3-2022

Applications of a Lightning Proxy to Generate Synthetic Lightning for Use in Physics-Based Image-Chain Models

Bryan G. Castro

APPLICATIONS OF A LIGHTNING PROXY TO GENERATE SYNTHETIC LIGHTNING FOR USE IN PHYSICS-BASED IMAGE-CHAIN MODELS

THESIS

Bryan G. Castro, Captain, USAF
AFIT-ENP-MS-22-M-082

DEPARTMENT OF THE AIR FORCE
AIR UNIVERSITY

Air Force Institute of Technology

Wright-Patterson Air Force Base, Ohio

AFIT-ENP-MS-22-M-082

APPLICATIONS OF A LIGHTNING PROXY TO GENERATE SYNTHETIC LIGHTNING FOR USE IN PHYSICS-BASED IMAGE-CHAIN MODELS

THESIS

Presented to the Faculty

Department of Aeronautics and Astronautics

Graduate School of Engineering and Management

Air Force Institute of Technology

Air University

Air Education and Training Command

In Partial Fulfillment of the Requirements for the

Degree of Master of Science in Atmospheric Science

Bryan G. Castro, BAS

Captain, USAF

March 2022

AFIT-ENP-MS-22-M-082

APPLICATIONS OF A LIGHTNING PROXY TO GENERATE SYNTHETIC LIGHTNING FOR USE IN PHYSICS-BASED IMAGE-CHAIN MODELS

Bryan G. Castro, BAS
Captain, USAF

Committee Membership:

Maj P. A. Saunders, PhD
Chair

A. M. Schinder, PhD
Member

Lt Col K. E. Fitch, PhD
Member

R. C. Tournay, PhD
Member

AFIT-ENP-MS-22-M-082

Abstract

A method of generating synthetic lightning through the use of a convective available potential energy (CAPE) times precipitation rate (P) proxy (CP) is applied over three distinct climatological zones of the world for a single warm season: central and southern Arizona of the United States, central Cuba, and North Korea. Global Forecast System (GFS) 0.25° by 0.25° forecast data for June, July, and August of 2019 is used to provide 6-hourly CAPE and precipitation rate, while Global Lightning Dataset (GLD360) data for the period 2016 to 2020 is used to provide observed lightning strokes. A five-year lightning climatology study is conducted on each region to identify lightning density and energy patterns. A correlation of monthly and seasonal lightning to the proxy is then conducted for each region at a 0.25°grid resolution to obtain correlation coefficients (*r*), p-values, and establish validity of the proxy. Linear regression is then applied to the two datasets to determine an appropriate constant of proportionality (*k*), enabling the proxy to be used to explicitly forecast lightning for each region. An additional correction factor (*cf)* is also calculated and applied to the proxy to correctly scale the number of strokes per area. The proxy is then used to forecast lightning at monthly and seasonal timescales and compared to observed lightning stroke densities at the same location and time to assess performance. The lightning climatology study results show that long-term variations in lightning patterns for each region are most influenced by warm-season convection, topography, and local effects. Correlation results show best *r* values at 0.799 (Arizona), 0.711 (Cuba), and 0.562 (North Korea), and poorest *r* values at 0.462 (Arizona), 0.260 (Cuba), and 0.005 (North Korea), indicating that while moderate correlations exist at

times over the monthly and seasonal timescales, a correlation over a longer time period encompassing several warm seasons would likely result in more accurate *r* values and constants of proportionality. Forecast results show that the CP proxy provides a fairly accurate depiction of lightning timing and location at monthly and seasonal timescales over Arizona, but performs poorly over much of Cuba and North Korea.

Thank you God for providing me so many blessings and opportunities.

I view science as simply Man's way of discovering what you have made.

For Racquel, Caroline, Melanie, and Brandon.

I love you all more than you can know.

Acknowledgements

I would like to thank my research advisor, Maj Peter Saunders, for his support, expertise, and encouragement. I would also like to thank the other members of my committee, Lt Col Kyle Fitch, Dr. Aaron Schinder, and Dr. Bob Tournay, for their assistance and feedback during the thesis research and writing stages. I have learned much from all of you during my time at AFIT. I would also like to thank the weather program faculty; I very much enjoyed the weather courses and your respective teaching styles/methods in each class.

Bryan G. Castro

Table of Contents

Page

List of Figures

Page

List of Tables

Page

List of Abbreviations

AFIT Air Force Institute of Technology

ASSET AFIT Sensor and Scene Emulation Tool

AZ Arizona

CAPE Convective Available Potential Energy

CCD Charged-Coupled Device

CG Cloud-to-Ground

CLDN Canadian Lightning Detection Network

CMIP5 Coupled Model Intercomparison Project Phase 5

CONUS Contiguous United States

CP CAPE × Precipitation Rate

DMZ Demilitarized Zone

ECMWF European Centre for Medium-Range Weather Forecasts

ERA ECMWF Reanalysis

FAR False Alarm Rate

GCM Global Climate Model

GEFS Global Ensemble Forecast System

GFS Global Forecast System

GLD360 Global Lightning Dataset

GLM Geostationary Lightning Mapper

GOES Geostationary Operational Environmental Satellite

GZ Gardiner/Ziegler

Heidke Skill Score (HSS)

hr Hour

Hz Hertz
ISS International Space Station

J Joule

JJA June, July, August

kA Kiloampere

keV Kiloelectron Volt

kg Kilogram

km Kilometer

LCFA Lightning Cluster Filter Algorithm

LEO Low Earth Orbiting

LFC Level of Free Convection

LIS Lightning Imaging Sensor

LMA Lightning Mapper Array

m meter

MPa Megapascal

mm Millimeter

ms Millisecond

MSE Mean Squared Error

NALDN North American Lightning Detection Network

NARR North American Regional Reanalysis

NASIC National Air and Space Intelligence Center

NLDN National Lightning Detection Network

NOAA National Oceanic and Atmospheric Administration

nm Nanometer

P Precipitation Rate

P-Value Probability Value

POD Probability of Detection

r Correlation Coefficient

RTEP Real-Time Event Processor

RR Riming Rate

S91 Saunders 1991

SAM Straka Atmospheric Model

SLM Stochastic Lightning Model

SP98 Saunders and Peck 1998

SPARC Stratosphere-troposphere Processes and their Role in Climate

TAK Takahashi

TRMM Tropical Rainfall Measuring Mission

TS Threat Score

VHF Very High Frequency

VLF Very Low Frequency

WRF ARW Weather Research and Forecasting Advanced Research

APPLICATIONS OF A LIGHTNING PROXY TO GENERATE SYNTHETIC LIGHTNING FOR USE IN PHYSICS-BASED IMAGE-CHAIN MODELS

Chapter 1: Introduction

1.1 Lightning Significance

On Safety

Thunderstorms are a common phenomenon worldwide. Both beautiful and dangerous, they are responsible for some of the most spectacular naturally occurring visual displays human beings can witness. They are also responsible for an average of 49 deaths each year in the United States ("Lightning Victims," 2021). Lightning is also of particular interest to the aviation community, both military and civilian.

Most commercial-sized aircraft are struck by lightning an average of once per year (WMO, 2014:657). While most modern commercial-sized aircraft are equipped to survive a lightning strike and protect its passengers, the threat of lightning on any aircraft is still a significant concern due to the damage it can cause to electrical components and the ensuing threat to humans mid-flight (Rash, 2010:23). Moreover, the bigger threat to aircraft is not usually the lightning itself, but the additional threats associated with the thunderstorms in which they occur: hail, strong winds, and wind shear.

Lightning also starts many forest fires and is responsible for over 30 % of all electric power line failures (WMO, 2014:657). Additionally, forecasting lightning with sufficient lead time (advance warning) is a vital and necessary skill required of all base weather stations within military installations (Department, 2021:58-59). Issuance of lightning watches and warnings ensure that sufficient actions can be taken to ensure the

safety of maintenance personnel working outside, as well as the base populace (Department, 2021:58-59). For these reasons and many more, accurate forecasts of lightning have been a chief desire of the general public, governments around the world, and their associated militaries.

AFIT Sensor and Scene Emulation Tool

The AFIT Sensor and Scene Emulation Tool (ASSET) is a physics-based image-chain model used to generate synthetic data sets for wide field of view electro-optical and infrared sensors (Young et al., 2017:2-3). It was originally developed for use at the Air Force Institute of Technology (AFIT) to model sensor response to irradiance and to generate data sets for algorithm development and testing (Young et al., 2017:3). Datasets include object detection and tracking, pattern of life trending, infrared seeker performance prediction, and multi-satellite viewing (Steward, 2020:2). Algorithm development and testing includes clutter and background suppression, image/signal processing, and sensor trade studies (Steward, 2020:14).

ASSET has operational uses as well, such as for target injection at the National Air and Space Intelligence Center (NASIC) (Steward, 2020:2, Young et al., 2017:3). ASSET works by taking a high-resolution image as the basis for a background scene, emulates remote sensing processes, and outputs an array of data frames (Young et al., 2017:3). It supports instances where absolute knowledge of an object's position and radiometric signature are needed, but the absolute radiometric representation and radiometric properties of a specific scene are not necessary (Young et al., 2017:3). ASSET works as a sensor emulator, not a simulator, and its value comes in its ability to reproduce the imperfections commonly associated with real sensors but not often

observed in simulated data (Young et al., 2017:3). ASSET currently has limited capability to reproduce lightning as viewed from wide field of view electro-optical and infrared sensors, under meteorologically sound conditions, spatially and temporally. This project will focus on creating a lightning model suitable for use within ASSET that accurately reproduces lightning under appropriate meteorological conditions, maintains low computational cost, and is user-friendly.

1.2 Problem Statement

The primary focus of this study is to investigate the feasibility of creating a global lightning model that accurately predicts high resolution cloud-to-ground (CG) and in-cloud (IC) lightning at seasonal and subseasonal timescales for use in ASSET.

1.3 Research Objectives

To create a global lightning model suitable for use within ASSET, this study will begin with a comprehensive literature review regarding thunderstorm electrification theories, methods of observing lightning, and modern approaches to lightning modeling currently in use in numerical weather prediction (NWP) models. Then, an observed lightning dataset will be chosen in order to conduct a small-scale lightning climatology study on three distinct regions in the world. Next, a lightning modeling method will be chosen based on successes from existing literature, and validated against the observed lightning dataset. Then, the lightning modeling method will we implemented to forecast lightning over three climatologically distinct regions. Overall, the goal of the study is to

produce a global lightning model that predicts lightning under meteorological conditions that are spatially and temporally sound for use within ASSET.

Chapter 2: Literature Review

2.1 Mechanics of Lightning Production

Thunderstorm Electrification

A lightning flash or discharge within a storm is the result of an electric field that has reached too high of a magnitude and neutralizes itself by means of a large electric current flowing through a channel between two distinct regions of different charges (MacGorman et al., 2001:460). The effect of a lightning discharge is to reduce the magnitude of the electric field, and to redistribute charge throughout the storm (MacGorman et al., 2001:460-461). The charge separation within a thunderstorm is initiated and maintained primarily by noninductive processes, with inductive processes playing a secondary role (Lopez, 2016:3057).

Noninductive charging refers to a charge transfer occurring through the collisions between rimed graupel/hail and unrimed snow/cloud-ice particles without the presence of a preexisting electric field (Lopez, 2016:3057). This process begins in the upper part of a storm, typically at a height at or above -20°C isotherm (National, 1986:123; Weidman, 2013:9). Large vertical motions bring saturated air to this level, resulting in graupel pellet formation, ice crystals, and super cooled water droplets. At some point in this upper region of the storm, the graupel will start to fall and collide with both super cooled water droplets and ice crystals (Reynolds et al., 1957:432; Weidman, 2013:9). The graupel will grow in size due to the accretion of the super cooled water droplets on its surface, but

collisions with the ice crystals results in the gain of positive or negative charge for either the graupel or the ice crystal (Reynolds et al., 1957:432; Weidman, 2013:9). Whichever sign of charge the graupel gains, the ice crystal gains the opposite charge (Reynolds et al., 1957:432; Weidman, 2013:9). The determination of whether the graupel or ice crystal gains positive or negative charge depends on the environmental temperature during the collision (Reynolds et al., 1957:430-433, Saunders et al., 1991:11012-11014, Takayashi, 1978:1541-1542). This dependence on temperature for charging results in typically three distinct charge regions of the thunderstorm: net positive charges above the -25°C isotherm, net negative charges between the -10°C and -25°C isotherms, net positive between the 0°C and -10°C isotherms, and net positive between the 0°C isotherm and the ground (Lopez, 2016:3058).

Inductive charging is a more minor charging process that is thought to occur after a thunderstorm has already achieved a separation of charges through noninductive charging, although both processes can occur simultaneously (Mansell et al., 2005:9; Lopez, 2016:3057). Inductive charging is a process by which graupel becomes polarized as falls through a preexisting uniform electric field, resulting in its upper half containing negative charges and its lower half containing positive charges (Brooks and Saunders, 1994:10627-10628; Weidman, 2013:5). As the graupel falls downward, it collides with smaller cloud particles of either supercooled liquid or ice due to differences in fall velocities (Brooks and Saunders, 1994: 10627-10628; Weidman, 2013:5-6). Collisional charge transfer only occurs when the particles do not stick together (Lamb and Verlinde, 2011:529-547). Thus, collisions from smaller ice particles are much more likely to result in a charge transfers as they are more likely to bounce off of the graupel, while liquid

particles tend to stick and freeze onto the graupel at freezing temperatures (Lamb and Verlinde, 2011:529-547). It is generally accepted that the noninductive charging mechanism is responsible for initial thunderstorm electrification and maintenance, while inductive charging plays a role in charge maintenance and distribution after the storm has been electrified (Lopez, 2016:3057). Once the magnitude of the electric field between these region reaches a certain limit, either an IC or CG discharge will occur to reduce its magnitude and neutralize the charges (MacGorman, 2001:460-462). Overall, lightning discharges involve a variety of microphysical processes.

Lightning Discharge

For a lightning discharge to occur, a conducting channel is created composed of ionized gas whose temperature reaches close to 30,000 K, or about five times the temperature of the surface of the sun (MacGorman et al., 2001:460, WMO, 2014:657). A lightning discharge (also called a flash) consists of individual elements called strokes, and contains anywhere from one to 26 strokes, although typically it contains three to five strokes (Wallace and Hobbs, 2006:255, WMO, 2014:657). The time interval between strokes is typically on the order of tens of milliseconds, and each stroke may branch off into numerous ground strike points (Wallace and Hobbs, 2006:255, WMO, 2014:658).

In general, each stroke is made up of a downward process called a leader, and an upward process called a return stroke (WMO, 2014:658). The leader can be further divided into stepped leaders and dart leaders categories. The first stroke leader has an intermittent appearance and is termed a stepped leader, while the tips of subsequent stroke leaders appear to move continuously and thus are called dart leaders. In a typical CG discharge, the leader is responsible for creating the conducting channel between the

charge region of the cloud and the ground, as well as distributes charges along the path (WMO, 2014:658). After the leader propagates downward, a return stroke propagates from the ground and back up to the charge region of the cloud along the same channel, and neutralizes the charge of the leader, resulting in an overall transfer of negative (or positive) charge to the ground (WMO, 2014:658).

Interestingly, the leader creates X-ray radiation with energies up to 250 keV, which is twice the energy of a chest X-ray (WMO, 2014:658). Furthermore, IC lightning discharges have been found to emit radiation at energies in the gamma range, as detected by sferic (atmospheric electric field change) waveform array stations (Stanley et al., 2006:1). The return stroke emits a high current of 15 kA to 30 kA which quickly heats the conducting channel to above 30,000 K (WMO, 2014:658). The rapid increase in temperature results in a rapid increase of channel pressure to over 1 MPa, resulting in channel explosion and a shockwave that humans perceive as thunder (WMO, 2014:659).

About 90 % of midlatitude CG lightning carries negative charge to the ground, while 10 % carries positive charge to the ground (Wallace and Hobbs, 2006:256). Positive CG lightning is much more dangerous than negative CG lightning, as it possesses a magnitude of charges of up to 10 times greater than negative CG lightning, resulting in greater damage to buildings, forests, and living things ("Severe," 2021, Wallace and Hobbs, 2006:256). Positive CG lightning also tends to originate from the edge of the cloud, resulting in lightning strikes that occur several miles away from the perceived edge of a storm ("Severe," 2021). For IC and cloud-to-cloud lightning, the discharge typically consists of a single leader that connects the negative and positive

regions of a cloud, resulting in a lower, but more continuous illumination of the cloud (Wallace and Hobbs, 2006:255).

2.2 Lightning Observation Methods

Lightning is a relatively small-scale feature, and defining its appearance through observation methods has been approached in several ways. Lightning emits substantial electromagnetic radiation in frequencies ranging from 1 Hz to near 300 MHz, as well as high electric current in tens of kA (WMO, 2014:657). It also produces optical energy ranging from 10^{14} to 10^{15} Hz (WMO, 2014:657). Over the past three decades, two distinct methods of defining observed lightning have emerged: as an area of optical energy viewed from a satellite (Goodman et al., 2013:35-37), or as a point of electrical current either IC or CG ("Vaisala," 2021).

National Lightning Detection Network

In the United States, the National Lightning Detection Network (NLDN) has operated continuously from 1989 until present day, and remains one of the longest running sources of lighting data to date ("Vaisala," 2021). The NLDN reports CG and IC lightning as point data in kA through the use of ground-based sensors that detect the electromagnetic signals in the ultra low frequency (ULF) to medium frequency (MF) range of 400 Hz to 400 kHz produced by lightning strikes (WMO, 2014:670; "Vaisala," 2021). For each lightning flash occurrence, multiple neighboring sensors use magnetic-direction-finding and time-of-arrival methods to observe the occurrence and determine the location, type, and energy associated with each lightning flash ("Vaisala," 2021). The

detection range of the NLDN includes all of the United States, out to approximately several hundred kilometers off-shore (Nag et al., 2014:2, Said et al., 2013:6906).

Continual upgrades to the NLDN in the past few decades have improved its detection efficiency and location accuracy, with a 2013 evaluation showing a median location accuracy of 0.15 to 0.25 km and >95% flash detection efficiency in the interior of the United States. Furthermore, a 0.25 to 0.5 km median location accuracy and 50% flash detection efficiency is noted at the outer edges of the network (Nag et al., 2014:1). The NLDN is part of the larger North American Lightning Detection Network (NALDN), which also consists of the Canadian Lightning Detection Network (CLDN) (Nag et al., 2014:1). Further NLDN upgrades from 2014 to 2018 included central processor updates to improve the classification of IC and CG strokes, as well as improve the detection efficiency (Murphy et al., 2021:574). The NLDN uses a flash algorithm to group a maximum of 15 strokes into a single flash, with specific spatial and temporal rules as to how strokes are integrated into flashes (Nag et al., 2014:4, Murphy et al., 2021:575). The deliverable product to the end user is either the post-processed flash or strokes.

Global Lightning Data

The Global Lightning Dataset (GLD360) is a very low frequency (VLF), ground-based, global lightning detection network that operates through the use of numerous VLF sensors placed around the world (WMO, 2014:673). It is operated by Vaisala, the same company that maintains the NLDN, and uses similar time-of-arrival and magnetic-detection-finding methods to detect CG and IC lightning, albeit with a lower location accuracy and detection efficiency due to the long-range Earth-ionosphere waveguide design of the sensors (Said and Murphy, 2016:1). It is able to capture lightning location,

type, and energy associated with each flash, operates in the 500 Hz to 50 kHz range, possesses a median location accuracy of 2.4 km to 1.8 km, a CG detection efficiency of 75% to 85%, and an IC detection efficiency of 30% to 50% (Said and Murphy, 2016:7). The deliverable product to the end user is usually the stroke data (Cramer, 2021). While the location accuracy and detection efficiency of the GLD360 is inferior to that of the NLDN, it has the advantage of being a truly global network that is not limited to continent use.

Geostationary Lightning Mapper

The Geostationary Lightning Mapper (GLM) is a lightning sensor onboard the current GOES-16 and GOES-17 and future GOES-T and GOES-U satellites that provides continuous, hemispheric monitoring of lightning (Goodman, 2012:10; "GOES-R," 2022). It is capable of detecting more than 70% of lightning flashes on a 24-hour basis, with a daytime detection efficiency of 70% and a nighttime detection efficiency of 90% (Bruning et al., 2019:14285; Rudlosky et al., 2019:1097). The sensor is a Charge-Coupled Device (CCD) imager with a set of optical filters that is able to detect lightning in a 1 nanometer (nm) spectral band centered on the 777.4 nm wavelength (Rudlosky et al., 2019:1098). The instrument consists of a single telescope with a variable pitch focal plane detector array of 1372 by 1300 pixels, which is further divided into 56 subarrays that are processed by 56 Real-Time Event Processors (RTEPs) (Rudlosky et al., 2019:1098). This allows the GLM to have a pixel resolution of 8 km at nadir, to 14 km at limb (Rudlosky et al., 2019:1097). In essence, the GLM is a high-resolution camera in a geostationary orbit around the Earth.

The GLM has a complex algorithm known as the lightning cluster filter algorithm (LCFA) that processes lightning data and clusters them into either events, groups, or flashes (Goodman et al., 2013:39). Events are defined as the triggered pixels whose brightness exceeds the background threshold during a single frame (Goodman et al., 2013:38). Events are the basic unit of data from the GLM and generally reflect a single optical pulse occurring within a 2 millisecond (ms) integration window. When multiple, spatially adjacent events occur within the 2 ms integration window, the result is a group (Goodman et al., 2013:39). When multiple groups occur within a 330 ms integration window and within 16.5 km of each other, the result is a flash (Goodman et al., 2013:39). The flash is the end goal of the algorithm and is meant to capture the light produced from a lightning channel discharge tree (Goodman et al., 2013:39).

Although the GLM has the advantage of providing spaceborne, optical characteristics of lightning, its main limitation is that its coverage is not global. Prior to the launch of the GLM aboard the GOES-R series satellites, the only source of spaceborne, lightning imagery was from low-earth-orbiting (LEO) satellites, such as from the Lightning Imaging Sensor (LIS) on the Tropical Rainfall Measuring Mission (TRMM) or the International Space Station (ISS) (WMO, 2014:893; "LIS," 2022). Lightning data from these LEO satellites was also limited in that it only captured the portion of the earth which it was traveling over, which made it sufficient for research but not ideal for operational forecast use ("LIS," 2022). So, while the GLM does not provide global coverage, it represents a vast improvement over its predecessors in that it provides continuous coverage of lightning over one portion of the earth, making it ideal for both research and operational uses.

2.3 Forecasting Lightning using Weather Models

Model Requirements

Lightning is a small-scale feature that can be defined as either a point or an area of strikes/flashes. Forecasting it through NWP models has largely been limited by resolution and an adequate representation of the microphysical processes involved. Lightning occurs with thunderstorms, which may be described as systems of cumulonimbus clouds with updrafts on the order of 1 km. Since cumulonimbus clouds themselves are relatively small-scale features with a horizontal diameter on the order of 5 to 10 km, methods of forecasting lightning have been focused on using high-resolution, cloud-resolving models (Bryan et al., 2003:2394; Fierro et al. 2013:2390). Generally speaking, cloud-resolving models possess a horizontal resolution of less than 5 km (Fierro et al., 2013:2390). It is widely accepted that a 1 km grid resolution is sufficient to resolve a thunderstorm's structure, but even finer resolutions on the order of 100 m are necessary to resolve deep moist convection (Bryan et al., 2003:2394, 2413). To forecast lightning, a model has to have a resolution high enough to explicitly resolve cumulus clouds and sufficient microphysics to model the electrification process. As such, most attempts at forecasting lightning in numerical prediction models have employed lightning parameterizations (Mansell et al., 2002:1).

Forecasting within Cloud-Resolving Models

Over the last three decades, two primary methods of forecasting lightning within cloud-resolving models have emerged: using electrification physics or microphysical proxy variables (Fierro et al., 2013:2390). For the explicit prediction of lightning through electrification physics, either a bulk flash model or those that can explicitly resolve

individual lightning channels ("branched lightning") are used (Fierro et al., 2013:2390-2391). A bulk flash model is one that attempts to represent the large-scale effects of lighting within a cloud, such as the charge distribution and the magnitude of the electric field at which a discharge will occur (Ziegler and MacGorman, 1994:837-838). For a bulk flash model, a horizontal resolution of less than 5 km is required to resolve individual cumulus clouds, although 3 km has found to work well (Fierro et al., 2013:2394). A branched lightning parameterization is one that attempts to simulate the actual channel of a lightning discharge, involving the stepped leader and dart leader propagations seen within strokes that result in that "branched" appearance commonly seen in still photos of lightning (Mansell et al., 2002:1-4, WMO, 2014:658). Although a stepped leader typically advances in lengths on the order of 50 m, a model resolution of 500 m has found to be sufficient to represent lightning channels while remaining computationally practical (Mansell et al., 2002:2-4).

A modern example of a bulk flash model is the advanced research version of the Weather Research and Forecasting (WRF ARW) model (Fierro et al., 2013:2391). This model parameterizes lightning by determining charge distributions resulting from noninductive and inductive processes between hydrometeor species, solving for the electric field, and then discharging lightning by determining locations where the electric field E_{mag} exceeds a prescribed threshold E_{crit} (Fierro et al., 2013:2392-2393). Discharging lightning involves identifying lightning initiation points at all grid cells where E_{mag} exceeds E_{crit}, centering the discharge around these initiation points, and forming a cylinder that extends vertically throughout the simulation domain (Fierro et al., 2013:2393). This discharge parameterization is intended to indicate the direction of

lighting trends, and not to accurately predict flash rates in a storm (Fierro et al., 2013:2394).

Another example of a bulk flash model is the non-hydrostatic Straka Atmospheric Model (SAM), which uses a microphysics package with a single-moment bulk scheme containing two liquid hydrometeor categories and ten ice categories (Mansell et al., 2005:3-5). In a 2005 study, Mansell and others used the SAM to investigate electrifications of continental multi-cell thunderstorms by testing the following noninductive electrification physics schemes within it: Takahashi 1978 scheme (TAK), Gardiner/Ziegler scheme (GZ), Saunders 1991 scheme (S91), Saunders and Peck 1998 (SP98) scheme, and Riming Rate (RR) scheme (Mansell et al., 2005:5-8). Overall, their results reinforced the theory that noninductive charge separation is the most important process that results in thunderstorm electrification, while inductive charging only plays a secondary role (Mansell et al., 2005:20). Furthermore, their work highlighted the vast uncertainty of processes that these parameterizations do not address, but may be vitally important to thunderstorm electrification: the best way of representing a mixed phased cloud, the effectiveness of noninductive charge separation in the -30° to -40°C temperature range, the dependence of charge separation on droplet size spectrum, or the effect of graupel particle temperature on charging (Mansell et al., 2005:21).

A modern example of a branched lightning parameterization is the dielectric breakdown model used by a study by Mansell et al. (2002), to simulate three-dimensional branched lightning in a numerical thunderstorm model (Mansell et al., 2002:2). This model was derived from the dielectric breakdown model first developed by Niemeyer et al. (1984) and Wiesmann & Zeller (1986), which involved simulating the macroscopic

properties of discharge channels (Nieymeyer et al., 1984:1034; Weismann and Zeller, 1984:1771). The model works by using a probability function to determine a discharge path between grid points within a cube (Mansell et al., 2002:3). The probability function finds the probability of a particular discharge path occurring by determining whether an electric field magnitude E_i meets or exceeds a critical electric field value E_{crit} between two grid points (Mansell et al., 2002:3). As discharge paths are found between grid points, additional discharge paths or "branches" are calculated iteratively based on electric potential between the new grid point and the previous (Mansell et al., 2002:3). This model is then applied as a stochastic lightning model (SLM) by simulating bidirectional discharges in regions of net charge density, such as those found in an electrified storm (Mansell et al., 2002:4). This model can successfully produce a variety of lightning structures as seen by lightning observing systems and high-speed photography, including CG and IC discharges with correct polarity, and branched leaders (Mansell et al., 2002:10). However, the biggest limitation with this model is that it is too computationally expensive to use in operational (NWP) models (Fierro et al., 2013:2391).

All parameterizations of thunderstorm electrification involve compromises that fail to reproduce some features of lightning (MacGorman et al., 2001:461). Whether they are bulk or branched lightning parameterizations, modern thunderstorm electrification physics schemes, while effective at simulating electrification, still do not account for an array of processes that may have a large effect on thunderstorm electrification. For this reason, proxy methods of forecasting lightning offer an indirect, yet effective way of predicting lightning that bypass some of the uncertainty involved in sufficiently representing lightning in an electrification scheme.

2.4 Proxy Methods of Forecasting Lightning

A proxy method of forecasting lightning selects variables that have a strong correlation with lightning for use in predicting future occurrences of lightning. Proxy methods for forecasting lightning have seen an increase in popularity in the last few decades due to their low computational cost and wide range of availability (Fierro et al., 2013:2391; Tippett et al., 2019:3933). The reason for their low computational cost is that they do not require any representation of electrification physics within the model (Fierro et al., 2013:2391). Furthermore, they also make possible the prediction of lightning in models that are too coarse to resolve clouds, convection, or that do not possess an electrification physics scheme.

Previous Proxy Methods and Correlations

During the 1990s and 2000s, the most popular lightning proxy method used in NWP models involved taking the fifth power of cloud top height and multiplying it by a constant in order to estimate the total flash rate (Rind and Price, 1994:10824; Moliné and Pontikis, 1995:1085-1087; Michalon et al., 1999:3097-3098; Yoshida et al., 2009:1-3). Proxy methods require that a strong correlation exists between thunderstorm variables and the occurrence of lightning. Numerous lightning correlation studies have previously been done utilizing precipitation rate, and convective available potential energy (CAPE) with varying levels of success. For example, a lightning study done by Soriano et al. (2001) over the Iberian Peninsula showed a 0.75 correlation coefficient (r) between monthly CG lightning flash density and monthly convective precipitation totals for semiarid zones, and a 0.65 correlation coefficient for these variables for humid zones during a warm season (Soriano et al., 2001:3001).

A 2014 study by Siingh et al. showed that monthly lightning flash density from the TRMM, CAPE, and convective precipitation rate all had varying degrees of correlation over India that depended on geographic location (Siingh et al., 2014:40). For southern India, lightning and CAPE had an $r = 0.81$, but only $r = 0.23$ and 0.24 in the mountainous northeast and northwest regions, respectively (Siingh et al., 2014:40). Meanwhile, lightning flash density and convective precipitation rate had an $r = 0.53$ in southern India, and slightly lower $r = 0.48$ and 0.49 in the mountainous northeast and northwest regions, respectively (Siingh et al., 2014:40).

In a 2009 study, McCaul et al. demonstrated two proxy methods within the WRF model, based only on ice-phased hydrometeors, to forecast lightning in northern Alabama (McCaul et al., 2009:709-729). The proxies were compared against observed lightning flashes derived from the north Alabama Lightning Mapper Array (LMA), a ground-based network that detects very high frequency (VHF) pulses from lightning channel segments (McCaul et al., 2009:712). The first method utilized upward fluxes of graupel in the mixed-phase region near the -15°C isotherm level of a storm, and it captured a majority of the temporal variability of lightning (McCaul et al., 2009:709). The second method utilized vertically integrated amounts of graupel, snow, and cloud ice and it predicted the areal coverage of lightning fairly accurately (McCaul et al., 2009:709-729). In 2015, Galanaki and others performed a ten-year study of CG lightning activity over the Eastern Mediterranean and found correlation coefficients between CAPE and CG lightning ranging from $r = 0.70$ to 0.85 (Galanaki et al., 2015:219).

Overall, these studies showed that thunderstorm variables such as upward graupel fluxes, vertically integrated graupel, snow, cloud ice, CAPE, and precipitation rates

exhibit strong correlations with lightning from both ground-based and space-based lightning networks and are well-suited as proxy variables for predicting lightning. However, using these variables by themselves as proxies pose some limitations, as they are not always guaranteed to occur with lightning. For example, not all precipitating clouds produce lightning, such as with stratiform-type clouds in stable environments. And vice versa, not all lightning is produced in precipitating clouds, such as with virga seen in high-based clouds in deserts. Furthermore, not all areas of increasingly high CAPE result in lightning discharges. For example, CAPE can be high in a region, but if the environmental lapse rate is not high enough to permit an air parcel to rise beyond the level of free convection (LFC), the thunderstorm will never have a chance to start. In a way, CAPE alone indicates the strength of thunderstorm, but only if the thunderstorm is able to form in the first place. For these reasons, examining CAPE and precipitation rate together as a proxy presents a novel way of estimating lightning flash rates that overcomes some of the individual limitations of CAPE and precipitation rate.

CAPE Times Precipitation Proxy Study by Romps et al., 2014

A study conducted by David M. Romps and others in 2014 proposed a new proxy theorizing that lightning flash rate was proportional to the product of CAPE and precipitation rate (P) (Romps et al., 2014:851). This study was done in order to find an improved proxy method of representing lightning in global climate models (GCMs) from the Coupled Model Intercomparison Project Phase 5 (CMIP5), as GCMs were unable to explicitly predict lightning (Magi, 2015:434). CAPE data was calculated from Stratosphere-troposphere Processes and their Role in Climate (SPARC) radiosonde observations, while P was taken from the National Oceanic and Atmospheric

Administration River Forecast Centers averaged on a 0.5°grid (Romps et al., 2014:852). Lightning data included CG lightning flashes taken from the NLDN binned along a 0.5°grid (Romps et al., 2014:852). Their results showed that the product of mean CAPE and P during 2011 over CONUS yielded a coefficient of determination (R^2) of 0.77 with mean observed CG lightning flash rates from the NLDN for the same time period (Romps et al., 2014:853). This was an impressive finding and showed that this proxy explained 77% of the variance in lightning flash rates over CONUS at an annual scale (Romps et al., 2014:853). The exact form of the equation for the proxy was given by:

$$F = \frac{\eta}{E} \times P \times CAPE \qquad (1)$$

Where F is the lightning flash rate ($m^{-2}\ s^{-1}$), η is a dimensionless conversion efficiency representing the ratio of power per area dissipated by lightning to the CAPE per area per time available to condensates, and E is the energy discharge per flash in joules (Romps et al., 2014:852). P is the precipitation rate ($kg\ m^{-2}\ s^{-1}$), and $CAPE$ is the convective available potential energy ($J\ kg^{-1}$) (Romps et al., 2014:852).

The rationale behind this proxy was that the product of P and CAPE would give the maximum rate at which kinetic energy was imparted to ascending water condensates (Romps et al., 2014:852). However, a constant of proportionality, $\frac{\eta}{E}$, was introduced in order to facilitate unit conversions and allow more direct comparisons between the proxy and observed lightning. In the 2014 study, the proxy was found to be $1.3 \times 10^{-11} J^{-1}$ by using the fact that 1 mm of precipitation was equivalent to $1\ kg\ m^{-2}$ of liquid water, but also varied according to the temporal range of the forecast period (Romps et al., 2014:853, Tippett and Koshak, 2018:10721).

CAPE Times Precipitation Proxy Study by Romps et al., 2018

A follow up study by Romps and others in 2018 tested this proxy with greater detail by expanding its use to include CONUS seasonal time periods, diurnal cycles, and global spatial scales (Romps et al., 2018:12623). CAPE and P fields were forecasted using several reanalysis datasets with resolutions ranging from 0.25° to 3°(Romps et al., 2018:12624). Observed lightning datasets included a combination of ground-based and space-based sensing systems, with resolutions ranging from 0.001° to 0.5°(Romps et al., 2018:12624).

One key difference with this study was that R^2 was found to be greater when the mean was taken after CAPE and P were multiplied (Romps et al, 2018:12626). This differed from the 2014 study, in that the 2014 study computed the means of CAPE and P first, and then the two quantities were multiplied (Romps et al., 2014:853). Both methods were tested over CONUS at annual time scales, and it was found that $\overline{CAPE \times P}$ resulted in an $R^2 = 0.83$, while $\overline{CAPE} \times \bar{P}$ resulted in an $R^2 = 0.65$ (Romps et al., 2018:12626). The reason for this difference was that $\overline{CAPE \times P}$ maintained the temporal covariance between CAPE and P, resulting in a more accurate spatial distribution of lightning (Romps et al., 2018:12626).

Another key difference was that the CAPE, P, and lightning datasets covered a period of nine to 10 years each, resulting in a smoothing out of interannual and interseasonal outliers, but performed well overall at capturing major diurnal trends, seasonal variations, and spatial distributions of lightning (Romps et al., 2018:1624-1628). Additionally, while the CAPE × P proxy performed well over land, it performed poorly over open ocean areas (Romps et al., 2018:12629). This finding indicated that CAPE and

P alone cannot explain the variance of lightning at all locations, and some other factor must have been responsible (Romps et al., 2018:12629). Romps et al. proposed several theories as to why this variance might exist: higher aerosol concentration over land leading to stronger convection and higher flash rates, lower relative humidity over land resulting in cloud structures that are less prone to convective entrainment, and greater differences in topography and land surface types resulting in a greater irregularity in buoyancy profiles (Romps et al., 2018:12629).

Lightning Proxies over Bangladesh, Dewan et al., 2017

A study by Dewan and others was conducted over Bangladesh to assess the performance of CAPE, precipitation rate (P), and CAPE × P as proxies for lightning (Dewan et al., 2017:1649). Observed lightning data came from the TRMM LIS, covering cloud-to-cloud and CG lightning flashes over a period of 17 years at a resolution of 4 km to 7 km (Dewan et al., 2017:1651). CAPE data was obtained from the European Centre for Medium-Range Weather Forecasts (ECMWF) Re-Analysis-Interim database at 0.5° by 0.5° resolution, while precipitation data in the form of total rain rate ($mm\ hr^{-1}$) was obtained from the TRMM's 3A12 dataset at a resolution of 0.5° by 0.5°; both the CAPE and precipitation data covered the same 17 year period as the lighting data (Dewan et al., 2017:1652).

To assess the performance of CAPE, P, and $CAPE \times P$ as lightning proxies, *r* was calculated at the 95% significance level (Dewan et al., 2017:1652). CAPE was averaged at a monthly timescale, while P was averaged at monthly, seasonal, and annual time timescales (Dewan et al., 2017:1652). The performance of all three variables as lightning proxies was strong, with CAPE performing best at monthly and annual, $CAPE \times P$

performing best at pre-monsoon and monsoon, and P performing similarly to $CAPE \times P$ at pre-monsoon, but poorer in all other cases (Dewan et al., 2017:1653). CAPE showed r = 0.902, 0.703, 0.550, and 0.702 at monthly, pre-monsoon, monsoon, and annual timescales, respectively (Dewan et al., 2017:1653). $CAPE \times P$ showed r = 0.895, 0.818, 0.686, and 0.375 at monthly, pre-monsoon, monsoon, and annual timescales, respectively (Dewan et al., 2017:1653). Finally, P showed r = 0.734, 0.701, 0.455 at monthly, pre-monsoon, and monsoon timescales, respectively (Dewan et al., 2017:1653).

This study showed that CAPE exhibited the strongest correlation with lightning at monthly and annual timescales over Bangladesh, although there seemed to be a lag period between the maximum CAPE and the maximum lightning activity (Dewan et al., 2017:1653). For example, the maximum mean monthly CAPE occurred in April, while the maximum mean flashes occurred in May (Dewan et al., 2017:1653). This association between CAPE and lightning made sense, as CAPE is a conditional instability parameter of the atmosphere whose increase generally indicates the formation of a favorable convective environment (Dewan et al., 2017:1653). However, the lag period could be explained by variety of factors unique to Bangladesh, including seasonal variations of surface temperature, increasing lapse rates due to region-unique warm and cold air advection sources at middle and lower tropospheric levels, and orographic effects (Dewan et al., 2017:1653). P had the lowest correlation with lightning overall, but this could be attributed to Bangladesh's unique monsoon characteristics and sources of precipitation, such as the domination of squall lines during pre-monsoon, the frequent occurrence of monsoonal troughs and depressions during the monsoon, and differences in cloud heights and ice content per flash that vary seasonally (Dewan et al., 2017:1657).

Meanwhile, $CAPE \times P$ showed strong correlation with lightning and was able to explain 89% of lightning variance on a monthly scale, but showed weaker correlations on annual timescales (Dewan et al., 2017:1657). In contrast to the study by Romps et al. in 2014 and 2018, Dewan's study included a longer period of 17 years of data, which shows that $CAPE \times P$ as a proxy for lightning loses utility at longer timescales, at least over Bangladesh (Dewan et al., 2017:1657).

CAPE Times Precipitation Proxy by Tippett and Koshak, 2018

Following the promising work of Romps et al. 2014 in using $CAPE \times P$ as a proxy for lightning, Tippett and Koshak conducted a more comprehensive study on the use of this proxy to forecast lightning over CONUS in 2018. Observed lightning came from the NLDN in the form of CG flashes summed along a 1° by 1° degree grid over a 13-year period (Tippett and Koshak, 2018:10720). CAPE and precipitation rate (P) was provided by the NCEP Global Ensemble Forecast System (GEFS) at a 34 km horizontal resolution for the first eight forecast days, and at 70 km resolution for days nine through 35, with retrospective forecasts constructed using NOAA Subseasonal Experiment (SubX) Climate Testbed (Tippett and Koshak, 2018:10720).

The lightning proxy was calculated using the product of collocated daily values of CAPE and P, with a constant of proportionality $\frac{\eta}{E}$ that varied according to lead time, but with no spatial or seasonal dependence (Tippett and Koshak, 2018:10721). This method of computing the lighting proxy differed from the Romps et al. 2014 study in that it used collocated daily values of CAPE and P, instead of CAPE and P averaged over CONUS, in order to preserve the resolution of regional features (Tippett and Koshak, 2018:10721).

The performance of the lightning proxy as a forecast tool was then assessed by applying it to three different types of forecasts: number of CG flashes, spatial extent of CG lighting occurrence, and lightning/no lightning maps, all at a 1° by 1° degree grid resolution (Tippett and Koshak, 2018:10721). The forecast number of CG flashes skill was evaluated using rank Spearman correlation to reduce the effect of outliers, as well as Pearson correlation for daily and weekly forecasts (Tippett and Koshak, 2018:10721). The lightning/no lightning maps were evaluated using two by two contingency tables using Threat Score (TS), Probability of Detection (POD), False Alarm Rate (FAR), and the Heidke Skill Score (HSS) (Tippett and Koshak, 2018:10721).

The forecasted spatial extent of the $CAPE \times P$ proxy over CONUS at an annual scale showed an uncentered pattern correlation of 0.95 and a centered pattern correlation of 0.89, which was overall very strong (Tippett and Koshak, 2018:10722). The forecasted rank correlation for the number of CG flashes was greater than 0.8 at all lead times when the annual cycle was included, but reduced to 0.6 for forecasts of short lead times when the annual cycle was not included, suggesting that the proxy performed better at predicting the spatial extent of CG lightning rather than the number of flashes (Tippett and Koshak, 2018:10724). It was also noted that forecast skill was not consistent throughout the year, with shorter lead times performing worse during the warm season of June through October (Tippett and Koshak, 2018:10724). This finding was believed to be related to the fact that United States precipitation forecasts tend to perform poorer when the major source of precipitation comes from small-scale convective events (Tippett and Koshak, 2018:10724). Overall, this study showed that using $CAPE \times P$ as a proxy to forecast lightning exhibited statistically significant skill for lead times of up to 15 days

(Tippett and Koshak, 2018:10726-10727). Additionally, greater levels of forecast skill would be expected from more realistic models, such as those with explicit lightning physics, the ability to resolve convection, or that are able to diagnose lightning from convective parameterization schemes (Tippett and Koshak, 2018:10727).

CAPE Times Precipitation Proxy by Tippett et al., 2019

Tippett and others performed a follow-up study in 2019 to further gauge the performance of the $CAPE \times P$ proxy over CONUS. Total precipitation rate (mm/day) and most unstable CAPE (J/kg) data were taken from the North American Regional Reanalysis (NARR) at three-hourly resolution, averaged onto a 1° by 1° grid, while CG lightning flashes came from the NLDN, summed at 1° by 1° grids for the period 2003 – 2016 (Tippett et al., 2019:3933-3934). The proxy (CP) was calculated by multiplying the three-hourly values of CAPE and precipitation rate, resulting in a proxy for the number of CG flashes occurring within each grid cell for each three-hourly period (Tippett et al., 2019:3934). The three-hourly CP values were then summed to form daily, monthly, seasonal, and annual values, and then scaled so that they may be correctly compared to the CG flashes (Tippett et al., 2019:3935). The scaling factor allowed the area-weighted sum of CP values during the 2003-2016 time period to match the number of CONUS flashes, and was computed by:

$$scaling\ factor = \frac{areaweighted\ sum\ (CP)}{sum\ (CG\ flashes)} \qquad (2)$$

The scaling factor was then applied to the CP for positively or negatively charged flashes on a 1°×1° grid in the following way:

$$CP\ (negative\ flashes) = 64.17 \times CAPE \times Precipitation \times \cos(\phi) \qquad (3)$$

$$CP\ (positive\ flashes) = 4.79 \times CAPE \times Precipitation \times \cos(\phi) \qquad (4)$$

Where ϕ was latitude in radians and $\cos(\phi)$ scales the varying grid cell area accordingly (Tippett and et al., 2019:3935). Performance of CP was then assessed spatially by use of pattern correlations along a latitude/longitude grid, and then temporally by use of a correlation and mean-squared error (MSE) (Tippett et al., 2019:3935). The error in the MSE was defined as the difference between CP and the number of CG flashes; MSE was also normalized by CG flash variance at the same temporal resolution (Tippett et al., 2019:3935).

Results showed that CP yielded an $r = 0.87$ with daily positive CG flashes and an $r = 0.89$ with daily negative CG flashes (Tippett et al., 2019:3937). This finding differed from the similar study done by Romps et al. (2014), where they found an $r = 0.72$ with CG flashes (positive and negative) over CONUS for a single year (Romps et al., 2014:853). Tippett and others theorized that this difference may have been due to NARR CAPE being more representative of the large-scale environment, whereas CAPE calculated by radiosonde data as Romps and others did in 2014 may contain small-scale variability (Tippett et al., 2019:3937). At monthly scales, CP and CG flashes correlated well, except during July through August where the CP values were too low, and September through October where the CP values were too high (Tippett et al., 2019:3937). At annual scales, CP explained 63 and 52% of the variance of negative and positive CG flash counts, respectively (Tippett et al., 2019:3937). Overall, this study showed that the CP proxy performed best at daily and monthly timescales and in the cool season months of November through April (Tippett et al., 2019:3943). The poorer performance of the CP during the warm season months of May through October may

indicate that CP is better associated with storm occurrence than with the total number of CG flashes within a storm (Tippett et al., 2019:3943).

Chapter 3: Data and Methods

3.1 GLD360 Lightning Data

GLD360 stroke data, consisting of both IC and CG lightning for the period 2016 – 2020, was chosen due to its global coverage and high resolution/accuracy (1.8 km to 2.4 km location accuracy). Stroke data was chosen because flash data was neither commercially available, nor regularly produced for this sensing network (Cramer, 2021). Additionally, stroke data typically provides more detailed spatial and temporal information about a lightning discharge compared to flash data, as flashes are comprised of strokes (Cramer, 2021). Although NLDN has a superior resolution/accuracy (0.15 to 0.25 km location accuracy), it was not utilized due its non-global coverage which was too limiting of a factor for ASSET's need. GLM data was also considered, due to its nature of originating from an electro-optical sensor in space, which closely mirrors the manner in which ASSET generates a scene. However, GLM data was ultimately not used for calculations in this study due to its lower resolution (8 km to 14 km) and non-global coverage.

3.2 0.25° Global Forecast System Model Description

The 0.25° Global Forecast System (GFS) was chosen for calculations within this study due to its global coverage, 0.25° resolution (on-par with other global models), 384 forecast hour capability, and ease of availability (GFS, 2021). It possesses a horizontal

resolution ranging from 13 km to 34 km depending on latitude for forecast days zero to 10, and a vertical resolution of 64 hybrid sigma-pressure levels, both of which are fairly high for operational uses (GFS, 2021). Although this model does not possess any lightning parameterizations and is incapable of outputting any form of lightning forecasts, it does regularly output variables important to thunderstorm and lightning production, such as CAPE, total precipitation, convective precipitation, precipitable water, temperature, ice water mixing ratio, and graupel (GFS, 2021). Additionally, it possesses both shallow and deep convection parameterizations, which should improve the accuracy of the aforementioned variables important to thunderstorm and lightning production (GFS, 2021). The GFS is initialized four times daily (00Z, 06Z, 12Z, and 18Z), with forecast hours available in three-hour increments (GFS, 2021).

3.3 Modeling Method - CAPE times Precipitation as a Proxy

To model lightning, the $CAPE \times P$ proxy method, demonstrated by Romps et al. (2014, 2018), Dewan et al. (2017), and Tippett et al. (2018, 2019), was chosen due to its large correlation with lightning flashes at daily and monthly timescales (Tippett et al., 2019:3932), availability of CAPE and P as common output variables in the GFS, and low computational cost. Since the GFS contains no lightning parameterizations and is unable to output a lightning forecast, the $CAPE \times P$ proxy presents an indirect method of forecasting lightning daily, monthly, and annual timescales that is limited by the resolution of the GFS.

Additionally, since previous studies showed that CAPE and P individually had greater variability when correlated with lightning across varying geographic locations and

seasons, $CAPE \times P$ appears to present an advantage of correlating more directly with lightning regardless of the situation (Soriano, 2001:3001, Siingh et al., 2014:38-39). After all, not all increases in CAPE result in thunderstorm development, and not all increases in precipitation rate indicate lightning. Increases in CAPE can indicate the potential development of a variety of cumuliform clouds capable of producing showery-type precipitation, but whose vertical extents do not reach high enough to allow mixed phase hydrometeors to interact and allow electrification to be achieved.

Meanwhile, increases in precipitation rate may be due to a variety of non-thunderstorm related events, such as the onset of nimbostratus clouds associated with a passing warm or occluded front, resulting in the sudden onset of continuous moderate precipitation without lightning ("Nimbostratus," 2021). As mentioned earlier, the $CAPE \times P$ proxy represents the theoretical maximum rate at which kinetic energy is imparted to ascending water condensates; a rate whose increase is closely tied to thunderstorm development and lightning production (Romps et al., 2014:852). Together, $CAPE \times P$ appears to present a more reliable proxy for lightning than CAPE or P alone. Henceforth, the proxy will be referred to as "CP," as coined by Tippett and Koshak in their 2018 study.

3.4 Lightning Climatology Study on Three Distinct Regions

Utilizing GLD360 data, a five-year lightning climatology study was conducted on the following 4°× 4° zones: central Arizona (Lat 32/36, Lon -110/-114), central Cuba (Lat 20/24, Lon -77/-81), and North Korea (Lat 38/42, Lon 125/129). Figures 1, 2, and 3 below show these zones on a terrain relief map. These three zones were chosen due to

their distinct climates according to the Köppen-Geiger climate classification system. Arizona is classified as Tropical and subtropical desert (BWh), Cuba is classified as Tropical wet-dry (Aw), and North Korea is Humid continental (Dwa) ("Climate," 2021, "Arnfield, 2020"). The lightning climatology study included lightning stroke density ($counts\ km^{-2}$), energy density ($kA\ km^{-2}$), stroke-energy distribution, stroke diurnal cycle, and mean energy diurnal cycle.

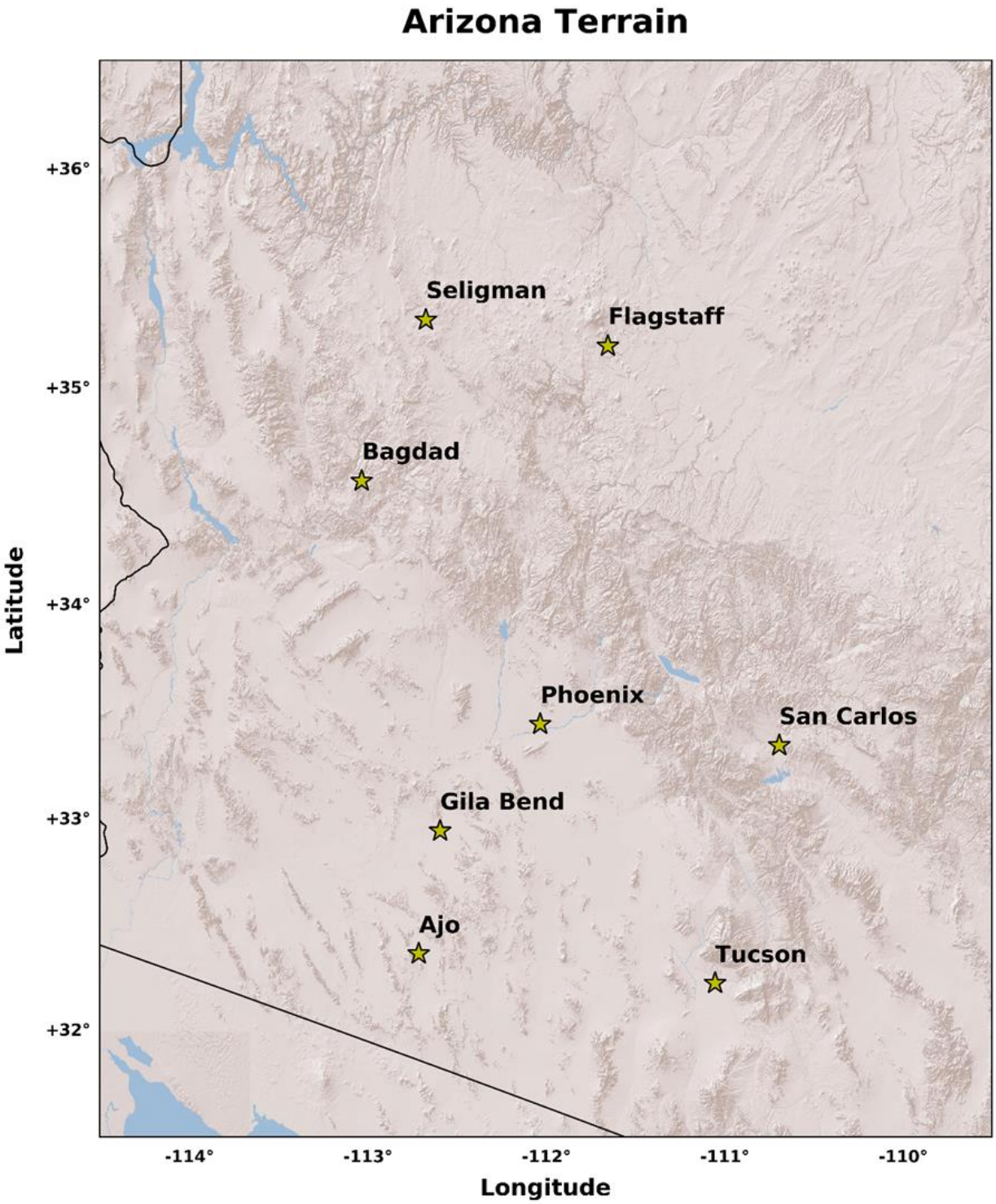

Figure 1. Arizona Terrain Relief Map

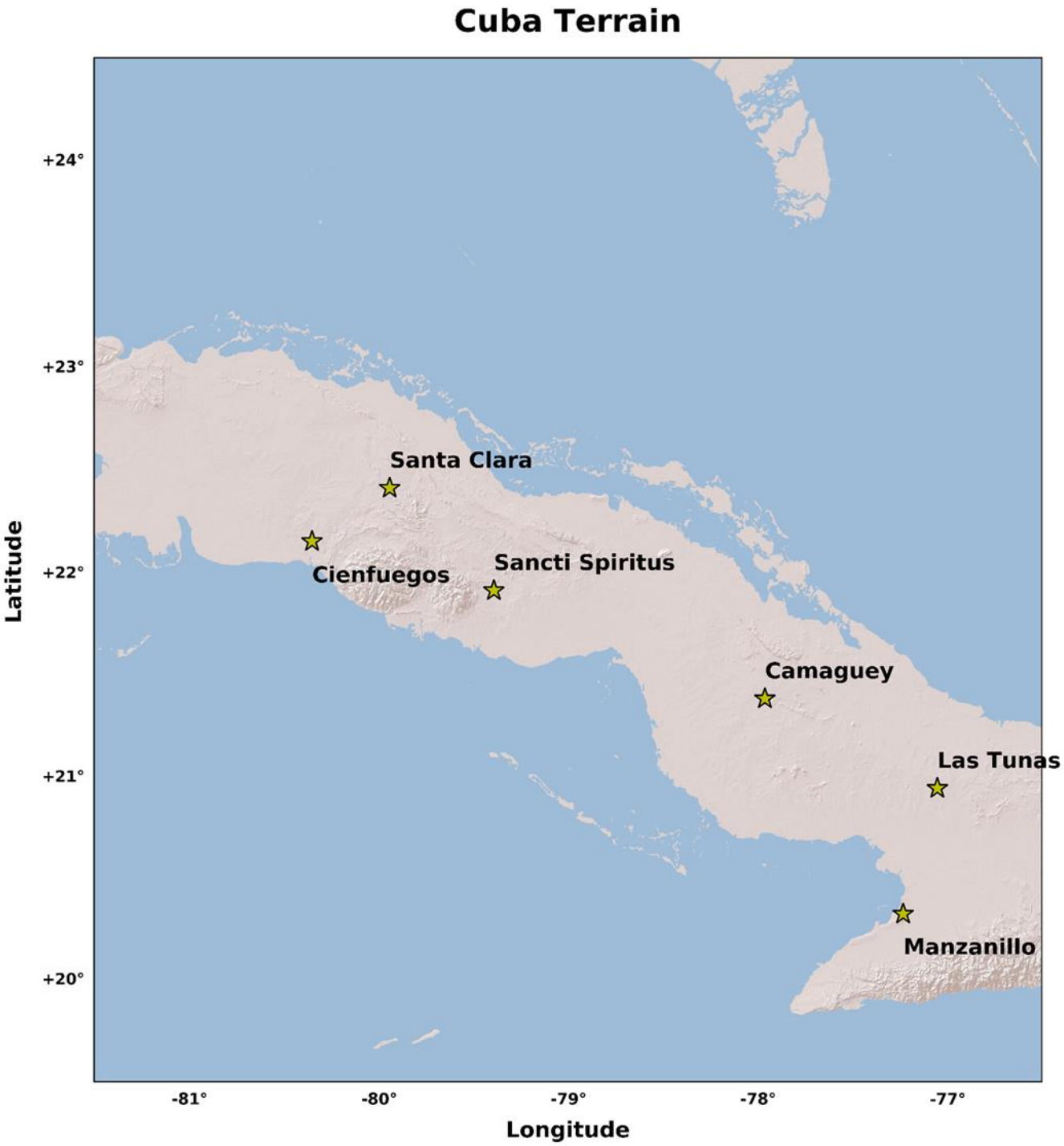

Figure 2. Cuba Terrain Relief Map

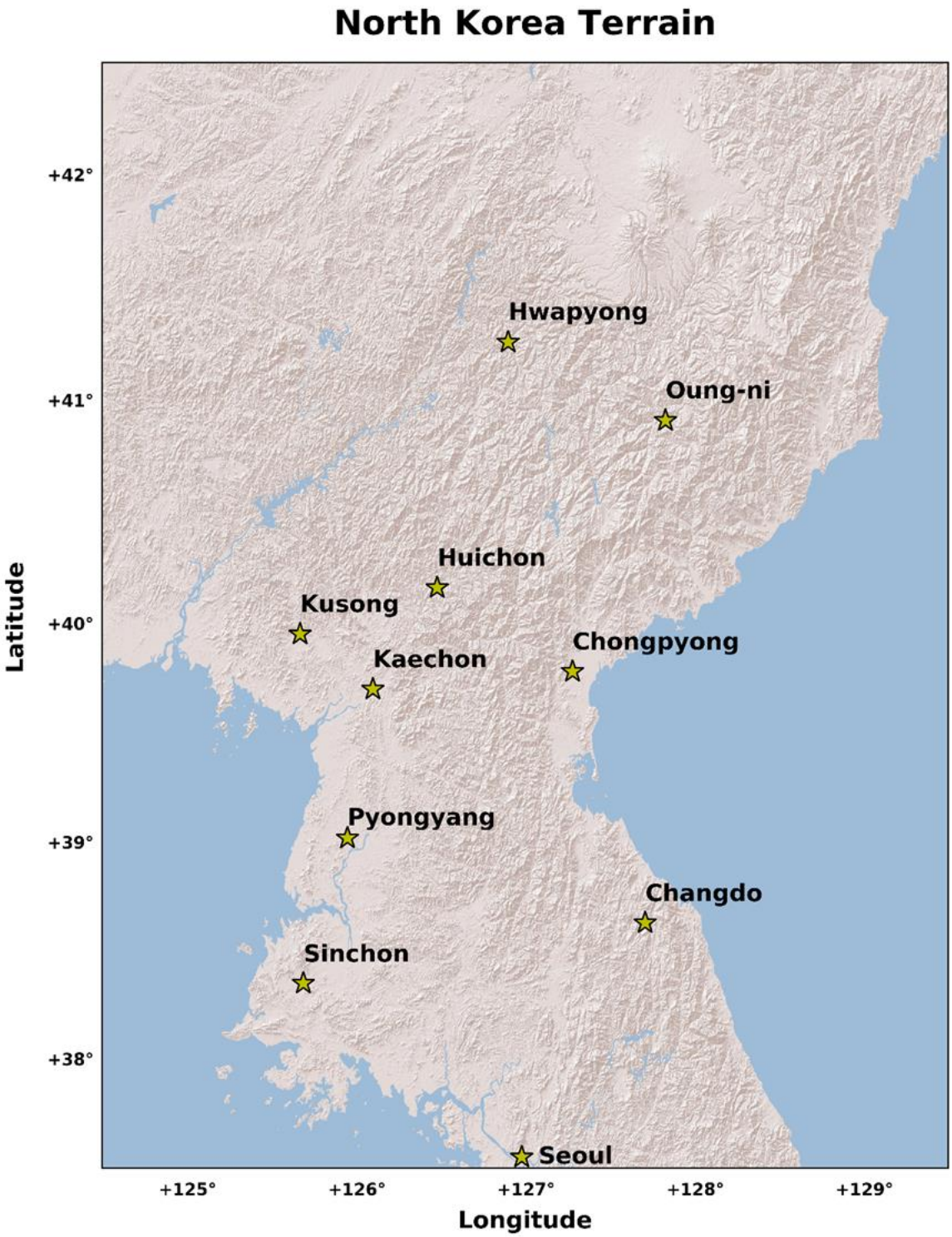

Figure 3. North Korea Terrain Relief Map

3.5 CP to Lightning Correlation Study Over Three Zones

For the months of June, July, and August of 2019, and their combined season (JJA), a Pearson correlation between CP and lightning stroke density was conducted for the three zones. A Pearson correlation coefficient (*r)* was found for each location in June, July, August, and JJA, along with corresponding p-values. For p-values <.001, the correlation was considered statistically significant. CP was calculated by taking the 6-hour forecast (f006) CAPE and 6-hour (f006) precipitation rate at the 00Z, 06Z, 12Z, and 18Z initialization times, and concatenated monthly and seasonally. The two arrays were then averaged spatially and daily and multiplied together, resulting in a one dimensional arrays of size 30 (June), 31 (July, August) and 92 (JJA) elements long.

Lightning stroke density was calculated by summing the daily strokes as elements over 0.25° grid areas within each 4° by 4° by zone, concatenating the elements at monthly and seasonal durations, then summing the daily total count over the spatial domain. This process resulted in one-dimensional arrays of size 30 (June), 31 (July, August) and 92 (JJA) elements long. A Pearson correlation was then performed on the CP and lightning arrays at the monthly and seasonal timescales to obtain correlation coefficient *r* values and p-values. The Pearson correlation was chosen due to its insensitivity to magnitude, as the goal of the correlation was to capture similarities in rates of change between the two quantities. Lightning and CP units were converted appropriately to units strokes $km^{-2}\ hr^{-1}$ and $J\ kg^{-1} mm\ hr^{-1}$, respectively, similar to the aforementioned studies conducted by Romps et al. and Tippett et al.

3.6 Calculation of the Constants of Proportionality and Correction Factor

The aforementioned Pearson correlation procedure was meant to find a relationship between lightning stroke rate and precipitation rate; however, these are two fundamentally different quantities, as shown by their units. Romps et al. and Tippett et al. demonstrated that these quantities increase and decrease together at predictable rates, as seen by the high *r* values in their studies, but a constant of proportionality must be applied to the CP proxy if it is to be used directly as a forecasting tool. To find this constant, a linear regression was performed on the CP and lightning stroke arrays at monthly and seasonal timescales, and the resulting slope was used to compute the constant of proportionality (*k*) in units of $J^{-1}kg\ mm^{-1}km^{-2}$. *k* was then used to facilitate the conversion of CP into units of strokes $km^{-2}hr^{-1}$, enabling the CP to be used to directly forecast lightning.

A correction factor (*cf*) was also calculated in order to scale the CP to the same magnitude as the observed strokes per grid box. The calculation for *cf* for each grid box was found in the following manner:

$$cf = \frac{maximum\ observed\ strokes\ (km^{-2}\ hr^{-1})}{maximum\ \overline{CAPE} \times \overline{P} \times k\ \ (km^{-2}\ hr^{-1})} \qquad (5)$$

3.7 Using CP to Forecast Lightning Over Zones

To test the proxy as a forecasting tool, the CP was calculated by taking the 6-hourly CAPE and precipitation rate at the four initialization times and concatenating the elements at monthly and seasonal timescales, resulting in arrays of dimensions 120 x 17 x 17 (June), 124 x 17 x 17 (July and August), and 364 x 17 x 17 (JJA). The CAPE and

precipitation arrays were then averaged at monthly and seasonal timescales and multiplied together, resulting in arrays of dimension 17 x 17. The CP arrays were then multiplied by *k* in order to convert from units of $J\ kg^{-1}mm\ hr^{-1}$ to strokes ($counts\ km^{-2}hr^{-1}$) and facilitate a true comparison to observed stroke densities. Next, the array were multiplied by a correction factor *cf* in order to scale their quantities to the comparable magnitude of the strokes for the same month and season. These forecast arrays were then plotted for each zone and visually compared to observed lightning stroke densities plotted at 0.25° grid resolution (17 x 17) for the same zone and time period. Difference plots consisting of the forecast CP minus the observed strokes on their respective 17 x 17 grids for each month and season was also generated to aid in the visual comparison. The CP equation used to forecast lightning strokes was given by:

$$CP = \overline{CAPE} \times \bar{P} \times k \times cf \qquad (6)$$

Where *CP* was the stroke rate in $counts\ km^{-2}\ hr^{-1}$, $\overline{CAPE}$ was the daily average convective available potential energy in $J\ kg^{-1}$, $\bar{P}$ was the daily average precipitation rate in $mm\ hr^{-1}$, *k* was the constant of proportionality in units $J^{-1}kg\ mm^{-1}km^{-2}$, and *cf* was the dimensionless correction factor.

Chapter 4: Results and Analysis

4.1 Lightning Climatology Study (2016 – 2020)

Arizona

Over Arizona, the greatest concentration of strokes occurred along the southern and eastern parts of the domain (Figure 4). A large concentration of strokes occurred just west of Tucson and south of Phoenix largely over the Tohono O'odham Nation

Reservation (Figure 4), which consists mostly of open desert with small mountains spread throughout (Figure 1). Another large concentration area occurred to the northeast of Tucson in the vicinity of several mountain chains known as the Santa Catalinas, Rincons, Galiuros, and Graham mountains (Figure 1). A third large concentration area occurred due east of Phoenix in the San Carlos Reservation (Figure 4), which is also largely open desert littered with several mountain chains (Figure 1). It is known that the majority of lightning over central and southern Arizona occurs during the North American Monsoon period, which occurs during the months of June through September (Adams and Comrie, 1997:2197). During this period, thunderstorms arise frequently as diurnally driven processes enhanced by higher terrain and mountains (Bieda et al., 2009:4214). The greatest concentration of stroke energy also coincided with the greatest stroke concentration area (Figure 5).

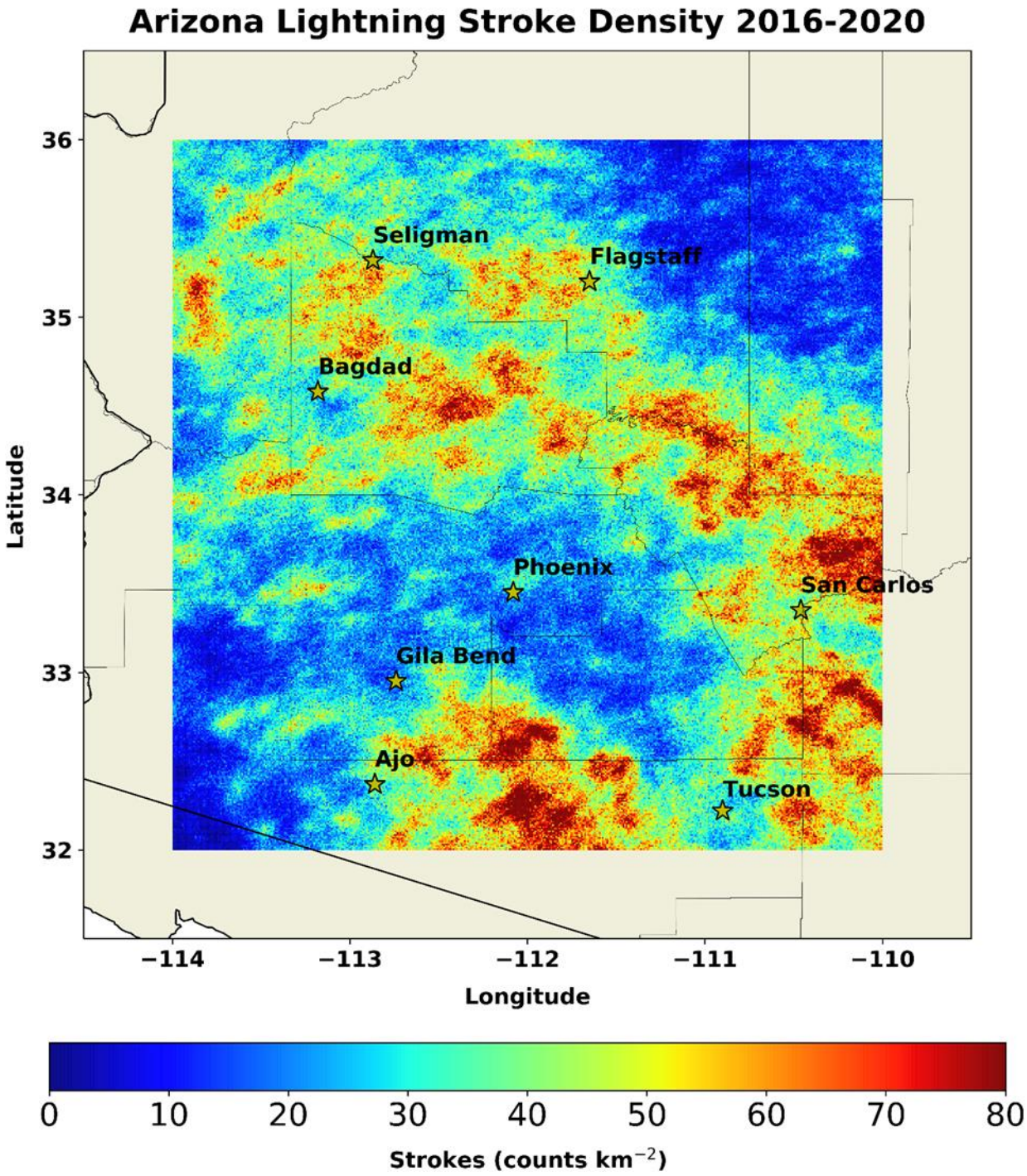

Figure 4. Arizona Stroke Density 2016-2020

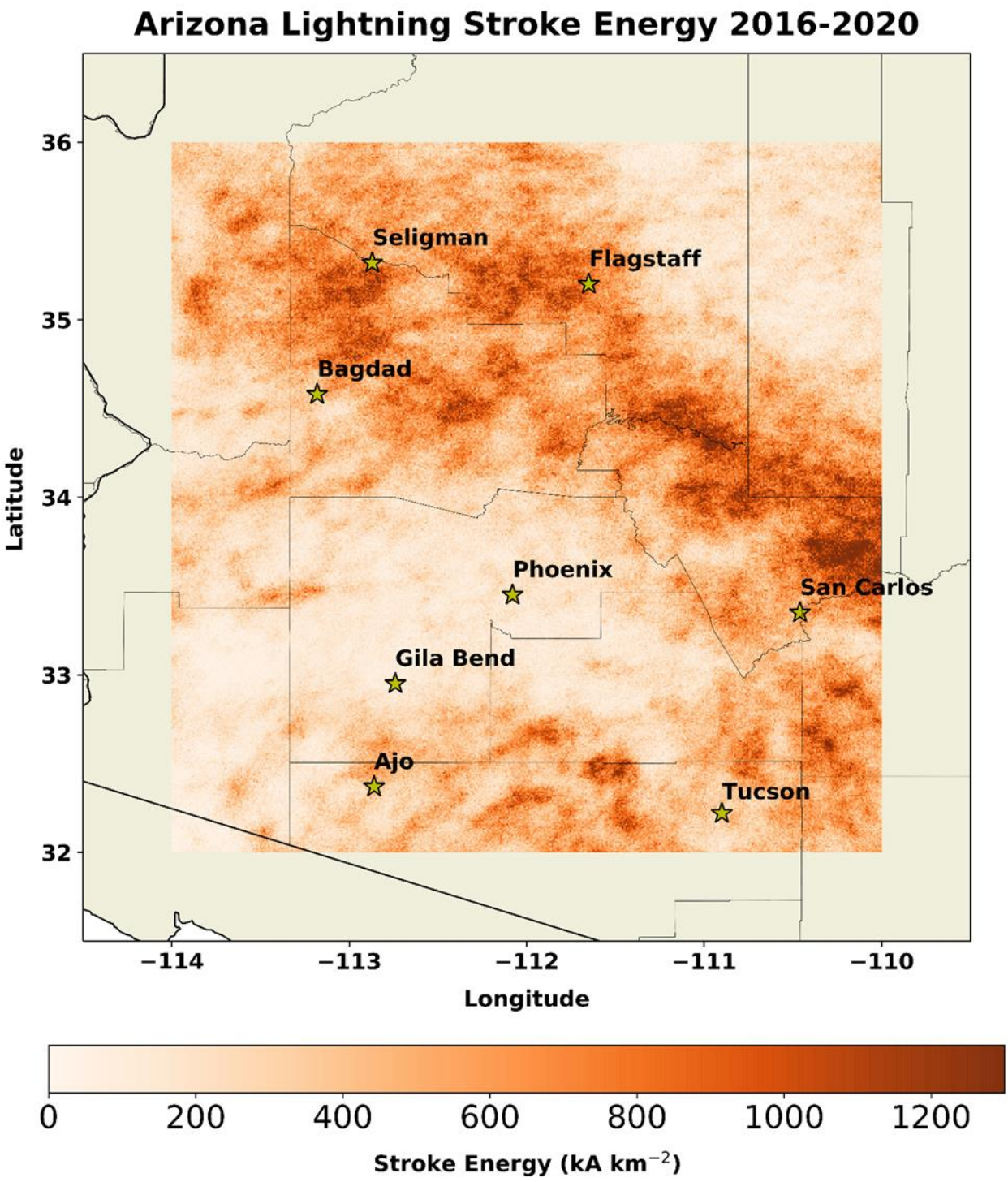

Figure 5. Arizona Stroke Energy 2016-2020.

The stroke counts followed a predictable diurnal trend, with a maximum of 678,411 strokes occurring during the late afternoon around 16L, and a minimum of 30,986 strokes occurring during the morning around 06L to 08L (Figure 7, top left). This diurnal pattern supports the fact that most thunderstorms in this region form in the mid-to-late afternoon due to strong surface heating and available moisture (Bieda et al., 2009:4214). Also of note is that the majority of strokes occurred in the lower end of the stroke energy range of less than 50 kA (Figure 6). Furthermore, the greatest stroke energies coincided with strokes that occurred during the morning at approximately 08L, with greatest mean stroke energies near 17 kA, and mean top 10% stroke energies near 63kA (Figure 7, top right and bottom right). This was an interesting relationship to find, because it indicates that while morning thunderstorms are rare, they seem to produce the most powerful lightning. Additionally, the mean bottom 10% stroke energies reached a minimum at approximately 10L with a value near 1.5 kA (Figure 7, bottom left), indicating that the highest and lowest energy thunderstorms tend to occur almost concurrently in the morning.

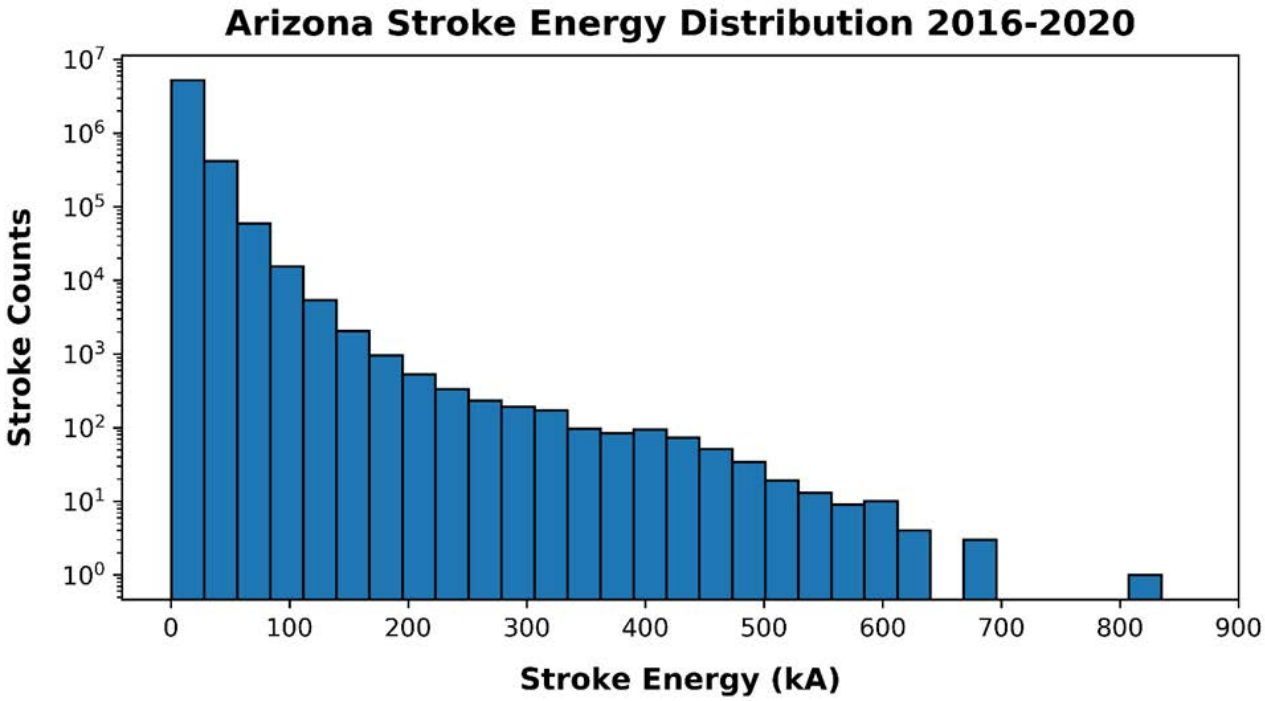

Figure 6. Arizona Stroke Energy Distribution 2016-2020.

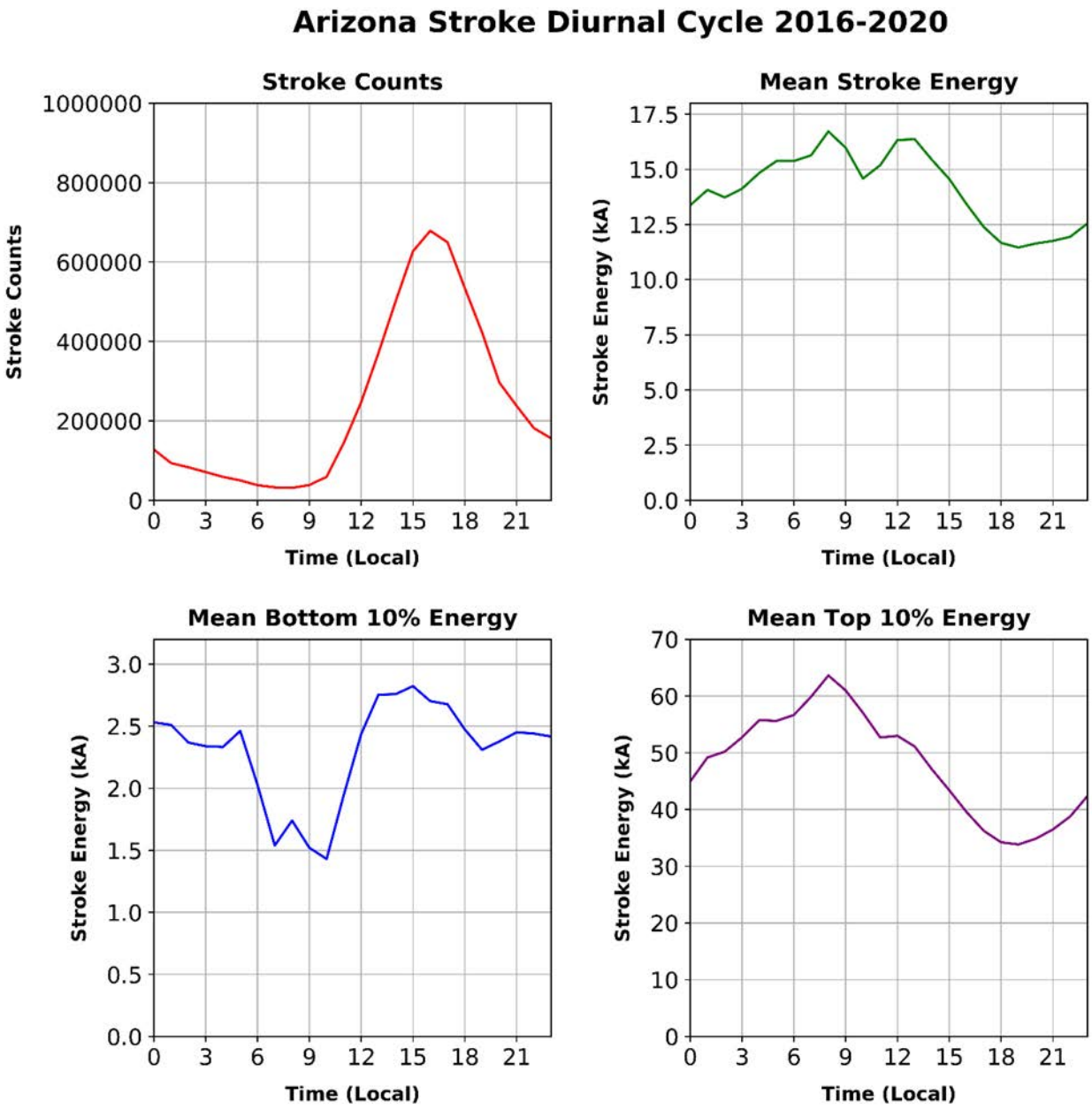

Figure 7. Arizona Stroke Diurnal Cycle 2016-2020: Stroke Counts (top left), Mean Stroke Energy (top right), Mean Bottom 10% Energy (bottom left), Mean Top 10% Energy (bottom right).

Cuba

Cuba is known to have a wet season occurring from late April to early November, with high convective activity occurring during July and August ("Climate-Cuba," 2022). During the convective season, thunderstorms frequently form in the late afternoon due to strong surface heating and sea breeze convergence (Mayor and Mesquita, 2015:1). While Cuba is a relatively flat island, it contains several notable topographical features: a small mountain chain in the south-central region just west of Sancti Spiritus, a set of mountains lining the far southeastern coast and far northeastern coast near Guantanamo, and another set of mountains on the western side, just west of Havana.

Over Cuba, the greatest stroke densities occurred in three distinct regions: north of Cienfuegos, southwest of Camaguey, and southeast of Manzanillo (Figure 8). Lightning stroke energies also peaked in these areas, owing to these large stroke densities (Figure 9). These high stroke density and energy areas also coincided with mountains that undoubtedly provided orographic lift, such those located far northwest of Cienfuegos, north and south of Camaguey, and southeast of Manzanillo (Figure 2). It is interesting to note that the majority of lightning strokes occurred on land, which supports the effect of the sea breeze convergence as a mechanism for triggering convection during the summer months. As Cuba is an island surrounded by water on all sides, strong surface heating of the land would induce onshore-flow from all directions in the late afternoon, resulting in low-level convergence and thunderstorm formation in the interior of the island. Additionally, the locations of the mountains likely modifies the sea breeze convergence, providing an area of greatest lift and thus, greatest stroke densities. Although the GLD360 sensors are land-based and exhibit some land-sea bias on the outer edge of the

sensor network (meaning that detection efficiency reduces the further one is from the sensor), Cuba is located well-within the reach of the GLD360 network (Said et al., 2013:6907, 6910-6911). Thus, the high concentration of stroke densities seen in the mainland of Cuba is most likely not due to GLD360 sensor limitations, but due to purely meteorological reasons.

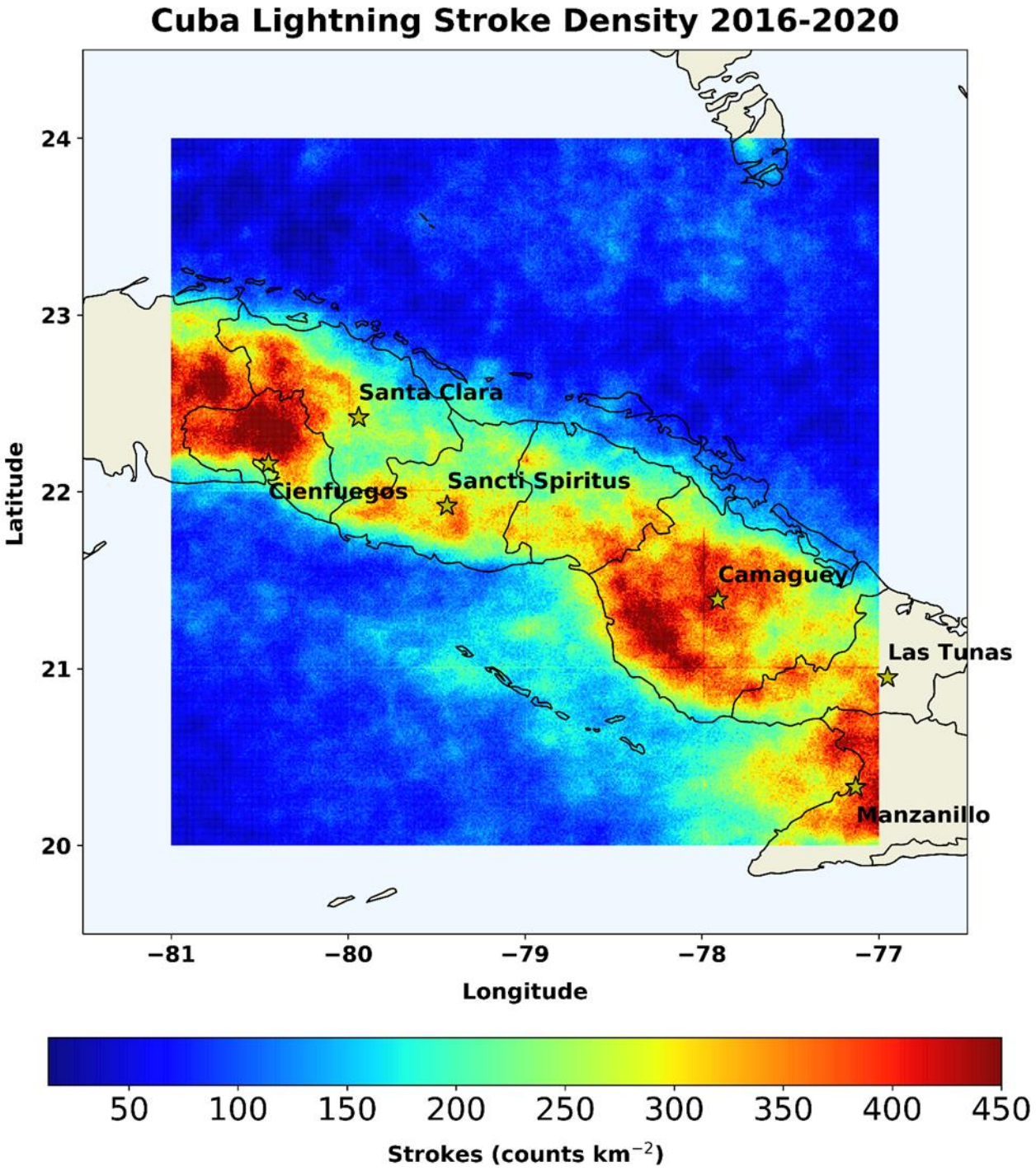

Figure 8. Cuba Stroke Density 2016-2020.

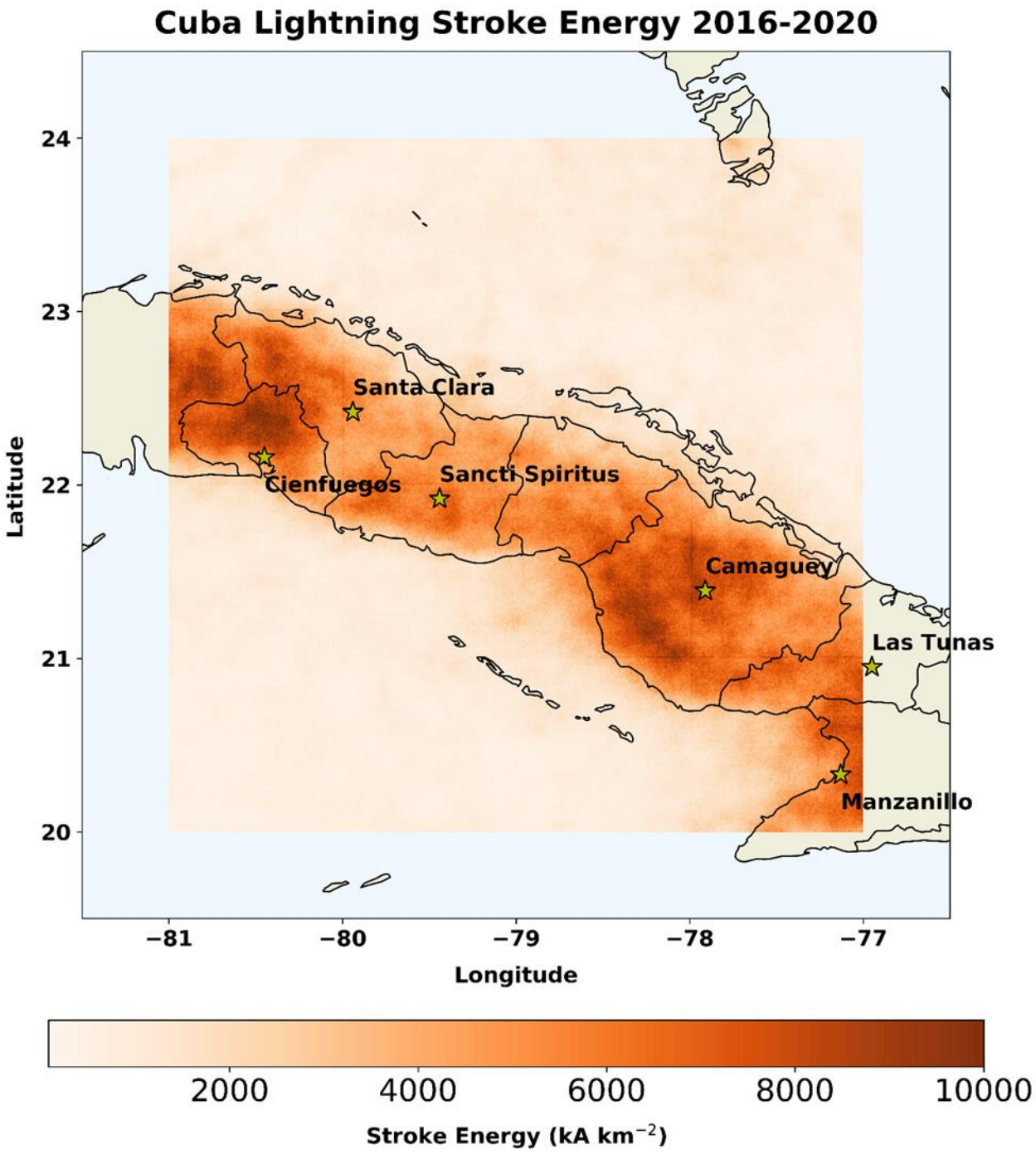

Figure 9. Cuba Stroke Energy 2016-2020.

The stroke counts also followed a predictable diurnal cycle, with a peak count of 3,725,711 strokes occurring during the late afternoon at approximately 16L, and a minimum of 238,814 strokes occurring during the early morning between 05L and 06L (Figure 11, top left). As these strokes appear to be diurnally driven, we can infer that the majority of strokes occur during the convective months of June, July, and August when diurnal effects drive thunderstorm onset ("Climate-Cuba," 2022). The mean stroke energy was approximately 24 kA and peaked around 04L, which coincides with the early morning when thunderstorms are rarest (Figure 11, top right). This parallels the finding from Arizona, in which the greatest stroke energies were also observed in the early morning when thunderstorms were least frequent. The stroke energy distribution for Cuba also paralleled Arizona, in which the majority of strokes occurred in the lower range of energy (Figure 10). Additionally, the mean top 10% energy of strokes peaked at approximately 04L with a value near 87.5 kA (Figure 11, bottom right), while the mean bottom 10% energy of strokes reached a minimum around 11L with a value near 3.9 kA (Figure 11, bottom left). This indicates that while the majority of thunderstorms occurred in the afternoon, they contained the smallest energies. In contrast, the fewest thunderstorms occurred in the early morning, but contained the greatest energies.

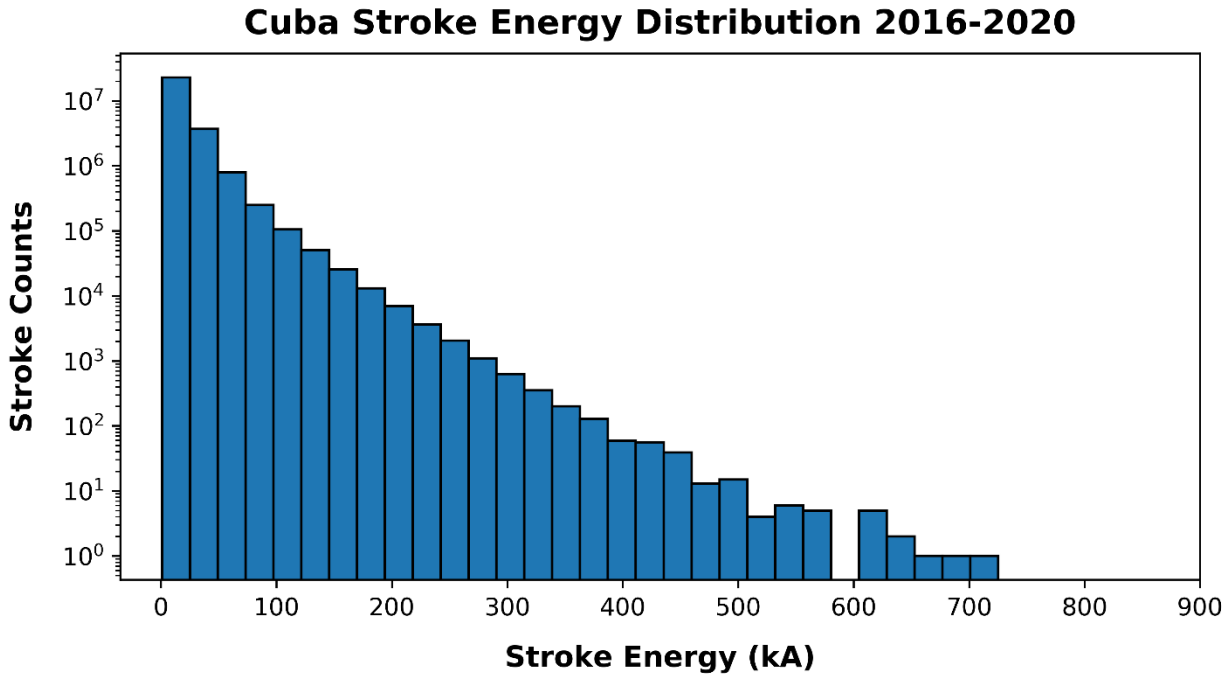

Figure 10. Cuba Stroke Energy Distribution 2016-2020.

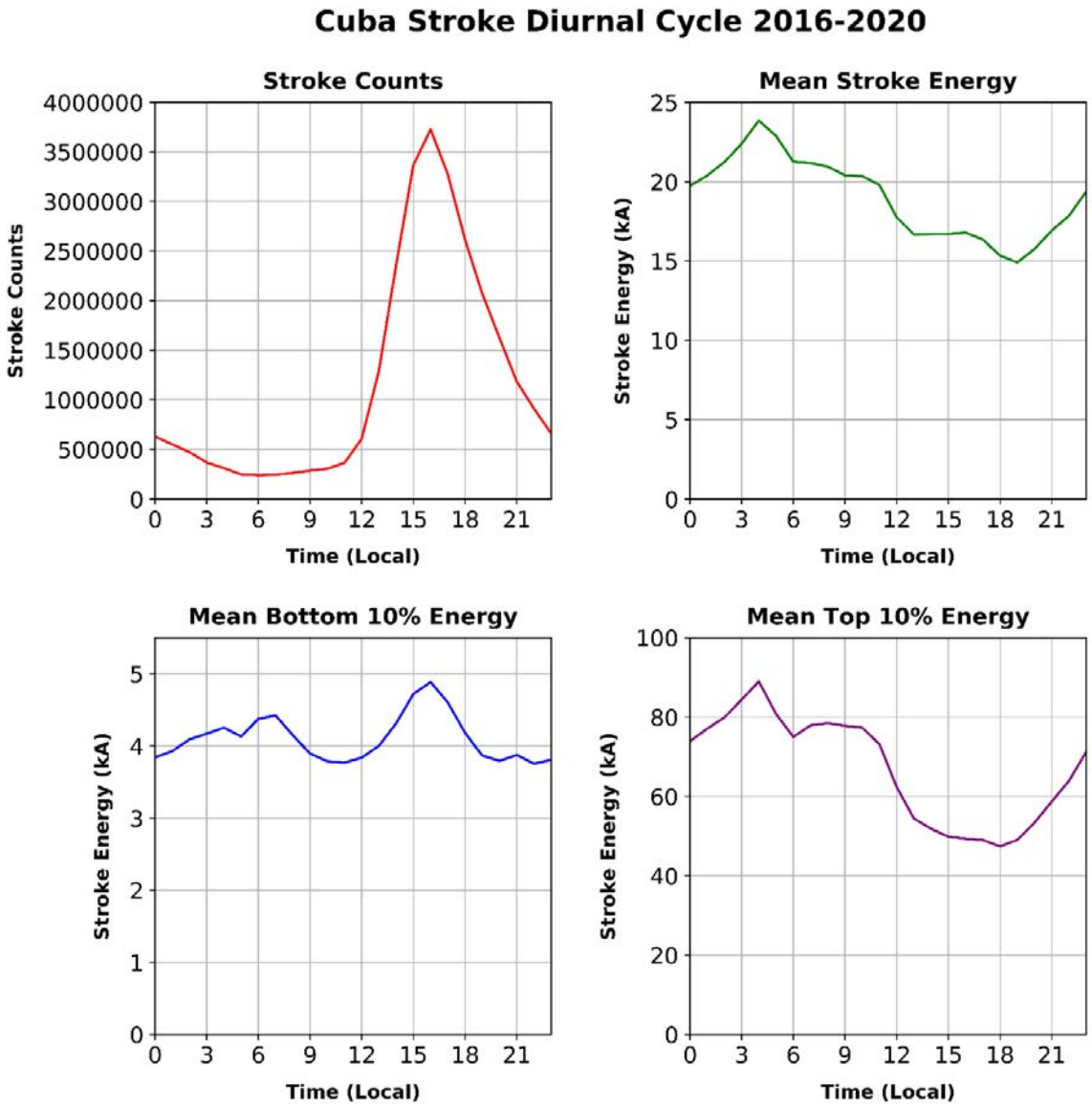

Figure 11. Cuba Stroke Diurnal Cycle 2016-2020: Stroke Counts (top left), Mean Stroke Energy (top right), Mean Bottom 10% Energy (bottom left), Mean Top 10% Energy (bottom right).

North Korea

North Korea possesses a continental climate, characterized by cold, sunny winters and hot, wet summers ("Climate – North Korea," 2022). Its wet season generally occurs during May through September, with the most active thunderstorm period occurring during July and August ("Climate - North Korea," 2022, Walters and Traxler, 1994:4-4,4-5). It is a geographically diverse region, with the eastern half dominated by mountainous terrain, the western side flatter and bordered by the Yellow Sea, and the far eastern side bordered by the Sea of Japan (Figure 3). The country experiences a true monsoon circulation with a prevailing cold northwesterly wind in the wintertime that switches to warm southerly during the summer ("Climate – North Korea," 2022).

Lightning density results illustrated overall sparse lightning compared to Arizona and Cuba, with the greatest stroke densities occurring along the west coast, just northwest of Pyongyang in the Korea Bay (Figure 12). The concave shape of the bay, the increasing slope of terrain towards the interior, as well as regular afternoon sea breezes likely results in upslope flow and enhanced lift in this region, resulting in frequently occurring thunderstorms in this area during the summer months (Figure 3, Figure 12). Other larger stroke density areas included the vicinity of a small mountain chain near the demilitarized zone (DMZ) just northwest of Seoul, the interior of a large mountain chain in the northeast corner just southeast of Oung-ni, in the mountains northeast of Chongpyong, and along the mountainous terrain in the far interior of the country, near the China border (Figure 3, Figure 12). Not surprisingly, the locations of the densest stroke energy areas coincided with the aforementioned high stroke density areas (Figure 12, Figure 13).

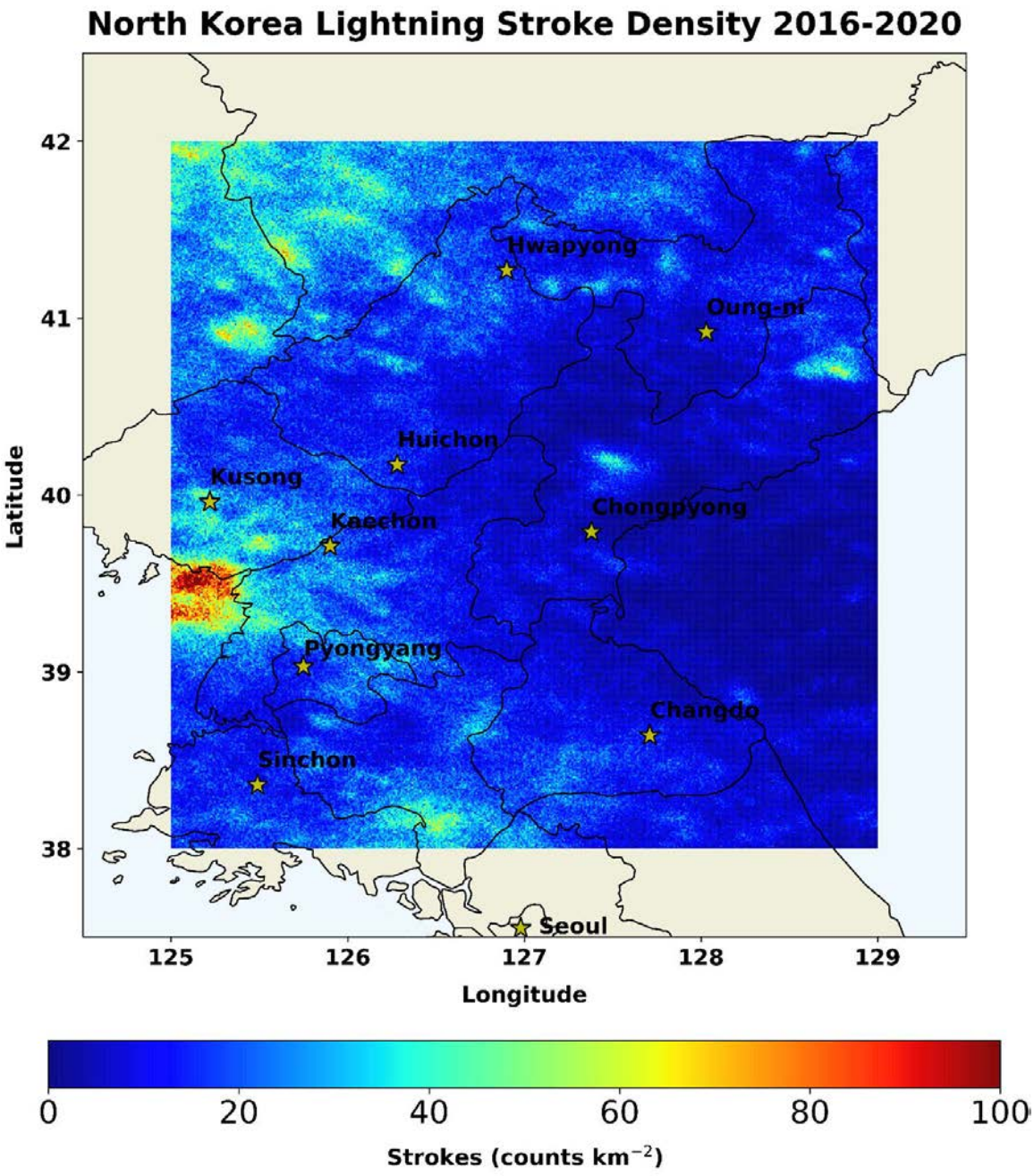

Figure 12. North Korea Stroke Density 2016-2020.

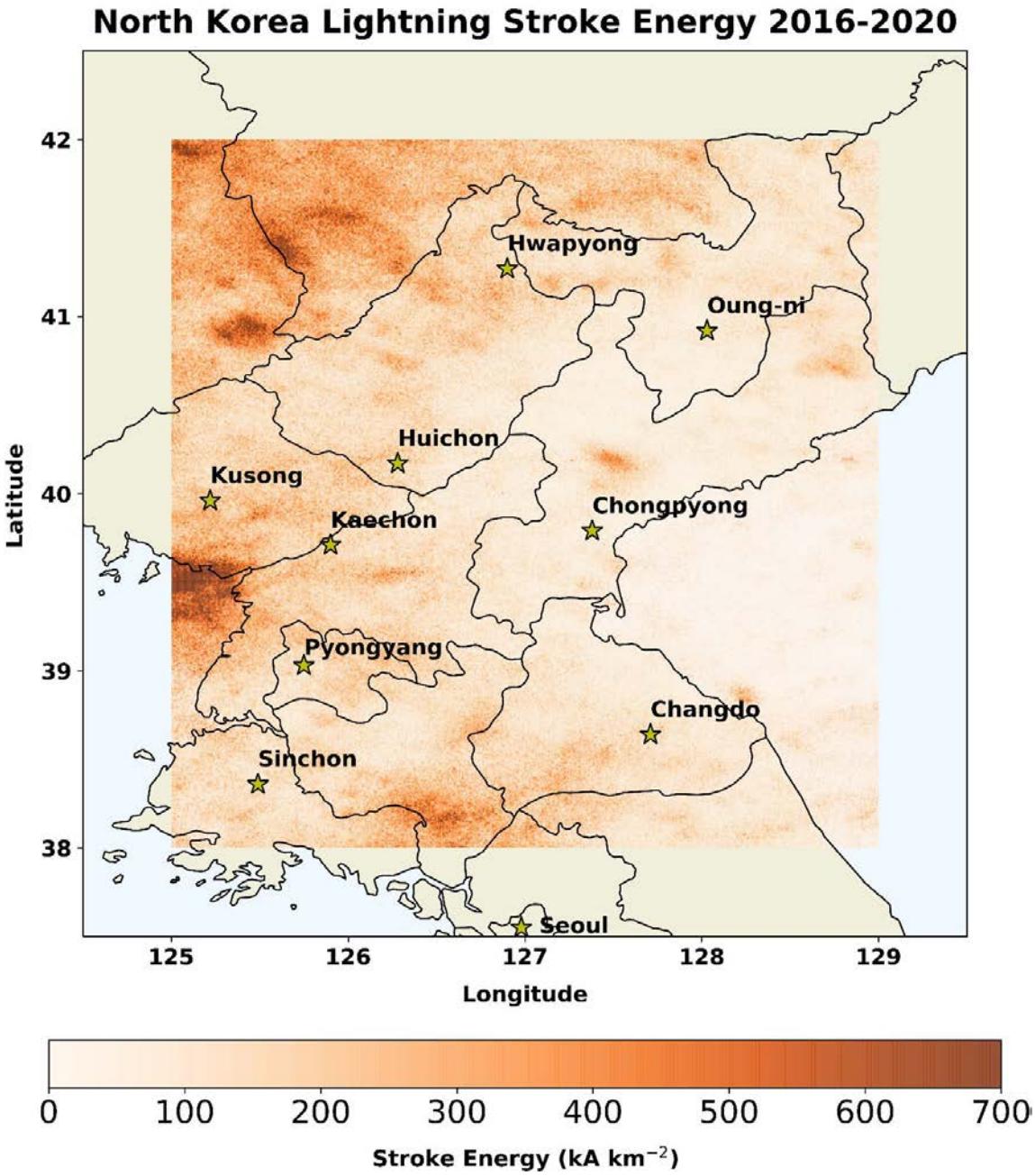

Figure 13. North Korea Stroke Energy 2016-2020.

North Korea had overall much fewer lightning strikes than either Arizona or Cuba, with maximum stroke counts of 203,676 strokes occurring at approximately 17L during the late afternoon, and minimum stroke counts of 21,687 strokes occurring at approximately 06L during the early morning (Figure 15). Again, the pattern of thunderstorm activity appears to be tied to diurnal trends, as with Arizona and Cuba. This is supported by the finding that most thunderstorms in North Korea occur in July and August and are primarily triggered by strong surface heating (Walters and Traxler, 1994:4-4). The high stroke area in the Korea Bay (Figure 12) is likely due to a tertiary circulation induced by surface heating, in which local pressure differences arise due to afternoon temperature differential between the land and water, driving onshore flow that interacts with the upsloping terrain to create an area of enhanced lift (Holle, 2022). The distribution of stroke energy to stroke counts followed suit with Arizona and Cuba, with the majority of strokes occurring in the lower end of stroke energy at less than 25 kA (Figure 14). The stroke energy over a diurnal cycle also showed similarity with Arizona and Cuba, in that the smallest mean stroke energy of 8 kA (Figure 15, top right) occurred during the afternoon during the period of most thunderstorm activity, while the greatest mean stroke energy of 19 kA (Figure 15, top right) occurred during the early morning during the period of least thunderstorm activity. The mean top 10% energy peaked at 73 kA during the early morning around 04L (Figure 15, bottom right), while the mean bottom 10% energy reached its minimum near 1.9 kA during the early evening around 19L (Figure 15, bottom left). Interestingly, a secondary mean bottom 10% energy minimum near 2.1 kA also occurred in the morning around 08L (Figure 15, bottom left).

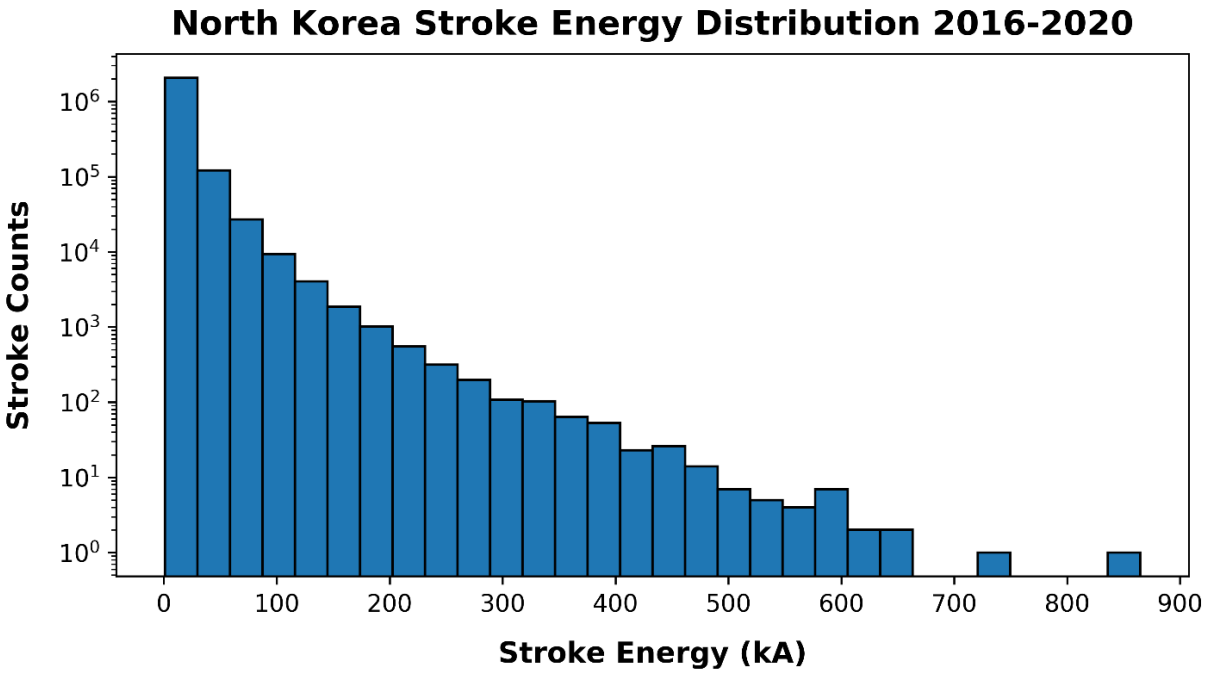

Figure 14. North Korea Stroke Energy Distribution 2016-2020.

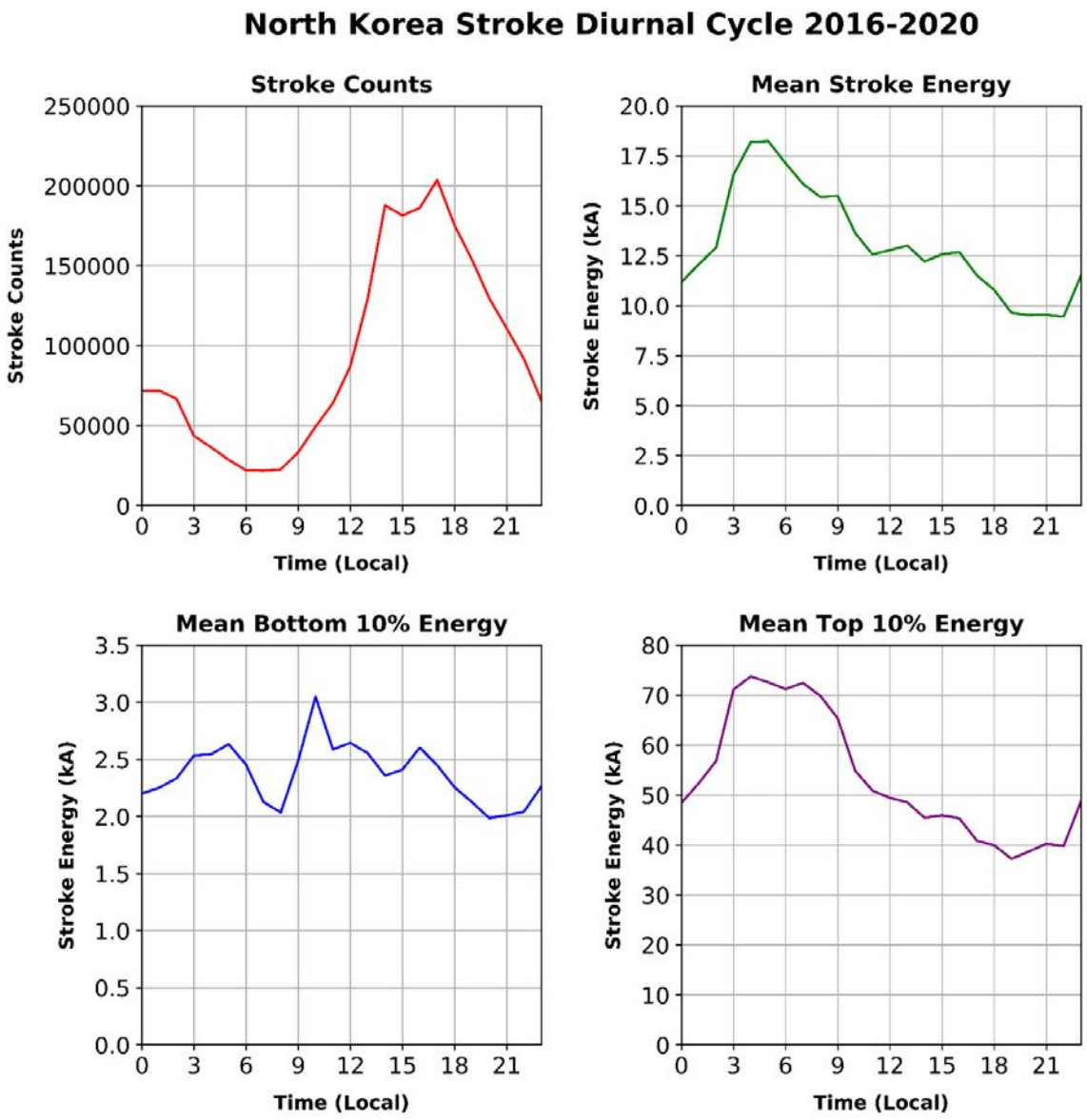

Figure 15. North Korea Stroke Diurnal Cycle 2016-2020: Stroke Counts (top left), Mean Stroke Energy (top right), Mean Bottom 10% Energy (bottom left), Mean Top 10% Energy (bottom right).

Overall, the 2016-2020 lightning climatology study resulted in three main findings: the occurrence of lightning at each location appears to be diurnally driven; most

of the annual variability in lightning appears to be due to the contribution of lightning from the warm season (June, July, and August), meaning that these three months produce the majority of lightning for the year; the greatest stroke energies tend to occur in the early morning when thunderstorms are least likely to occur, while lowest stroke energies tend to occur in the afternoon when thunderstorms are most likely to occur. These findings were consistent across Arizona, Cuba, and North Korea in this study. Additionally, the strange inverse of stroke frequency and energy is actually a common phenomenon that has been observed worldwide, and is known to occur due to reduced atmospheric mixing overnight (Woollaston, 2015). As the surface cools overnight and stability increases throughout an air column, atmospheric mixing is reduced, resulting in fewer available particles to create a separation of charge between regions of a cloud and the ground (Woollaston, 2015). Fewer particles means that larger distances exist between charge regions of a cloud, and so a more powerful buildup of charge is necessary for a discharge to occur (Woollaston, 2015).

4.2 CP to Lightning Correlation Study

Arizona

Arizona showed fairly strong correlation between CP and lightning strokes for June, with $r = 0.718$ and p-value <0.001, indicating that the correlation was statistically significant (Figure 16, top left). For July, Arizona exhibited an even stronger, statistically significant correlation, with $r = 0.779$ and a p-value<0.001 (Figure 16, top right). For August, the correlation was much weaker, with $r = 0.462$ and a p-value = 0.01, indicating slightly less statistical significance (Figure 16, bottom left). For JJA, the correlation was

moderately strong, at $r = 0.558$ and a p-value<0.001, indicating strong statistical significance (Figure 16, bottom right).

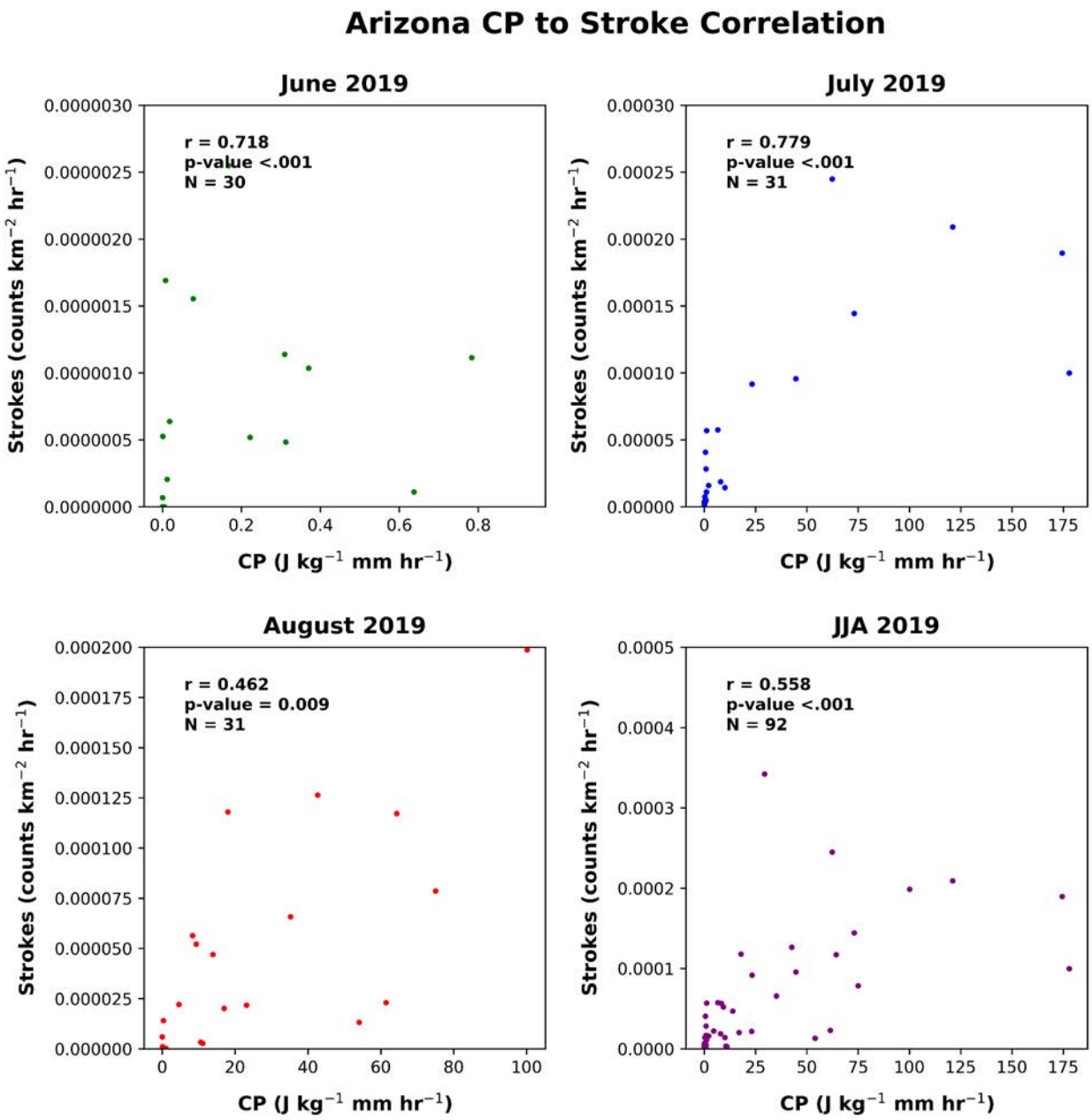

Figure 16. Arizona CP to Stroke Correlation: June 2019 (top left), July 2019 (top right), August 2019 (bottom left), JJA 2019 (bottom right).

Cuba

Cuba for June exhibited moderately strong, statistically significant correlation, at $r = 0.684$ and a p-value<0.001 (Figure 17, top left). For July, the correlation was even stronger, with $r = 0.711$ and a p-value<0.001, indicating statistical significance (Figure 17, top right). For August, the correlation was much weaker, with $r = 0.26$ and a p-value = 0.158, indicating it was not statistically significant and that there was a strong possibility that $r = 0.26$ was due to random chance (Figure 17, bottom left). For JJA, the correlation was moderately strong, with $r = 0.597$ and a p-value<0.001, indicating statistical significance (Figure 17, bottom right).

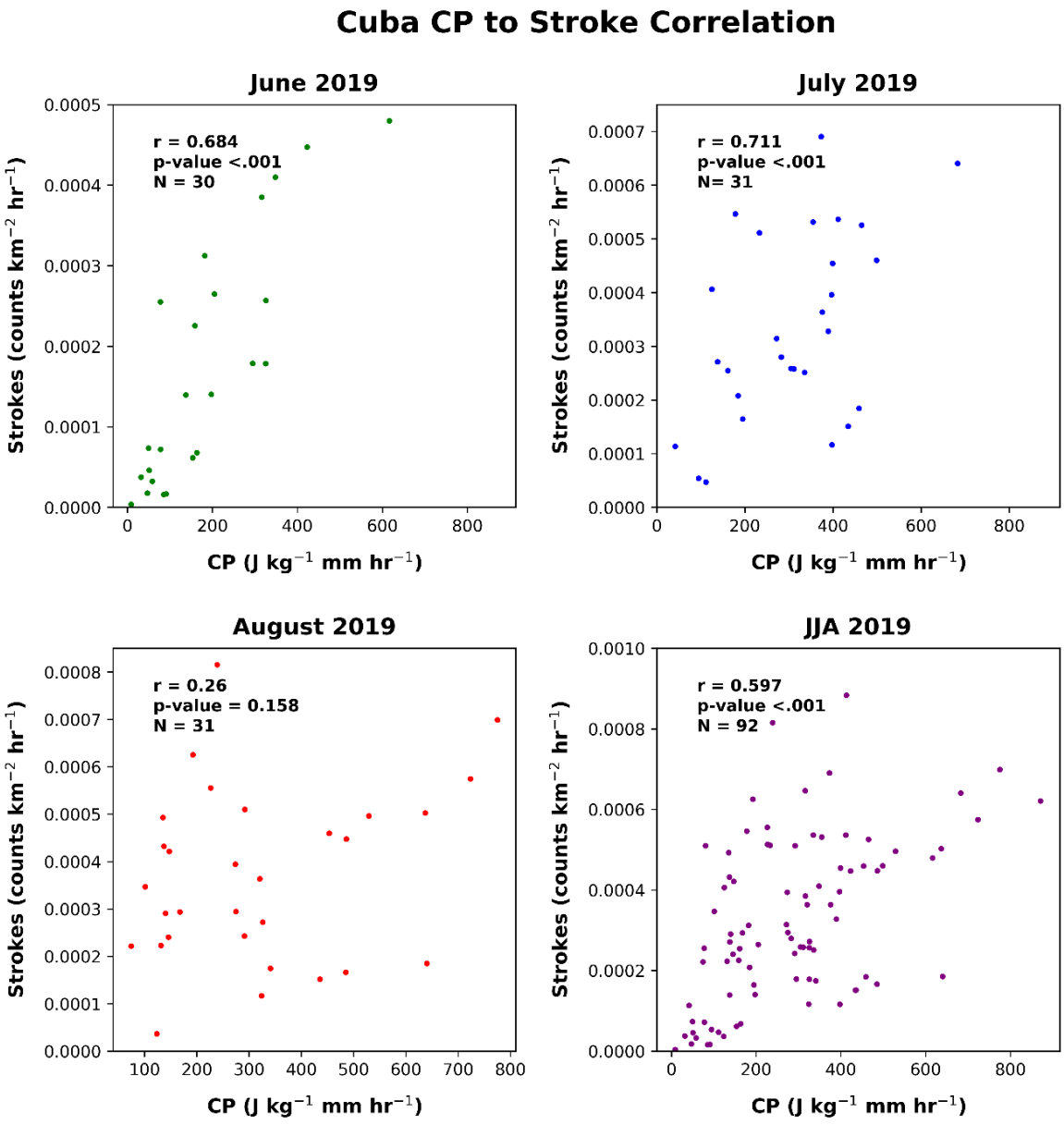

Figure 17. Cuba CP to Stroke Correlation: June 2019 (top left), July 2019 (top right), August 2019 (bottom left), JJA 2019 (bottom right).

North Korea

North Korea for June exhibited moderately strong, statistically significant correlation, with $r = 0.562$ and a p-value = 0.001 (Figure 18, top left). For July, the correlation degraded to a very weak $r = 0.005$, and p-value = 0.977, indicating that it was not statistically significant (Figure 18, top right). For August, the correlation improved to a reasonably strong $r = 0.412$ and a p-value = 0.02, indicating slightly less statistical significance (Figure 18, bottom left). For JJA, the correlation was very weak and statistically insignificant, with $r = 0.164$ and a p-value = 0.12 (Figure 18, bottom right).

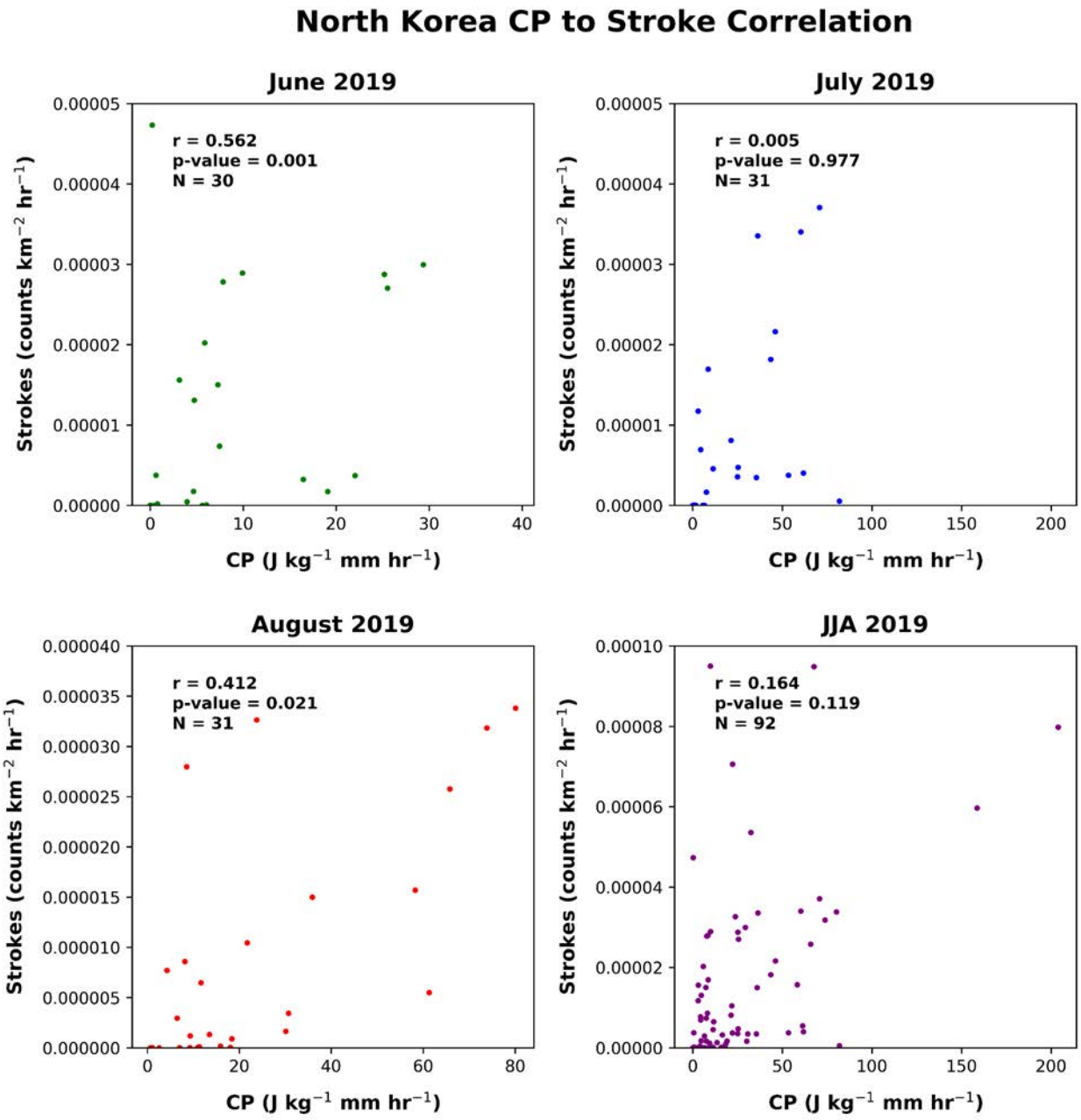

Figure 18. North Korea CP to Stroke Correlation: June 2019 (top left), July 2019 (top right), August 2019 (bottom left), JJA 2019 (bottom right).

Limitations of the Pearson Correlation

The nature of the Pearson correlation is that it is insensitive to magnitude and only captures the relative changes between quantities over time. Thus, in this study, the correlation was likely high because the CP was performing well at predicting lightning strokes in the sense of "occurred vs not-occurred" instances, but it was not always accurately capturing the relative stroke changes over time when lightning was occurring. For example, on days when lightning strokes were zero, the CP was correctly predicting values of zero. For days when lightning was actually occurring, the CP correctly predicted positive values, but these values did not always increase and/or decrease at the same rate as actual strokes. However, as stated before in the methods section, the Pearson correlation was specifically chosen since it is insensitive to magnitude. The intent of using the Pearson correlation was to find a relationship in changes between two quantities over time. In other words, the hope was that increases in CP would correspond to increases in stroke rate, and vice versa.

Calculating Constants of Proportionality (k) and Correction Factors (cf)

Linear regression was performed on the correlations between observed stroke counts and the CP proxy for June, July, August, and JJA for each location in order to find the slope. This slope was then used as constants of proportionality for CP in order to convert CP into units of strokes so that the CP proxy may now be used directly as a forecasting tool and compared against actual observed lightning. The slopes obtained for each location and time period can be seen in Table 1 below. The constants of proportionality have units of $J^{-1}kg\ mm^{-1}km^{-2}$ and were multiplied with the CP proxy arrays in order to convert the proxy from units $J\ kg^{-1}mm\ hr^{-1}$ to strokes

($counts\ km^{-2}hr^{-1}$). The resulting CP proxy arrays were then plotted as forecasts and compared against observed lightning densities in the next section.

Table 1. Constants of Proportionality (*k*)

Location	Time Period	k	Units
Arizona	June 2019	9.163×10^{-6}	$J^{-1}kg\ mm^{-1}km^{-2}$
Arizona	July 2019	1.078×10^{-6}	$J^{-1}kg\ mm^{-1}km^{-2}$
Arizona	August 2019	2.138×10^{-6}	$J^{-1}kg\ mm^{-1}km^{-2}$
Arizona	JJA 2019	1.411×10^{-6}	$J^{-1}kg\ mm^{-1}km^{-2}$
Cuba	June 2019	9.493×10^{-7}	$J^{-1}kg\ mm^{-1}km^{-2}$
Cuba	July 2019	1.258×10^{-6}	$J^{-1}kg\ mm^{-1}km^{-2}$
Cuba	August 2019	2.737×10^{-7}	$J^{-1}kg\ mm^{-1}km^{-2}$
Cuba	JJA 2019	8.430×10^{-7}	$J^{-1}kg\ mm^{-1}km^{-2}$
North Korea	June 2019	4.389×10^{-6}	$J^{-1}kg\ mm^{-1}km^{-2}$
North Korea	July 2019	1.528×10^{-8}	$J^{-1}kg\ mm^{-1}km^{-2}$
North Korea	August 2019	1.489×10^{-6}	$J^{-1}kg\ mm^{-1}km^{-2}$
North Korea	JJA 2019	5.176×10^{-7}	$J^{-1}kg\ mm^{-1}km^{-2}$

As mentioned in 3.6, a calculation of *cf* was needed in order to correct the magnitude of the CP stroke forecast to more closely match that of the observed strokes. *cf* was found by dividing the maximum number of strokes $km^{-2}\ hr^{-1}$ by the maximum number of CP strokes $km^{2}\ hr^{-1}$ for each grid box and can be seen in Table 2 below.

Table 2. Correction Factors (*cf*)

Location	Time Period	cf	Units
Arizona	June 2019	3.522	Dimensionless
Arizona	July 2019	114.371	Dimensionless
Arizona	August 2019	70.273	Dimensionless
Arizona	JJA 2019	87.128	Dimensionless
Cuba	June 2019	32.819	Dimensionless
Cuba	July 2019	60.135	Dimensionless
Cuba	August 2019	265.757	Dimensionless
Cuba	JJA 2019	53.621	Dimensionless
North Korea	June 2019	14.387	Dimensionless
North Korea	July 2019	7211.689	Dimensionless
North Korea	August 2019	61.782	Dimensionless
North Korea	JJA 2019	126.137	Dimensionless

4.3 CP Forecasts and Performance

Arizona

Comparing the CP proxy to observed GLD360 lightning densities (Figures 19, 20, 21 below) for June 2019 over Arizona, the proxy captured the general outline of frequent thunderstorm areas near the town of San Carlos, which sits in a valley between the Graham mountains to the south and the rocky San Carlos reservation to the north. It overforecasted a high-density area of strokes to the northeast of Tucson over the Santa

Catalina mountains. The proxy also captured well the area of strokes to the west, south, northeast, and east of Flagstaff, but missed the area of strokes to the east of Bagdad and northeast of Seligman.

For July 2019 (Figures 22, 23, 24 below), the proxy performed well at capturing the large area of strokes in the open desert area between Ajo and Tucson, the stroke area northeast of Tucson in the Santa Catalina mountains, and the general stroke areas north of San Carlos. It overforecasted an area of strokes to the northwest of Flagstaff near Bagdad.

For August 2019 (Figures 25, 26, 27 below), the proxy did well at capturing the area of strokes east of San Carlos and east of Tucson, but grossly underforecasted a large area of strokes in the open desert between Ajo and Tucson. The proxy also overforecasted a large of area of strokes to the northwest of Tucson in the open desert area that is bordered on its northeast side by mountains. To the north, the proxy decently captured the area of strokes near Flagstaff, but completely missed an area of strokes around Seligman.

For JJA 2019 (Figures 28, 29, 30 below), the proxy did a fair job at capturing the spatial extent of strokes in the open desert area between Ajo and Tucson, although it underforecasted the northern extension. It also captured well the area of strokes to the northeast of Tucson in the Santa Catalina mountains, albeit shifted the strokes slightly too far to the west. It captured well the general area of strokes to the north of San Carlos, but misplaced it slightly too far to the west. Finally, the proxy captured well the area of strokes to the northwest of Flagstaff, overforecasted an area of strokes west of Bagdad, and missed an area of strokes near Seligman.

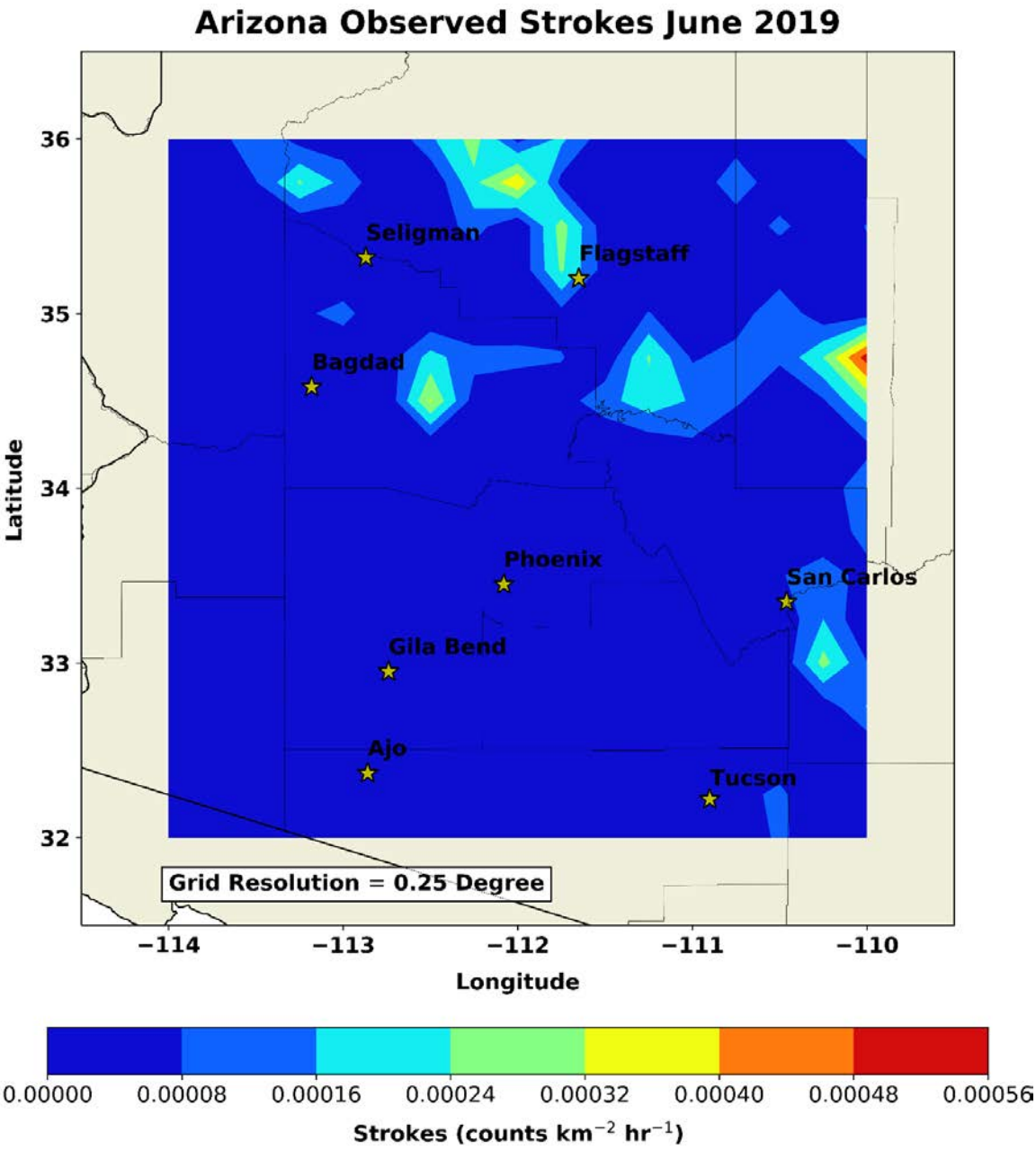

Figure 19. Arizona Observed Strokes June 2019.

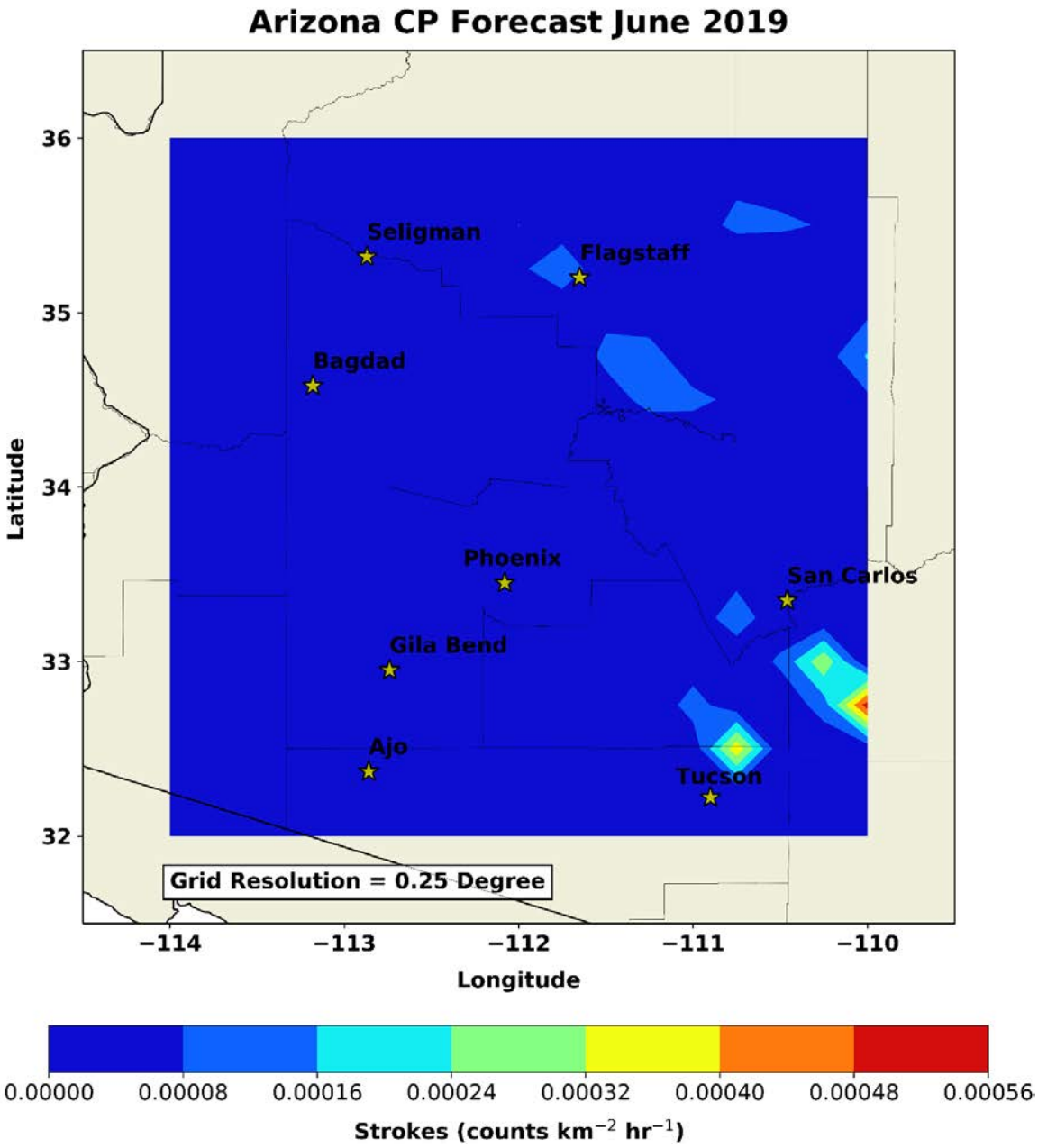

Figure 20. Arizona CP Forecast June 2019.

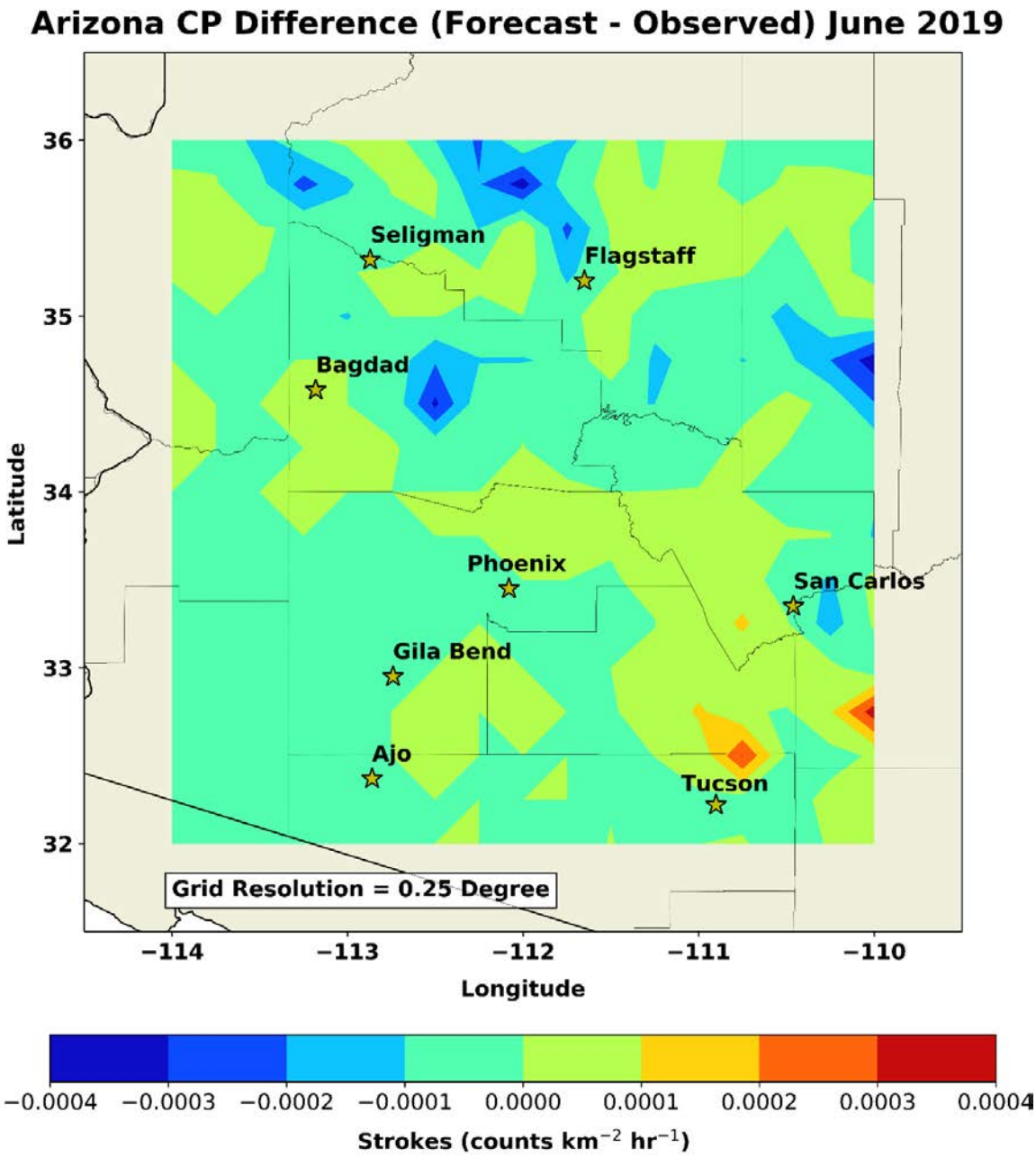

Figure 21. Arizona CP Difference (Forecast – Observed) June 2019.

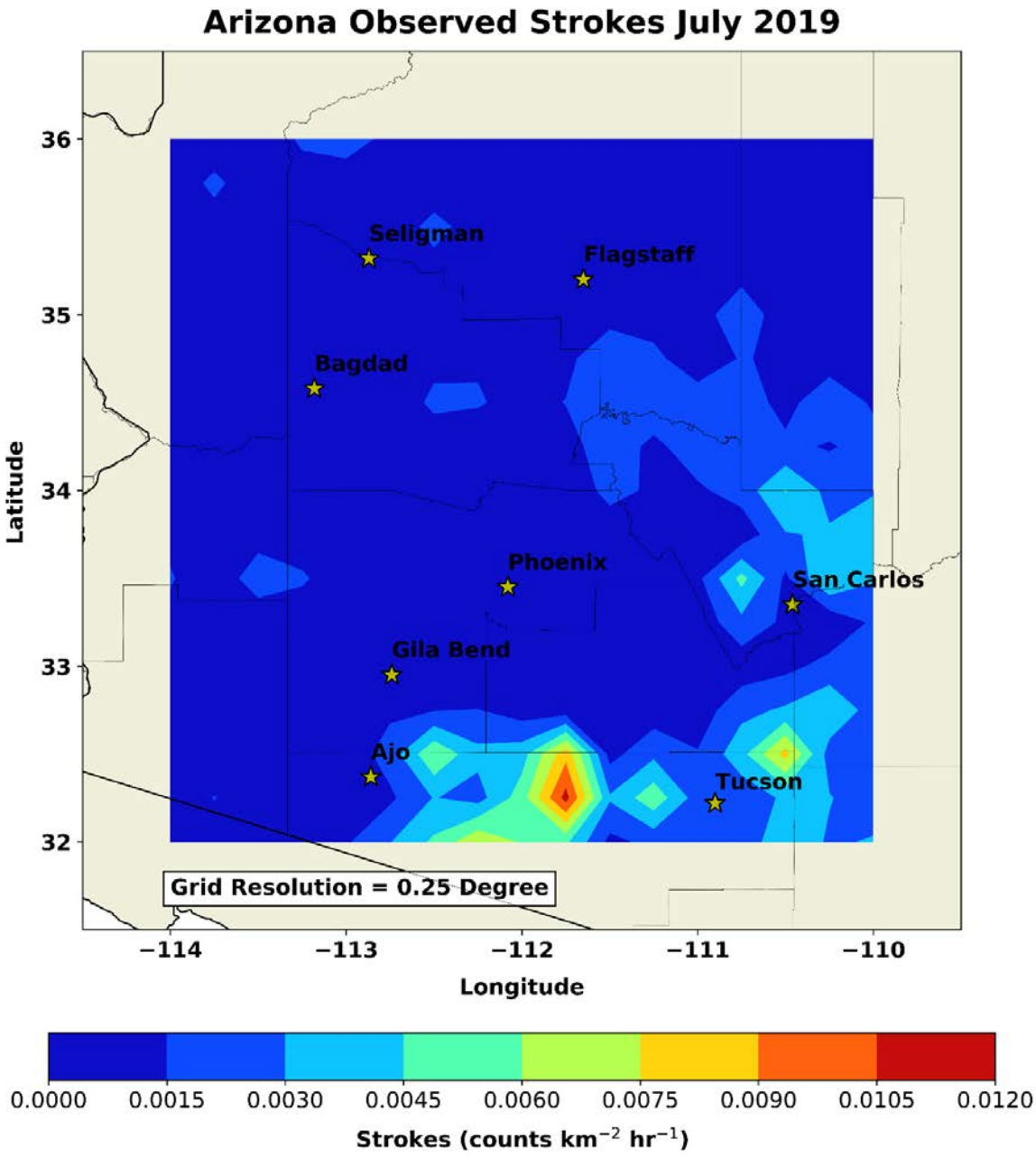

Figure 22. Arizona Observed Strokes July 2019.

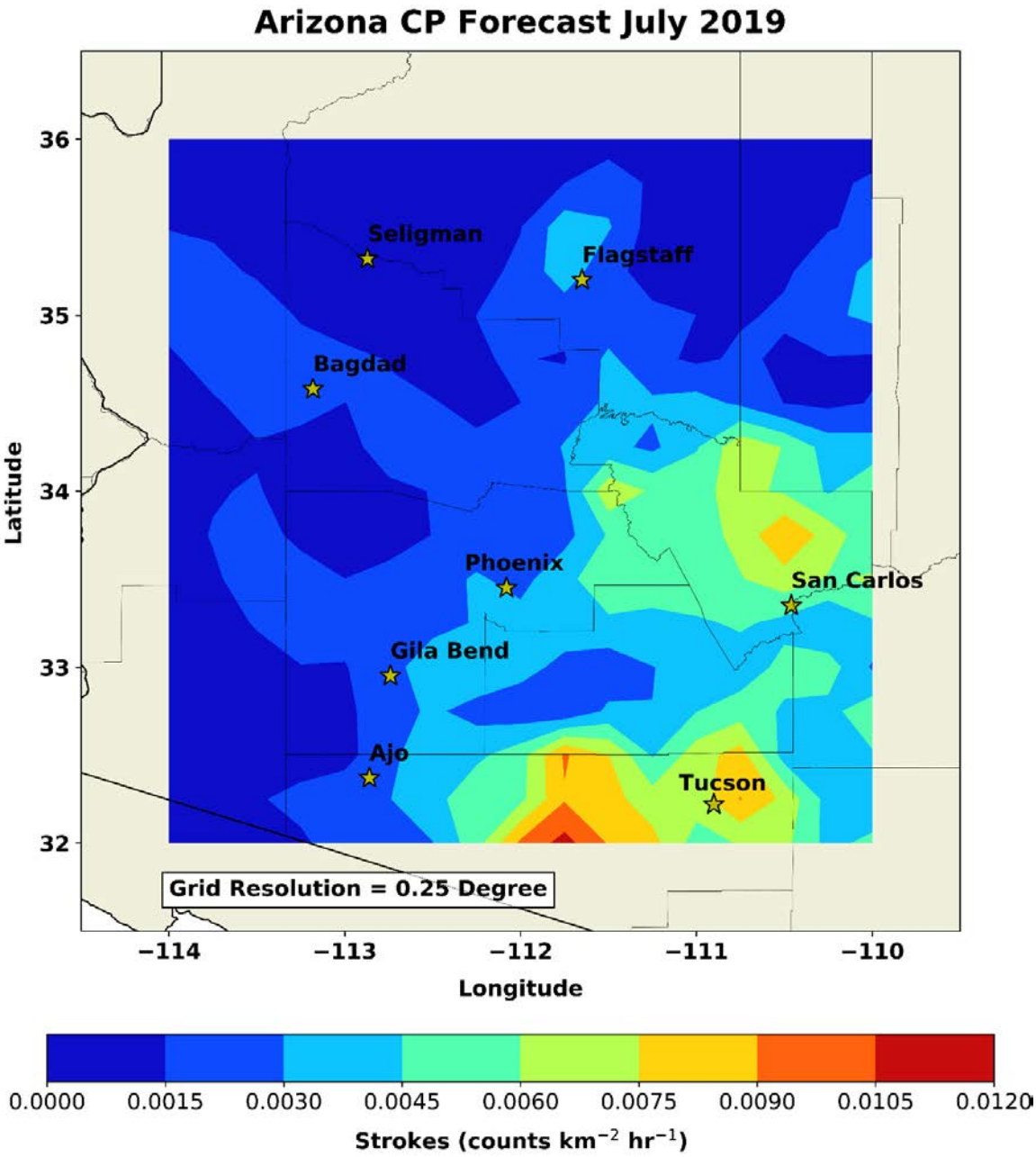

Figure 23. Arizona CP Forecast July 2019.

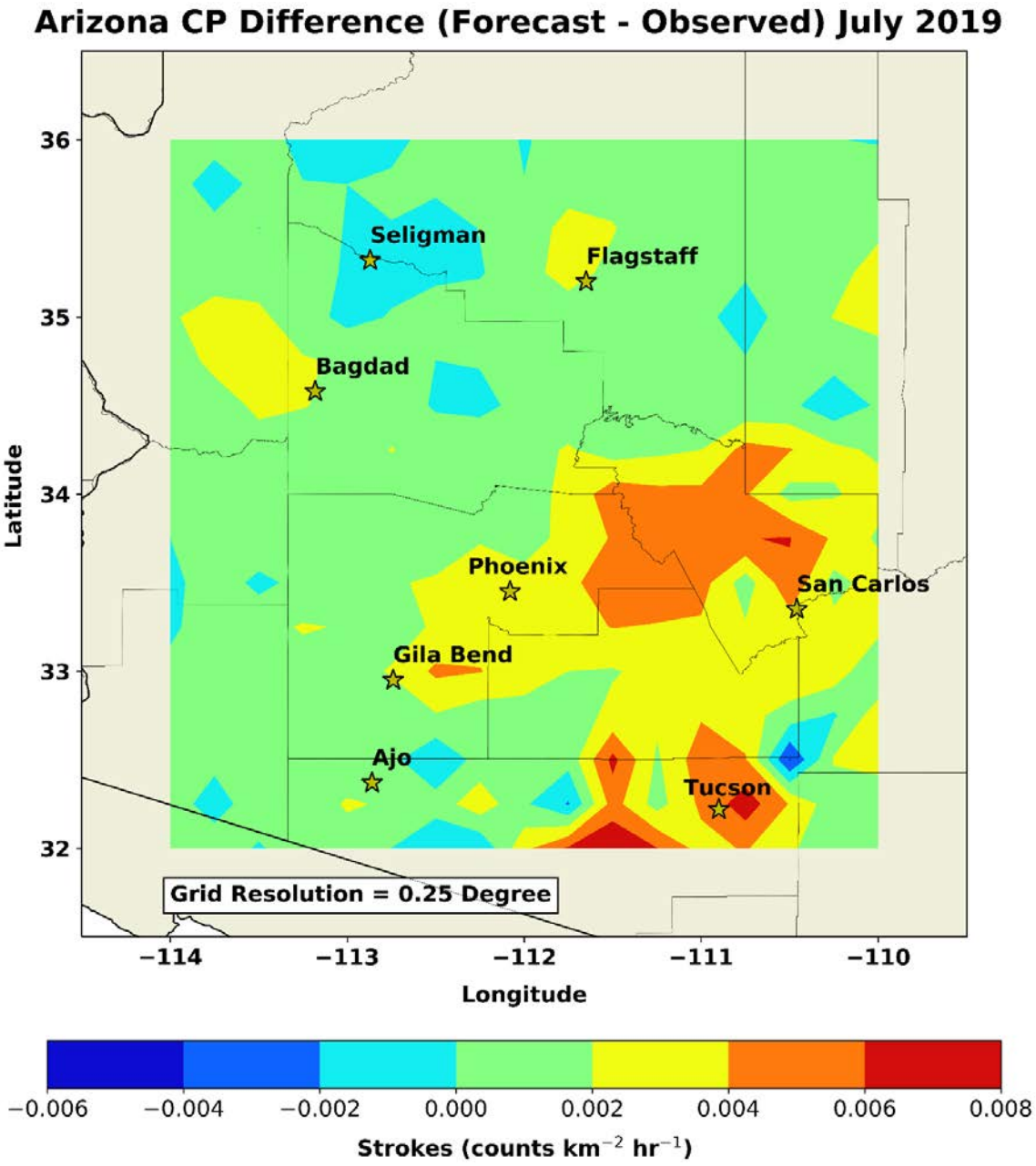

Figure 24. Arizona CP Difference (Forecast – Observed) July 2019.

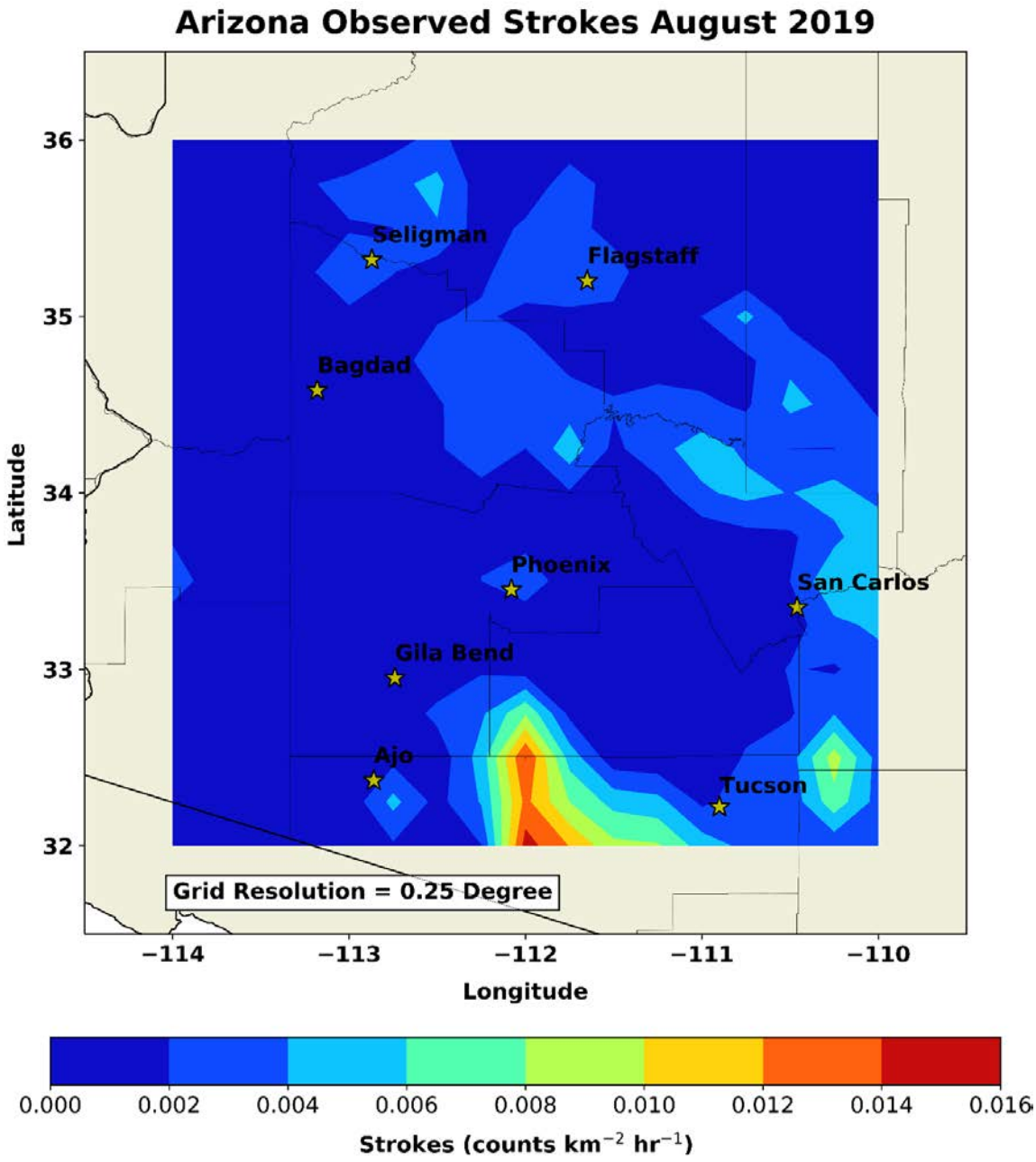

Figure 25. Arizona Observed Strokes August 2019.

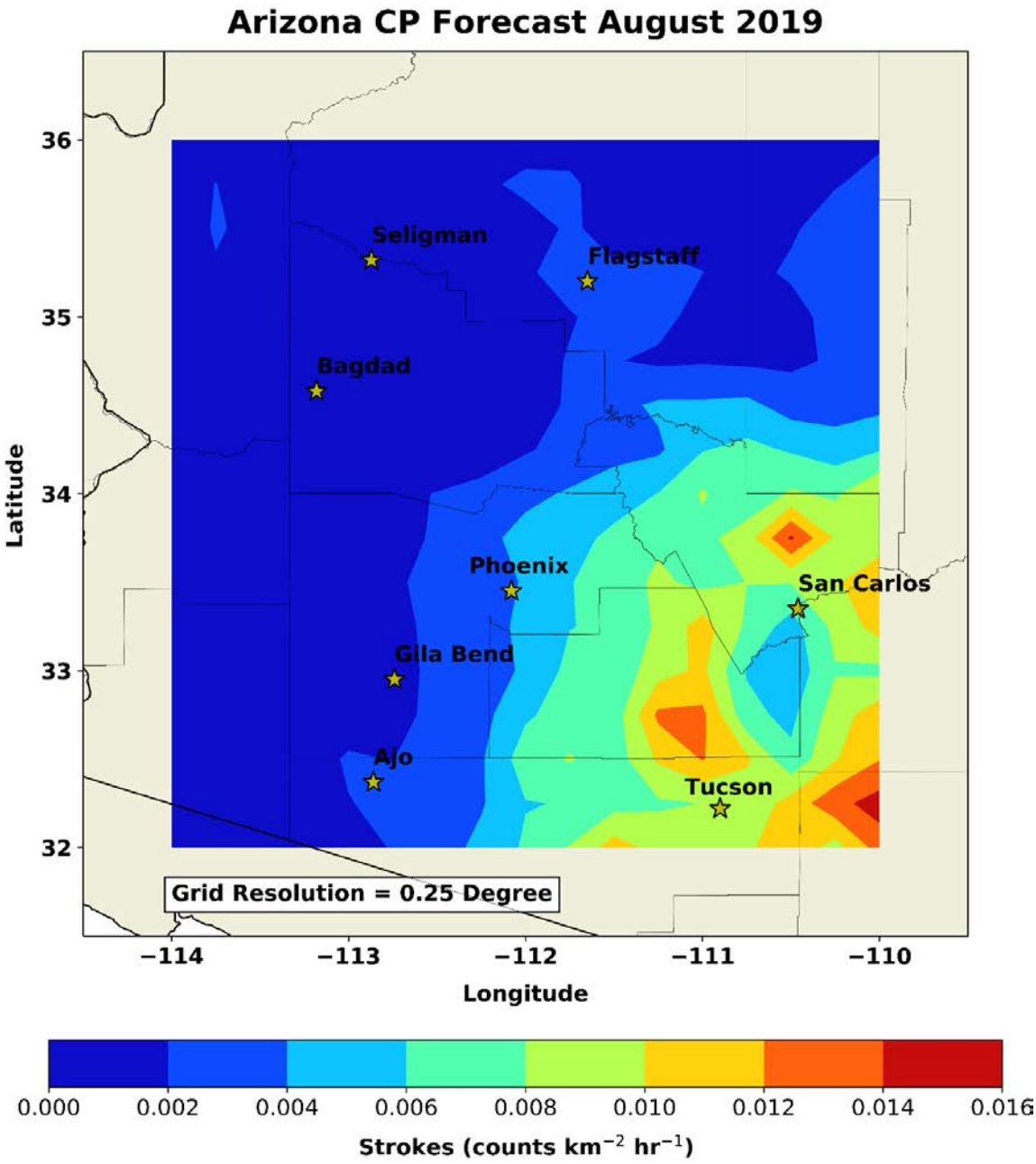

Figure 26. Arizona CP Forecast August 2019.

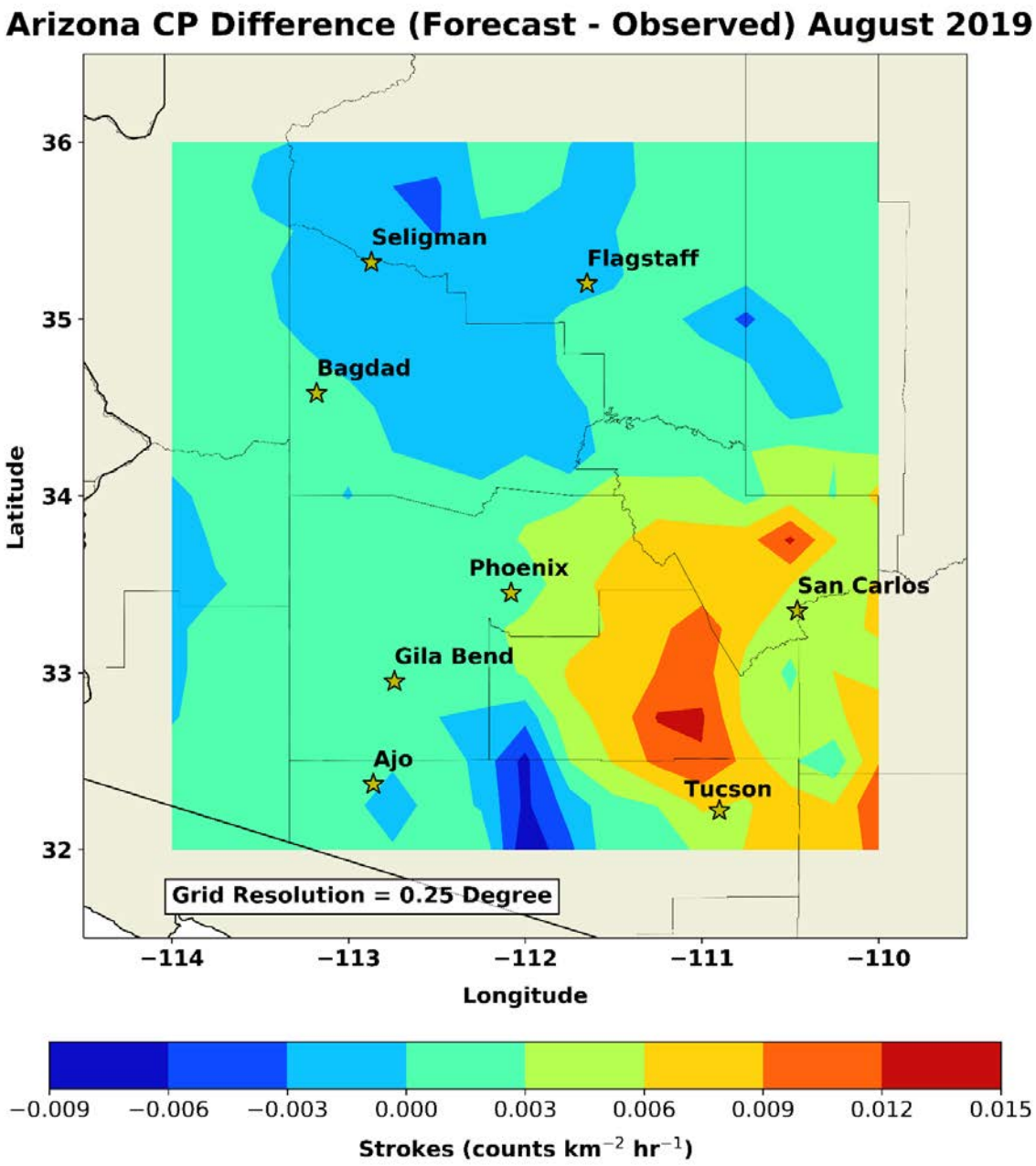

Figure 27. Arizona CP Difference (Forecast – Observed) August 2019.

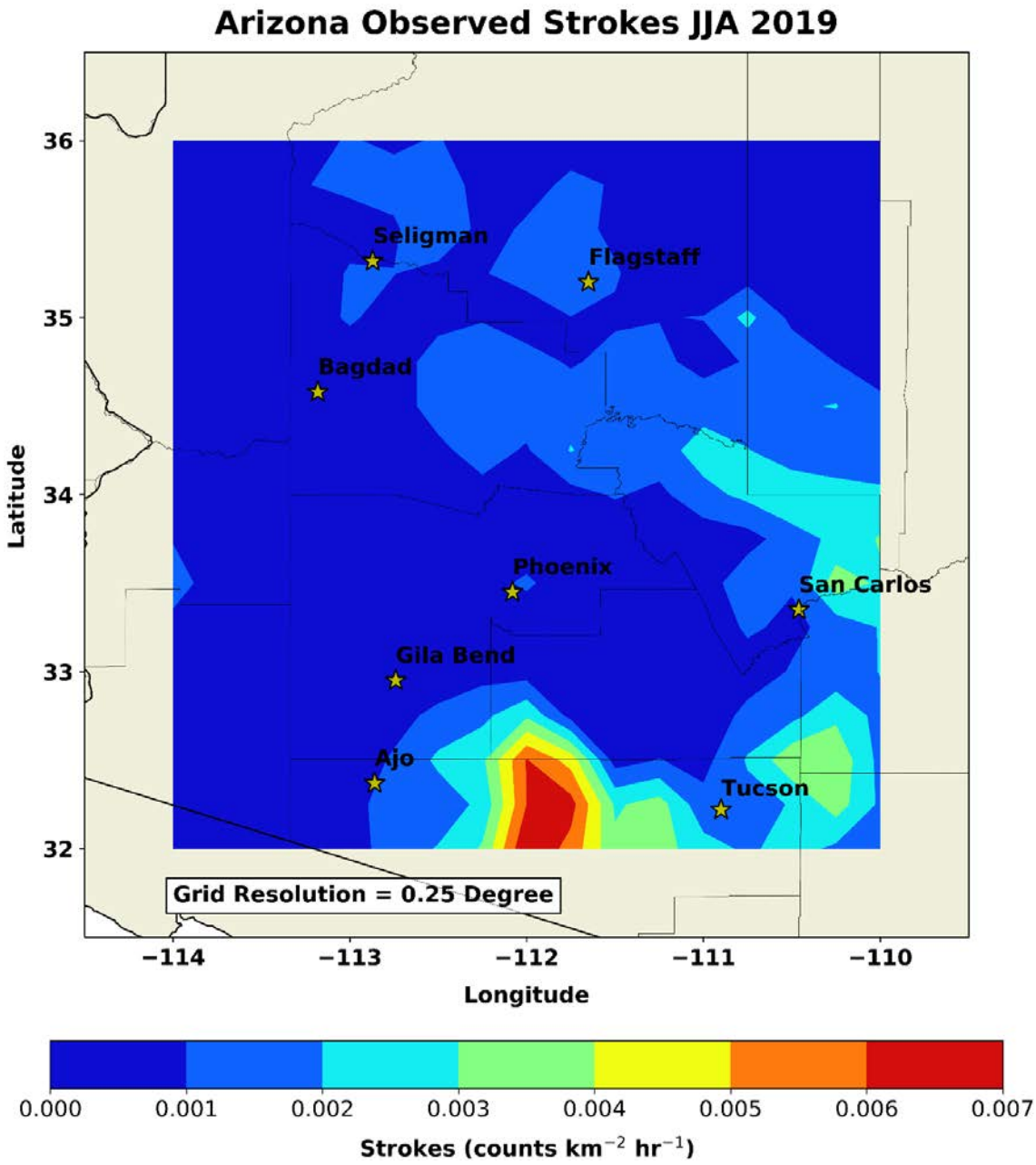

Figure 28. Arizona Observed Strokes JJA 2019.

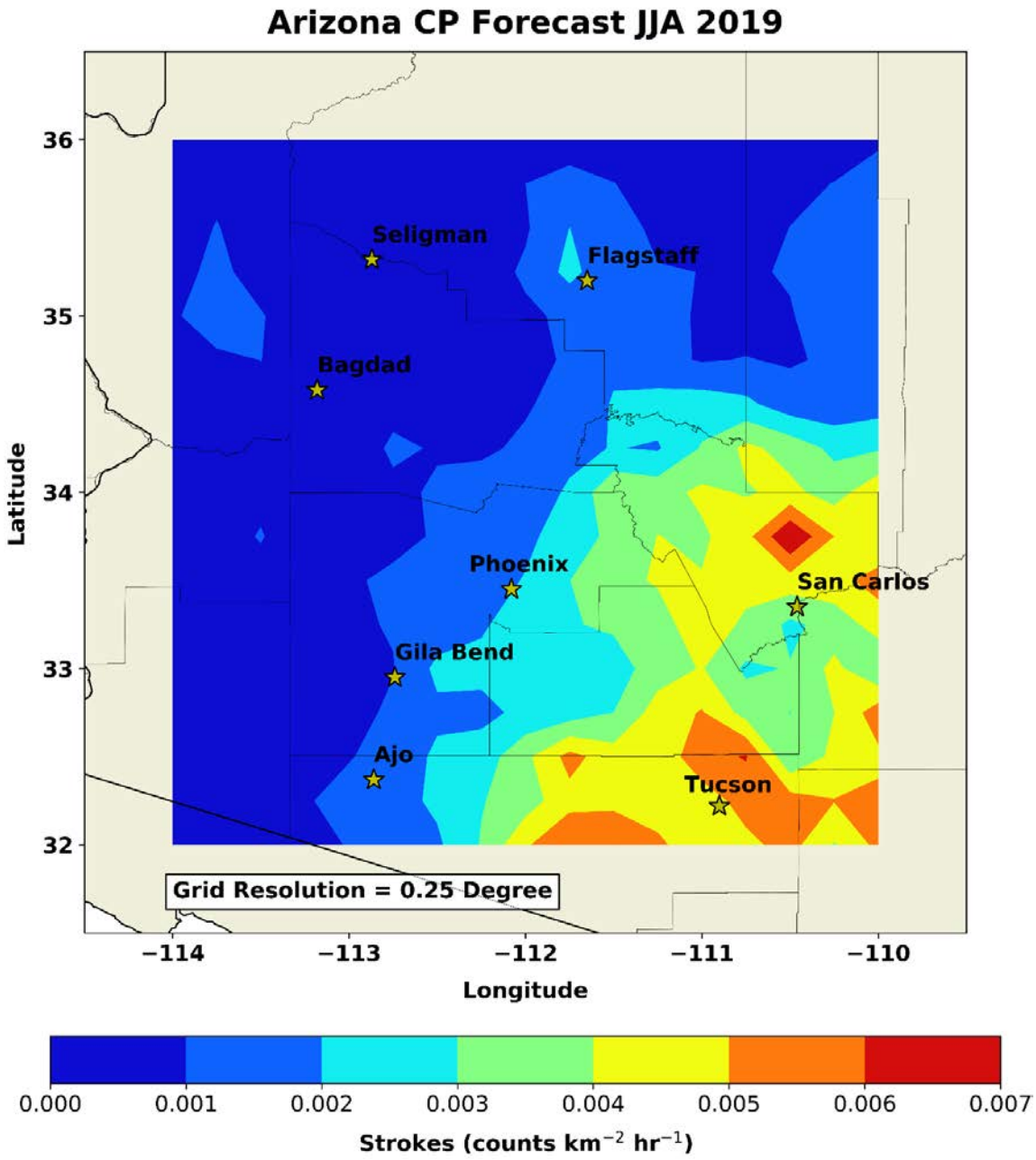

Figure 29. Arizona CP Forecast JJA 2019.

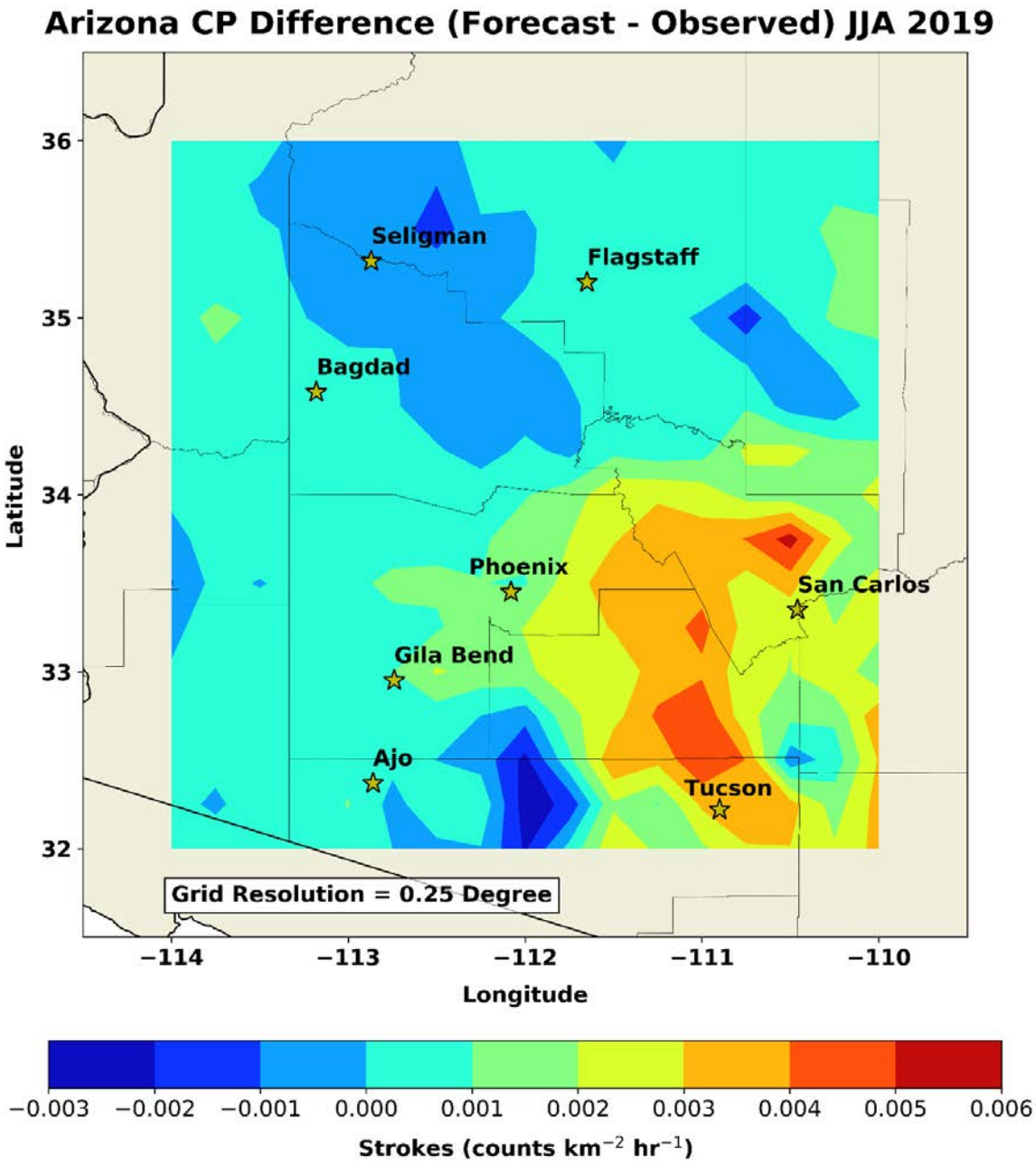

Figure 30. Arizona CP Difference (Forecast – Observed) JJA 2019.

Cuba

Over Cuba in June 2019 (Figures 31, 32, 33 below), the proxy performed well overall. It decently captured the area of strokes near Camaguey, but it overforecasted an area of strokes just north of Sancti Spiritus. It also missed a significant area of strokes north of Cienfueges, but decently captured the area of strokes to the northeast, near the tip of the Bahamas. Additionally, the proxy correctly captured the spatial extent of most of the large no-stroke areas on the southwest and northeast sides of the island.

For July 2019 (Figures 34, 35, 36 below), the proxy performed poorly overall, and either missed or misplaced stroke areas throughout the continent. For example, it grossly overforecasted an area of strokes over Cienfuegos, over the shores south of Sancti Spiritus, and over the shores just north of Santa Clara. It did a decent job of capturing the general of strokes near the Bahamas to the northeast, as well as the elongated area of strokes just west of Manzanillo.

For August 2019 (Figures 37, 38, 39 below), the proxy overall performed poorly. It missed a large area of strokes over Camaguey, Las Tunas, as well as grossly overforecasted an area of strokes south of Cienfueges. The only area that was captured decently well was the area of strokes to the northeast near the Bahamas.

For JJA 2019 (Figures 40, 41, 42 below), the proxy overall captured the spatial extent of strokes over the continent and surrounding waters fairly. It correctly captured the area to the northeast near the Bahamas, but overforecasted its concentration. It significantly underforecasted the extent and concentration of strokes around Camaguey and the area northwest of Cienfuegos, but overforecasted the number of strokes over Cienfuegos.

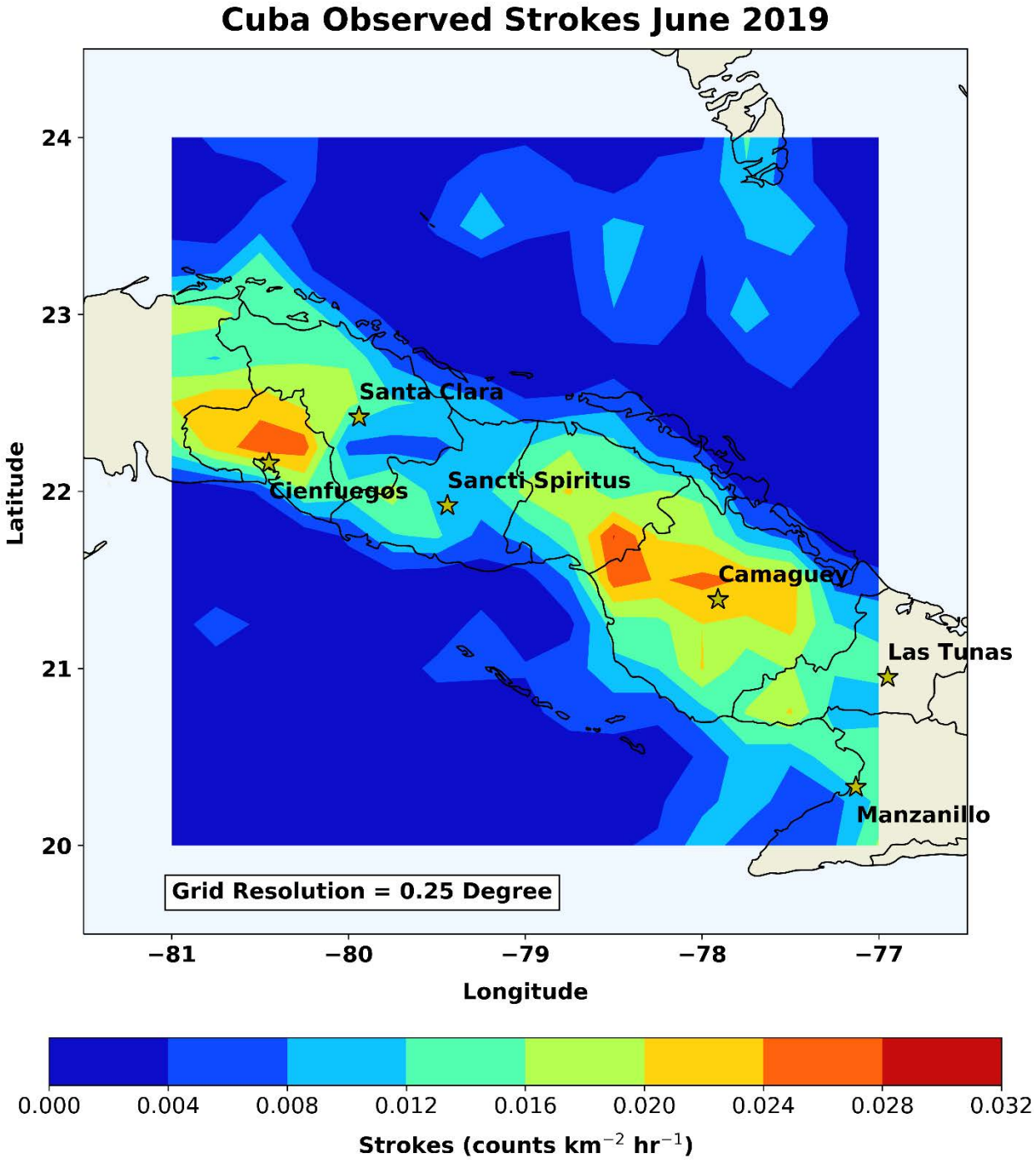

Figure 31. Cuba Observed Strokes June 2019.

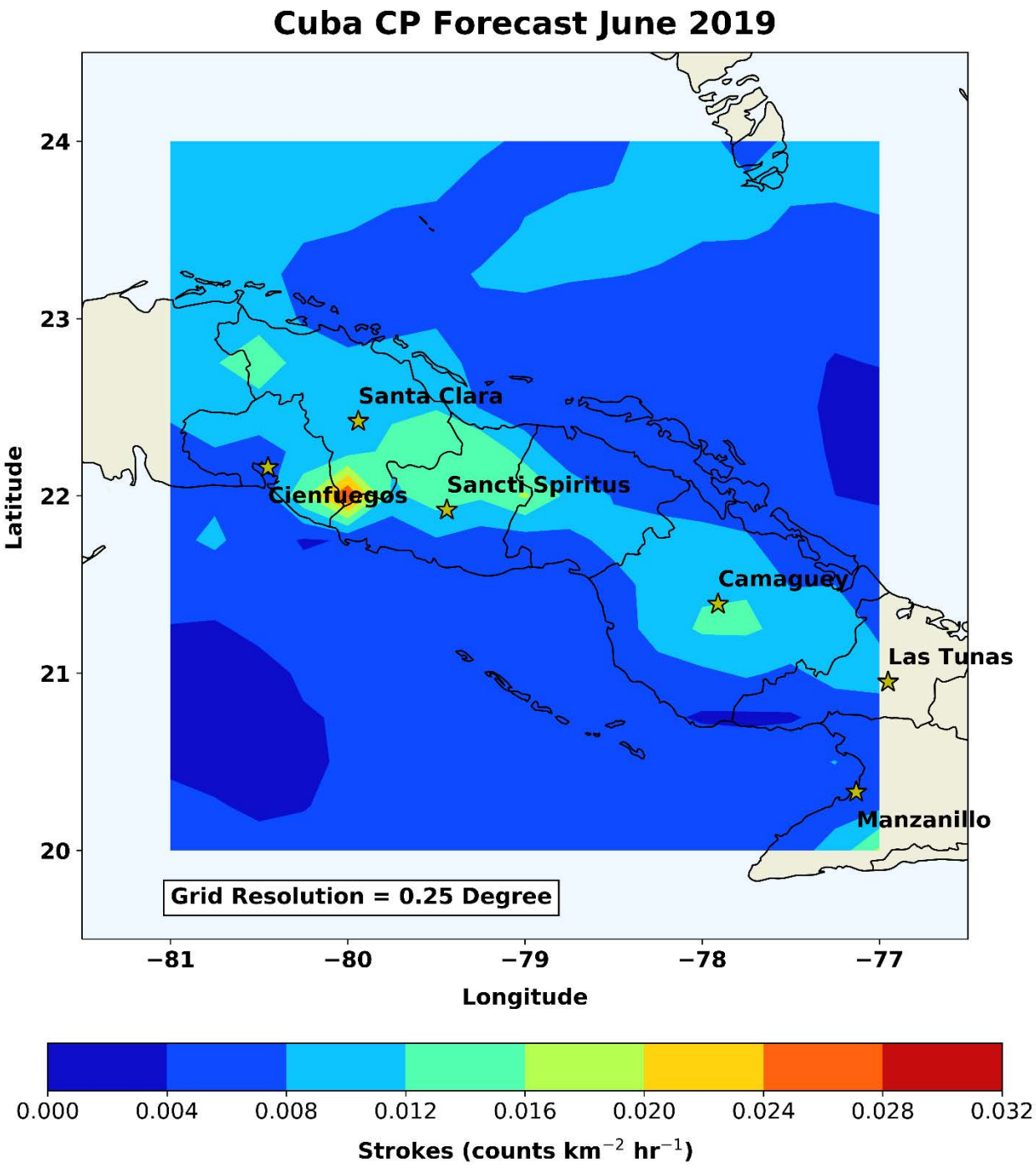

Figure 32. Cuba CP Forecast June 2019.

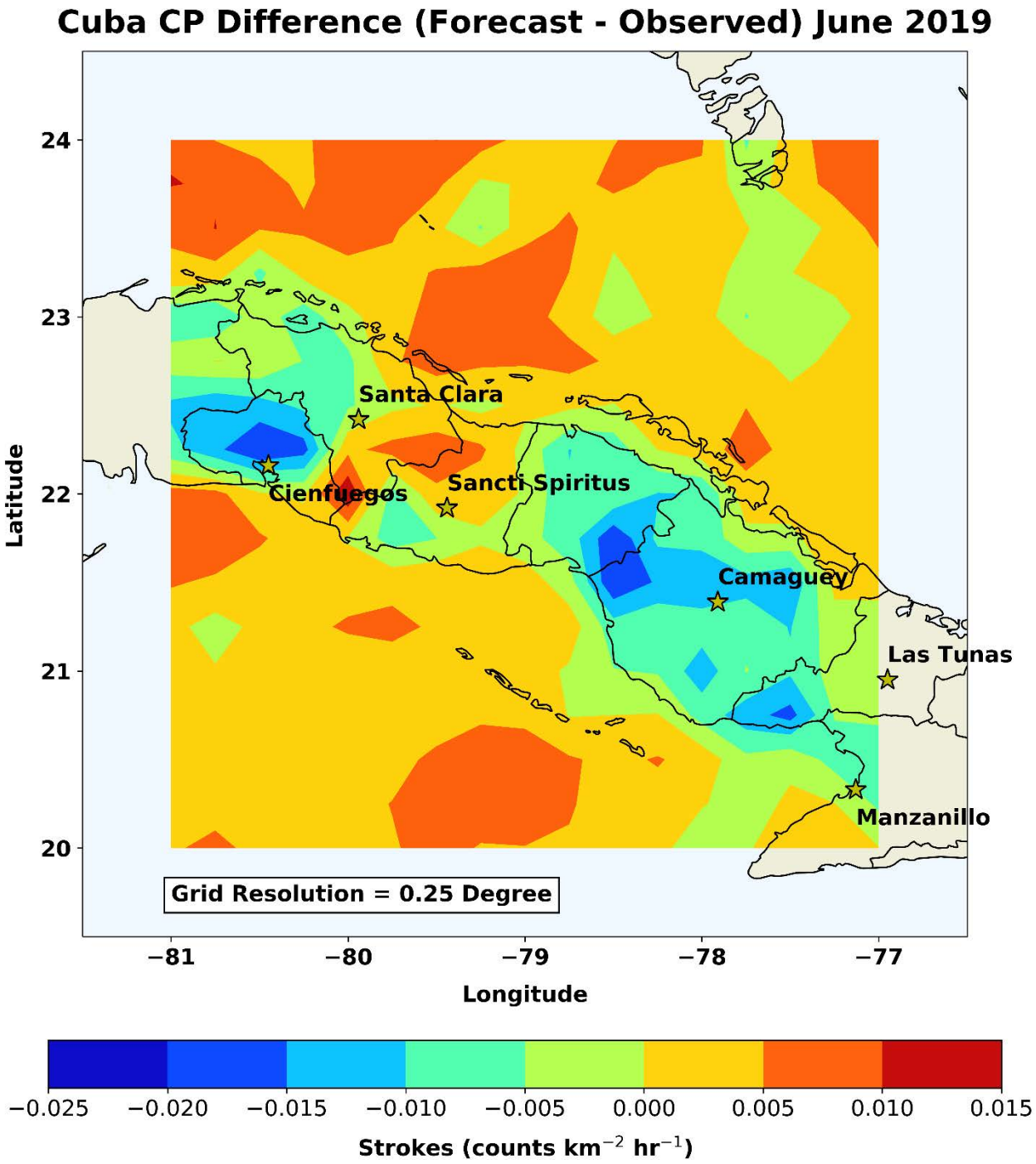

Figure 33. Cuba CP Difference (Forecast – Observed) June 2019.

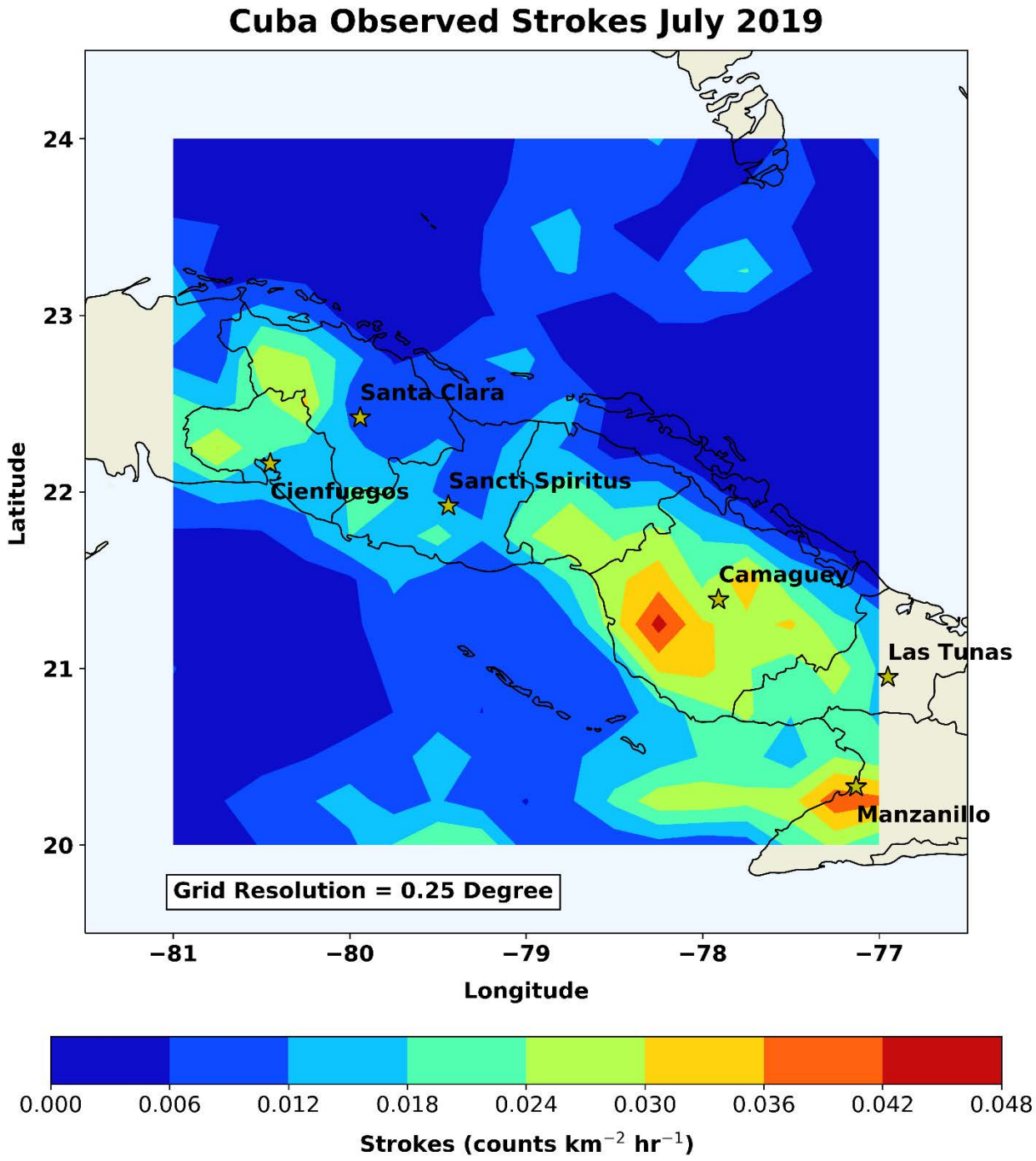

Figure 34. Cuba Observed Strokes July 2019.

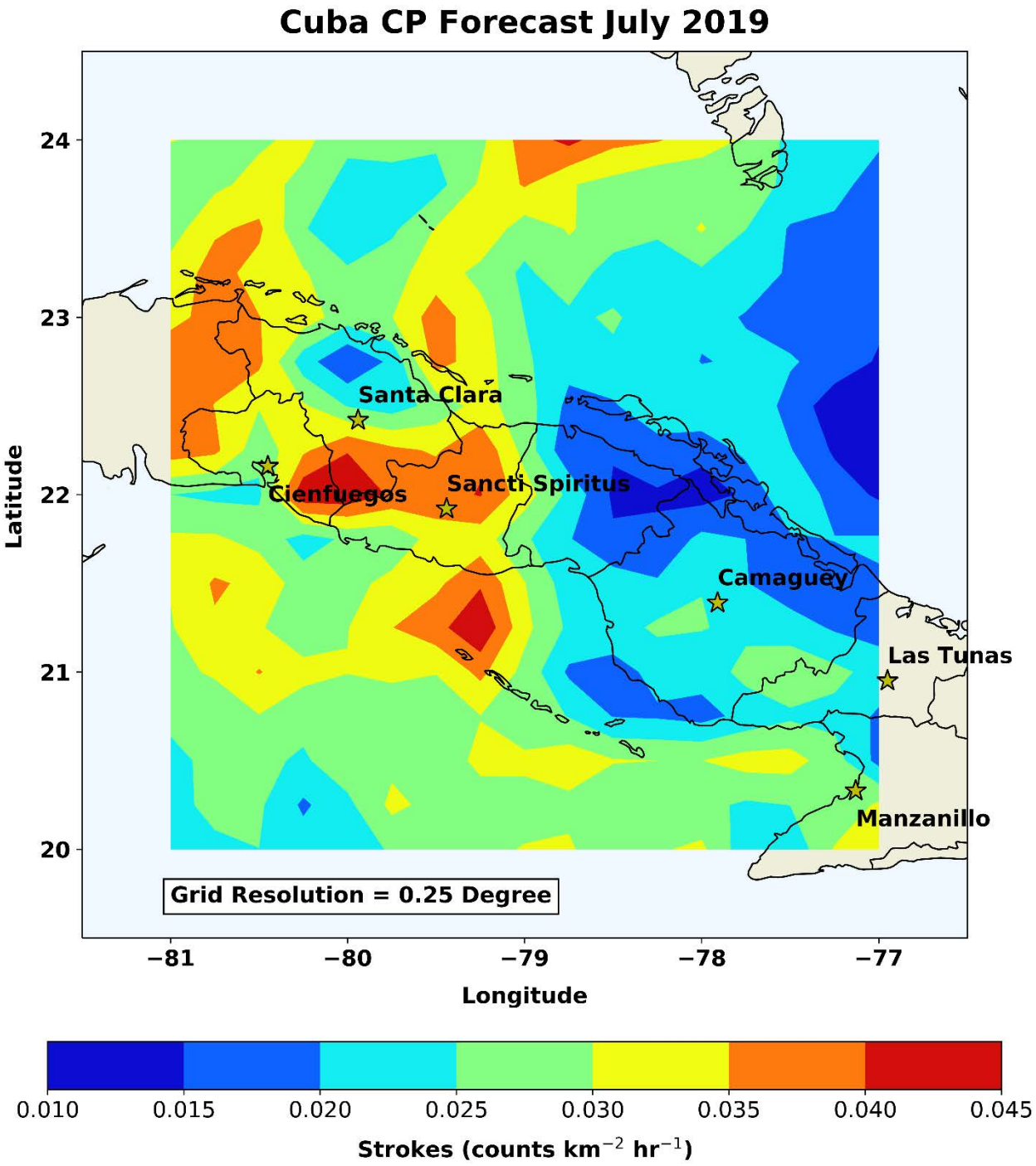

Figure 35. Cuba CP Forecast July 2019.

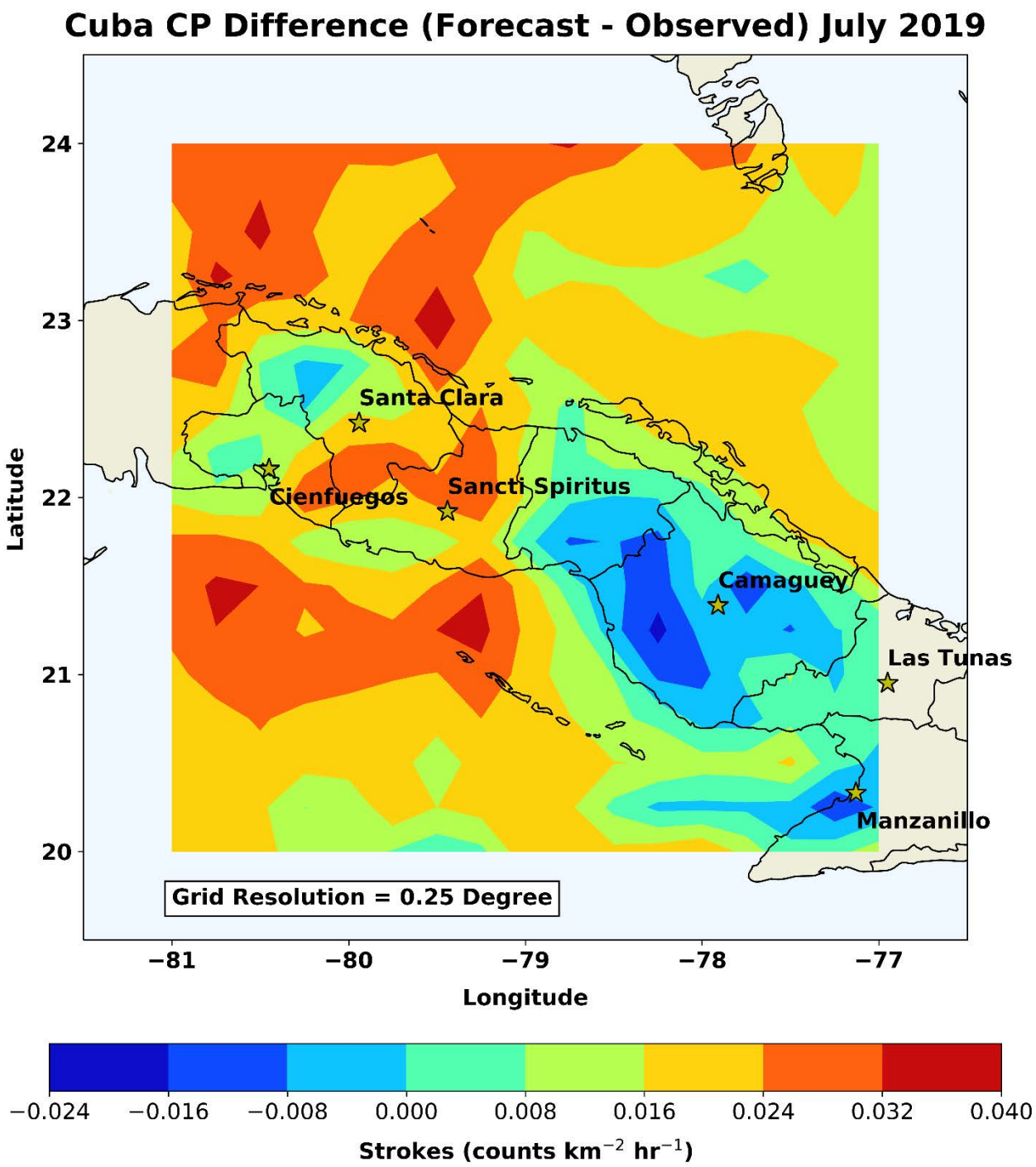

Figure 36. Cuba CP Difference (Forecast – Observed) July 2019.

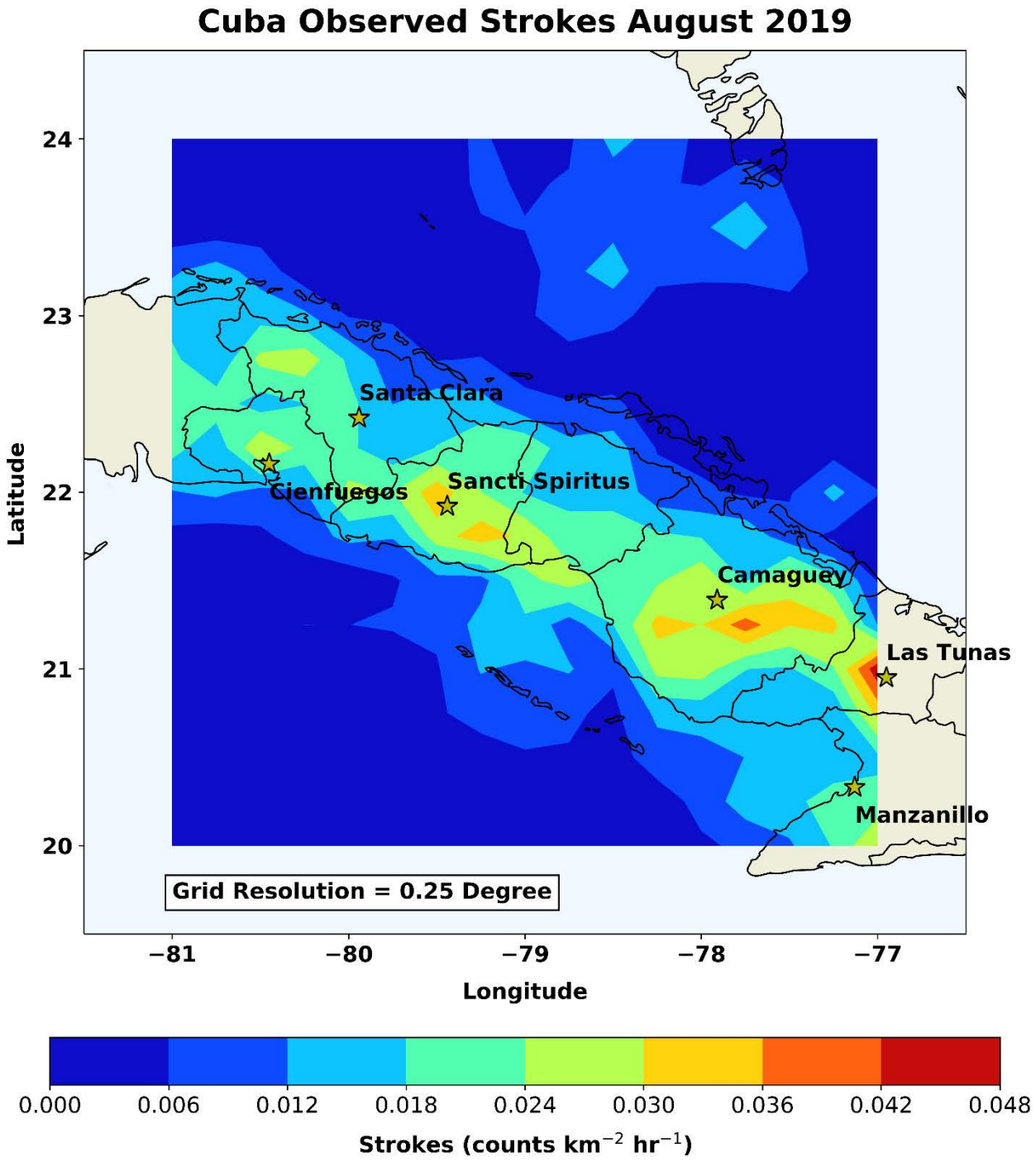

Figure 37. Cuba Observed Strokes August 2019.

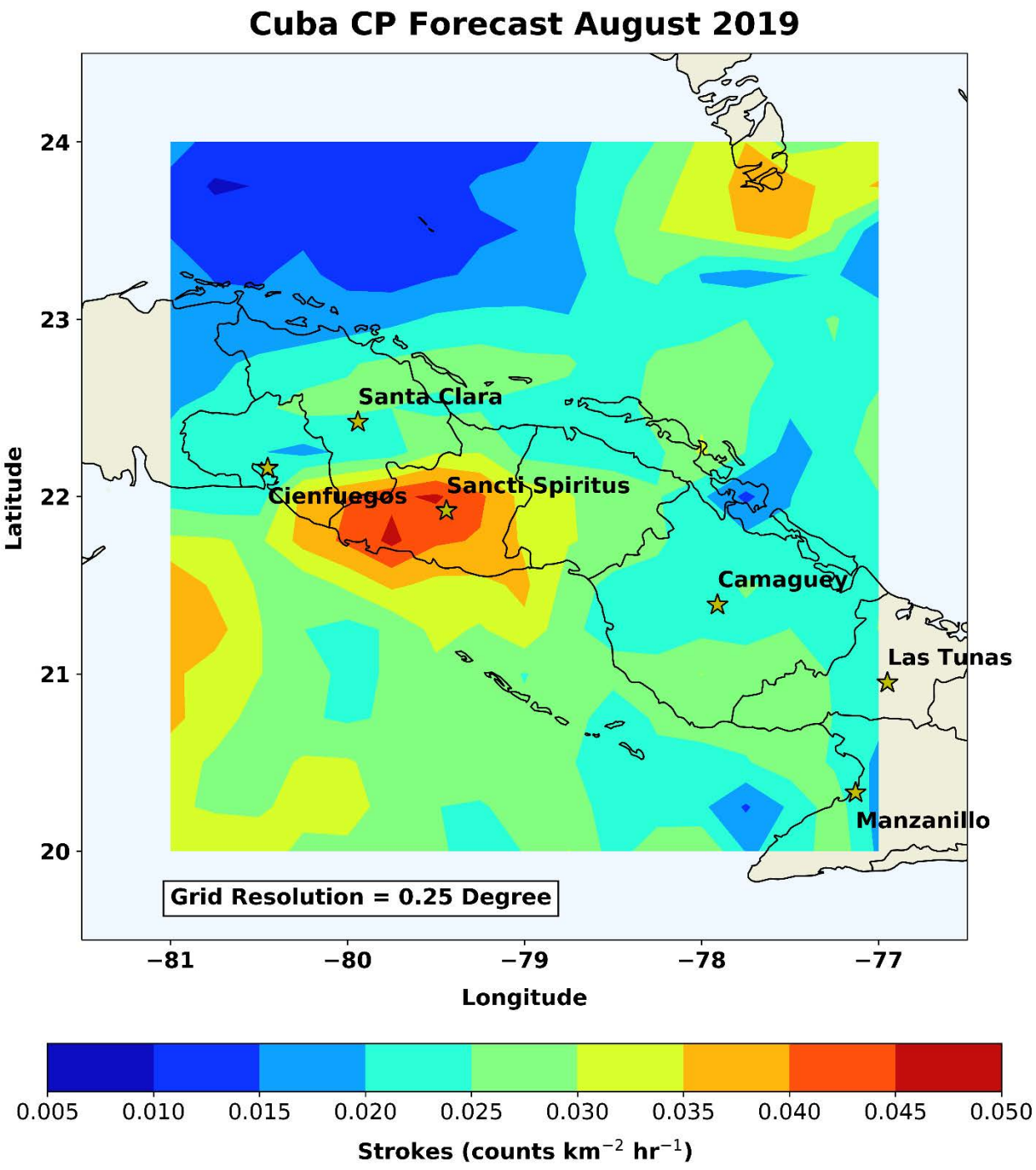

Figure 38. Cuba CP Forecast August 2019.

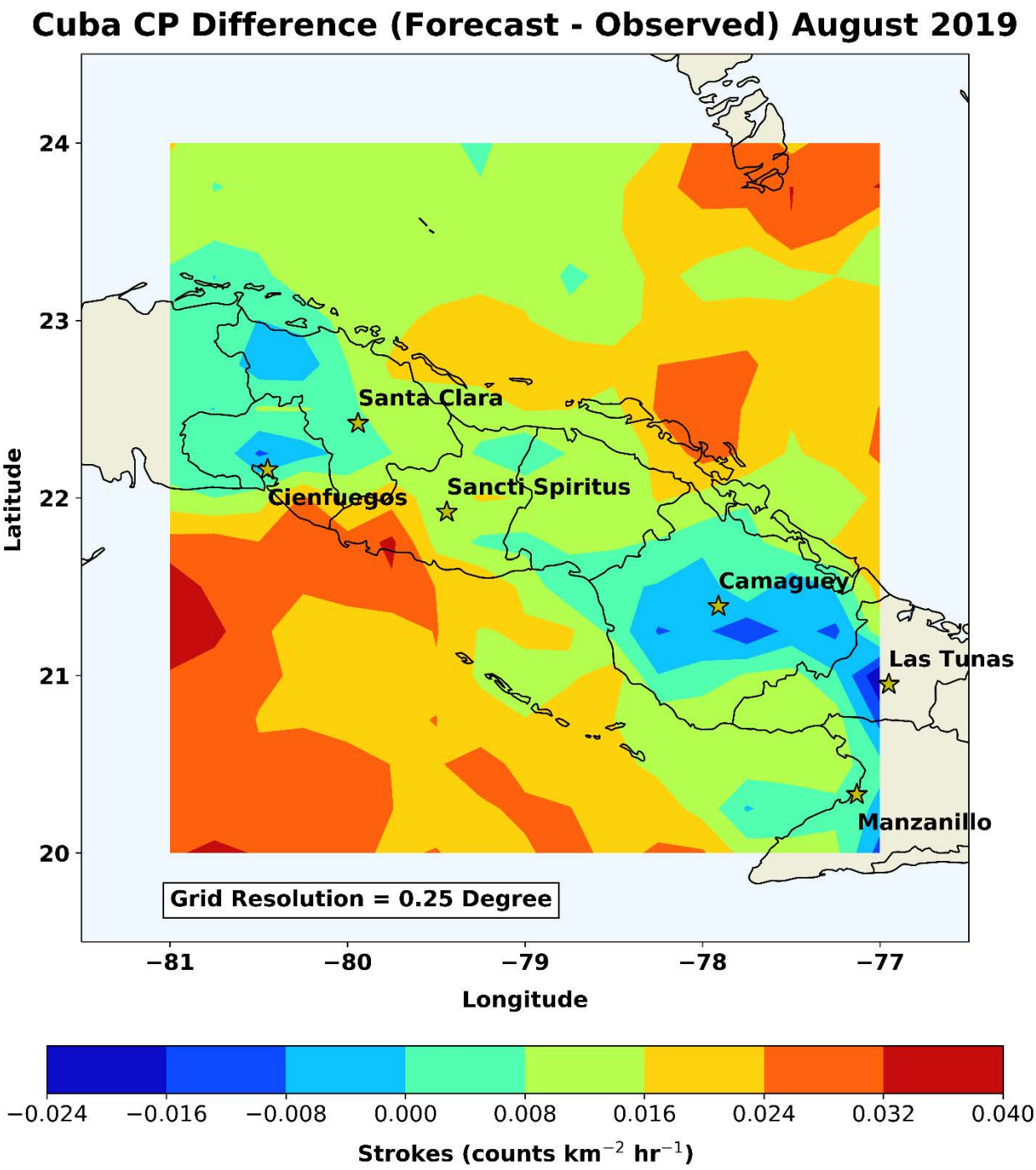

Figure 39. Cuba CP Difference (Forecast – Observed) August 2019.

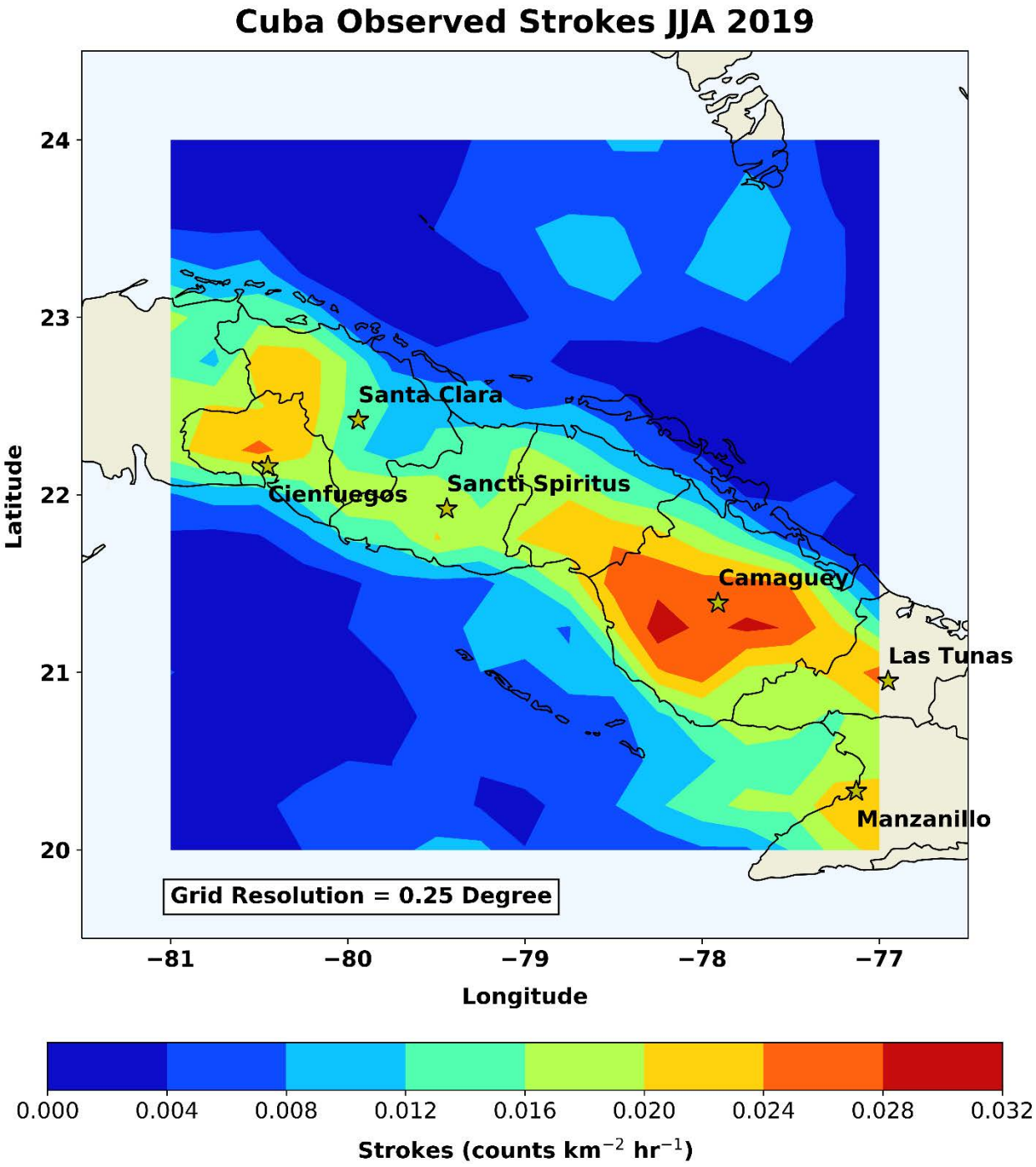

Figure 40. Cuba Observed Strokes JJA 2019.

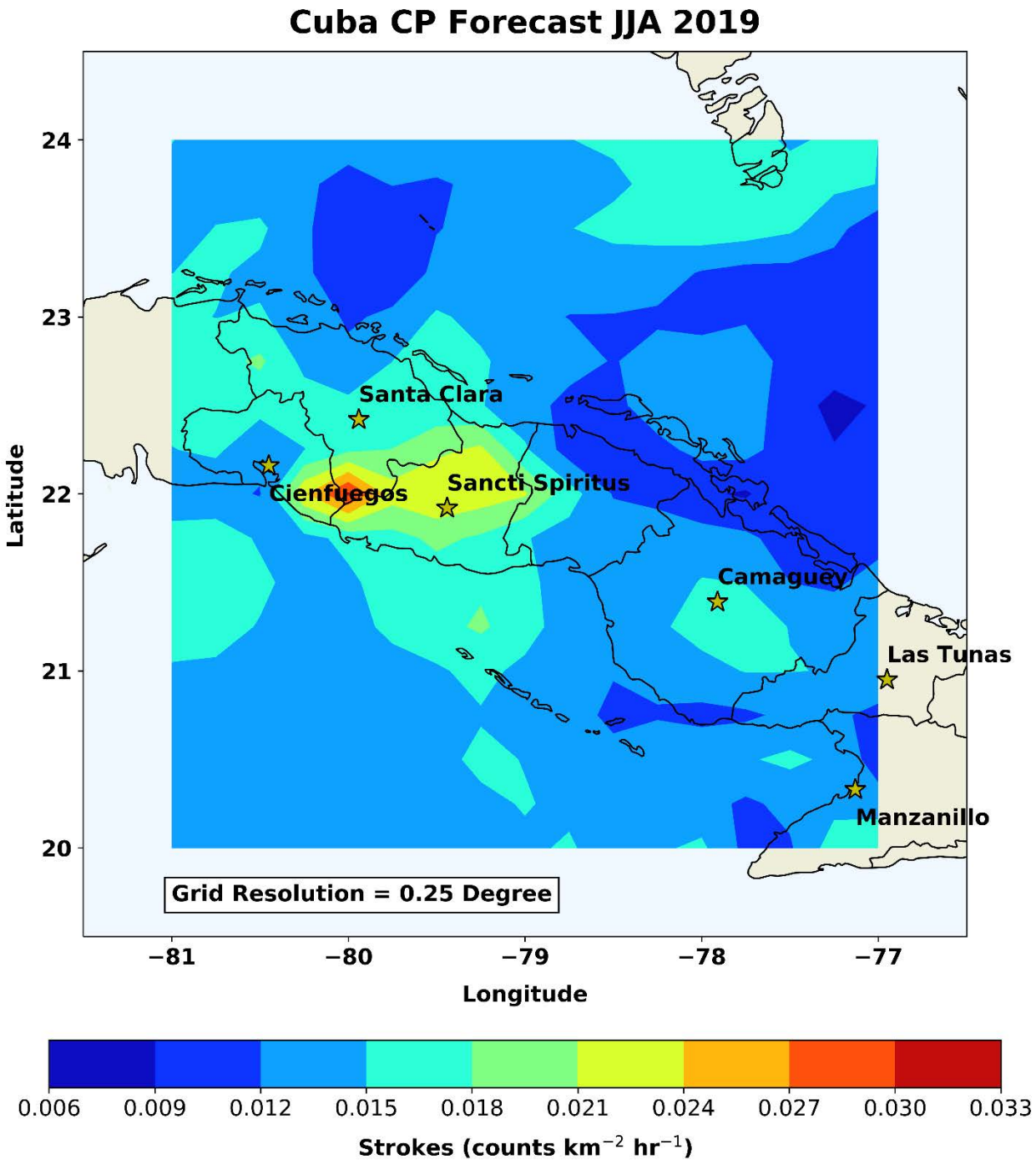

Figure 41. Cuba CP Forecast JJA 2019.

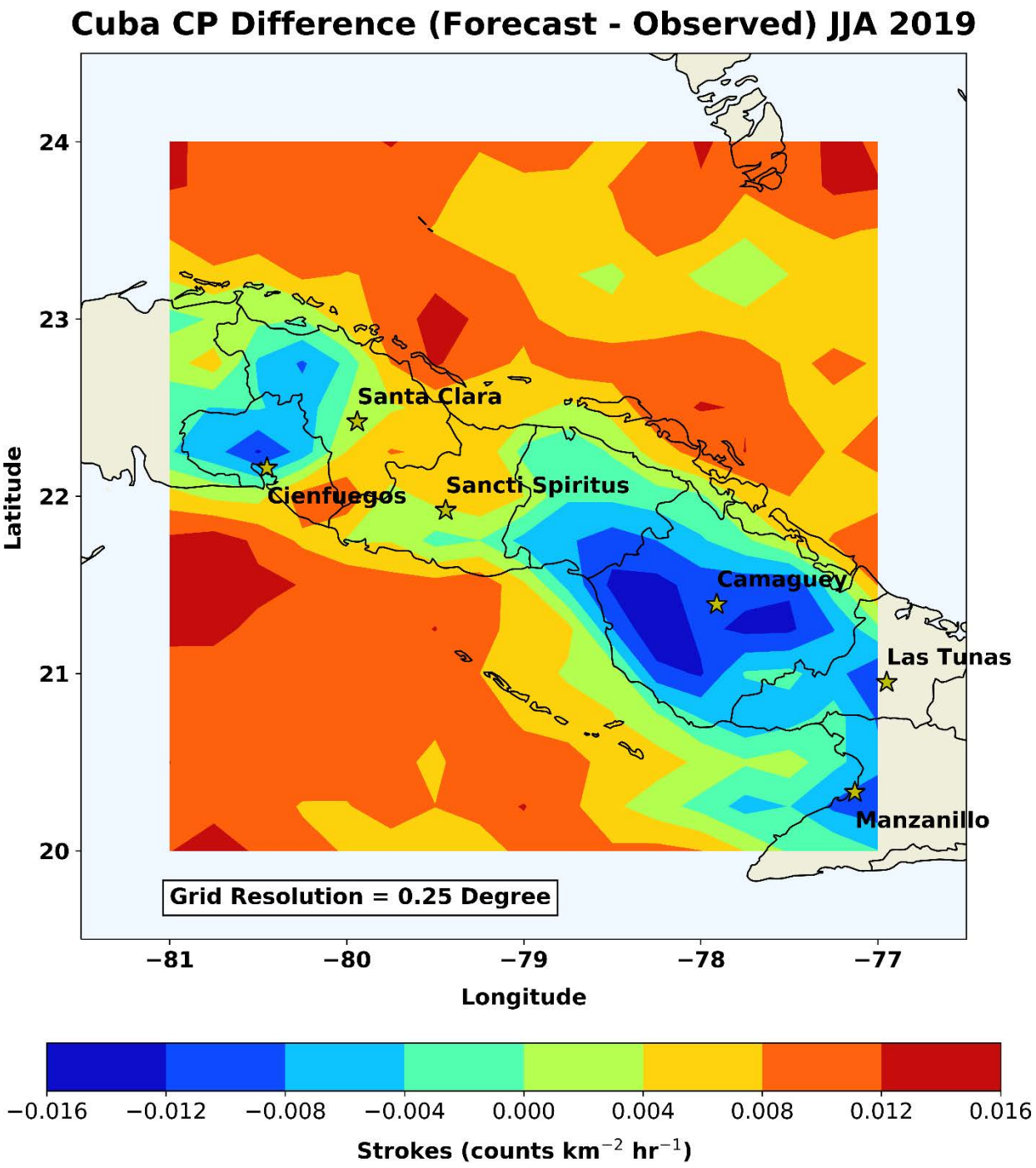

Figure 42. Cuba CP Difference (Forecast – Observed) JJA 2019.

North Korea

Over North Korea in June 2019 (Figures 43, 44 ,45 below), the proxy did a decent job of capturing the area of strokes in the mountainous areas to the north and northwest of Hwapyong near the border of China, but overforecasted strokes along the large set of mountains between Chonpyong and Huichon. It also overforecasted an area of strokes in the mountainous area to the northwest of Huichon, as well as completely missed an area of strokes in the Korea Bay area northwest of Pyongyang.

For July 2019 (Figures 46, 47, 48 below), the proxy did a decent job of capturing the spatial extent of strokes in the mountains in the far northwest along the China border, but grossly overforecasted strokes for the majority of the interior of North Korea. The proxy overforecasted strokes along much of the mountainous areas between Kaechon, Huichon, Hwapyong, and Oung-pi, as well as along the border with South Korea.

For August 2019 (Figures 49, 50, 51 below), the proxy performed poorly for the majority of North Korea, completely underforecasting an area of strokes west of Kaechon and along the coastal area south of Chongpyong. It also overforecasted much of the area of strokes along the mountainous areas along the China border, but captured the spatial extent decently.

For JJA 2019 (Figures 52, 53, 54 below), the proxy decently captured the large concentration of strokes in the mountainous area northwest of the China border, as well as the west of Kaechon. Otherwise, it overforecasted strokes for the majority of the mountains stretching from Oung-ni in the northeast to Pyongyang in the southwest, as well as along the mountains stretching southward to the DMZ. It also overforecasted a large area of strokes in the mountains between Kusong, Huicho, and Hwapyong.

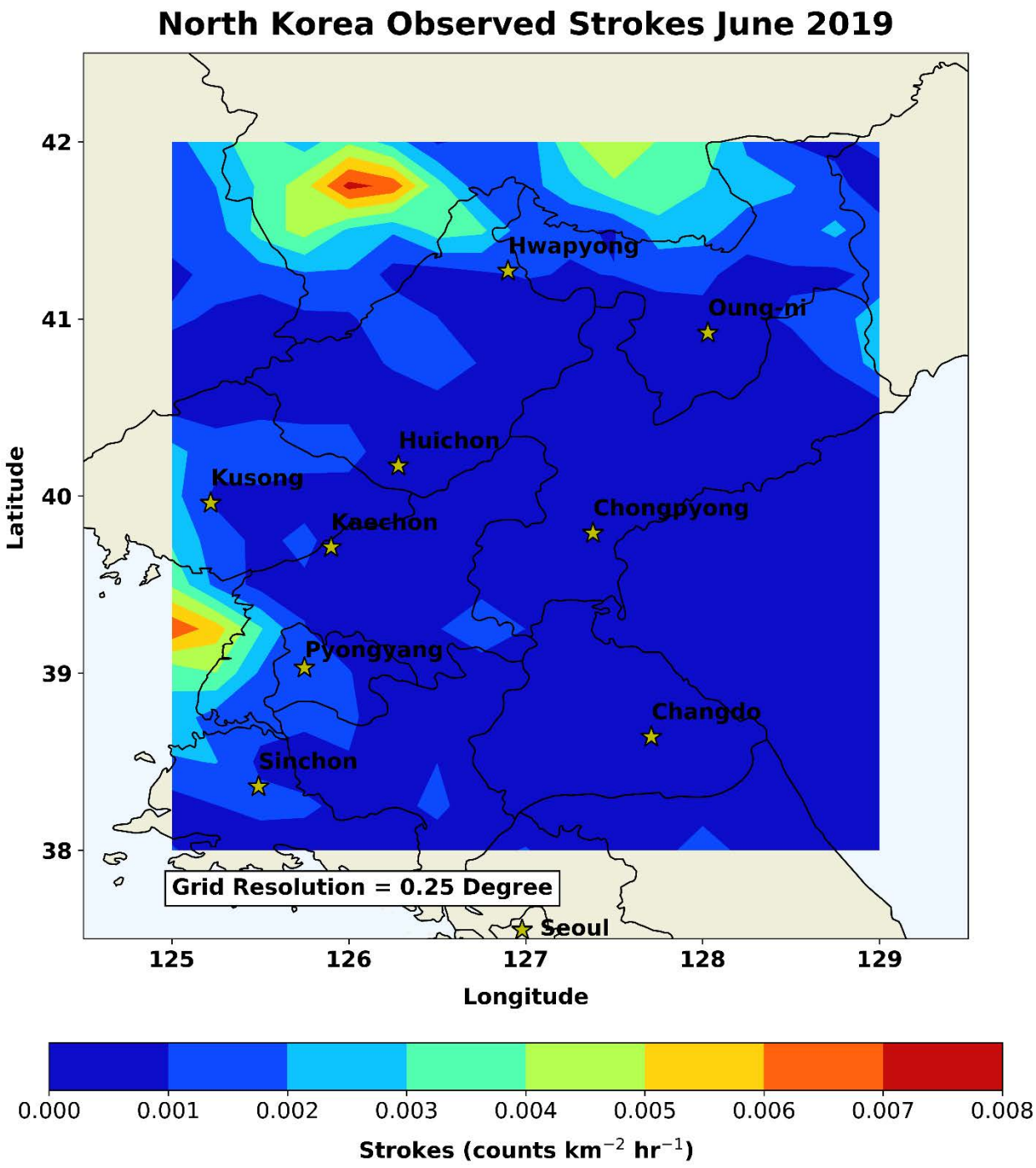

Figure 43. North Korea Observed Strokes June 2019.

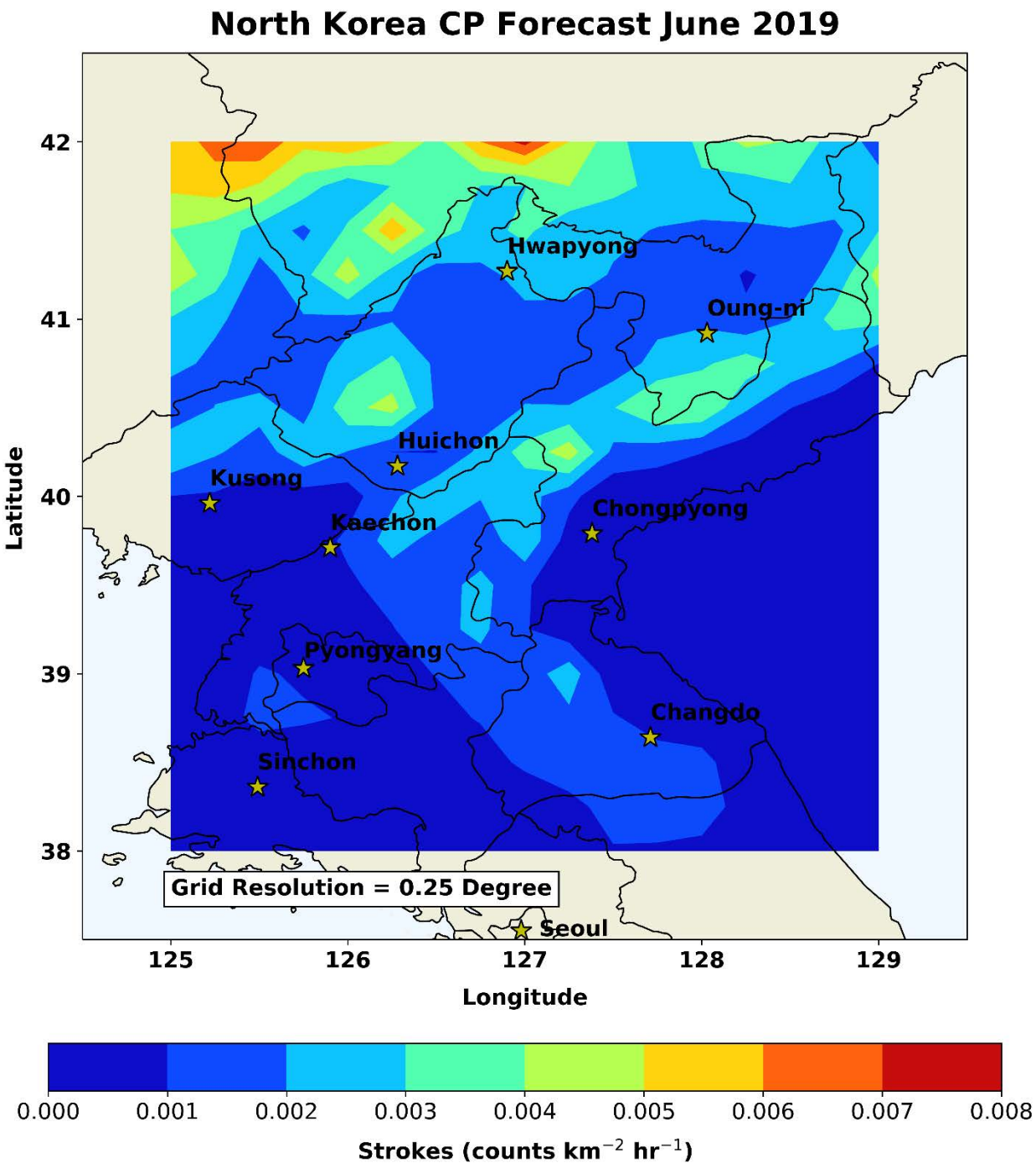

Figure 44. North Korea CP Forecast June 2019.

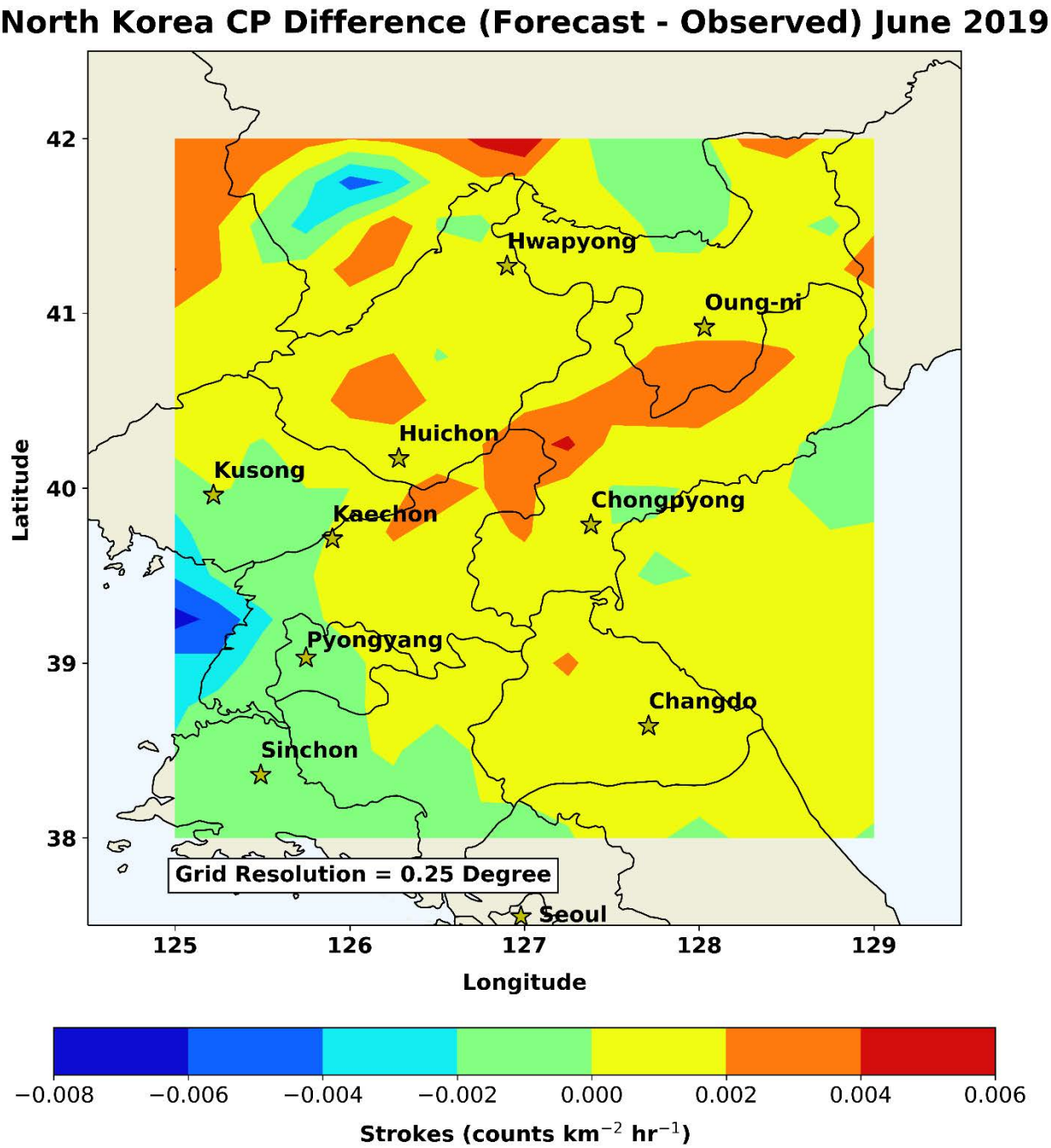

Figure 45. North Korea CP Difference (Forecast – Observed) June 2019.

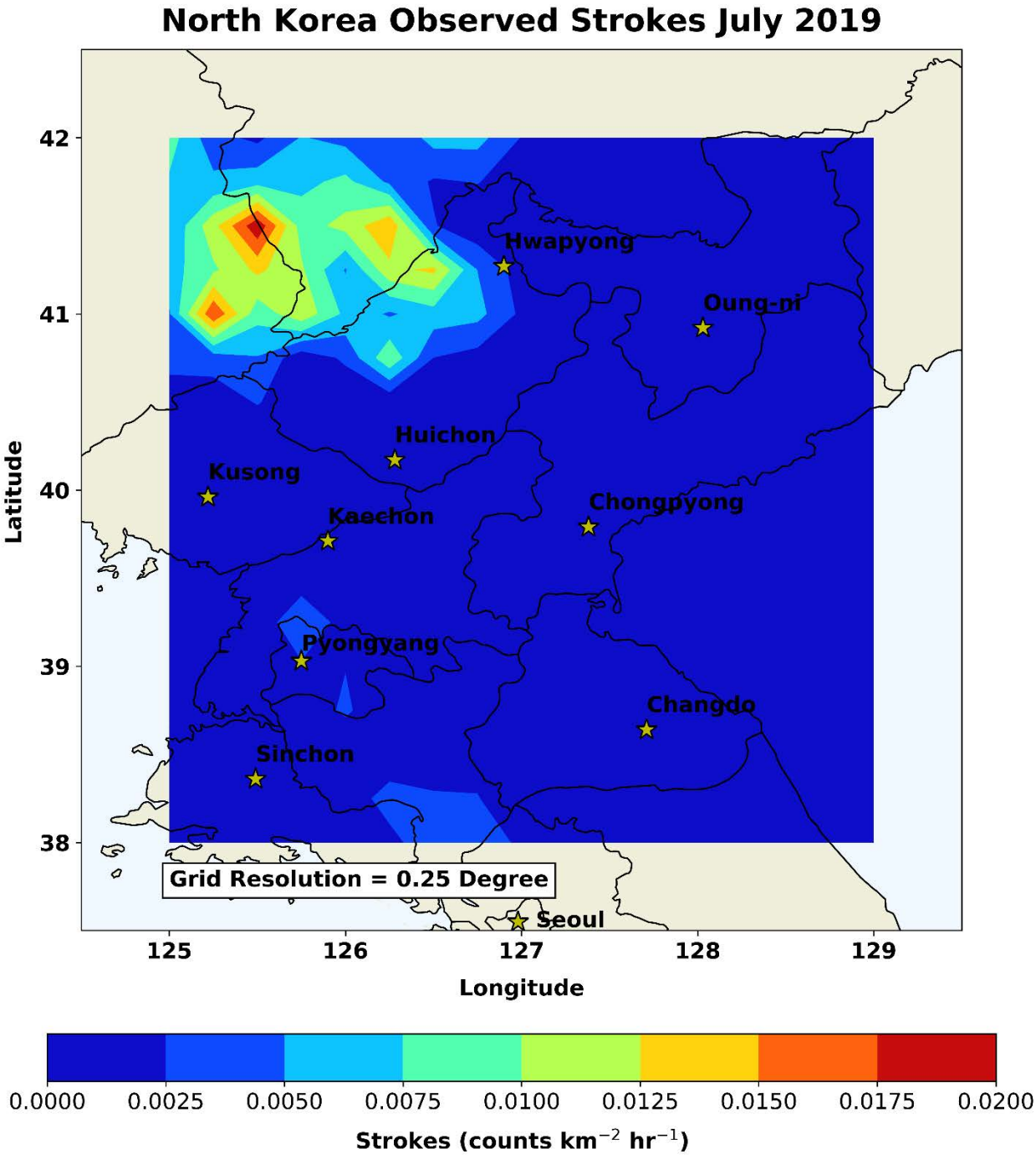

Figure 46. North Korea Observed Strokes July 2019.

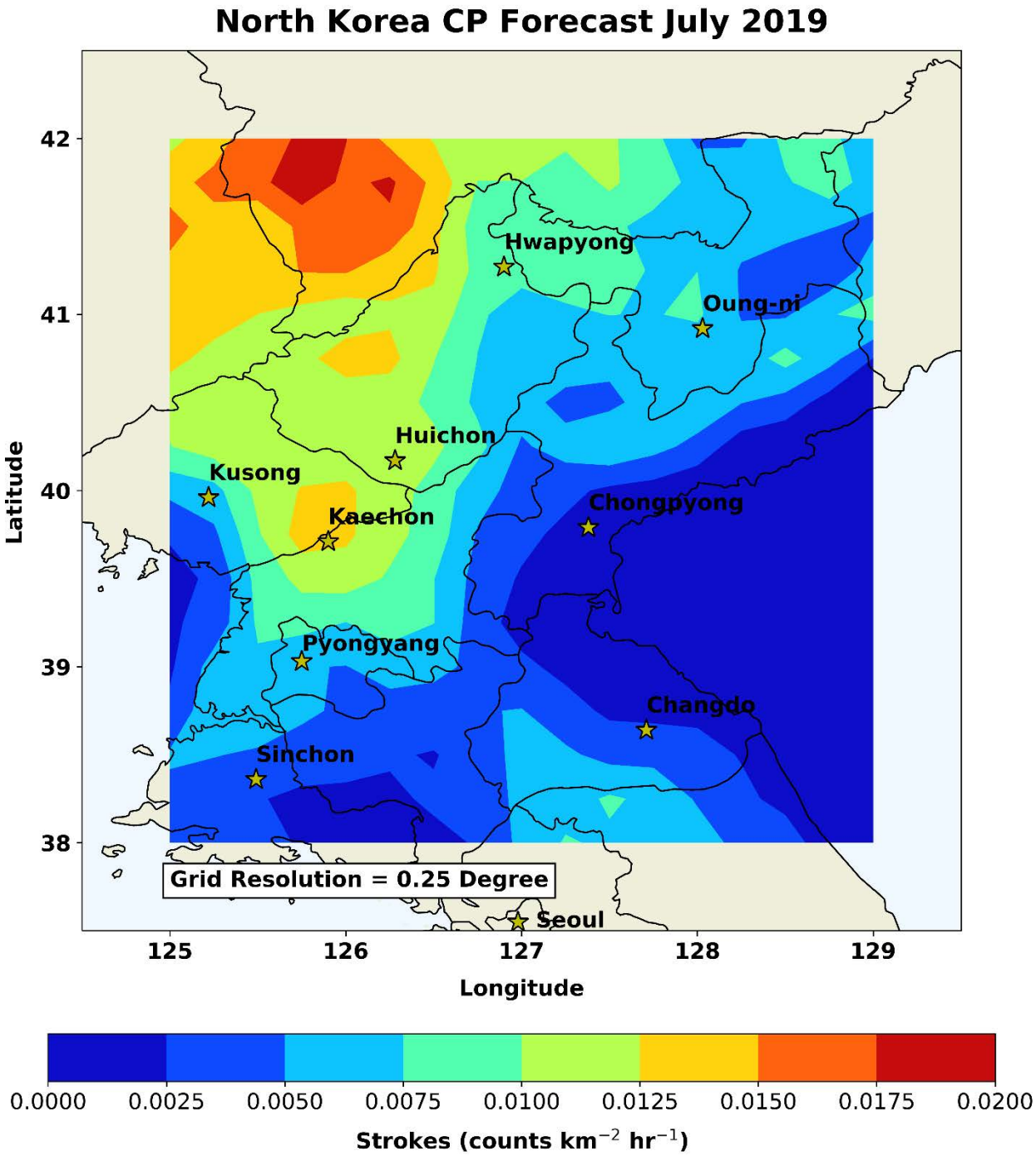

Figure 47. North Korea CP Forecast July 2019.

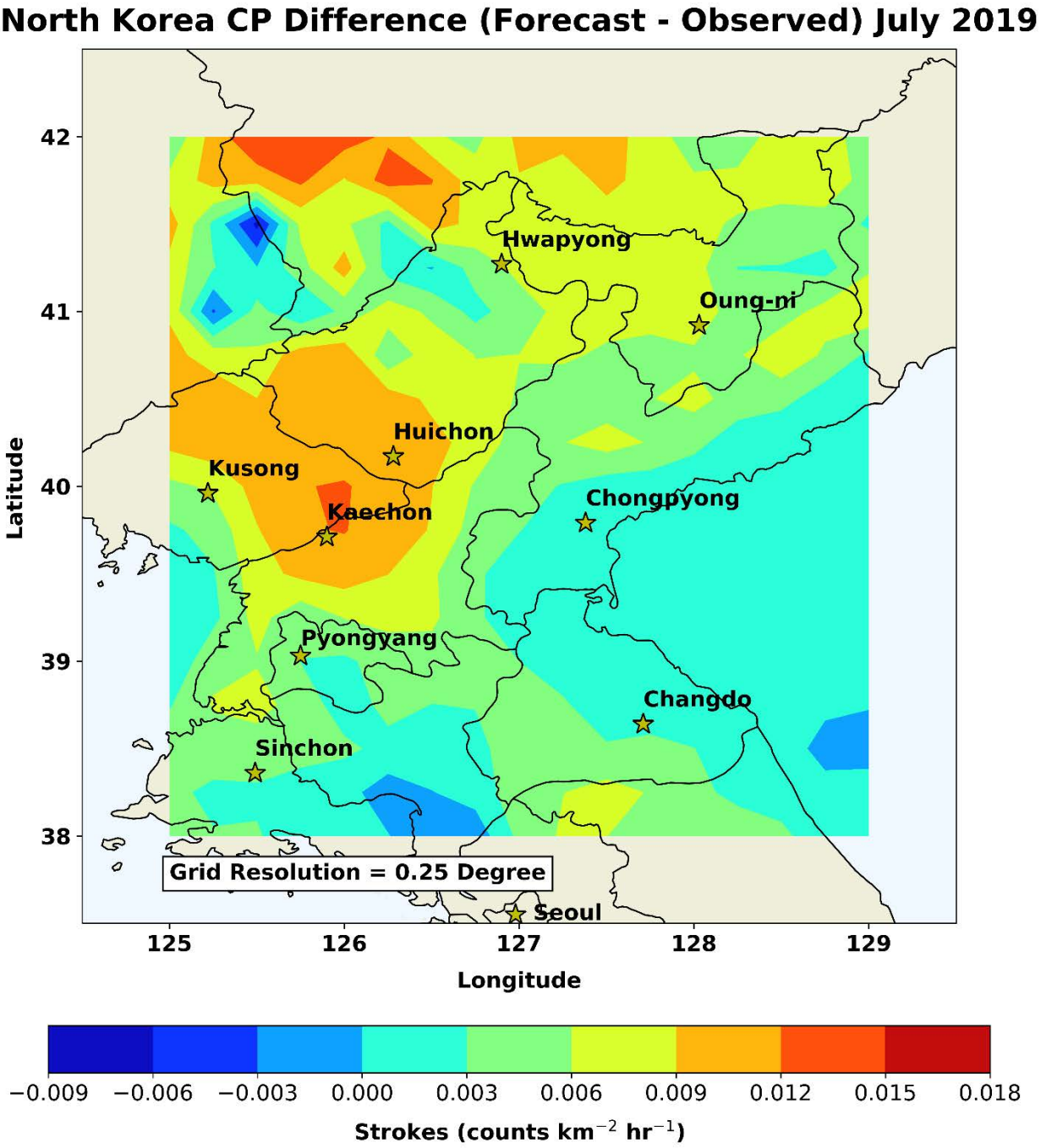

Figure 48. North Korea CP Difference (Forecast – Observed) July 2019.

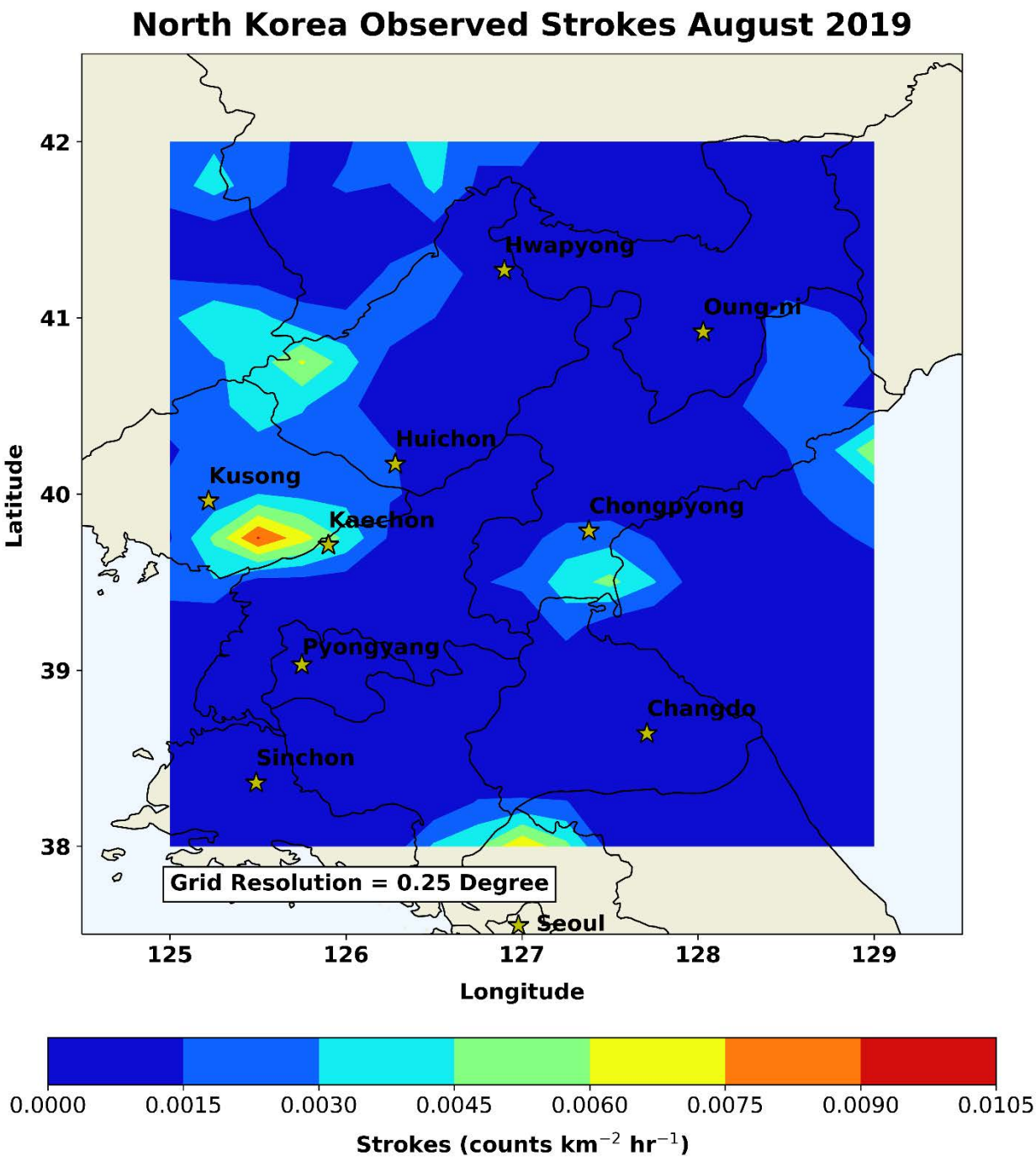

Figure 49. North Korea Observed Strokes August 2019.

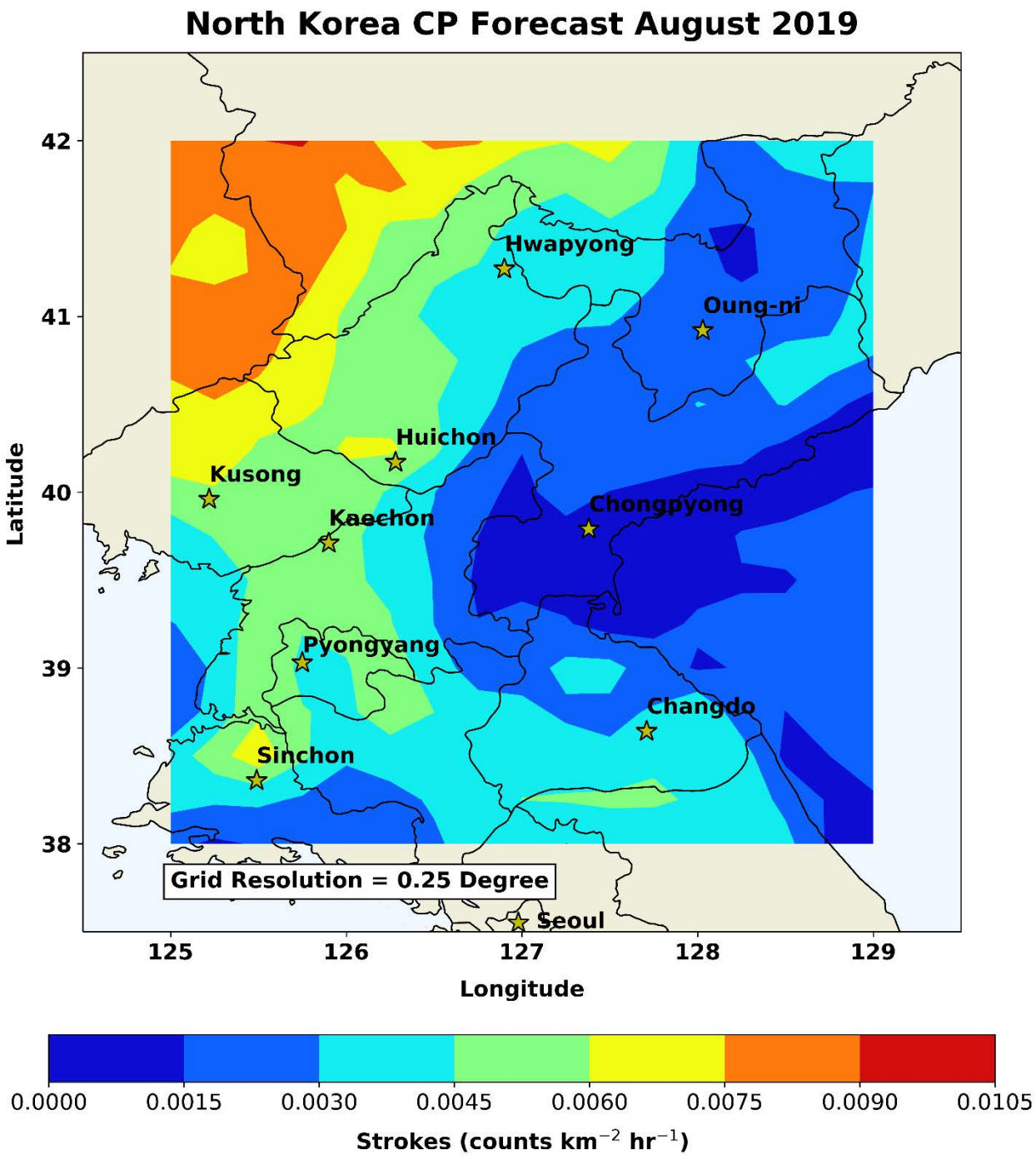

Figure 50. North Korea CP Forecast August 2019.

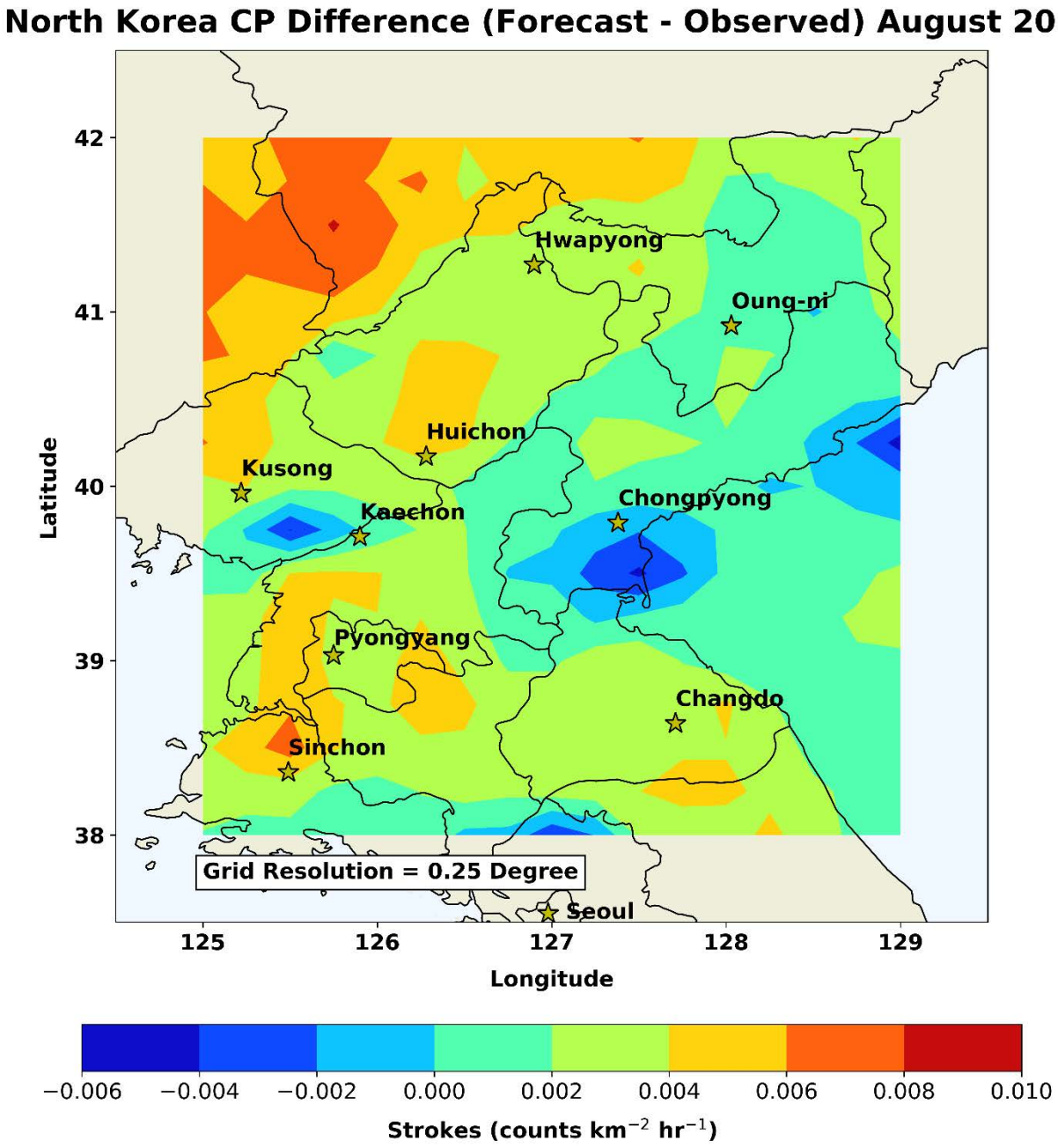

Figure 51. North Korea CP Difference (Forecast – Observed) August 2019.

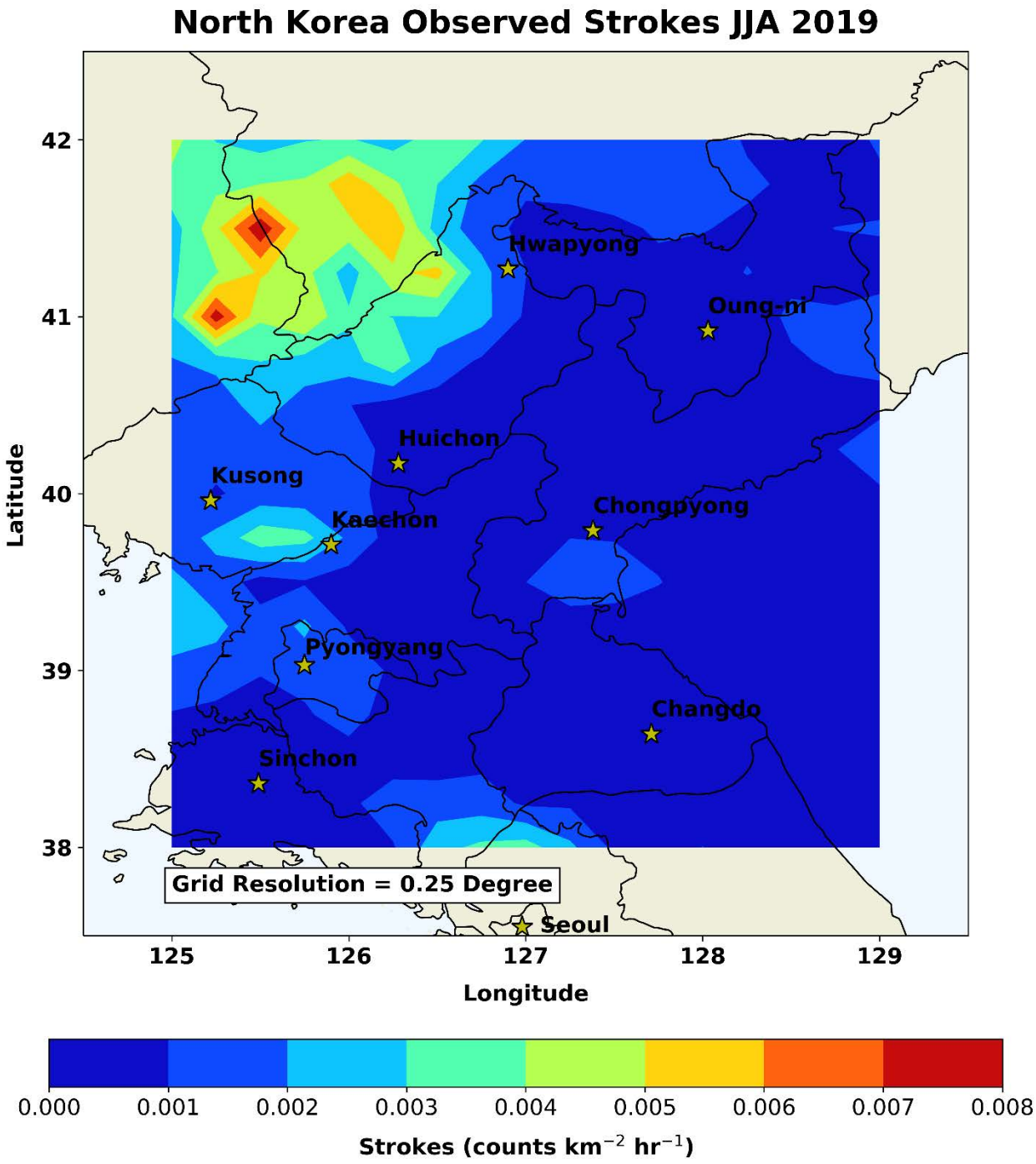

Figure 52. North Korea Observed Strokes JJA 2019.

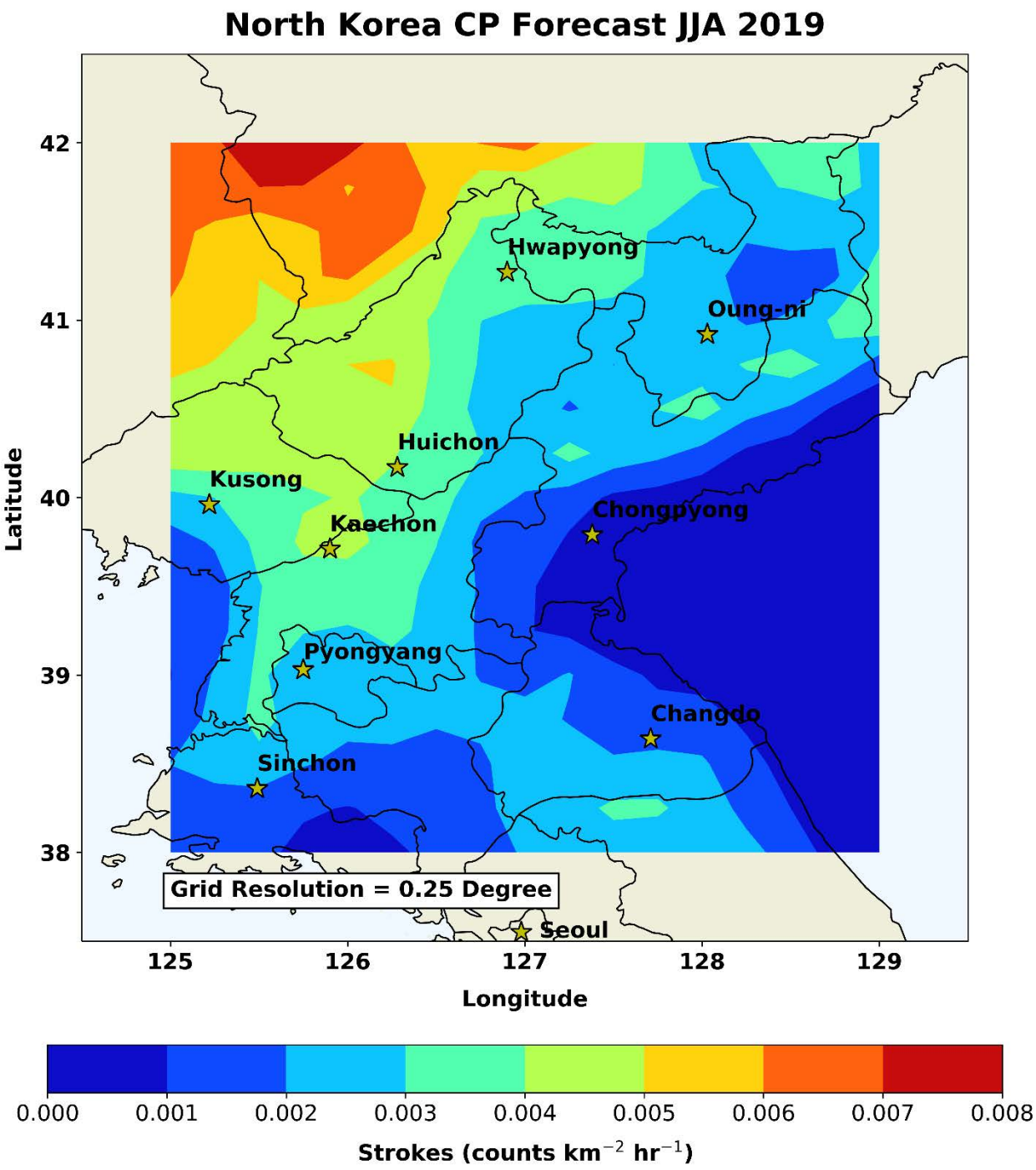

Figure 53. North Korea CP Forecast JJA 2019.

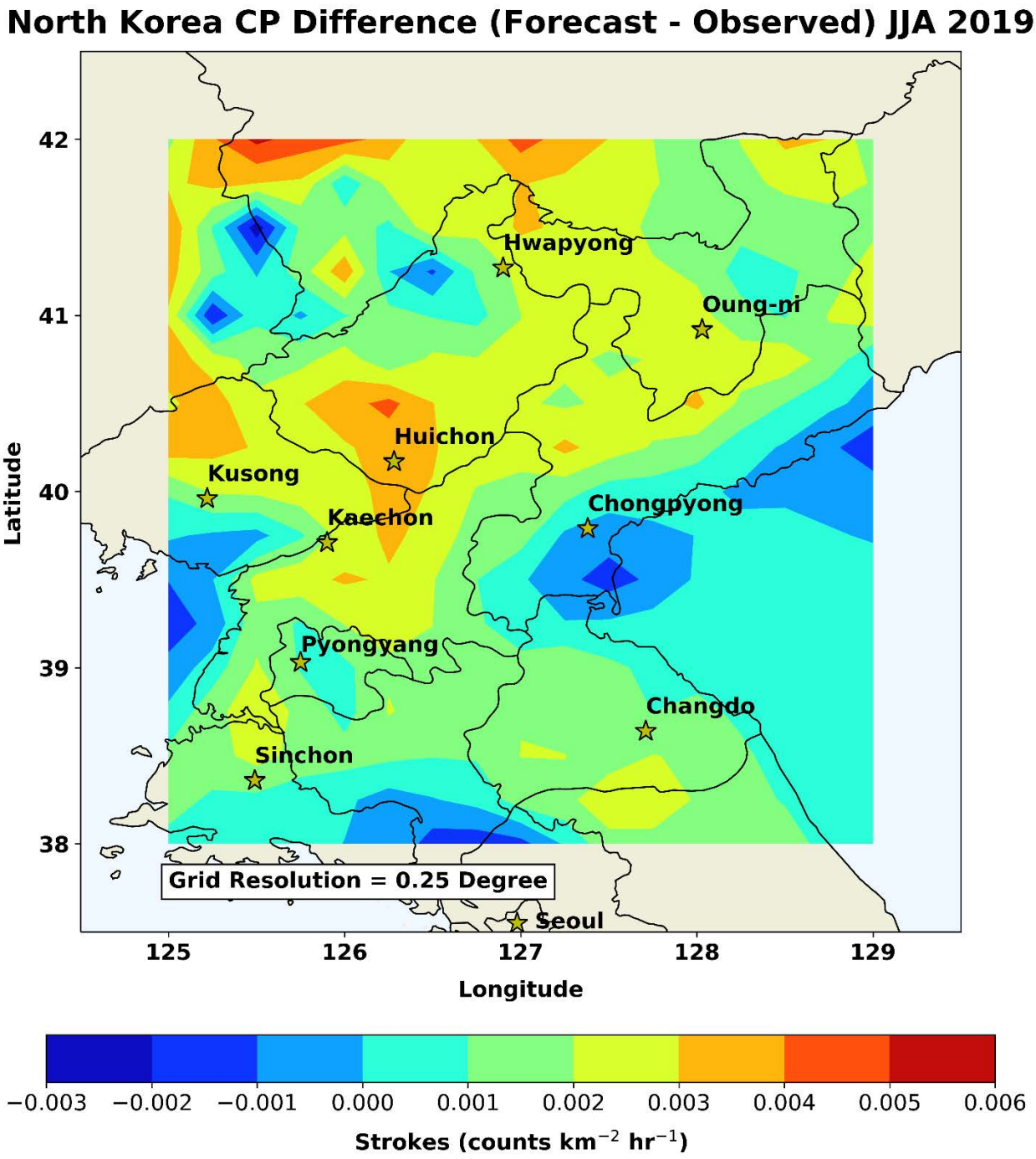

Figure 54. North Korea CP Difference (Forecast – Observed) JJA 2019.

Chapter 5: Discussion

5.1 Proxy Effectiveness

Arizona

The CP proxy performed well over Arizona at capturing the spatial extent and location of the stroke areas, and coincided with the high *r* values found from the correlation study. For example, in June and July, Arizona exhibited statistically significant *r* values of 0.718 and 0.779, respectively, and these months were also the best performing in terms of forecasting the location and spatial extent of the strokes. August and JJA of Arizona exhibited lower *r* values of 0.462 and 0.558, respectively, which coincided with slightly poorer performances in the forecasted location and spatial extent of the strokes. In instances where the general location and spatial extent of the thunderstorms were fairly well-predicted, the magnitude of the number of strokes were also fairly matched. An example of this was in July 2019 (Figures 22, 23, 24), in the open desert area between Ajo and Tucson, as well as in the mountains just northeast of Tucson. However, there were several instances where the CP proxy failed to capture the general extent and magnitude of some of the thunderstorm areas, or forecasted an area of thunderstorms were none occurred. An example of this was in July 2019 just northwest of Flagstaff, where the proxy was predicting a well-defined area of thunderstorms in the 0.0030 to 0.0045 strokes $km^{-2}hr^{-1}$ range, but the observed strokes showed nothing.

Cuba

In Cuba, the CP proxy exhibited fairly strong *r* values of 0.684 and 0.711 for June and July, respectively, but did not do as well at predicting stroke concentration locations as well it had for Arizona for those same months. In other words, while the correlations

due to variations in temporal occurrence were strong, the spatial placement of the thunderstorms did not follow suit in the actual forecast. As noted from the climatology study done on the region, June, July, and August are the busiest convective months for Cuba, and this time period accounts for the majority of stroke variability exhibited on annual timescales. However, most of the lightning produced in this region for this time period is driven by diurnal heating and local effects, such as the sea breeze convergence that regularly occurs in the afternoon (Mayor and Mesquita, 2015:1). Sea breeze convergence results in thunderstorm formation in the interior of an island, but the locations can be further modified and complicated by changes in the larger scale environmental flow, which affects the location of the sea breeze convergence areas and subsequent thunderstorm formation (Atkins and Wakimoto, 1997:2112). It is possible that the overall poor spatial accuracy of the CP proxy for Cuba was due to the 0.25° GFS not adequately resolving the sea breeze convergence effects.

North Korea

North Korea exhibited the poorest performance of the CP proxy out of all three regions, both in *r* values and location accuracy. The best *r* values were 0.562 in June, and 0.412 for August, which accounts for the variance due to the temporal occurrence of lightning. In all three months and the seasonal time period, the CP proxy consistently misplaced the thunderstorm locations for the majority of North Korea, except along the China border to the north. The majority of North Korea is covered in mountains and valleys, whose separations occur on scales much less than the 0.25° GFS is capable of resolving. June, July, and August are the busiest convective months for North Korea, and the triggering of thunderstorms is most strongly influenced by diurnal heating and local

terrain effects summer ("Climate – North Korea," 2022). Although the 0.25° GFS utilizes sigma coordinates in the lower levels to correctly follow terrain height, the model's resolution is likely too coarse to adequately capture local effects induced by terrain. The exception to this appears to be the frequent areas of thunderstorms that occurred in the far north along the China border. This large area of frequent thunderstorms during the summer months is located deep in the continent's interior, in an area of downward sloping terrain that flattens out to the northwest. It is likely that this large reoccurring area of thunderstorms is fueled by the large scale moist flow from the South China Sea, which is characteristic of the East Asian Monsoon that would be active during this time (Clift et al. 2020). Additionally, the reoccurring area of observed strokes in the Korea Bay, northwest of Pyongyang that was present in June 2019 (Figures 43, 44, 45) was completely missed by the CP forecast. In Chapter 4.2, it was hypothesized that this area of strokes was caused by reoccurring seabreeze effects aided by the shape of the coast (Holle, 2022). It is possible that these convective events were of small scale and were short-lived, although their cumulative effects were visible at a larger scale, as seen by the large number of strokes in the bay. It is possible that the 0.25° GFS was unable to spatially or temporally resolve these small, transient thunderstorms, even if they were reoccurring.

5.2 Limitations of the Proxy

Due to 0.25 GFS

In Tippett et al.'s 2019 study, they noted that the CP proxy performed worse during the warm season months of May to October when most lightning occurs, and

performed better during the cool season months of November to April (Tippett et al, 2019:3932). Additionally, as noted by Tippett and Koshak from their 2018 study, the skill of United States precipitation forecasts tends to be lower during warm months when the precipitation source is dominated by small-scale convective events (Tippett and Koshak, 2018:10724). Since warm season convective events are diurnally driven, it is possible that the poorer performance of the proxy in the current work was simply due to the 0.25° GFS being too coarse to capture the precipitation produced by thunderstorm cells, as they are too fine in horizontal scale (~5km or ~0.05°) and develop and dissipate too quickly (<1 hr). The 0.25° GFS is only able to output a forecast every three hours, which is too long of a time period to capture individual thunderstorm cells. However, it is likely that the CP proxy would perform well during the cool season as noted by Tippett et al. in their 2019 study, as this allows the possibility of convection to be more wide-spread and driven by synoptically forced events such as frontal passages and upper level trough/low passes. The 0.25° GFS excels at capturing the synoptic patterns of most regions, and would have no trouble resolving precipitation produced by along fronts or troughs. Since the proxy relies on both the presence of precipitation and CAPE to predict lightning, the absence of one or the other would preclude the forecast of lightning. By necessity of these two variables, this formulation sharply reduces the possibility of incorrectly forecasting lightning during stratiform or prolonged precipitation events, which is a weakness of using precipitation alone as a proxy. Overall, it is very likely that the instances of poorer performance of the CP proxy in the present study was due to the majority of precipitation being produced by small-scale convective events of short duration that the 0.25° GFS simply could not resolve spatially or temporally.

Due to Nature of the Stroke Data

An additional consideration to explain the inaccuracies of the CP proxy in this study is the physical differences between strokes and flashes. This study used strokes, which are essentially components of flashes. A single lightning flash (or discharge) contains anywhere from one to 26 strokes (Wallace and Hobbs, 2006:255, WMO, 2014:657). Also, the previous studies by Romps et al. (2014, 2018) and Tippett et al. (2018, 2019) used NLDN flashes and CG lightning only, whereas the present work used GLD360 strokes and total lightning (IC and CG). Flashes can be thought of as clusters or areas of strokes, which may have made them more suitable for use with the coarse 0.25° resolution models used in those previous studies. Although strokes convey more information than flashes, this information may not have aided the CP proxy effectiveness due to the resolution differences between the strokes and the 0.25° model data. Strokes were used because flashes were neither commercially available nor regularly produced by the GLD360 lightning network. Total lightning was used because the desire of ASSET was to have a lightning model that generates any lightning, indiscriminate of IC versus CG. Since the Romps and Tippett studies only used CG lightning, the addition of IC in the present study may have contributed negatively to the CP's performance. Additionally, the NLDN network has a much higher location accuracy of 0.15 km to 0.25 km, compared to the GLD360 network, which has a location accuracy of 1.8 km to 2.4 km. Furthermore, Romps et al. and Tippett et al. used 10 to 13 years of reanalysis model data to compute their correlations, while the present work used only three months. All these factors combined likely contributed to the CP proxy's frequent misplacement of thunderstorm areas, which was more apparent in the Cuba and North Korea zones.

5.3 Practical Application of the CP Proxy

Diagnostic Tool

The question now is, what would be a practical use of the CP proxy generated by model that is too coarse to resolve convective clouds? The CP proxy could possibly be used as a long-term thunderstorm diagnostic outlook tool. For example, since the 0.25° GFS outputs forecasts up to 384 hours out (16 days), the CP proxy could be modified to produce a thunderstorm probability/coverage plot over a large area. Operationally, this could be used by forecasters to give a long-term thunderstorm outlook to mission planners. This type of forecast would fall under the category of a "planning forecast," as it would be too coarse for use an "execution forecast," which is a high fidelity/fine scale forecast used to give pilots definite "go/no-go" weather conditions for takeoff and landing.

Ballparking Lightning

As far as a tool to be used within ASSET, this CP proxy method could be utilized if a separate statistical method were developed to simulate lightning down to approximately 0.01° (~1 km). This method would require fine scale modeling and spatial correlation to determine the most likely locations that individual thunderstorm cells would form within the broad (0.25°) area of thunderstorms originally predicted by the CP proxy. The CP proxy method does a decent job at capturing the general concentration of thunderstorms and location that are likely to appear over an area, and it does take into account the appropriate environmental conditions that would lead to a thunderstorm, such as high CAPE, precipitable water, and large scale orography (such as the Santa Catalinas

see in Arizona). For all intents and purposes, the CP proxy does a good job of placing thunderstorm concentrations at the appropriate place and time, albeit coarsely.

5.4 Conclusion

Summary

This work sought to create a global lightning model for use within ASSET that would forecast lightning strokes at the correct time and place, under meteorologically appropriate conditions. A thorough review of existing literature in lightning modeling methods currently in use in numerical weather prediction models was conducted, and ultimately a proxy method of forecasting lightning was chosen, due to successes in recent studies that demonstrated its effectiveness in predicting lightning with numerical weather models that cannot explicitly predict lightning. Since the need of ASSET was for a global lightning model, a global numerical weather prediction model, the 0.25° GFS, as well as a global lightning dataset, the GLD360, were chosen for this project. Results showed that the proxy method performed well at capturing the spatial extent, magnitude, and timing of strokes in the CONUS location, AZ, but performed generally poorly over Cuba and North Korea. The previous studies that demonstrated the effectiveness of the proxy method were performed over CONUS only, utilized a finer scale observed lightning network, the NLDN, and used flashes, as opposed to strokes. To the author's knowledge, the present work is the first of its kind to employ the proxy method over Cuba and North Korea. For practical use, the CP proxy method accomplishes the task of estimating the location of stroke areas at a monthly and seasonal timescale, but does so at a coarse 0.25° resolution. It does predict lightning areas under meteorologically appropriate conditions,

such as during times of high CAPE and increasing precipitation rates. As such, the method is currently best suited as a lightning diagnostic tool.

Future Work

For use within ASSET, the following future work is recommended. First, a statistical method should be employed to simulate the occurrence of individual thunderstorms within the 0.25° thunderstorm grid areas originally forecasted by the CP proxy. Since ASSET requires lightning to be simulated at a fine resolution (~0.01°), this would simulate lightning at the appropriate resolution. Next, a Heidke Skill Score (HSS) should be employed to more objectively test the performance of the CP proxy. An HSS, also known as a contingency table or confusion matrix, measures the fractional improvement of the forecast over a standard forecast (Warner, 2011:298). Its formula is given by:

$$HSS = 2\frac{(ad - bc)}{[(a + c)(c + d) + (a + b)(b + d)]} \qquad (7)$$

Where *a, b, c,* and *d* represent various outcomes for forecasted vs observed events. Other recommended future work includes performing the correlations over longer periods of time, using different versions of CAPE and precipitation, testing of the proxy method over cool seasons, comparing the performance of flashes vs strokes, and the addition of an altitude factor to correct for terrain dependencies on thunderstorm formation.

Bibliography

1. Adams, David K., and Andrew C. Comrie. “The North American Monsoon,” *Bulletin of the American Meteorological Society,* 78(10):2197-2214 (October 1997).

2. Arnfield, A. John. "Köppen climate classification," *Encyclopedia Britannica*, 11 Nov. 2020. Accessed 28 December 2021. https://www.britannica.com/science/Koppen-climate-classification

3. Atkins, Nolan T., and Roger M. Wakimoto. “Influence of the Synoptic-Scale Flow on Sea Breezes Observed during CaPE,” *Monthly Weather Review,* 125:2112-2130 (September 1997).

4. Ausick, Paul. “7 Aircraft Disasters Caused by Lightning Strikes.” *24/7wallst.com*, 24/7 Wall St. LLC. 11 January 2020. Retrieved on 3 December 2021. https://247wallst.com/transportation/2019/02/21/7-aircraft-disasters-caused-by-lightning-strikes/

5. Bieda, Stephen W. III, Christopher L. Castro, and Steven L. Mullen. “The Relationship of Transient Upper-Level Trough to Variability of the North American Monsoon,” *Journal of Climate*, 22:4213-4227 (August 2009).

6. Brooks, I. M. and C. P. R. Saunders. “An experimental investigation of the inductive mechanism of thunderstorm electrification,” *Journal of Geophysical Research*, 99:10627-10632 (May 1994).

7. Bruning, Eric C. et al.. “Meteorological Imagery for the Geostationary Lightning Mapper,” *Journal of Geophysical Research*, 124: 14258-14309 (December 2019).

8. Bryan, George H., John C. Wyngaard, and J. Michael Fritsch. “Resolution Requirements for the Simulation of Deep Moist Convection,” *American Meteorological Society Monthly Weather Review*, 131:2394-2416 (October 2003).

9. Clift, P. D., A. Holbourn, C. France-Lanord, H. Zheng. “Evolution of the Asian monsoon,” *Eos,* 101 (June 2020). https://doi.org/10.1029/2020EO146198

10. “Climate – Cuba.” *climatestotravel.com*. Climates To Travel: World Climate Guide. Retrieved 3 January 2022. https://www.climatestotravel.com/climate/cuba

11. “Climate – North Korea.” *climatestotravel.com*. Climates To Travel: World Climate Guide. Retrieved 3 January 2022. https://www.climatestotravel.com/climate/north-korea

12. “Climate Zones.” *weather.gov*. National Weather Service National Oceanic and Atmospheric Administration. Retrieved 26 August 2021. https://www.weather.gov/jetstream/climates

13. Cramer, John A. Vaisala Inc, Tucson, AZ. Personal Correspondence. 23 December 2021.

14. Department of the Air Force. *Air and Space Weather Operations*. AFMAN 15-129. 16 June 2021. https://static.e-publishing.af.mil/production/1/af_a3/publication/afman15-129/afman15-129.pdf

15. Dewan, Ashraf, Emmanuel T. Ongee, M. Rafiuddin, Md. Masudur Rahman, and Rezaul Mahmood. “Lightning activity associated with precipitation and CAPE over Bangladesh,” *International Journal of Climatology*, 38:1649-1660 (September 2017).

16. Fierro, Alexandre O., Edward R. Mansell, Donald R. MacGorman, and Conrad L. Ziegler. “The Implementation of an Explicit Charging and Discharge Lightning Scheme within the WRF-ARW Model: Benchmark Simulations of a Continental Squall Line, a Tropical Cyclone, and a Winter Storm,” *Monthly Weather Review*, 141:2390-2415 (January 2013).

17. Galanaki, E., V. Kotroni, K. Lagouvardos, and A. Argiriou. “A ten-year analysis of cloud-to-ground lightning activity over the Eastern Mediterranean region,” *Atmospheric Research*, 166:213-222 (July 2015).

18. “Global Forecast System (GFS).” *emc.ncep.noaa.gov.* National Oceanic and Atmospheric Administration, U.S. Department of Commerce. Retrieved 30 June 2021.

19. “GOES-R Mission Overview.” *goes-r.gov.* National Oceanic and Atmospheric Administration/National Aeronautics and Space Administration. Retrieved 1 February 2022. https://www.goes-r.gov/mission/mission.html

20. Goodman, Steven J. et al.. *GLM Lightning Cluster-Filter Algorithm, Algorithm Theoretical Basis Document Version 3.0*. NOAA NESDIS Center for Satellite Applications and Research. 30 July 2012.

21. Goodman, Steven J. et al.. “The GOES-R Geostationary Lightning Mapper (GLM),” *Atmospheric Research*, 125-126:34-39 (January 2013).

22. Holle, Ronald L. National Lightning Safety Council, Oro Valley, AZ. Personal Correspondence. 1 February 2022.

23. Lamb, Dennis, and Johannes Verlinde. *Physics and Chemistry of Clouds*. Cambridge: Cambridge University Press, 2011.

24. "Lightning Imaging Sensor (LIS)." *ghrc.nsstc.nasa.gov*. Global Hydrology Resource Center. Retrieved 1 February 2022. https://ghrc.nsstc.nasa.gov/lightning/overview_lis_instrument.html

25. "Lightning Victims." *weather.gov*. National Weather Service, National Ocean and Atmospheric Administration. Retrieved 15 March 2021. https://www.weather.gov/safety/lightning-victims

26. Lopez, Philippe. "A Lightning Parameterization for the ECMWF Integrated Forecasting System," *Monthly Weather Review,* 144:3057-3075 (September 2016).

27. MacGorman, Donald R., Jerry M. Straka, and Conrad L. Ziegler. "A Lightning Parameterization for Numerical Cloud Models," *Journal of Applied Meteorology*, 40:459-478 (March 2001).

28. Magi, Brian I. "Global Lightning Parameterization from the CMIP5 Climate Model Output," *Journal of Atmospheric and Oceanic Technology*, 32:434-452 (March 2015).

29. Mansell, Edward R., Donald R. MacGorman, Conrad L. Ziegler, and Jerry M. Straka. "Simulated three-dimensional branched lightning in a numerical thunderstorm model," *Journal of Geophysical Research*, 107, D9, 4075:1-13 (May 2002).

30. Mansell, Edward R., Donald R. MacGorman, Conrad L. Ziegler, and Jerry M. Straka. "Charge structure and lightning sensitivity in a simulated multicell thunderstorm," *Journal of Geophysical Research*, 110, D12101:1-24 (June 2005).

31. Mayor, Yandy G. and Michel D. S. Mesquita. "Numerical Simulations of the 1 May 2012 Deep Convection Event over Cuba: Sensitivity to Cumulus and Microphysical Schemes in a High-Resolution Model," *Advances in Meteorology*, 2015:1-17 (July 2015).

32. McCaul, Eugene W. Jr., Steven J. Goodman, Katherine M. LaCasse, and Daniel J. Cecil. "Forecasting Lightning Threat Using Cloud-Resolving Model Simulations, *Weather and Forecasting*, 24:709-729 (June 2009).

33. Michalon, N., A. Nassif, T. Saouri, J.F. Royer, and C. A. Pontikis. "Contribution to the climatological study of lightning," *Geophysical Research Letters*, 26:3097-3100 (October 1999).

34. Molinié, J., and C. A. Pontikis. "A climatological study of tropical thunderstorm clouds and lightning frequencies on the French Guyana coast," *Geophysical Research Letters,* 22:1085-1088 (May 1995).

35. Murphy, Martin J., John A. Cramer, and Ryan K. Said. "Recent History of Upgrades to the U.S. National Lightning Detection Network," *Journal of Atmospheric and Oceanic Technology*, 38:573-585 (March 2021).

36. Nag, Amitabh, Martin J. Murphy, Kenneth L. Cummins, Alburt E. Pifer, and John A. Cramer. "Recent Evolution of the U.S. National Lightning Detection Network," *23rd International Lightning Detection Conference/5th International Lightning Meteorology Conference*, Tucson, AZ. 2014. https://www.vaisala.com/sites/default/files/documents/Nag%20et%20al-Recent%20Evolution%20of%20the%20U.S.%20National%20Lightning%20Detection%20Network-2014-ILDC-ILMC.pdf

37. National Research Council. *The Earth's Electrical Environment (1986)*. Washington: National Academy Press, 1986. https://www.nap.edu/download/898#

38. "Nimbostratus clouds." *metoffice.gov.uk*. The Met Office. Retrieved 29 December 2021.

39. Niemeyer, L., L. Pietronero, and H. J. Wiesmann. "Fractal Dimension of Dielectric Breakdown," *Physical Review Letters*, 52:1033-1037 (March 1984).

40. Rind, David and Colin Price. "Possible implications of global climate change on global lightning distributions and frequencies," *Journal of Geophysical Research*, 99:10823-10831 (May 1994).

41. Rash, Clarence E. "When Lightning Strikes." *AeroSafety World.* Flight Safety Foundation. June 2010. Retrieved 3 December 2021. https://flightsafety.org/wp-content/uploads/2016/11/asw_jun10_p18-23.pdf

42. Reynolds, S. E., M. Brook, and Mary Foulks Gourley. "Thunderstorm Charge Separation," *Journal of Meteorology*, 14:426-436 (October 1957).

43. Romps, David M., Jacob T. Seeley, David Vollaro, and John Molinari. "Projected increase in lightning strikes in the United States due to global warming," *Sciencemag.org*, 346:851-853 (November 2014).

44. Romps, David M., Alexander B. Charn, Robert H. Holzworth, William E. Lawrence, John Molinari, and David Vollaro. "CAPE Times P Explains Lightning Over Land But Not the Land-Ocean Contrast," *Geophysical Research Letters,* 45:12623-12630 (November 2018).

45. Rudlosky, Scott D. et al.. "Initial Geostationary Lightning Mapper Observations," *Geophysical Research Letters*, 46: 1097-1104 (January 2019).

46. Saunders, C.P.R., W. D. Keith, and R. P. Mitzeva. "The Effect of Liquid Water on Thunderstorm Charging," *Journal of Geophysical Research*, 96:11007-11017 (June 1991).

47. Said, R. K., M. B. Cohen, and U. S. Inan. "Higher intense lightning over the oceans: Estimated peak currents from global GLD360 observations," *Journal of Geophysical Research*, 118:6905-6915 (July 2013).

48. Said, Ryan, and Martin Murphy. "GLD360 Upgrade: Performance Analysis and Applications," *24th International Lightning Detection Conference/6th International Lightning Meteorology Conference*, San Diego, CA. 2016.

49. "Severe Weather Safety Guide: Lightning." *weather.gov.* NOAA, National Weather Service. Paduca, Kentucky. Retrieved 27 December 2021. https://www.weather.gov/media/pah/WeatherEducation/lightningsafety.pdf

50. Siingh, Devendraa, P.S. Buchunde, R.P. Singh, Asha Nath, Sarvan Kumar, and R.N. Ghodpage. "Lightning and convective rain study in different parts of India," *Atmospheric Research*, 137:35-48 (September 2014).

51. Soriano, Luis Rivas, Fernando De Pablo, and Eulogio García Díez. "Relationship between Convective Precipitation and Cloud-to-Ground Lightning in the Iberian Peninsula," *Monthly Weather Review*, 129:2998-3003 (May 2001).

52. Stanley, Mark A. et al.. "A link between terrestrial gamma-ray flashes and intracloud lightning discharges," *Geophysical Research Letters*, 33:L06803:1-5 (March 2006).

53. Steward, Bryan J. "AFIT Sensor & Scene Emulation Tool (ASSET) Overview." Introduction and Orientation Briefing. Air Force Institute of Technology, Wright-Patterson AFB OH. March 2020.

54. Takahashi, Tsutomu. "Riming Electrification as a Charge Generation Mechanism in Thunderstorms," *Journal of the Atmospheric Sciences*, 35:1536-1548 (August 1978).

55. Tippett, Michael K. and William J. Koshak. "A Baseline for the Predictability of U.S. Cloud-to-Ground Lightning," *Geophysical Research Letters,* 45:10719-10728 (October 2018).

56. Tippett, Michael K., Chiara Lepore, William J. Koshak, Themis Chronis, and Brian Vant-Hull. "Performance of a simple reanalysis proxy for U.S. cloud-to-ground lightning," *International Journal of Climatology*, 39:3932-3946 (March 2019).

57. "Vaisala US NLDN Lightning Stroke Data." *ghrc.nsstc.nasa.gov*. Global Hydrology Resource Center. Retrieved 22 October 2021. https://ghrc.nsstc.nasa.gov/opendap/doc/vaistroke/vaistroke_dataset.html

58. Wallace, John M., and Peter V. Hobbs. *Atmospheric Science: An Introductory Survey, Second Edition*. Burlington: Elsevier, 2006.

59. Walters, Kenneth R. and Kathleen M. Traxler. *North Korea: A Climatological Study*. USAF Environmental Technical Applications Center, TN-94/003. Scott Air Force Base, Illinois, 29 August 1994.

60. Warner, Thomas Tomkins. *Numerical Weather and Climate Prediction*. New York: Cambridge University Press, 2011.

61. Weidman, Chuck. "Cloud Electrification Processes," Class Lectures, ATMO 589. University of Arizona, Tucson, AZ. March 2013. http://www.atmo.arizona.edu/students/courselinks/spring13/atmo589/lecture_notes/feb19/feb19.html

62. Wiesmann, H. J. and H. R. Zeller. "A fractal model of dielectric breakdown and prebreakdown in solid dielectrics," *Journal of Applied Physics*, 60:1770-1773 (September 1986).

63. Woollaston, Victoria. "Lightning is at its most powerful at 8am in the morning - but more storms occur in the afternoon," *dailymail.co.uk*, DailyMail.com. 17 March 2015. https://www.dailymail.co.uk/sciencetech/article-2998781/Lightning-powerful-8am-morning-storms-occur-afternoon.html

64. World Meteorological Organization (WMO). *Guide to Meteorological Instruments and Methods of Observation, 2014 Edition*. WMO-No. 8. Geneva: World Meteorological Organization, 2014. https://repository.oceanbestpractices.org/handle/11329/365

65. Yoshida, Satoru, Takeshi Morimoto, Tomoo Ushio, and ZenIchiro Kawasaki. "A fifth-power relationship for lightning activity from Tropical Rainfall Measuring Mission satellite observations," *Journal of Geophysical Research*, 114, D09104:1-10 (May 2009).

66. Young, Shannon R., Bryan J. Steward, and Kevin C. Gross. "Development and validation of the AFIT scene and sensor emulator for testing (ASSET)," Proc. SPIE 10178, Infrared Imaging Systems: Design, Analysis, Modeling, and Testing XXVIII, 101780A (3 May 2017).

67. Ziegler, Conrad L., and Donald R. MacGorman. "Observed Lightning Morphology Relative to Modeled Space Charge and Electric Field Distributions in a Tornadic Storm," *Journal of the Atmospheric Sciences*, 51:834-851 (March 1994).

REPORT DOCUMENTATION PAGE	*Form Approved* *OMB No. 074-0188*

1. REPORT DATE *(DD-MM-YYYY)*	**2. REPORT TYPE**	**3. DATES COVERED** *(From – To)*
25-03-2022	Master's Thesis	March 2021 – March 2022

TITLE AND SUBTITLE	**5a. CONTRACT NUMBER**
APPLICATIONS OF A LIGHTNING PROXY TO GENERATE SYNTHETIC LIGHTNING FOR USE IN PHYSICS-BASED IMAGE-CHAIN MODELS	**5b. GRANT NUMBER**
	5c. PROGRAM ELEMENT NUMBER
6. AUTHOR(S) Castro, Bryan, G., Captain, USAF	**5d. PROJECT NUMBER**
	5e. TASK NUMBER
	5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAMES(S) AND ADDRESS(S)	**8. PERFORMING ORGANIZATION REPORT NUMBER**
Air Force Institute of Technology Graduate School of Engineering and Management (AFIT/EN) 2950 Hobson Way, Building 640 WPAFB OH 45433-7765	AFIT-ENP-MS-22-M-082

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)	**10. SPONSOR/MONITOR'S ACRONYM(S)**
AGENCY (spelled out) ADDRESS PHONE and EMAIL ATTN: POC (no sponsor enter: Intentionally left blank)	**AFRL/RHIQ (example)**
	11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION/AVAILABILITY STATEMENT

DISTRUBTION STATEMENT A. APPROVED FOR PUBLIC RELEASE; DISTRIBUTION UNLIMITED.

13. SUPPLEMENTARY NOTES

This material is declared a work of the U.S. Government and is not subject to copyright protection in the United States.

14. ABSTRACT

A method of generating synthetic lightning through the use of a convective available potential energy (CAPE) times precipitation rate (P) proxy (CP) is applied over three distinct climatological zones of the world for a single warm season: central and southern Arizona of the United States, central Cuba, and North Korea. Global Forecast System (GFS) 0.25° by 0.25° forecast data for June, July, and August of 2019 is used to provide 6-hourly CAPE and precipitation rate, while Global Lightning Dataset (GLD360) data for the period 2016 to 2020 is used to provide observed lightning strokes. A five-year lightning climatology study is conducted on each region to identify lightning density and energy patterns. A correlation of monthly and seasonal lightning to the proxy is then conducted for each region at a 0.25°grid resolution to obtain correlation coefficients (r), p-values, and establish validity of the proxy. Linear regression is then applied to the two datasets to determine an appropriate constant of proportionality (k), enabling the proxy to be used to explicitly forecast lightning for each region. An additional correction factor (cf) is also calculated and applied to the proxy to correctly scale the number of strokes per area. The proxy is then used to forecast lightning at monthly and seasonal timescales and compared to observed lightning stroke densities at the same location and time to assess performance. The lightning climatology study results show that long-term variations in lightning patterns for each region are most influenced by warm-season convection, topography, and local effects. Correlation results show best r values at 0.799 (AZ), 0.711 (Cuba), and 0.562 (North Korea), and poorest r values at 0.462 (Arizona), 0.260 (Cuba), and 0.005 (North Korea), indicating that while moderate correlations exist at times over the monthly and seasonal timescales, a correlation over a longer time period encompassing several warm seasons would likely result in more accurate r values and constants of proportionality. Forecast results show that the CP proxy provides a fairly accurate depiction of lightning timing and location at monthly and seasonal timescales over Arizona, but performs poorly over much of Cuba and North Korea.

15. SUBJECT TERMS

Lightning, CAPE, Precipitation, CP, proxy, synthetic

16. SECURITY CLASSIFICATION OF:			**17. LIMITATION OF ABSTRACT**	**18. NUMBER OF PAGES**	**19a. NAME OF RESPONSIBLE PERSON**
					Saunders, Peter A, Maj, USAF, AFIT/ENP
a. REPORT	**b. ABSTRACT**	**c. THIS PAGE**			**19b. TELEPHONE NUMBER** *(Include area code)*
U	**U**	**U**	**UU**	137	(937) 255-6565, ext 4743 (peter.saunders@afit.edu)

Standard Form 298 (Rev. 8-98)

Prescribed by ANSI Std. Z39-18