1)/2 = 2144$ for 64 inputs. The standard (e.g., `scipy.linalg.lstsq`) approach to solving the generalized least squares problem is through applying singular value decomposition to the normal equations. However, this involves inverting a matrix that has the shape of the number of inputs (64) by the number of data points (8e+7 for CONUS2), which produces memory exhaustion. The standard approach for linear regression when the dataset is too large for memory is Stochastic Gradient Descent (e.g., `sklearn.linear_model.SGDRegression.partial_fit`). However, this was found to be prohibitively slow, and many passes over the dataset are required for convergence. Thus, the approach used to solve Equation (1) is employing the linear minimum mean squared error (MMSE) estimator. From the orthogonality principle (Papoulis and Pillai 2002), the solution of (1) is given by

$$w^T = C_x^{-1} C_{yx} \quad (2)$$

where C_x is the autocovariance matrix and C_{yx} is the cross-correlation matrix. Note that the memory required goes as the number of inputs and does not depend on the number of data points. These matrices can be calculated with just one pass over the data, accumulating the sums of x_i, y, $x_i x_j$, and $y x_i$ for inputs i,j. This approach also has the advantage that ablation studies can be conducted without needing to re-fit the model: you simply drop the rows and columns in C_x and C_{yx} for the inputs you want to remove and recalculate w using (2). One might raise the question whether this type of model is properly called "machine learning". I would argue that this approach can be considered to be the simplest form of machine learning, and that calculating the sums required to solve for the weights in (2) can only be performed by machine. One important difference with ML is that this approach does not have an optimizer since weights are calculated explicitly, rather than through an iterative process.

One aspect of the MMSE approach is that it allows us to observe that the autocovariance matrix is nearly singular. This shouldn't be entirely surprising given the correlation between different infrared channels on ABI and the correlation between different levels of the image pyramid. To what degree this can explain the limitations of ML noted in Section 2 is an open question. Reducing the number of inputs does reduce the condition number and increase the determinant, but also produces a significant loss of skill. Is it possible that use of numerous inputs provides value through a "Wisdom of Crowds" (Tetlock and Gardner 2015) ensemble averaging mechanism?

For initializing NWP, three-dimensional fields of radar reflectivity are required. I'm treating this problem as separable, where GREMLIN is used to derive horizontal maps of composite (vertical maximum) radar reflectivity, and a separate model is used to specify the vertical profile. The profile model uses the composite reflectivity to set the scale (maximum value) of the profile and determines the shape of the profile based on the composite reflectivity, the C13 brightness temperature, and whether the pixel is convective or stratiform. This model was derived and validated by Lee et al. (2022) but the details are outside of the scope of my dissertation.

The GREMLIN technique was originally developed for the purpose of *initializing* convection in NWP, which is the mechanism in HRRR. However, in the next generation of modeling systems, actual *assimilation* of radar reflectivity using BECs from ensembles is being considered. In that situation, uncertainty estimates become important, allowing the modeling system to blend between actual radar data and this synthetic radar data, based on the uncertainty. Moreover, it is clear that in certain situations (e.g., hurricane central dense overcast) there is simply little to no information content, and ideally our model would be able to flag situations where estimates lack information content. Initial results using a log likelihood loss function (Barnes and Barnes 2021) in GREMLIN produces uncertainty estimates that scale with the magnitude of the predictions. On one hand, these results make sense (e.g., Tian et al. 2013), however, even though convective echoes have high values, GREMLIN predictions are often very good. Meanwhile stratiform echoes typical have low to mid-range values but can be very uncertain due to a lack of cloud top information. It is possible that accounting for the non-Gaussian character of radar echoes using the approach in Barnes et al. (2021) might improve these results. Another potential approach is to screen scenes based on the magnitude of texture in the scene (Goldenstern 2022). One important test that uncertainty estimates should meet is whether conditioning the ML predictions on the uncertainty results in improved performance statistics.

Errors are typically characterized with a norm such as the mean squared error, but this only tells us about the accuracy in a mean sense and does not tell us whether the predictions have temporal variability that is consistent with the truth dataset. Hilburn et al. (2021c) assessed temporal consistency using a combination of time-histograms over limited areas and using Fourier Transforms. The time-histogram approach is useful for assessing individual storms, because if the area-of-interest becomes too large, jumpiness in different parts of the image can "cancel out". The Fourier approach is good for a global characterization of spatio-temporal characteristics and shows that MRMS has more small-scale "noise", while GREMLIN has more temporal variability on larger spatial scales. Examining loops of GREMLIN predictions shows that it can exhibit jumpy (on/off) behavior, particularly in stratiform areas. While that jumpiness could be due lightning, which naturally has jumpy temporal behavior, the jumpy behavior has been observed in areas without lightning as well. This inconsistency is less important for DA applications, however getting accurate time variability is important for nowcasting applications where meteorologists examine time trends in radar observables. The dissertation will examine temporal consistency of GREMLIN predictions and evaluate how that depends on model complexity (i.e., type of model and the number of parameters), whether certain inputs are primarily responsible for jumpy behavior, and whether there is evidence for noise amplification by nonlinearity. Initial results indicate the linear model is more temporally consistent than the CNN. Temporal consistency is an important metric for determining the fitness of a method for certain applications, but it is an area of ML that has received little attention until recently. Some of the

latest nowcasting models (e.g., Ravuri et al. 2021) incorporate temporal consistency into the loss function, which is an approach that may be considered.

4. Research Plan

This section outlines the dissertation content, describing the planned research and explaining how each step contributes to the outcome.

- **Chapter 1: GREMLIN Proof-of-Concept**
  - Can ML be used as an observation operator to assimilate/initialize GOES-R observations in convection allowing models in cloudy/precipitating scenes? What are the results from retrospective forecasts assimilating GREMLIN (in place of MRMS) in the RRFS? Where does GREMLIN have the largest impact?
  - How does GREMLIN make skillful predictions? What XAI tools are best for image-to-image translation problems? How can additional tools such as ablation studies and synthetic datasets be used to evaluate models?
  - Why is the loss function so important? How do you use the loss function to focus the attention of the network on the most important aspects of one's dataset?

- **Chapter 2: Interpretable GREMLIN**
  - What exactly is the "spatial context" that CNNs utilize to make skillful predictions? How does one know where there is sufficient spatial information to trust predictions? How do I interpret the nearly singular covariance matrix of GREMLIN inputs, and what implications does that have for interpretability, generalizability, and predictor importance assessments?
  - What are the relative contributions of different channels (ABI vs GLM), multiresolution information (pyramid levels), image gradients (kernels), and nonlinearity to GREMLIN skill?
  - How does the temporal consistency of GREMLIN CNN compare to the temporal consistency of Interpretable GREMLIN linear model relative to an MRMS baseline? What conclusions can I draw regarding the potential of GREMLIN to be used for nowcasting applications?

- **Chapter 3: Global GREMLIN Validation**
  - Does GREMLIN perform similarly on the 2 km FD grid as the 3 km HRRR grid? How important are additional GOES inputs from C02, C05, and lightning area? Does inclusion of NWP fields of humidity and temperature improve false alarms that occur for GREMLIN in winter scenes?
  - Can I see evidence of a covariance shift between land and ocean scenes? Does validation using DPR-derived composite reflectivity provide evidence that GREMLIN is trustworthy over the ocean? Are any changes to lightning accumulation period required over the ocean where lightning is relatively less frequent?
  - How does GREMLIN skill vary as a function of precipitation type (convective vs stratiform) and as a function of scale (mesoscale vs synoptic scale)? Are validation results consistent with uncertainty estimates made by GREMLIN?

5. Work Plan

I am working on my dissertation at a level commensurate with 7-8 credits per semester. Chapter 1 of the dissertation has been published (Hilburn et al. 2021a). Chapter 2 is in preparation and will soon be submitted for publication (May 2022). Chapter 3 research will be conducted Summer 2022-Spring 2023 with planned submission date May 2023. Two key aspects supporting the feasibility of this work are: 1) GOES data archive at CIRA, and 2) high performance computing resources at NOAA.

6. Expected Benefits

The research conducted for this dissertation will have several expected benefits. The first benefit will be to increase and improve utilization of GEO observations, thereby enhancing the value of the investments in GOES-R made by NOAA. Even over CONUS, there are locations with gaps in radar coverage, and there are cases where radars "go down" during periods of interesting weather, and GREMLIN can provide forecasters information in those cases. The second benefit will be to demonstrate the ability of ML to improve operational DA through improved utilization of remotely sensed observations. This is expected to improve NWP forecasts of high-impact weather hazards because of the high spatial- and temporal-resolution information required to correctly forecast these situations. The third benefit will be to provide an explanation for the basis of ML to make better use of remotely sensed observations than traditional techniques, thereby supporting development of interpretable and trustworthy AI models. Improved understanding of spatial context in ML can foster development of advanced algorithms for other phenomena, such as dust detection, gravity waves, and cloud properties (Miller et al. 2017, 2015, 2014). And finally, this dissertation will publish (with DOIs) the datasets developed for this research in CSU Digital Repository (Mountain Scholar), which will support and encourage additional ML research in the atmospheric sciences.

7. Budget

Kyle Hilburn's GREMLIN research is supported by the GOES-R Program through the project "Assimilating GOES-R Latent Heating in FV3 using Machine Learning", funded through 2019 NASA ROSES A.33 Solicitation NNH19ZDA001N-ESROGSS: Earth Science Research from Operational Geostationary Satellite Systems. Period of performance: 11/01/2020 – 10/31/2023. Award number: NA19OAR4320073. Principal Investigator: Kyle Hilburn.

Kyle Hilburn's Linux Workstation, used for data download, preparation, and running real-time GREMLIN predictions, was purchased with funding from the project "CIRA Support to Connecting GOES-R with Rapid-Update Numerical Forecast Models for Advance Short-Term Prediction and Data Fusion Capabilities" supported by GOES-R Program Office. Principal Investigator: Christian Kummerow.

I thank NOAA RDHPCS for access to the Fine Grain Architecture System on Hera, without which this research would not be possible. GREMLIN training was performed on Hera.

My tuition and fees are partly (45%) supported by CSU Employee Study Privilege. I thank Steve Miller for covering the remainder of my tuition and fees using his Start-Up Package.

References

Adelson, E. H., Anderson, C. H., Bergen, J. R., Burt, P. J., and Ogden, J. M., 1984: Pyramid methods in image processing. *RCA Engineer*, 33–41.

Alexander, C., J. Carley, P. L. Heinselman, and L. Harris, 2020: Advancements of the FV3 stand-alone regional model. *American Meteorological Society Annual Meeting*.

Apodaca, K., and M. Zupanski, 2018: Variational and hybrid (EnVar) methodologies to add the capability to assimilate GOES-16/GLM observations into GDAS. *JCSDA Quarterly*, **58**, 12-20.

Bach, S., Binder, A., Montavon, G., Klauschen, F., Müller, K. R., & Samek, W., 2015: On pixel-wise explanations for non-linear classifier decisions by layer-wise relevance propagation. *PloS One*, **10**, doi: 10.1371/journal.pone.0130140.

Back, A., A. Kliewer, J. R. Mecikalski, K. Hilburn, Y. Lee, E. Sebok, D. Dowell, E. C. Bruning, M. Xue, R. Kong, S. Benjamin, E. P. James, C. R. Alexander, G. Ge, K. Pederson, and S. Weygandt, 2022: Novel Convection-Indicating Satellite Products Assimilated in Experimental Rapid Refresh Systems, *American Meteorological Society Annual Meeting*.

Barnes, E. A., and R. J. Barnes, 2021: Controlled abstention neural networks for identifying skillful predictions for regression problems. *arXiv:2104.08236*.

Barnes, E. A., R. J. Barnes, and N. Gordillo, 2021: Adding uncertainty to neural network regression tasks in the geosciences. *arXiv:2109.07250*.

Benjamin, S. G., S. S. Weygandt, J. M. Brown, M. Hu, C. R. Alexander, T. G. Smirnova, J. B. Olson, E. P. James, D. C. Dowell, G. A. Grell, H. Lin, S. E. Peckham, T. L. Smith, W. R. Moninger, and J. S. Kenyon, 2016: A North American hourly assimilation and model forecast cycle: The Rapid Refresh. *Mon. Wea. Rev.*, **144**, 1669-1694, doi: 10.1175/MWR-D-15-0242.1.

Burt, P. J. and Adelson, E. H., 1983: The Laplacian pyramid as a compact image code. *IEEE Trans. Communications*, COM-31(4), 532–540.

Connell, B., E. Dagg, and M. Bowlan, 2020: Day Cloud Phase Distinction RGB Quick Guide. https://rammb.cira.colostate.edu/training/visit/quick_guides/QuickGuide_DayCloudPhaseDistinction_final_v2.pdf.

DeMaria, M., 2022, personal communication, March 29.

Ebert-Uphoff, I., and K. A. Hilburn, 2020: Evaluation, tuning and interpretation of neural networks for meteorological applications. *Bull. Amer. Meteorol. Soc.*, **101**, E2149-E2170, doi: 10.1175/BAMD-D-20-0097.1.

Flora, M., 2021: Developing machine learning-based severe weather hazard guidance for the Warn-on-Forecast System (WoFS), *NOAA Workshop on Leveraging AI in Environmental Sciences*.

Geer, A., and Coauthors, 2018: All-sky satellite data assimilation at operational weather forecasting centres. *Quart. J. Roy. Meteor. Soc.*, **144**, 1192-1217, doi: 10.1002/qj.3202.

Goldenstern, E., 2022: Assessing the state-dependency of infrared satellite precipitation errors. Master's Thesis, Colorado State University, 81 pp.

Grasso, L. D., and T. J. Greenwald, 2004: Analysis of 10.7-μm brightness temperatures of a simulated thunderstorm with two-moment microphysics. *Mon. Wea. Rev.*, **132**, 815-825.

Grecu, M. and Anagnostou, E. N., 2001: Overland precipitation estimation from TRMM passive microwave observations. *J. Appl. Meteor.*, **40**, 1367–1380.

Guilloteau, C. and Foufoula-Georgiou, E., 2020: Beyond the pixel: Using patterns and multiscale spatial information to improve the retrieval of precipitation from spaceborne passive microwave imagers. *J. Atmos. Oceanic Technol.*, **37**, 1571–1591, doi: 10.1175/JTECH-D-19-0067.1.

Gustafsson, N., and Coauthors, 2018: Survey of data assimilation methods for convective-scale numerical weather prediction at operational centres. *Quart. J. Roy. Meteor. Soc.*, **144,** 1218-1256, doi: 10.1002/qj.3179.

Heaven, D., 2019: Deep trouble for deep learning. *Nature,* **574**, 163-166, doi: 10.1038/d41586-019-03013-5.

Hilburn, K., M. Marchand, Y. Lee, C. Kummerow, and C. Alexander, 2018: Using GOES-16 to Improve Short Term Forecasts. *American Meteorological Society Annual Meeting*.

Hilburn, K., S. Miller, M. Marchand, and A. Libardoni, 2019: Using High-Resolution Observations from GOES-16 to Improve Operational Short-Range Forecasting. *AMS / EUMETSAT / NOAA Joint Satellite Conference*.

Hilburn, K., Y. Lee, M. Zupanski, and T.-C. Wu, 2020: Assimilating GOES-R latent heating in FV3 using machine learning. *American Geophysical Union Fall Meeting*.

Hilburn, K. A., I. Ebert-Uphoff, and S. D. Miller, 2021a: Development and interpretation of a neural network-based synthetic radar reflectivity estimator using GOES-R satellite observations. *J. Appl. Meteor. Climatol.*, **60**, 3-21, doi: 10.1175/JAMC-D-20-0084.1.

Hilburn, K., Y. Lee, and I. Ebert-Uphoff, 2021b: Improving GREMLIN: A Case Study in AI Application Development, *NOAA Workshop on Leveraging AI in Environmental Sciences*.

Hilburn, K., Y. Lee, and I. Ebert-Uphoff, 2021c: GREMLIN: GOES radar estimation via machine learning to inform NPW. *American Geophysical Union Fall Meeting*.

Jones, T. A., D. Stensrud, L. Wicker, P. Minnis, and R. Palikonda, 2015: Simultaneous radar and satellite data storm-scale assimilation using an ensemble Kalman filter approach for 24 May 2011. *Mon. Wea. Rev.*, **143**, 165-194, doi: 10.1175/MWR-D-14-00180.1.

Jones, T. A., P. Skinner, N. Yussouf, K. Knopfmeier, A. Reinhart, X. Wang, K. Bedka, W. Smith, Jr., and R. Palikonda, 2020: Assimilation of GOES-16 radiances and retrievals into the Warn-on-Forecast system. *Mon. Wea. Rev.*, **148,** 1829-1859, doi: 10.1175/MWR-D-19-0379.1.

Kong, R., M. Xue, A. O. Fierro, Y. Jung, C. Liu, E. R. Mansell, and D. R. MacGorman, 2020: Assimilation of GOES-R Geostationary Lightning Mapper flash extent density data in GSI EnKF for the analysis and short-term forecast of a mesoscale convective system. *Mon. Wea. Rev.*, **148**, 2111-2133, doi: 10.1175/MWR-D-19-0192.1.

Kummerow, C., Y. Hong, W. S. Olson, S. Yang, R. F. Adler, J. McCollum, R. Ferraro, G. Petty, D.-B. Shin, and T. T. Wilheit, 2001: The evolution of the Goodard profiling algorithm (GPROF) for rainfall estimation from passive microwave sensors. *J. Appl. Meteor.*, **40**, 1801-1820.

Lapuschkin, S., S. Wäldchen, A. Binder, G. Montavon, W. Samek, and K.-R. Müller, 2019: Unmasking Clever Hans predictors and assessing what machines really learn. *Nature Communications*, **10**, 1-8, doi: 10.1038/s41467-019-08987-4.

Lee, Y., C. D. Kummerow, and M. Zupanski, 2022: Latent heating profiles from GOES-16 and its impacts on precipitation forecasts. *Atmos. Measurement Techs.*, submitted.

Lin, J., S. S. Weygandt, S. G. Benjamin, and M. Hu, 2017: Satellite radiance data assimilation within the hourly updated Rapid Refresh. *Wea. Forecasting*, **32**, 1273-1287, doi: 10.1175/WAF-D-16-0215.1.

Line, W. E., T. J. Schmit, D. T. Lindsey, and S. J. Goodman, 2016: Use of geostationary super rapid scan satellite imagery by the Storm Prediction Center. *Wea. Forecasting*, **31**, doi: 10.1175/WAF-D-15-0135.1, 483-494.

Mamalakis, A., I. Ebert-Uphoff, and E. A. Barnes, 2022: Neural network attribution methods for problems in geoscience: A novel synthetic benchmark dataset. *Envir. Data Science*, submitted.

McGovern, A., R. Lagerquist, D. J. Gagne II, G. E. Jergensen, K. L. Elmore, C. R. Homeyer, and T. Smith, 2019: Making the black box more transparent: Understanding the physical implications of machine learning. *Bull. Amer. Meteor. Soc.*, **100**, 2175-2199, doi: 10.1175/BAMS-D-18-0195.1.

Miller, S. D., Y.-J. Noh, and A. K. Heidinger, 2015: Liquid-top mixed-phase cloud detection from shortwave-infrared satellite radiometer observations: A physical basis. *J. Geophys. Res. Atmos.*, **119**, 8245-8267, doi: 10.1002/2013JD021262.

Miller, S. D., W. C. Straka III, J. Yue, S. M. Smith, M. J. Alexander, L. Hoffman, M. Setvák, and P. T. Partain, 2015: Upper atmospheric gravity wave details revealed in nightglow satellite imagery. *Proc. Nat. Acad. Sci.*, **112**, doi: 10.1073/pnas.1508084112.

Miller, S. D., R. L. Bankert, J. E. Solbrig, J. M. Forsythe, Y.-J. Noh, and L. D. Grasso, 2017: A dynamic enhancement with background reduction algorithm: Overview and application to satellite-based dust storm detection. *J. Geophys. Res. Atmos.*, **112**, 12938-12959, doi: 10.1002/2017JD027365.

Montavon, G., Samek, W., & Müller, K. R., 2018: Methods for interpreting and understanding deep neural networks. *Digital Signal Processing*, *73*, 1-15. doi: 10.1016/j.dsp.2017.10.011.

Olson, W. S., Y. Hong, C. D. Kummerow, and J. Turk, 2001: A texture-polarization method for estimating convective-stratiform precipitation area coverage from passive microwave radiometer data. *J. Appl. Meteor.*, **40**, 1577-1591.

Papoulis, A., and S. U. Pillai, 2002: Probability, Random Variables, and Stochastic Processes, 4th Ed., McGraw Hill, 852 pp.

Pfreundschuh, S. 2022: personal communication, March 10.

Ravuri, S., K. Lenc, M. Willson, D. Kangin, R. Lam, P. Mirowski, M. Fitzsimons, M. Athanassiadou, S. Kashem, S. Madge, R. Prudden, A. Mandhane, A. Clark, A. Brock, K. Simoyan, R. Hadsell, N. Robinson, E. Clancy, A. Arribas, and S. Mohamed, 2021: Skilful precipitation nowcasting using deep generative models of radar. *Nature*, **597**, 672-677, doi: 10.1038/s41586-021-03854-z.

Roebber, P., 2009: Visualizing multiple measures of forecast quality. *Wea. Forecasting*, **24**, 601-608, doi: 10.1175/2008WAF2222159.1.

Ronneberger, O., P. Fischer, and T. Brox, 2015: U-Net: Convolutional networks for biomedical image segmentation. *MICCAI*, 234-241.

Rutledge, S. A., K. Hilburn, A. Clayton, B. Fuchs, and S. D. Miller, 2020: Evaluating Geostationary Lightning Mapper flash rates within intense convective storms. *J. Geophys. Res. Atmos.*, **125**, e2020JD032827, doi: 10.1029/2020JD032827.

Tetlock, P. E., and D. Gardner, 2015: Superforecasting: The Art and Science of Prediction. Crown Publishers, 340 pp.

Tian, Y., G. J. Huffman, R. F. Adler, L. Tang, M. Sapiano, V. Maggioni, and H. Wu, 2013: Modeling errors in daily precipitation measurements: Additive or multiplicative? *Geophys. Res. Letts.,* **40**, 2060-2064, doi: 10.1002/grl.50320.

Toms, B. A., Barnes, E. A., & Ebert-Uphoff, I., 2019: Physically interpretable neural networks for the geosciences: Applications to Earth system variability. *arXiv:1912.01752*.

Walther, A., W. Straka, and A. K. Heidinger, 2013: ABI algorithm theoretical basis document for daytime cloud optical and microphysical properties (DCOMP). Version 3.0, 11-June-2013, 66 pp.

Wilks, D. S., 2006: Statistical methods in the atmospheric sciences, 2nd Ed., Academic Press, 627 pp.

Zhang, Y., F. Zhang, and D. J. Stensrud, 2018: Assimilating all-sky infrared radiances from GOES-16 ABI using an ensemble Kalman filter for convection-allowing severe thunderstorms prediction. *Mon. Wea. Rev.*, **146**, 3363-3381, doi: 10.1175/MWR-D-18-0062.1.

Zhang, Y., D. J. Stensrud, and F. Zhang, 2019: Simultaneous assimilation of radar and all-sky satellite infrared radiance observations for convection-allowing ensemble analysis and prediction of severe thunderstorms. *Mon. Wea. Rev.*, **147**, 4389-4409, doi: 10.1175/MWR-D-19-0163.1.

Doctoral Preliminary Examination Research Prospectus

Environmental Controls on Tropical Land-Aerosol-Cloud Interactions

Gabrielle R. Leung

Department of Atmospheric Science
Colorado State University

24 March 2023

Doctoral Committee:

Advisor: Susan C. van den Heever

Shantanu Jathar
Sonia M. Kreidenweis
Steven D. Miller

1. Overview and Goals

Extensive land cover changes (LCC) in the tropics are driven by anthropogenic activities like deforestation (**Fig. 1**), agricultural expansion, and urbanization (Lambin et al. 2003; Turubanova et al. 2018). These perturbations propagate to the atmosphere via surface energy, moisture, and momentum budgets (Mahmood et al. 2014). Where LCC coincide with strong land-atmosphere coupling—as in the tropics, where much convection is driven by mesoscale flows sensitive to surface properties (Yang and Slingo 2001)—they exert as strong an influence on atmospheric temperature as CO_2 concentrations do (Pitman et al. 2009; Avila et al. 2012).

Quantifying how LCC impact clouds, precipitation, and radiation is difficult given the complex, non-linear interactions involved (Laguë et al. 2019; Drager et al. 2022). LCC often occur alongside changes to aerosols (Heald and Spracklen 2015) that further modify the surface energy budget (Grant and van den Heever 2014; Park and van den Heever 2022). Land-aerosol-cloud interactions further vary due to differences in meteorology (Findell and Eltahir 2003a; Lawrence and Vandecar 2015). As a result, *disagreement exists not only about the magnitude of LCC impacts on convection, but also their overall sign* (Chen and Avissar 1994; Takahashi et al. 2017).

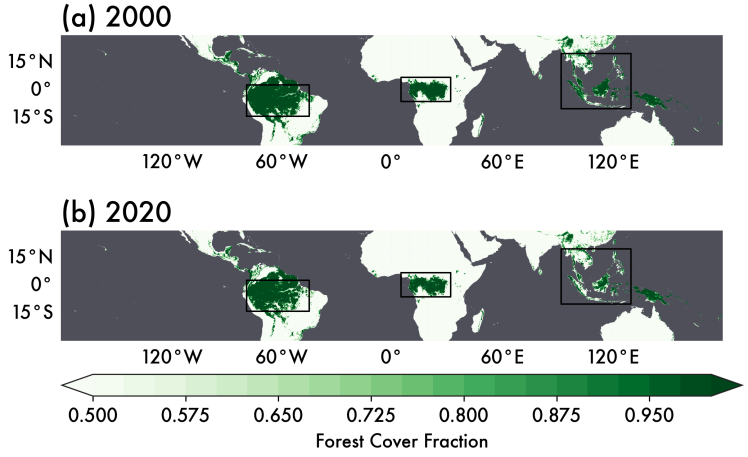


Figure 1. Fraction of forest cover in (a) 2000 and (b) 2020. The three boxed regions are frontiers of tropical deforestation (west to east: the Amazon, the Congo, and Southeast Asia). Data from UMD Global Forest Change dataset (Hansen et al., 2013).

With current scientific understanding, it is difficult to determine even whether a given LCC would likely lead to increases or decreases in precipitation, much less how this might vary in different regions. Such answers are *crucial to fully assess the impact of anthropogenically-driven LCC on the earth system*. Local impacts on clouds and precipitation are important, as almost half the global population lives in the tropics (Kummu and Varis 2011). Furthermore, tropical clouds are an essential driver of global circulations and hold outsized importance for global climate (Riehl and Malkus 1958). It is therefore essential to quantify the impacts of LCC on clouds across environments and improve our understanding of the processes driving these impacts.

As such, this work seeks to **describe how the impacts of land surface perturbations on tropical convection varies across thermodynamic and aerosol environments**. We will use both *satellite observations* and *idealized simulations* to answer three science questions (SQs):

SQ1: Which individual or combined surface properties are the primary drivers of tropical land surface-cloud interactions and their impacts on precipitation and radiation?

SQ2: How do aerosol concentrations and radiative characteristics modulate land surface impacts on convection?

SQ3: What are the key meteorological parameters affecting land-aerosol-cloud interactions?

2. Background

2.1. Land Surface-Cloud Interactions

The surface energy budget is a balance between net radiation and surface heat fluxes (**Fig. 2a**) (Mahmood et al. 2014; Gentine et al. 2019). LCC can change this budget first by modifying the net radiation at the surface: increasing albedo (brighter surface, e.g., dark forest to bright concrete) increases reflected shortwave radiation. Second, surface properties control the partition of sensible to latent heat fluxes (Bowen ratio): decreasing evaporative resistance (more moisture, e.g., broadleaf to needleleaf trees) increases latent heat flux at the expense of sensible heat flux. Finally, surface properties shift the balance of radiation and heat fluxes: decreased roughness (smoother surface, e.g., tall buildings to short crops) hinders heat removal through turbulent fluxes in favor of radiating energy via outgoing thermal emission.

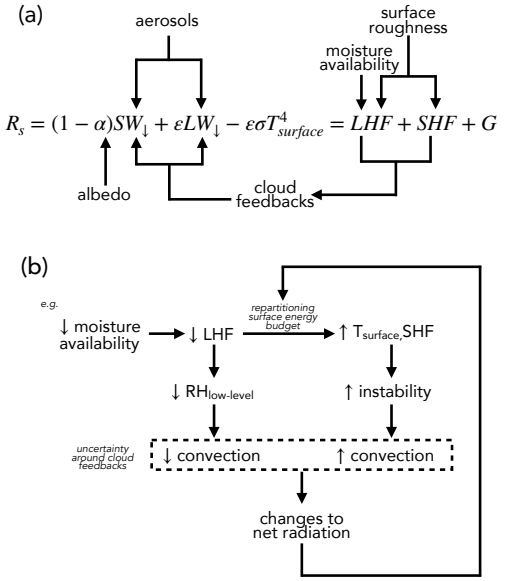


Figure 2. Conceptual diagram illustrating how land surface and aerosol perturbations influence the surface energy budget. The surface energy budget (a) is a balance between net radiation (R_s=net shortwave + net longwave), latent heat fluxes (LHF), sensible heat fluxes (SHF), and heat storage (G; negligible for land over long timescales). (b) shows how a surface perturbation can drive contrasting feedbacks on clouds depending on the surface energy budget partitioning (adapted from Chen et al. 2019).

Changes to the surface energy budget modify the atmosphere above the perturbed surface and even downstream (Davin and de Noblet-Ducoudré 2010). In particular, changes to surface fluxes can have critical impacts on the formation and development of convection. Clouds that form then shield the surface from incoming radiation and trap outgoing radiation, further modifying the surface energy budget. Due to these feedbacks, the cloud response to a given surface perturbation is not straightforward. A decrease in soil moisture (**Fig. 2b**) may reduce latent heat fluxes and moisture (suppressing convection), but this is compensated by increased sensible heat fluxes and instability (supporting convection) (Chen et al. 2019). Because of this sensitivity to the partitioning between heat fluxes, the net convective response depends to some extent on cloud type. Deep, precipitating clouds driven by large-scale forcing tend to be less impacted by surface perturbations compared to shallow, non-precipitating clouds, which can still be important for surface radiation (Baidya Roy and Avissar 2002; Gentine et al. 2019). LCC may promote the development of shallow clouds into deeper ones (Cioni and Hohenegger 2017), though research into this is relatively limited. Simulations with detailed representations of cloud physics and land surface-cloud processes are required to assess the impact of a given LCC on the cloud field and its precipitation and radiative effects.

2.2. Establishing mechanisms behind land surface-cloud interactions

Incorporating the cloud response to surface perturbations in numerical models is increasingly important as LCC accelerates (Lambin et al. 2003) and plays a growing impact on weather and climate (Schneider et al. 2019; Pongratz et al. 2021). However, representing these

interactions is non-trivial, and understanding the processes driving convective responses to LCC is crucial. Most models represent LCC as a step change between categories (e.g., cropland to urban, forest to bare soil). In reality though, LCC influence many distinct properties at the same time (Davin and de Noblet-Ducoudré 2010; Grant and van den Heever 2014; Laguë et al. 2019).

The impacts of each surface parameter on convection are complex, even in the simplest case where only one parameter is changed (e.g., intensifying irrigation leading to increased surface moisture). Recent work by Drager et al. (2022) showed isolated perturbations to soil moisture drive non-monotonic changes to rainfall, implying that a step change in surface moisture—as is commonly used in modeling—may not capture the true precipitation response.

To further complicate the matter, interactions between parameters are possible, either dampening the atmospheric response to LCC (Bell et al. 2015; Takahashi et al. 2017; Duveiller et al. 2018) or heightening it (Grant and van den Heever 2014; Laguë et al. 2019). The specific combinations of surface properties involved in a given LCC are therefore important. Even studies on the same LCC over the same region can give opposing results depending on which properties are represented (e.g., Schneck and Mosbrugger 2011; Takahashi et al. 2017), and it is difficult to determine which is most credible. This complexity also means it is not clear whether mechanisms driving cloud responses from a given LCC can be applied across the board. For example, Chen et al. (2019) found *decreased* roughness length due to deforestation tends to increase precipitation, whereas Dado and Narisma (2019) found *increasing* roughness length due to urbanization tends to increase precipitation as well.

Without separately perturbing surface properties over the full parameter space, it is impossible to isolate which property or combination of properties is responsible for—and what physical processes are involved in—changes to the cloud field. This leads to our first science question: Which individual or combined surface properties are the primary drivers of tropical land surface-cloud interactions and their impacts on precipitation and radiation? We hypothesize that *convection is more sensitive to non-radiative surface parameters (e.g., surface moisture availability, roughness length) than to radiative ones (e.g., albedo) in the high-radiation tropics. Surface radiation is most sensitive to roughness length via impacts on shallow clouds driven by thermals, whereas precipitation is most sensitive to synergistic interactions between moisture availability and roughness length that support the transition from shallow to deeper convection.*

2.3. Aerosol Impacts on Land Surface-Cloud Interactions

Perturbations to the land surface tend to coincide with changes to the aerosol environment (Heald et al. 2008; Heald and Spracklen 2015). Drivers of LCC may drive transient aerosol perturbations, as when wildfires clear forested land and emit smoke (Tosca et al. 2011; Sena et al. 2013). LCC can also lead to persistent changes in aerosol magnitude and type. For instance, urbanization is associated with combustion-related aerosol, while bare soil is susceptible to dust lofting (Wu et al. 2012; Heald and Spracklen 2015; Bukowski and van den Heever 2022).

Aerosols modify clouds both through microphysics (indirect effect) and radiation (direct effect) (Twomey 1977; Albrecht 1989; Tao et al. 2012). Taken separately, aerosol perturbations

and LCC can influence clouds to a similar order of magnitude (Grant and van den Heever 2014; Schneider et al. 2019). In combination, aerosol-cloud interactions can either dampen or amplify surface impacts on convection (Ten Hoeve et al. 2011; Grant and van den Heever 2014; Park and van den Heever 2022), though the factors controlling these interactions are understudied.

There is abundant observational and modeling evidence of the microphysical impacts of aerosols on cloud lifetime, rain onset and duration, and precipitation efficiency (e.g., Berg et al. 2008; Li et al. 2013; Altaratz et al. 2014; Dagan et al. 2015). Surface perturbations associated with cities have been shown to change the likelihood of cloud formation; clouds that do form can subsequently be impacted by emissions of urban aerosol particles, which act as additional cloud condensation nuclei (CCN) and lead to changes in cloud and storm characteristics (van den Heever and Cotton 2007; Carrió et al. 2010; Sarangi et al. 2018).

Aerosol particles also lead to the extinction of downwelling radiation and change the surface energy budget (**Fig. 2a**) (Jiang and Feingold 2006). This can dampen cloud responses to surface perturbations, as shown by Park and van den Heever (2022) for soil moisture. The extent to which the surface energy budget is impacted depends on aerosol loading, since synergistic interactions between moderate surface and aerosol perturbations can lead to the opposite cloud response versus perturbing them separately (Grant and van den Heever 2014). Aerosol type also matters for the surface energy budget. Particles like ammonium sulfate primarily prevent radiation from reaching the surface through scattering, thereby reducing the flux of radiation into the surface. On the other hand, particles such as black carbon and dust do reduce incoming shortwave radiation, but simultaneously absorb and reemit longwave radiation back towards the surface (and atmosphere)—in that case, aerosol impacts on the net surface radiation may be smaller due to competing shortwave and longwave effects (Zhang et al. 2008; Lee et al. 2014).

Given that aerosol and land surface perturbations are often concurrent, it is valuable to understand the relative magnitude of their impacts and how the two interact. This leads to our second science question: How do aerosol concentrations and radiative characteristics modulate land surface impacts on convection? We hypothesize that *aerosol impacts on land surface-cloud coupling are non-monotonic. Moderate aerosol loadings (i.e., middle tercile of observations) enhance surface influence on clouds through synergistic interactions, while extreme aerosol loadings (i.e., top tercile of observations) diminish the importance of surface properties by reducing insolation. These impacts are stronger for scattering than for absorbing aerosol since the latter's shortwave and longwave effects on net radiation partially compensate.*

2.4. The Influence of Background Meteorology

Although LCC happen all over the tropics, their impacts are not always the same across locations. The few existing studies examining LCC at global scales show that differences in background meteorology across regions drive different emergent responses to LCC (Davin and de Noblet-Ducoudré 2010; Seneviratne et al. 2010; Winckler et al. 2017). Understanding the environmental dependence of land-aerosol-cloud interactions is therefore important to identifying where and when LCC and associated aerosol perturbations have the most influence.

The thermodynamic environment varies substantially from region to region, even within the tropics. Though the Amazon, the Congo, and Southeast Asia have similar tropical rainforests (**Fig. 1**), their distribution of key environmental parameters is different (**Fig. 3**). This may explain discrepancies in past studies about the precipitation response to deforestation. Deforestation in the Congo has been shown to drive decreased rainfall, while the opposite is true for Southeast Asia. It is possible that once forests are removed as a moisture source, the air in the Congo is otherwise too dry (**Fig. 3d**) to support rain formation, whereas most of Southeast Asia is close enough to oceans that additional moisture can be fluxed in (Lawrence and Vandecar 2015; Winckler et al. 2017). Other studies have similarly shown that the precipitation response to urban growth depends strongly on whether a city is coastal or inland (Zhang et al. 2022).

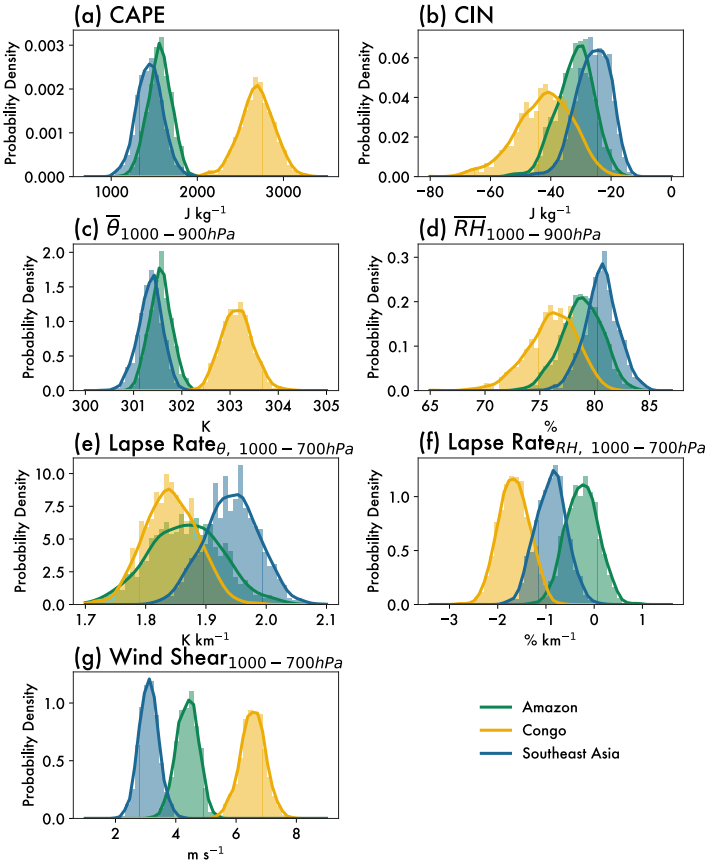


Figure 3. Probability distributions of monthly mean environments for three tropical forests: (a) convective available potential energy (CAPE), (b) convective inhibition (CIN), (c) wind speed shear from 1000-700 hPa, (d) mean low-level potential temperature (θ), (e) mean low-level relative humidity (RH), (f) θ lapse rate, and (g) RH lapse rate. Low-level means (d,e) are taken from 1000-900hPa, and lapse rates are the mean from 1000-700hPa, Histograms are populated with bootstrapped estimates of the mean sampled (n=1000) from all land ERA-5 data (Δx=0.25º) for 2000-2020. Bounding boxes for each region are in Fig. 1.

Meteorology also varies substantially between seasons, especially in monsoon regions. Multiple studies have noted the impacts of LCC such as deforestation or urbanization depend on the season as a response to changes in humidity throughout the column (Durieux et al. 2003; Ten Hoeve et al. 2011; Lazzarini et al. 2015). This further points to environmental-dependence of land-aerosol-cloud interactions.

Though not much is known about the environmental sensitivities of convective responses to surface perturbations in general (as opposed to specific types of LCC), most of what research does exist revolves around the wet/dry soil advantage problem (i.e., does it rain more over wet or dry soils?). Dry soils support stronger sensible heat fluxes, which trigger rain by deepening the boundary layer until it reaches the level of free convection. On the other hand, wet soils support stronger latent heat fluxes and moister boundary layers, which can lower the lifted condensation level to trigger convection more easily. Which mechanism prevails appears to depend on environmental properties such as near-surface temperature, low-level humidity, and the ratio of

potential temperature to moisture lapse rates (Findell and Eltahir 2003a,b; Gentine et al. 2013; Cioni and Hohenegger 2017). These factors, along with other variables such as vertical wind shear, also play a role in modulating cloud responses to aerosol perturbations (Lee et al. 2008; Carrió and Cotton 2011; Yamaguchi et al. 2019; Sokolowsky et al. 2022).

The need to test the sensitivity of land-aerosol-cloud interactions across a range of thermodynamic conditions leads to our third science question: What are the key meteorological parameters affecting land-aerosol-cloud interactions? We hypothesize based on past literature that *land-aerosol-cloud interactions are most sensitive to boundary layer humidity, followed by the low-level temperature and moisture lapse rates, then wind shear. LCC has the largest impact on clouds and precipitation when environmental conditions only marginally support convection (e.g., moderate CAPE or humidity).*

3. Proposed Work

To address the uncertainties surrounding land-aerosol-cloud interactions and their environmental controls, a systematic assessment of such interactions is required in both real-world scenarios and across the full range of potential environments. We therefore propose a combined satellite and modeling approach consisting of two primary tasks:

Task 1: estimating the impacts of deforestation and aerosol loading on cloud properties across three tropical rainforest basins using satellite observations

Task 2: describing mechanisms of and environmental controls on land-aerosol-cloud interaction using idealized LES ensembles and variance-based sensitivity analysis

3.1. Impacts of land use and aerosol on cloudiness across regions from satellites

For Task 1, we will use **long-term, high-spatial resolution satellite data to estimate the impact of deforestation on cloud properties across three tropical rainforest regions**: the Amazon, the Congo, and Southeast Asia. These three regions are major frontiers of tropical forest loss (**Fig. 1**), and also span a range of aerosol and thermodynamic environments (**Fig. 3**). Comparisons across the three regions will allow us to evaluate real-world atmospheric responses to a prevalent LCC while giving a sense of the potential variability in these responses. Interannual variability in cloud cover is difficult to separate from the cloud response following deforestation. However, it is still possible to estimate deforestation impacts given a sufficiently large sample size and an appropriate statistical method.

Satellite data are well-suited to providing the large number of sample points needed. We will take annual forest extent and change from the UMD Global Forest Change (GFC) product (Δx=30m, annual coverage from 2000-2020, **Fig. 1**), which is derived from Landsat 7 Enhanced Thematic Mapper Plus (ETM+) and Landsat 8 Operational Land Imager (OLI) imagery (Hansen et al. 2013). Over the same period, MODIS Terra and Aqua provide near-daily cloud, aerosol, and atmosphere measurements at four times spanning the diurnal cycle (Aqua: 1:30AM/PM, Terra: 10:30AM/PM) with relatively high spatial resolution (Δx=1km for cloud top height, cloud fraction, and near-infrared water vapor, Δx=3km for Dark Target aerosol optical depth (AOD),

Δx=5km for atmospheric profiles). The GFC forest cover will be scaled to the same resolution as MODIS measurements by taking the mean annual forest cover within each MODIS pixel, such that the two datasets can be collocated to create a sample population with millions of pixels.

To separate the influence of surface changes from interannual variability, we will adapt a statistical method from Crompton et al. (2021). The "difference-in-differences" metric ε for a given property (e.g., annual cloud fraction) describes the change following a deforestation event that can be *attributed* to that LCC and not interannual variability. Deforestation events will be identified in the GFC dataset as pixels which lose >50% of forest cover in a given year. Any changes in the cloud field can be quantified using the temporal difference around the deforestation event for a given property ($\Delta C_{deforested}=C_{deforested,\ y+1}-C_{deforested,\ y-1}$). Nearby intact forest pixels may also experience cloud changes ($\Delta C_{control}=C_{control,\ y+1}-C_{control,\ y-1}$), but these are due to other climatic variations (e.g., ENSO phase). The difference in cloud changes experienced by these two areas will thus be the change due to deforestation only ($\varepsilon_c = \Delta C_{deforested} - \Delta C_{control}$). One major assumption this depends on is that the control pixels are near enough to deforested pixels to experience the same variability in synoptic-scale meteorology, but far enough that they are not directly impacted by the surface perturbation. Crompton et al. (2021) defined this distance as pixels that are 10-25km away from any deforestation event, but we will also test the sensitivity of our results to the threshold used.

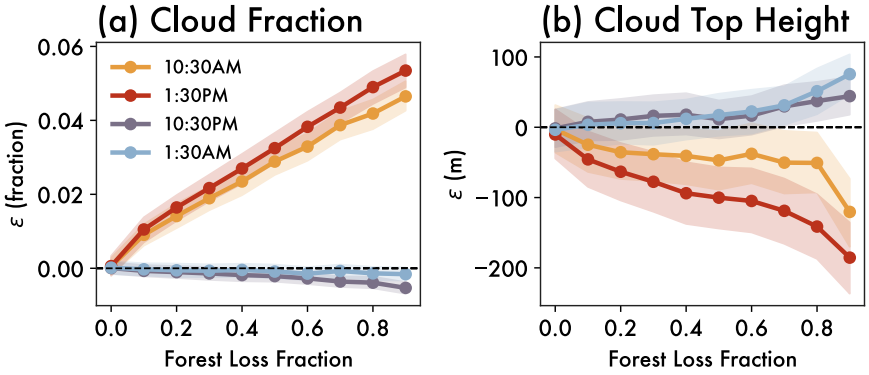


Figure 4. Estimated cloud response to mean forest loss in 1km radius for annual mean (a) cloud fraction (fraction of sky) and (b) cloud top height (m). Different colored lines indicate swaths from different times of day. Dark line indicates the bootstrapped estimate of the mean (n=1000) and shaded area spans the 25th to 75th percent confidence intervals.

Analysis has already been conducted for Southeast Asia (**Fig. 4**). We find deforestation leads to greater daytime cloud fractions and a shift to shallower clouds. This is consistent with Chen et al. (2019), which proposes reduced surface roughness and soil moisture due to deforestation drives anomalous moisture flux from the surrounding maritime regions to support clouds. We also find there is little to no change in nighttime clouds, as expected if the signal is indeed driven by solar heating.

After repeating the above analysis for all three forest regions, we will address our science questions by sub-setting the sample using ancillary data to capture different environments. Compositing **Fig. 4** according to AOD tercile addresses SQ2 on the aerosol modulation of LCC impacts. If our hypothesis is correct, the composite of pixels with AODs falling into the middle tercile of observations would have a magnified response (steeper slope of $\frac{d\varepsilon}{dForest\ Loss}$) versus the lower tercile, while the opposite would hold for the upper tercile. Similarly, compositing according to meteorological variables such as precipitable water and atmospheric stability will

address SQ3 and determine which thermodynamic parameters the cloud response is most sensitive to. Finally, we can also compare the response across the three regions while holding aerosol and meteorology constant to address SQ1. Although all three regions have tropical rainforests, there are nuances in the character of deforestation across the regions which lead to differences in the specific surface properties changed. For example, deforestation in Southeast Asia is heavily driven by conversion of pristine forest to oil palm plantations, which results in smaller changes to albedo, moisture, and roughness length when compared to deforestation in the Congo or the Amazon which are driven by expansion of croplands or livestock feeding grounds, respectively (Lambin et al. 2003). Thus, differences in the cloud responses between these three regions (while holding other variables constant) can give an idea of which surface properties are the primary drivers of cloud responses in the real world.

3.2. Environmental controls on land-aerosol-cloud processes from idealized models

The real-world assessment of how aerosols, meteorology, and the land surface impact clouds proposed in Task 1 will be complemented with a process-focused analysis of a large model ensemble in Task 2 (**Fig. 5**). We plan to use **an ensemble of high-resolution idealized simulations and Gaussian process emulation to elucidate the mechanisms of land-aerosol-cloud interactions and to determine the key properties driving these interactions**.

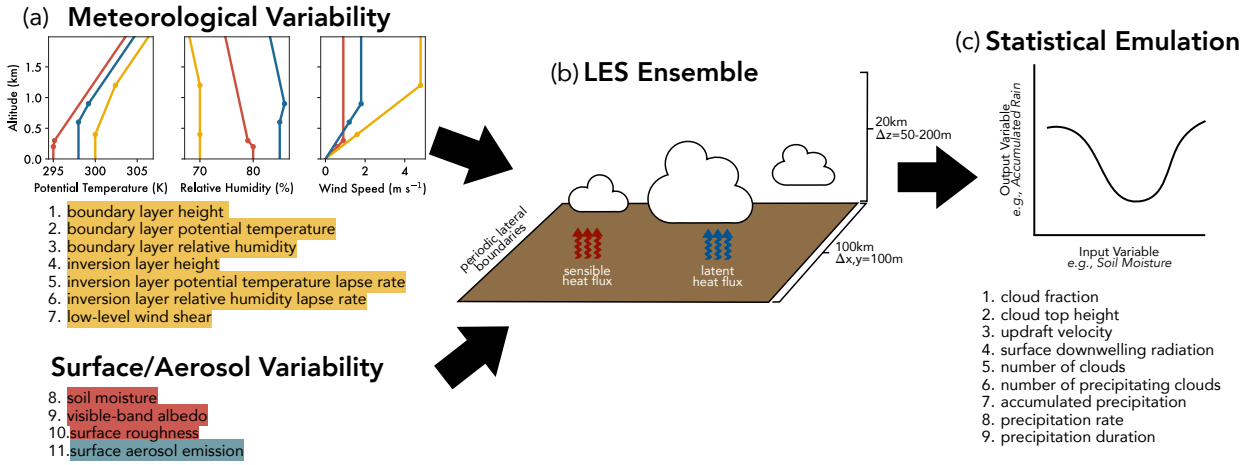


Figure 5. Schematic illustrating the Gaussian process emulation framework. (a) Input parameters consist of surface, aerosol, and thermodynamic variables that are covaried to efficiently fill the physically realistic parameter space. Note that input parameters are color coded according to the relevant science questions. Unique combinations of input parameters are used to construct a set of initial conditions, shown here for three combinations of meteorological variables, as in Igel et al. (2018). Each set of inputs are used to initialize (b) an LES to create a full ensemble. The (c) output parameters are then predicted from the input parameters using a statistical emulator.

We will use the Regional Atmospheric Modeling System (RAMS, current version 6.3.03) to run an ensemble of large eddy simulations (LES) designed to test the sensitivity of convection to a range of surface, aerosol, and meteorological properties. RAMS is a non-hydrostatic, cloud-resolving model with two-moment bin-emulating microphysics and sophisticated aerosol schemes (Cotton et al. 2003; Saleeby and van den Heever 2013). It is coupled to the Land Ecosystem-Atmosphere Feedback (LEAF-3) model (Walko et al. 2000), which allows for dynamic interactions between the land and atmosphere through the two-way exchange of surface

moisture, energy, and momentum fluxes. RAMS also uses a two-stream radiation scheme that incorporates aerosol radiative effects (Harrington 1997; Saleeby and van den Heever 2013). These simulations (**Fig. 5b**) will be run over a 100x100km domain with periodic lateral boundaries, at spatial resolutions fine enough to resolve a range of shallow to deep convection (Δx=100m, Δz=50-200m) for 24 hours (Δt=1s) to fully represent the diurnal cycle of convection. The land surface will be uniform and the initial thermodynamic conditions will be horizontally homogenous (but vertically varying) within each simulation, though these will vary between ensemble members. Convection will then be initialized using random thermal perturbations within the boundary layer. In similar LES we have conducted in the past, realistic cloud fields develop under this set-up within a matter of hours (Leung and van den Heever 2022).

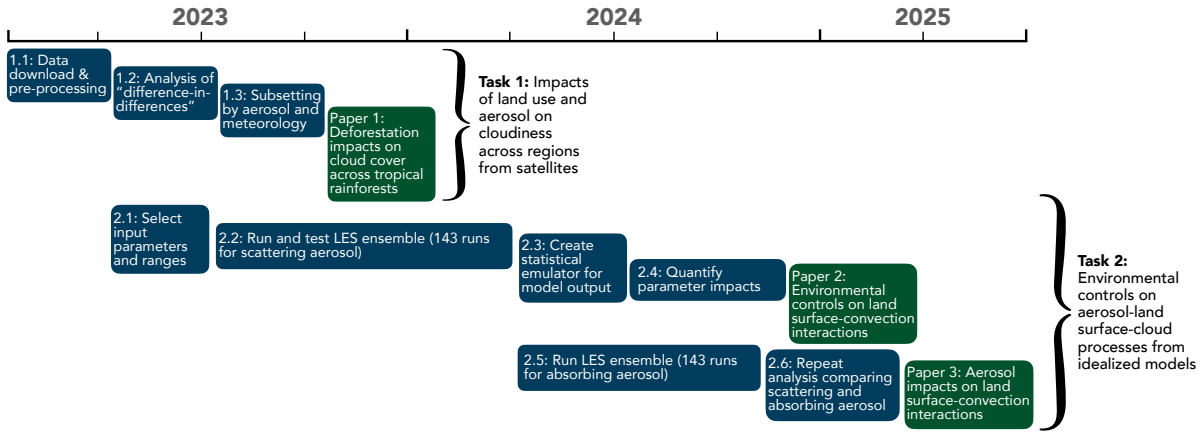
Estimating the individual and combined effects of surface, meteorology, and aerosol parameters requires careful ensemble design, for which we will use a Gaussian process emulation framework (**Fig. 5**) (Lee et al. 2011; Johnson et al. 2015; Igel et al. 2018; Park et al. 2020). First, based on past studies and the hypotheses described earlier, we have already identified eleven input parameters (**Fig. 5a**) that are likely to impact land-aerosol-cloud interactions. A physically realistic range will then be assigned to each input parameter based on past literature and climatology (**Fig. 3**). Second, a set of initial conditions will be designed to capture the range of possible values for all input parameters, to ensure any non-monotonic effects are depicted. Given the large number of input parameters, each additional condition tested adds significant computational costs. For example, if even just three values are used for each of the eleven input parameters, a total of 3^{11}=177,147 simulations are needed—obviously, this is not realistic. Instead, we will use a space-filling Latin hypercube algorithm (Lee et al. 2011) in order to efficiently span the parameter space: only 143 simulations are needed (110 for training, 33 for testing) (Bastos and O'Hagan 2009). Third, we will run our ensemble of LES (**Fig. 5b**) as already described. Fourth, we will identify a set of output parameters (**Fig. 5c**) pertaining to convective properties. These output parameters can consist of domain-/time-averaged values, specific percentiles of values, or cloud counts. Fifth, we will construct an emulator to estimate the relationship between the input and output parameters. This emulator can then be tested for accuracy; if the emulator is not able to predict the output parameters accurately (e.g., if not enough of the ensemble members produce deep convection), further training simulations can be added. Finally, variance-based sensitivity analysis will allow us to estimate the impact of each individual parameter and multi-parameter interaction (Oakley and O'Hagan 2004).

The results of the sensitivity analysis will then be used to address our science questions. We will quantify how much of the variance in output cloud properties can be explained by a given surface parameter or combination of parameters, thereby directly answering SQ1. For example, Park et al. (2020) used the same method to show that 75% of the variance in median updraft speed in sea breeze convection could be attributed to soil moisture. Next, we will identify which meteorological input factors address the most variance, paying particular attention to how the sensitivity of clouds to surface parameters is impacted by differences in initial thermodynamics, thereby addressing SQ3. This analysis will also allow us to assess the relative contributions of

aerosol and surface parameters, which addresses the first half of SQ2 on aerosol loading. We also hypothesized that the radiative properties of the aerosol (whether it is light-scattering or absorbing) are important for land surface interactions. To answer the second half of SQ2, we plan to run two versions of the ensemble: one with ammonium sulfate and one with absorbing carbon particles. Comparing the two ensembles will allow us to confirm or refute our hypothesized responses as a function of aerosol type. For all three science questions, once the input parameters responsible for the most variance are identified, we can then select specific simulations to investigate in further detail (e.g., comparing the evolution of the boundary layer throughout the day in low- and high-surface roughness simulations) to better understand the processes driving these land-aerosol-cloud interactions.

3.3. Work plan

We expect the work proposed here will be completed by the end of Fall 2025. Download and analysis scripts for Task 1 have already been written and tested for Southeast Asia; these are straightforward to reapply for the Amazon and Congo. This task should be completed and result in a paper outlining the regional differences in deforestation impacts on cloud properties from satellite data to be submitted by the end of 2023. Most of the time and computational cost involved in this proposal consists of building the ensemble in Task 2. We have already begun selection of input parameters and uncertainty ranges, and can start running the first set of LES (scattering aerosol) by mid-2023 in parallel to the analysis in Task 1. The analysis procedure for the statistical emulator and variance-based sensitivity analysis will then be done for the first ensemble while the second set of LES (absorbing aerosol) is run. Given the large number of simulations and parameters to test, we expect to write two papers: one on environmental controls and one on aerosol (loading and type) controls on land surface-convection interactions.



In summary, we anticipate that the combined satellite and modeling approach we propose here will offer *novel insight into the processes driving land-aerosol-cloud interactions and how those processes vary across environments*. In doing so, we will not only provide a systematic estimate of how LCC and their associated aerosol perturbations impact cloud properties, but also **improve understanding of where and when these interactions are most probable and therefore essential to weather and climate**.

Albrecht, B. A., 1989: Aerosols, Cloud Microphysics, and Fractional Cloudiness. *Science*, **245**, 1227–1230, https://doi.org/10.1126/science.245.4923.1227.

Altaratz, O., I. Koren, L. A. Remer, and E. Hirsch, 2014: Review: Cloud invigoration by aerosols—Coupling between microphysics and dynamics. *Atmospheric Research*, **140–141**, 38–60, https://doi.org/10.1016/j.atmosres.2014.01.009.

Avila, F. B., A. J. Pitman, M. G. Donat, L. V. Alexander, and G. Abramowitz, 2012: Climate model simulated changes in temperature extremes due to land cover change. *J. Geophys. Res.*, **117**, https://doi.org/10.1029/2011JD016382.

Baidya Roy, S., and R. Avissar, 2002: Impact of land use/land cover change on regional hydrometeorology in Amazonia. *Journal of Geophysical Research: Atmospheres*, **107**, LBA 4-1-LBA 4-12, https://doi.org/10.1029/2000JD000266.

Bastos, L. S., and A. O'Hagan, 2009: Diagnostics for Gaussian Process Emulators. *Technometrics*, **51**, 425–438.

Bell, J. P., A. M. Tompkins, C. Bouka-Biona, and I. S. Sanda, 2015: A process-based investigation into the impact of the Congo basin deforestation on surface climate. *Journal of Geophysical Research: Atmospheres*, **120**, 5721–5739, https://doi.org/10.1002/2014JD022586.

Berg, W., T. L'Ecuyer, and S. van den Heever, 2008: Evidence for the impact of aerosols on the onset and microphysical properties of rainfall from a combination of satellite observations and cloud-resolving model simulations. *Journal of Geophysical Research: Atmospheres*, **113**, https://doi.org/10.1029/2007JD009649.

Bukowski, J., and S. C. van den Heever, 2022: The Impact of Land Surface Properties on Haboobs and Dust Lofting. *Journal of the Atmospheric Sciences*, **79**, 3195–3218, https://doi.org/10.1175/JAS-D-22-0001.1.

Carrió, G. G., and W. R. Cotton, 2011: Urban growth and aerosol effects on convection over Houston. Part II: Dependence of aerosol effects on instability. *Atmospheric Research*, **102**, 167–174, https://doi.org/10.1016/j.atmosres.2011.06.022.

Carrió, G. G., W. R. Cotton, and W. Y. Y. Cheng, 2010: Urban growth and aerosol effects on convection over Houston: Part I: The August 2000 case. *Atmospheric Research*, **96**, 560–574, https://doi.org/10.1016/j.atmosres.2010.01.005.

Chen, C.-C., and Coauthors, 2019: Thermodynamic and Dynamic Responses to Deforestation in the Maritime Continent: A Modeling Study. *Journal of Climate*, **32**, 3505–3527, https://doi.org/10.1175/JCLI-D-18-0310.1.

Chen, F., and R. Avissar, 1994: Impact of Land-Surface Moisture Variability on Local Shallow Convective Cumulus and Precipitation in Large-Scale Models. *Journal of Applied Meteorology and Climatology*, **33**, 1382–1401, https://doi.org/10.1175/1520-0450(1994)033<1382:IOLSMV>2.0.CO;2.

Cioni, G., and C. Hohenegger, 2017: Effect of Soil Moisture on Diurnal Convection and Precipitation in Large-Eddy Simulations. *Journal of Hydrometeorology*, **18**, 1885–1903, https://doi.org/10.1175/JHM-D-16-0241.1.

Cotton, W. R., and Coauthors, 2003: RAMS 2001: Current status and future directions. *Meteorol Atmos Phys*, **82**, https://doi.org/10.1007/s00703-001-0584-9.

Dagan, G., I. Koren, and O. Altaratz, 2015: Aerosol effects on the timing of warm rain processes. *Geophysical Research Letters*, **42**, 4590–4598, https://doi.org/10.1002/2015GL063839.

Davin, E. L., and N. de de Noblet-Ducoudré, 2010: Climatic Impact of Global-Scale Deforestation: Radiative versus Nonradiative Processes. *Journal of Climate*, **23**, 97–112, https://doi.org/10.1175/2009JCLI3102.1.

Drager, A. J., L. D. Grant, and S. C. van den Heever, 2022: A Non-Monotonic Precipitation Response to Changes in Soil Moisture in the Presence of Vegetation. *Journal of Hydrometeorology*, https://doi.org/10.1175/JHM-D-21-0109.1.

Durieux, L., L. A. T. Machado, and H. Laurent, 2003: The impact of deforestation on cloud cover over the Amazon arc of deforestation. *Remote Sensing of Environment*, **86**, 132–140, https://doi.org/10.1016/S0034-4257(03)00095-6.

Duveiller, G., J. Hooker, and A. Cescatti, 2018: The mark of vegetation change on Earth's surface energy balance. *Nat Commun*, **9**, 679, https://doi.org/10.1038/s41467-017-02810-8.

Findell, K. L., and E. A. B. Eltahir, 2003a: Atmospheric Controls on Soil Moisture–Boundary Layer Interactions. Part I: Framework Development. *Journal of Hydrometeorology*, **4**, 552–569, https://doi.org/10.1175/1525-7541(2003)004<0552:ACOSML>2.0.CO;2.

——, and ——, 2003b: Atmospheric controls on soil moisture-boundary layer interactions: Three-dimensional wind effects. *Journal of Geophysical Research: Atmospheres*, **108**, https://doi.org/10.1029/2001JD001515.

Gentine, P., A. A. M. Holtslag, F. D'Andrea, and M. Ek, 2013: Surface and Atmospheric Controls on the Onset of Moist Convection over Land. *Journal of Hydrometeorology*, **14**, 1443–1462, https://doi.org/10.1175/JHM-D-12-0137.1.

——, A. Massmann, B. R. Lintner, S. Hamed Alemohammad, R. Fu, J. K. Green, D. Kennedy, and J. Vilà-Guerau de Arellano, 2019: Land–atmosphere interactions in the tropics – a review. *Hydrology and Earth System Sciences*, **23**, 4171–4197, https://doi.org/10.5194/hess-23-4171-2019.

Grant, L. D., and S. C. van den Heever, 2014: Aerosol-cloud-land surface interactions within tropical sea breeze convection. *Journal of Geophysical Research: Atmospheres*, **119**, 8340–8361, https://doi.org/10.1002/2014JD021912.

Hansen, M. C., and Coauthors, 2013: High-Resolution Global Maps of 21st-Century Forest Cover Change. *Science*, https://doi.org/10.1126/science.1244693.

Harrington, J. Y., 1997: Effects of radiative and microphysical processes on simulated warm and transition season Arctic stratus.

Heald, C. L., and D. V. Spracklen, 2015: Land Use Change Impacts on Air Quality and Climate. *Chem. Rev.*, **115**, 4476–4496, https://doi.org/10.1021/cr500446g.

Heald, C. L., and Coauthors, 2008: Predicted change in global secondary organic aerosol concentrations in response to future climate, emissions, and land use change. *Journal of Geophysical Research: Atmospheres*, **113**, https://doi.org/10.1029/2007JD009092.

van den Heever, S. C., and W. R. Cotton, 2007: Urban Aerosol Impacts on Downwind Convective Storms. *Journal of Applied Meteorology and Climatology*, **46**, 828–850, https://doi.org/10.1175/JAM2492.1.

Igel, A. L., S. C. van den Heever, and J. S. Johnson, 2018: Meteorological and Land Surface Properties Impacting Sea Breeze Extent and Aerosol Distribution in a Dry Environment: Factors Impacting Sea Breezes. *J. Geophys. Res. Atmos.*, **123**, 22–37, https://doi.org/10.1002/2017JD027339.

Jiang, H., and G. Feingold, 2006: Effect of aerosol on warm convective clouds: Aerosol-cloud-surface flux feedbacks in a new coupled large eddy model. *Journal of Geophysical Research: Atmospheres*, **111**, https://doi.org/10.1029/2005JD006138.

Johnson, J. S., Z. Cui, L. A. Lee, J. P. Gosling, A. M. Blyth, and K. S. Carslaw, 2015: Evaluating uncertainty in convective cloud microphysics using statistical emulation. *Journal of Advances in Modeling Earth Systems*, **7**, 162–187, https://doi.org/10.1002/2014MS000383.

Kummu, M., and O. Varis, 2011: The world by latitudes: A global analysis of human population, development level and environment across the north–south axis over the past half century. *Applied Geography*, **31**, 495–507, https://doi.org/10.1016/j.apgeog.2010.10.009.

Laguë, M. M., G. B. Bonan, and A. L. S. Swann, 2019: Separating the Impact of Individual Land Surface Properties on the Terrestrial Surface Energy Budget in both the Coupled and Uncoupled Land–Atmosphere System. *Journal of Climate*, **32**, 5725–5744, https://doi.org/10.1175/JCLI-D-18-0812.1.

Lambin, E. F., H. J. Geist, and E. Lepers, 2003: Dynamics of Land-Use and Land-Cover Change in Tropical Regions. *Annual Review of Environment and Resources*, **28**, 205–241, https://doi.org/10.1146/annurev.energy.28.050302.105459.

Lawrence, D., and K. Vandecar, 2015: Effects of tropical deforestation on climate and agriculture. *Nature Climate Change*, **5**, 27–36, https://doi.org/10.1038/nclimate2430.

Lazzarini, M., A. Molini, P. R. Marpu, T. B. M. J. Ouarda, and H. Ghedira, 2015: Urban climate modifications in hot desert cities: The role of land cover, local climate, and seasonality. *Geophysical Research Letters*, **42**, 9980–9989, https://doi.org/10.1002/2015GL066534.

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**, 12253–12273, https://doi.org/10.5194/acp-11-12253-2011.

Lee, S. S., L. J. Donner, V. T. J. Phillips, and Y. Ming, 2008: The dependence of aerosol effects on clouds and precipitation on cloud-system organization, shear and stability. *Journal of Geophysical Research: Atmospheres*, **113**, https://doi.org/10.1029/2007JD009224.

——, G. Feingold, A. McComiskey, T. Yamaguchi, I. Koren, J. V. Martins, and H. Yu, 2014: Effect of gradients in biomass burning aerosol on shallow cumulus convective circulations. *Journal of Geophysical Research: Atmospheres*, **119**, 9948–9964, https://doi.org/10.1002/2014JD021819.

Leung, G. R., and S. C. van den Heever, 2022: Controls on the Development and Circulation of Terminal versus Transient Congestus Clouds and Implications for Midlevel Aerosol Transport. *Journal of the Atmospheric Sciences*, **79**, 3083–3101, https://doi.org/10.1175/JAS-D-21-0314.1.

Li, X., W.-K. Tao, H. Masunaga, G. Gu, and X. Zeng, 2013: Aerosol Effects on Cumulus Congestus Population over the Tropical Pacific: A Cloud-Resolving Modeling Study. *Journal of the Meteorological Society of Japan*, **91**, 817–833, https://doi.org/10.2151/jmsj.2013-607.

Mahmood, R., and Coauthors, 2014: Land cover changes and their biogeophysical effects on climate. *International Journal of Climatology*, **34**, 929–953, https://doi.org/10.1002/joc.3736.

Oakley, J. E., and A. O'Hagan, 2004: Probabilistic sensitivity analysis of complex models: a Bayesian approach. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, **66**, 751–769, https://doi.org/10.1111/j.1467-9868.2004.05304.x.

Park, J. M., and S. C. van den Heever, 2022: Weakening of tropical sea breeze convective systems through interactions of aerosol, radiation, and soil moisture. *Atmospheric Chemistry and Physics*, **22**, 10527–10549, https://doi.org/10.5194/acp-22-10527-2022.

Park, J. M., S. C. van den Heever, A. L. Igel, L. D. Grant, J. S. Johnson, S. M. Saleeby, S. D. Miller, and J. S. Reid, 2020: Environmental Controls on Tropical Sea Breeze Convection and Resulting Aerosol Redistribution. *Journal of Geophysical Research: Atmospheres*, **125**, e2019JD031699, https://doi.org/10.1029/2019JD031699.

Pitman, A. J., and Coauthors, 2009: Uncertainties in climate responses to past land cover change: First results from the LUCID intercomparison study. *Geophysical Research Letters*, **36**, https://doi.org/10.1029/2009GL039076.

Pongratz, J., C. Schwingshackl, S. Bultan, W. Obermeier, F. Havermann, and S. Guo, 2021: Land Use Effects on Climate: Current State, Recent Progress, and Emerging Topics. *Curr Clim Change Rep*, https://doi.org/10.1007/s40641-021-00178-y.

Riehl, H., and J. S. Malkus, 1958: On the heat balace in the equatorial trough zone. *Geophysica*, **6**, 503–558.

Saleeby, S. M., and S. C. van den Heever, 2013: Developments in the CSU-RAMS Aerosol Model: Emissions, Nucleation, Regeneration, Deposition, and Radiation. *Journal of Applied Meteorology and Climatology*, **52**, 2601–2622, https://doi.org/10.1175/JAMC-D-12-0312.1.

Sarangi, C., S. N. Tripathi, Y. Qian, S. Kumar, and L. Ruby Leung, 2018: Aerosol and Urban Land Use Effect on Rainfall Around Cities in Indo-Gangetic Basin From Observations and Cloud Resolving Model Simulations. *Journal of Geophysical Research: Atmospheres*, **123**, 3645–3667, https://doi.org/10.1002/2017JD028004.

Schneck, R., and V. Mosbrugger, 2011: Simulated climate effects of Southeast Asian deforestation: Regional processes and teleconnection mechanisms. *Journal of Geophysical Research: Atmospheres*, **116**, https://doi.org/10.1029/2010JD015450.

Schneider, L., C. Barthlott, C. Hoose, and A. I. Barrett, 2019: Relative impact of aerosol, soil moisture, and orography perturbations on deep convection. *Atmospheric Chemistry and Physics*, **19**, 12343–12359, https://doi.org/10.5194/acp-19-12343-2019.

Sena, E. T., P. Artaxo, and A. L. Correia, 2013: Spatial variability of the direct radiative forcing of biomass burning aerosols and the effects of land use change in Amazonia. *Atmos. Chem. Phys.*, **13**, 1261–1275, https://doi.org/10.5194/acp-13-1261-2013.

Seneviratne, S. I., T. Corti, E. L. Davin, M. Hirschi, E. B. Jaeger, I. Lehner, B. Orlowsky, and A. J. Teuling, 2010: Investigating soil moisture–climate interactions in a changing climate: A review. *Earth-Science Reviews*, **99**, 125–161, https://doi.org/10.1016/j.earscirev.2010.02.004.

Sokolowsky, G. A., S. W. Freeman, and S. C. van den Heever, 2022: Sensitivities of Maritime Tropical Trimodal Convection to Aerosols and Boundary Layer Static Stability. *Journal of the Atmospheric Sciences*, **79**, 2549–2570, https://doi.org/10.1175/JAS-D-21-0260.1.

Takahashi, A., T. Kumagai, H. Kanamori, H. Fujinami, T. Hiyama, and M. Hara, 2017: Impact of Tropical Deforestation and Forest Degradation on Precipitation over Borneo Island. *Journal of Hydrometeorology*, **18**, 2907–2922, https://doi.org/10.1175/JHM-D-17-0008.1.

Tao, W.-K., J.-P. Chen, Z. Li, C. Wang, and C. Zhang, 2012: Impact of aerosols on convective clouds and precipitation. *Reviews of Geophysics*, **50**, https://doi.org/10.1029/2011RG000369.

Ten Hoeve, J. E., L. A. Remer, and M. Z. Jacobson, 2011: Microphysical and radiative effects of aerosols on warm clouds during the Amazon biomass burning season as observed by MODIS: impacts of water vapor and land cover. *Atmos. Chem. Phys.*, **11**, 3021–3036, https://doi.org/10.5194/acp-11-3021-2011.

Tosca, M. G., J. T. Randerson, C. S. Zender, D. L. Nelson, D. J. Diner, and J. A. Logan, 2011: Dynamics of fire plumes and smoke clouds associated with peat and deforestation fires in Indonesia. *Journal of Geophysical Research: Atmospheres*, **116**, https://doi.org/10.1029/2010JD015148.

Turubanova, S., P. V. Potapov, A. Tyukavina, and M. C. Hansen, 2018: Ongoing primary forest loss in Brazil, Democratic Republic of the Congo, and Indonesia. *Environ. Res. Lett.*, **13**, 074028, https://doi.org/10.1088/1748-9326/aacd1c.

Twomey, S., 1977: The Influence of Pollution on the Shortwave Albedo of Clouds. *Journal of the Atmospheric Sciences*, **34**, 1149–1152, https://doi.org/10.1175/1520-0469(1977)034<1149:TIOPOT>2.0.CO;2.

Walko, R. L., and Coauthors, 2000: Coupled Atmosphere–Biophysics–Hydrology Models for Environmental Modeling. *Journal of Applied Meteorology and Climatology*, **39**, 931–944, https://doi.org/10.1175/1520-0450(2000)039<0931:CABHMF>2.0.CO;2.

Winckler, J., C. H. Reick, and J. Pongratz, 2017: Robust Identification of Local Biogeophysical Effects of Land-Cover Change in a Global Climate Model. *Journal of Climate*, **30**, 1159–1176, https://doi.org/10.1175/JCLI-D-16-0067.1.

Wu, S., L. J. Mickley, J. O. Kaplan, and D. J. Jacob, 2012: Impacts of changes in land use and land cover on atmospheric chemistry and air quality over the 21st century. *Atmospheric Chemistry and Physics*, **12**, 1597–1609, https://doi.org/10.5194/acp-12-1597-2012.

Yamaguchi, T., G. Feingold, and J. Kazil, 2019: Aerosol-Cloud Interactions in Trade Wind Cumulus Clouds and the Role of Vertical Wind Shear. *Journal of Geophysical Research: Atmospheres*, **124**, 12244–12261, https://doi.org/10.1029/2019JD031073.

Yang, G.-Y., and J. Slingo, 2001: The Diurnal Cycle in the Tropics. *Monthly Weather Review*, **129**, 784–801, https://doi.org/10.1175/1520-0493(2001)129<0784:TDCITT>2.0.CO;2.

Zhang, W., J. Yang, L. Yang, and D. Niyogi, 2022: Impacts of City Shape on Rainfall in Inland and Coastal Environments. *Earth's Future*, **10**, e2022EF002654, https://doi.org/10.1029/2022EF002654.

Zhang, Y., R. Fu, H. Yu, R. E. Dickinson, R. N. Juarez, M. Chin, and H. Wang, 2008: A regional climate model study of how biomass burning aerosol impacts land-atmosphere interactions over the Amazon. *Journal of Geophysical Research: Atmospheres*, **113**, https://doi.org/10.1029/2007JD009449.