

PNNL-30832

Boundary Layer Climatology at ARM Southern Great Plains

January 2021

Raghavendra Krishnamurthy
Rob K Newsom
Duli Chand
William J Shaw

DISCLAIMER

This report was prepared as an account of work sponsored by an agency of the United States Government. Neither the United States Government nor any agency thereof, nor Battelle Memorial Institute, nor any of their employees, makes **any warranty, express or implied, or assumes any legal liability or responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights.** Reference herein to any specific commercial product, process, or service by trade name, trademark, manufacturer, or otherwise does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States Government or any agency thereof, or Battelle Memorial Institute. The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States Government or any agency thereof.

PACIFIC NORTHWEST NATIONAL LABORATORY
operated by
BATTELLE
for the
UNITED STATES DEPARTMENT OF ENERGY
under Contract DE-AC05-76RL01830

Printed in the United States of America

Available to DOE and DOE contractors from the
Office of Scientific and Technical Information,
P.O. Box 62, Oak Ridge, TN 37831-0062;
ph: (865) 576-8401
fax: (865) 576-5728
email: reports@adonis.osti.gov

Available to the public from the National Technical Information Service
5301 Shawnee Rd., Alexandria, VA 22312
ph: (800) 553-NTIS (6847)
email: orders@ntis.gov <<https://www.ntis.gov/about>>
Online ordering: <http://www.ntis.gov>

Boundary Layer Climatology at ARM Southern Great Plains

January 2021

Raghavendra Krishnamurthy
Rob K Newsom
Duli Chand
William J Shaw

Prepared for
the U.S. Department of Energy
under Contract DE-AC05-76RL01830

Pacific Northwest National Laboratory
Richland, Washington 99354

Summary

Operational since 1992, the Atmospheric Radiation Measurement (ARM) Southern Great Plains (SGP) site in Oklahoma, USA, has become a reference research site for meteorological studies. Because of an open data policy the ARM data are used by researchers all over the world. In this report, we review the long-term climatology of the atmospheric boundary layer, the SGP instrumentation, the site and some site-specific atmospheric conditions, which potentially affect wind turbines in the region. Because the atmospheric boundary layer is bounded and influenced by the land surface, observations of surface radiation components and heat fluxes are crucial to understanding land–atmosphere interactions. The entrainment of air, updrafts, downdrafts, and boundary layer height characteristics is needed for understanding the structure and growth of the atmospheric boundary layer. Therefore, measurements from both ground surface in situ and remote-sensing observations at the SGP site provide an overall climatology and their interactions from ground surface up to the boundary layer. Measurements from a 60 m meteorological tower, surface flux stations, disdrometers, soil temperature and moisture flux plates, coherent Doppler lidar, Raman lidar, radiosondes, and satellite data at the SGP central facility were analyzed. All the measurements were generally made within a few square kilometers of each other at the central facility.

This report focuses on data from January 2010 to June 2020 at the SGP central facility. The various sections describe the ARM SGP site and surrounding wind turbines; in situ and remote-sensing instrumentation used in the report; mathematical equations to analyze fluxes, turbulence, and other boundary layer parameters; a climatological analysis of surface winds, fluxes and thermodynamic parameters for several years; an analysis of observed winds in the framework of Monin-Obukhov Similarity Theory (MOST); an analysis of the boundary layer winds and direction from a Doppler lidar; multi-year turbulence estimates through the boundary layer from a Doppler lidar; atmospheric boundary layer water vapor and relative humidity profiles from Raman lidar; cloud base height and boundary layer height from multiple sensors and satellite data; and finally site-specific atmospheric conditions, such as nocturnal low-level jets.

Diurnal, seasonal, and yearly variations of surface, subsurface and boundary layer quantities, such as wind speed, direction, temperature, atmospheric stability, soil temperature, and various atmospheric fluxes at the SGP site, showed distinct trends useful for focused modeling studies. The applicability of surface similarity theory to ARM SGP data was also evaluated, which showed that, compared to southerly flows, northerly flows are aligned with MOST estimates. Boundary layer winds and direction profiles for several years from a Doppler lidar showed a consistent presence of a nocturnal low-level jet and predominant southerly wind directions through the boundary layer at the SGP site. The interannual variability at the SGP site is low (<3.5 percent), and has a mean annual wind speed of approximately 7 m s^{-1} at 100 m AGL. Boundary layer turbulence and moisture transport from Doppler and Raman lidars are evaluated, which provides evidence of increased water vapor mass flux into the great plains during nocturnal low-level jets. The moisture flux from nocturnal low-level jets is observed to be at its maximum during summer periods. A novel machine learning algorithm is implemented to accurately estimate the planetary boundary layer height, providing further insights into the growth and destruction of the convective boundary layer height during various seasons and land–atmosphere conditions. The high frequency of low-level clouds during winter, spring, and fall seasons is validated using the multi-sensor array and satellite estimates of cloud top height. Satellite vegetative fraction data provide insight into seasonal surface roughness and vegetation variability around the SGP site.

The analyses contained in this report provide a great deal of new information about atmospheric conditions at the SGP site. It is expected to aid future field campaigns around the SGP site and associated modeling studies. In addition, perspectives from ARM instrument mentors, translators, and the experience gained will inform the configuration and analysis of data for future analyses at the SGP site.

The SGP site staff captured a photo of a spectacular "shelf" cloud on May 13, 2005, during the area's annual severe storm season (Figure S.1). This storm was reported to drop nickel-sized hail and carry winds up to 75 mph.



Figure S.1. Spectacular "shelf" cloud image captured at the SGP site. Visible in the foreground is the ARM user facility's 60-meter tower and Radiometer Calibration Facility. (Image courtesy of the U.S. Department of Energy ARM user facility.)

Acknowledgments

The authors acknowledge the ARM field technicians, scientist and mentors for collecting, processing and archiving the field data. Their work is instrumental in any research involving ARM data. The authors also thank the AWAKEN (American Wake Experiment) project and ARM for supporting this research work. Pacific Northwest National Laboratory is operated by Battelle Memorial Institute for the DOE under Contract DE-AC05-76RL01830.

Acronyms and Abbreviations

AGL	above ground level
ARM	atmospheric radiation measurement
AWAKEN	American wake experiment
CBH	cloud base height
CBL	convective boundary layer
DJF	December, January, February
DLWSTATS	Doppler lidar vertical velocity (W) statistics
DLWIND	Doppler lidar wind speed
DOE	U.S. Department of Energy
EBBR	energy balance Bowen ratio
IAV	interannual variability
IEC	International Electrotechnical Commission
JJA	June, July, August
Lidar	light detection and ranging
LES	large-eddy simulation
LLJ	low-level jet
LHF	latent heat flux
MAM	March, April, May
MCS	mesoscale convective system
MF	momentum flux
MODIS	moderate resolution imaging spectroradiometer
MOST	Monin–Obukhov similarity theory
PBL	planetary boundary layer
PBLH	planetary boundary layer height
PECAN	Plains Elevated Convection at Night
PPI	planned position indicator
RF	Random Forest
RH	relative humidity
RHI	range height indicator
RMS	root mean square
RMSE	root mean square error
SGP	Southern Great Plains
SHF	sensible heat flux
SNR	signal-to-noise ratio
SMOS	Surface Meteorological Observation System
SON	September, October, November

TKE	turbulence kinetic energy
TI	turbulence intensity
UTC	Coordinated Universal Time
VAP	value-added product
VC	vegetation cover
WRF	weather research forecasting

Contents

Summary	ii
Acknowledgments.....	iv
Acronyms and Abbreviations.....	v
Contents	vii
1.0 Introduction	1
2.0 ARM SGP Site.....	3
2.1 Site Layout.....	3
2.2 Wind Farms Near the SGP Site	3
3.0 Instrumentation.....	5
3.1 Surface Instruments.....	5
3.2 Doppler Lidars	6
3.3 Raman Lidars	8
3.4 Radiosondes.....	9
3.5 Satellite Products.....	10
4.0 Mathematical Preliminaries and Algorithms	11
4.1 Surface Fluxes and Atmospheric Stability	11
4.2 Similarity Functions.....	12
4.3 Remote-Sensing Turbulence	13
4.4 Boundary Layer Height Using Machine Learning	14
5.0 Atmospheric Boundary Layer at the SGP	18
5.1 Subsurface and Surface Layer.....	18
5.2 Boundary Layer Horizontal Wind Profiles.....	32
5.3 Boundary Layer Turbulence.....	38
5.4 Boundary Layer Water Vapor and Relative Humidity	44
5.5 Cloud Base and Boundary Layer Height	54
6.0 Site-specific Atmospheric Conditions.....	65
6.1 Nocturnal Low-level Jet.....	65
6.2 Spatial Variability	67
6.3 Possible Effect of Wind Farms at the SGP Site.....	71
7.0 Perspectives and Future Work	74
8.0 References.....	75
Appendix A – Instrument List	A.1
Appendix B – Code and Plot Repository	B.1

Figures

Figure 1.	(Left) ARM Southern Great Plains central facility (C1) measurement site aerial view and sample set of instrument locations, and (right) the location of the central facility (C1) and extended facilities (E32, E37, E39, and E41).....	3
Figure 2.	Wind farms surrounding the ARM SGP central facility and other satellite sites (approximate locations in yellow star). The filled circles represent surrounding wind turbines in the area. The name of each wind farm (number of turbines), starting year of production, and total wind farm wattage are also provided for each wind farm.	4
Figure 3.	Suite of subsurface and surface measurements at the SGP central facility. (a) 60 m meteorological tower, (b) soil temperature and moisture flux measurements, (c) disdrometer, (d) energy balance Bowen ratio (EBBR) station, (e) surface meteorological observation system (SMOS), and (f) eddy-correlation (ECOR) flux station. (Images courtesy of the U.S. Department of Energy ARM user facility.)	6
Figure 4.	ARM Doppler lidar (inset) installed at the SGP central facility. (Image courtesy of the U.S. Department of Energy ARM user facility.)	7
Figure 5.	Yearly Doppler lidar range availability from 2012 to 2019 from filtered wind profile VAP estimates at the SGP C1.....	8
Figure 6.	Raman lidar installed on top of a shipping container at the SGP central facility in 2007. (Image courtesy of the U.S. Department of Energy ARM user facility.).....	9
Figure 7.	New generation radiosondes released on October 12, 2013, at the SGP site. (Image courtesy of the U.S. Department of Energy ARM user facility.)	10
Figure 8.	Graphical representation of a standard random forest algorithm (recreated from online sources and observed sample trees from current RF model).....	16
Figure 9.	Vertical velocity variance on June 8, 2019, at the SGP C1 site with boundary layer height estimates from the lidar thresholding technique (orange dotted line, Tucker et al. 2009), random forest model (black dotted line), and radiosondes (filled circles, Sivaraman et al. 2013). Cloud base height estimates from the lidar are also shown (filled squares).....	17
Figure 10.	Hourly averaged (UTC time) surface (a) air temperature, (b) friction velocity (u^*) and turbulence kinetic energy (TKE), and (c) moisture and heat flux at 10 m AGL from 2010 to 2019 at the SGP C1 site.	19
Figure 11.	Friction velocity vs atmospheric stability (z/L) at the SGP C1 site for stable stratification.	20
Figure 12.	Seasonal hourly averaged surface air temperature and, TKE for four seasons—DJF, MAM, JJA and SON—from 2010 to 2019 at the SGP C1 site.	21
Figure 13.	Turbulence intensity versus average wind speed at 25 m AGL from 2012 to 2019 at the SGP C1 site. The dashed lines show IEC curves for various input turbulence intensity estimates defined in IEC 61400-1.....	22

Figure 14. Atmospheric stability classification from 2012 to 2019 (a) monthly averaged, (b) hourly averaged, (c) wind speed bins, and (d) and (e) hourly averaged for southerly and northerly wind direction sectors, respectively at the SGP C1 site.....24

Figure 15. Atmospheric stability as a function of wind direction from 2012 to 2019 at the SGP C1 site.25

Figure 16. Dimensionless wind shear ϕ_m as a function of dimensionless stability z/L from 2010 to 2014. Data with $u^* < 0.1$ and $U < 3$ m/s were filtered. The average of all data points is given by ϕ_m26

Figure 17. Same as Figure 16, but for the southerly and northerly wind direction regimes.....27

Figure 18. Average surface roughness (z_o) in near-neutral conditions ($|L| > 500$) at the SGP C1 site versus (a) wind direction and (b) monthly at 25 m AGL from 2003 to 2014.28

Figure 19. (a), (c) Hourly and (b), (d) monthly averaged soil temperature from 2012 to 2019 and soil specific water content profiles from 2012 to 2015 at the SGP C1 site.29

Figure 20. Difference between soil temperature and air temperature (a) hourly and (b) monthly from 2015 to 2019 at the SGP C1 site.30

Figure 21. (a) Monthly and (b) hourly median precipitation rate from 2012 to 2019 at the SGP C1 site.31

Figure 22. Hourly averaged longwave (L), shortwave (K) and Net (Q) radiation from 2012 to 2019 at the SGP C1 site.....32

Figure 23. Hourly averaged wind speed (U) and direction from (a) 2014, (b) 2015, (c) 2016, (d) 2017, (e) 2018, and (f) 2019 at the SGP central facility. Colored contours of U for every 1 ms^{-1} are displayed and the arrows depict mean wind direction. Data availability below 5 percent is not shown. Apart from year 2016 (subplot c), which had 15 percent less data availability compared to other years, all other years had a minimum data availability greater than 95 percent at the lowest range-gate of 90 m.33

Figure 24. Hourly averaged wind speed (U) and direction seasonal variability from 2013 to 2019 at the SGP central facility. Colored contours of U for every 1 ms^{-1} are displayed and the arrows represent mean wind direction. Measurements at higher altitudes are irregular due to reduced data availability.34

Figure 25. Time series of hourly, daily, monthly, and yearly averaged windspeeds (U) at 90 m AGL from 2013 to 2019 at the SGP central facility.....35

Figure 26. Average wind rose at (a) 90 m and (b) 500 m AGL from 2013 to 2019 at the SGP central facility.....36

Figure 27. Monthly averaged (January to December [a to l]) wind rose at 90 m AGL from 2013 to 2019 at the SGP central facility.37

Figure 28. Wind speed and direction distribution at 90 m AGL from 2013 to 2019 at the SGP central facility.....38

Figure 29. Hourly averaged vertical velocity variance (σ_w^2) from 2014 to 2019 (a-f) at the SGP central facility. Colored contours of σ_w^2 for every $0.1 \text{ m}^2\text{s}^{-2}$ are displayed. Data availability below 5 percent are not displayed. Apart

from year 2016 (subplot c), which had 15 percent less data availability compared to other years, all other years had a minimum data availability greater than 95 percent at the lowest range-gate of 90 m.40

Figure 30. Hourly averaged vertical velocity variance (σ_w^2) estimates for (a) DJF, (b) MAM, (c) JJA, and (d) SON seasonal variability from January 2016 to May 2020 at the SGP central facility. Colored contours of σ_w^2 for every 0.1 ms^{-1} are displayed. Data availability below 5 percent are not displayed.41

Figure 31. (a) Average horizontal wind speed and (b) average vertical velocity variance from 20 June to 27 June, 2018 at the SGP central facility.42

Figure 32. Average vertical velocity variance rose at (a) 105 m and (b) 500 m AGL from 2012 to 2019 at the SGP central facility.43

Figure 33. Hourly averaged eddy dissipation rate from vertical stares from January 2013 to May 2020 at 100 m AGL for seasons (a) DJF, (b) MAM, (c) JJA, and (d) SON at the SGP central facility.43

Figure 34. Diurnal mean (a) water vapor mixing ratio and (b) RH as observed by the Raman lidar at the SGP central facility for December, January, and February (DJF), 2016 through 2020. Panel (c) shows the diurnal mean mixing ratio (black) and RH (red) from the met station at 10 m AGL. Sunrise (~13:30 UTC) and sunset (~23:10 UTC) times are indicated by the dashed vertical lines in each panel.45

Figure 35. Diurnal mean (a) water vapor mixing ratio and (b) RH as observed by the Raman lidar at the SGP central facility for March, April, and May (MAM), 2016 through 2020. Panel (c) shows the diurnal mean mixing ratio (black) and RH (red) from the met station at 10 m AGL. Sunrise (~11:50 UTC) and sunset (~01:00 UTC) times are indicated by the dashed vertical lines in each panel.46

Figure 36. Diurnal mean (a) water vapor mixing ratio and (b) RH as observed by the Raman lidar at the SGP central facility for June, July, and August (JJA), 2016 through 2020. Panel (c) shows the diurnal mean mixing ratio (black) and RH (red) from the met station at 10 m AGL. Sunrise (~11:20 UTC) and sunset (~01:45 UTC) times are indicated by the dashed vertical lines in each panel.47

Figure 37. Diurnal mean (a) water vapor mixing ratio and (b) RH as observed by the Raman lidar at the SGP central facility for September, October, and November (SON), 2016 through 2020. Panel (c) shows the diurnal mean mixing ratio (black) and RH (red) from the met station at 10 m AGL. Sunrise (~12:30 UTC) and sunset (~23:10 UTC) times are indicated by the dashed vertical lines in each panel.48

Figure 38. Diurnal mean (a) water vapor mass flux and (b) wind direction as observed by the Raman and Doppler lidars at the SGP central facility for DJF, 2016 through 2020. The dotted lines indicate sunset (~23:10 UTC) and sunrise (~13:30 UTC).49

Figure 39. Diurnal mean (a) water vapor mass flux and (b) wind direction as observed by the Raman and Doppler lidars at the SGP central facility for MAM, 2016 through 2020. The dotted lines indicate sunrise (~11:50 UTC) and sunset (~01:00 UTC).50

Figure 40. Diurnal mean (a) water vapor mass flux and (b) wind direction as observed by the Raman and Doppler lidars at the SGP central facility for JJA, 2016 through 2020. The dotted lines indicate sunrise (~11:20 UTC) and sunset (~01:45 UTC).....51

Figure 41. Diurnal mean (a) water vapor mass flux and (b) wind direction as observed by the Raman and Doppler lidars at the SGP central facility for SON, 2016 through 2020. The dotted lines indicate sunrise (~12:30 UTC) and sunset (~23:10 UTC).....52

Figure 42. Median water vapor mass flux as a function of wind direction at 10 m AGL for (a) DJF, (b) MAM, (c) JJA, and (d) SON at the SGP central facility. Statistics were computed using data from January 2016 to June 2020.53

Figure 43. Cloud base height distributions as observed from a Doppler lidar from 2012 to 2019 at the SGP central facility.55

Figure 44. Hourly averaged cloud base height distributions over four seasons: (a) DJF, (b) MAM, (c) JJA, and (d) SON from 2012 to 2019 at the SGP central facility.56

Figure 45. Box plot seasonal distribution of (a) cloud fraction and (b) cloud top minimum altitude from 1° resolution pixel over the ARM SGP site. The statistics are developed using 20 years of MODIS monthly averaged data. The center line within the box indicates the median value of the distribution, the box represents the first and third quartile of the distribution, the error bars represent the minimum and maximum values in the distribution, and the markers indicate outliers not used in the calculation of the median value.57

Figure 46. Median diurnal CBH from RL for four seasons (a) DJF, (b) SON, (c) JJA, and (d) MAM from 2016 to 2020 at the SGP central facility. The dotted lines indicate sunset and sunrise for each season.58

Figure 47. Distributions of cloud base height estimates from the Raman lidar during the period from January 2016 to June 2020 at the SGP central facility.59

Figure 48. Hourly averaged RF-estimated boundary layer heights from 2016 to 2019, overlaid on vertical velocity variance estimates (shown in Figure 29) at the SGP central facility.....60

Figure 49. Hourly averaged PBLH from RF algorithm from 2015 to 2019 at the SGP central facility.61

Figure 50. Hourly averaged boundary layer height over four seasons at the SGP central facility from 2015 to 2019 at the central facility.62

Figure 51. Daily maximum boundary layer height averaged from 2015 to 2019 for each season at the SGP central facility. The red horizontal line indicates the median value, the blue box represents the 25th and 75th percentile levels, and the + signs indicate possible outliers not used to calculate the median value.63

Figure 52. Partial dependence between PBL height, RH, and soil temperature using RF regression.64

Figure 53. Daily median low-level jet height distributions at the SGP central facility from May 2011 to May 2020 for (a) DJF, (b) MAM, (c) JJA, and (d) SON.....66

Figure 54. Daily maximum low-level jet wind speed distributions at the SGP central facility from May 2011 to May 2020 for (a) DJF, (b) MAM, (c) JJA, and (d) SON.....67

Figure 55. Temporal variation in vegetation cover (%) from Fraction of Photosynthetically Active Radiation using MODIS satellite observations for the year 2018. The observations at an interval of ~8 days are compiled over 1x1 deg domain with the SGP C1 site at the center.68

Figure 56. Spatial distribution of vegetation cover (%) from Fraction of Photosynthetically Active Radiation using MODIS satellite observation for (a) a day in the first week of January and (b) a day in last week of June 2018 (right panel). The VC>100 filled values are from water, rocks, or other non-vegetated reflected pixels. The SGP site is shown in the middle of map by a black star enclosed in a square.68

Figure 57. Spatial variability at four satellite ARM sites (E32, E37, E39, and E41) showing monthly averaged (a) soil moisture, (b) soil temperature, and (c) soil-air temperature difference from 2014 to 2019.70

Figure 58. Hourly averaged random forest estimated PBLH in 2019 at E32, E37, E39, and C1.....71

Tables

Table 1. Wind direction sectors, distance, and turbine characteristics affecting measurements at the SGP.....4

Table 2. Stability classification based on L thresholds12

Table 1-A: SGP Instrument list, data stream, parameters used for the analysis. A.1

1.0 Introduction

The objective of Atmospheric Radiation Measurement (ARM) User Facility and U.S. Department of Energy (DOE) Office of Science is to engage the research community to aid in understanding critical processes associated with aerosol, boundary layer, clouds, and precipitation using long-term observational data sets. As a source of high-quality, long-term atmospheric measurements, the ARM Southern Great Plains (SGP) site in Oklahoma, USA, is a huge resource to the research community. The ARM data are predominantly used to study aerosol and cloud processes in the atmosphere. A recent triennial review of the ARM program highlighted the need for collaboration with a broader user base beyond the DOE Office of Biological and Environmental Research and the Atmospheric System Research program. Because of the recent growth of wind farms around the SGP site, the ARM data set is an invaluable asset for wind farm developers in the region and elsewhere. The long-term archive provides profiles of wind and other thermodynamic parameters essential for accurately assessing the annual energy production of a wind farm. The data also provide key information for studying the impact of local conditions, such as low-level jet, nocturnal convection initiation, mesoscale convective storms, down bursts, etc., on wind farms and for validating microscale and mesoscale models.

Oklahoma ranks third in U.S. for installed wind capacity, providing over 8,000 MW of electricity, which represents more than \$15 billion in capital investment (AWEA 2019). In 2019, the state generated approximately 35 percent of its electricity from wind energy, the third highest in the country, and provided enough electricity to power the equivalent of 2.6 million average U.S. homes. Although the landscape and topographic flows around the SGP site are relatively simple compared to complex terrain sites, wind forecasting at the SGP has other challenges (Yang et al. 2006; Gibbs et al. 2011; Berg et al. 2015; Berg and Lamb 2016). Very few studies have focused on wind power forecasting and the effect of wind farm microclimate on the surrounding region (Wharton et al. 2013). The upcoming American Wake Experiment (AWAKEN), funded by the DOE, is focused on studying the dynamics of wind farm/turbine wakes. This multi-institutional field campaign is expected to begin in 2022 near the SGP. This report provides a multi-year assessment of the atmospheric boundary layer at the SGP site to aid future field campaigns and modeling studies conducted around the SGP site.

The ARM SGP site has one central facility (C1) and several extended/satellite measurement locations (Mather and Voyles 2013). The four cornerstone satellite measurement locations (E32, E37, E39, and E41) used in this study are shown in Figure 1. The satellite measurement locations are approximately 50 miles from the central facility. Instruments ranging from in situ sensors, such as eddy covariance flux systems, to advanced remote-sensing systems, such as scanning Ka-band Doppler radar, Raman lidar, and ceilometers, have been deployed and operational for more than 20 years. The first scanning coherent Doppler lidar was acquired in 2010 and was deployed at C1 on October 15, 2010. Since then, four other scanning Doppler lidars were deployed in all the satellite sites around the SGP site in May 2016. Measurements from Doppler lidars have been instrumental in understanding various boundary layer processes, especially updrafts and downdrafts during convective conditions, and the vertical structure of the vertical velocity, winds, and direction from the land surface up to the boundary layer (Mather and Voyles 2013; Berg et al. 2017; Geerts et al. 2017; Lareau et al. 2018). Raman lidars have been operational at the SGP site since 1996. Raman lidars provide accurate height- and time-resolved measurements of water vapor mixing ratio, temperature, aerosol, and cloud optical properties with good vertical and temporal resolution. Ground-based thermodynamic profiles are essential for improving the skill of weather forecasting models (Lareau 2019; Liu 2019). There is

a shortage of humidity and temperature measurements to accurately characterize mesoscale processes, which is needed to better understand the water and energy cycles (Newsom et al. 2019). Data from satellites above the SGP site (for example, from the Moderate Resolution Imaging Spectroradiometer), have also been used to explain the spatial variability of key surface and cloud properties (Wang and Zhao 2017). In this report, measurements from several in situ sensors, Doppler lidars, Raman lidars, radiosondes, and satellite data centered around the SGP central measurement facility (C1) are used to study the long-term evolution of the atmospheric boundary layer. Key atmospheric variables, such as surface fluxes, atmospheric stability, surface radiation, winds and direction from ground surface to top of the boundary layer, boundary layer height, eddy dissipation rate, vertical velocity variance profiles, water vapor mixing ratio, and temperature profiles were evaluated. Yearly, seasonal, and diurnal profiles of various atmospheric profiles provide a holistic view of long-term variability at the SGP site. This will provide appropriate information for future field campaign organizers to target certain atmospheric features specific to the SGP site, such as the nocturnal low-level jet, nocturnal convection initiation, and to understand the seasonal and diurnal variability of planetary boundary layer heights, turbulence, atmospheric stability, etc. These profiles will also provide the necessary boundary and initial conditions required for developing idealized simulations around the SGP site.

Data from the ARM SGP site are publicly available for download and near real-time visualization using the ARM Data Discovery tool (<https://adc.arm.gov/discovery/>). All the data have been quality controlled and are available in standard NetCDF formats, which can now be read by many standard processing software (such as MATLAB, Python, IDL etc). NetCDF data sets also contains the metadata for each site along with the nearest flux tower, meteorological station, and ceilometer measurements, and some value-added products (VAPs), which have similar products and supporting information that would aid in filtering the data. We have also made available relevant MATLAB codes used in this analysis to read and process ARM data sets on GitHub (see Appendix B for more details).

This report is organized as follows: Section 2 provides details about the ARM SGP site; Section 3 provides details about the instrument data being used in this analysis; and mathematical expressions of certain atmospheric parameters are introduced in Section 4. A multi-year assessment of atmospheric boundary layer at the SGP site is provided in Section 5, site-specific atmospheric conditions are discussed in Section 6, and perspectives and conclusions are provided in Section 7.

2.0 ARM SGP Site

The SGP site was the first field measurement site established by DOE's ARM program in 1992. It is considered the largest and most extensive climate research field site in the world. Scientists are using measurements obtained at the SGP site to improve the understanding and representation of the atmospheric boundary layer, clouds, and aerosols, as well as their interactions, in climate and Earth system models.

2.1 Site Layout

The site has a relatively homogeneous terrain and experiences a wide variety of cloud types and surface flux properties. This area also experiences widespread seasonal variation in temperature and humidity. Surrounding federal and state research centers in Oklahoma also provide a good opportunity to collaborate and make use of the large network of weather and climate research instrumentation at the SGP site. As shown in Figure 1 the central facility provides a variety of atmospheric variables from a suite of remote-sensing and surface measurements.

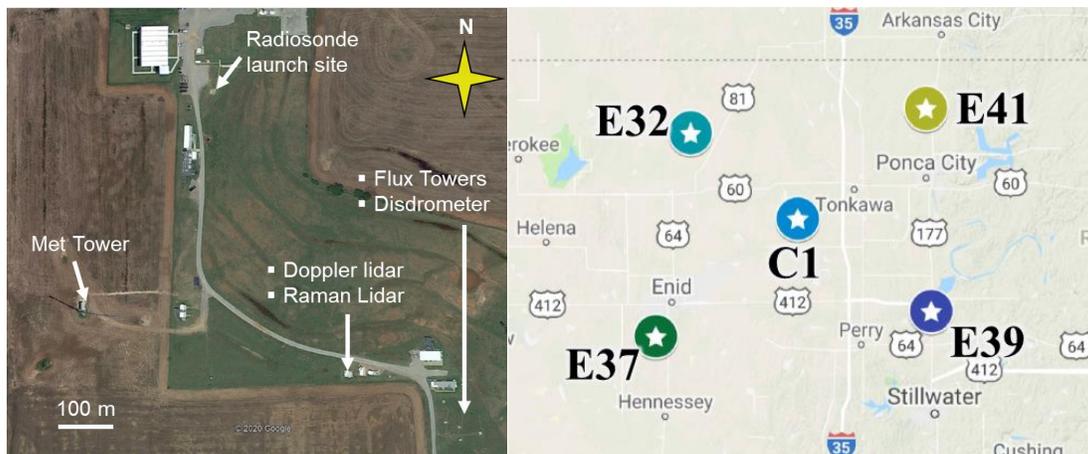


Figure 1. (Left) ARM Southern Great Plains central facility (C1) measurement site aerial view and sample set of instrument locations, and (right) the location of the central facility (C1) and extended facilities (E32, E37, E39, and E41).

2.2 Wind Farms Near the SGP Site

Although the ARM data sets are predominantly used for studying aerosol and cloud interaction, the surrounding wind farms near the SGP site makes it an interesting location at which to study wind farm interactions and the modification of the atmospheric boundary layer caused by wind farm wakes (Newsom and Krishnamurthy 2020). Figure 2 shows the locations of all the wind turbines surrounding the SGP site, and Table 1 shows the approximate wind directions, distances from the nearest wind turbine to the lidar locations, and wind turbine characteristics within each wind direction sector. Except for lidars at E37 and E39, the other three Doppler lidars in the SGP network are within reasonable distances to be affected by surrounding wind farm wakes.

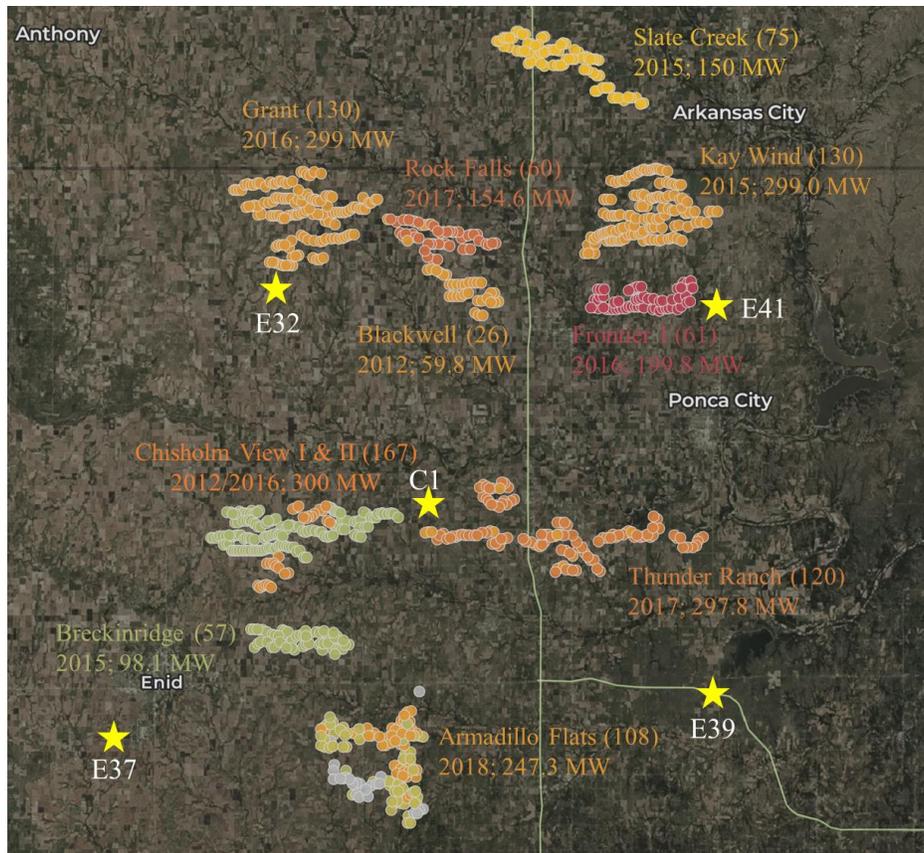


Figure 2. Wind farms surrounding the ARM SGP central facility and other satellite sites (approximate locations in yellow star). The filled circles represent surrounding wind turbines in the area. The name of each wind farm (number of turbines), starting year of production, and total wind farm wattage are also provided for each wind farm.

Table 1. Wind direction sectors, distance, and turbine characteristics affecting measurements at the SGP site. No information is provided when the nearest wind farm is at least 20 km away from the measurement site.

Site	Wind Direction Sectors (degrees)	Approximate Distance of the Nearest Wind Turbine to the Site (m)	Common Wind Turbine Height (m)	Rotor Diameter (m)	Type of Turbine	Built Year
C1	67 - 93	6700	90	116	GE 2.5 MW	2017
	112 - 196	3500	80	116	GE 2.3 MW	2017
	243 - 270	4600	80	82.5	GE 1.68 MW	2012
E32	45 - 60	11500	80	108	Siemens 2.3 MW	2016
E37	--	> 20000	--	--	--	--
E39	--	> 20000	--	--	--	--
E41	205 - 255	2500	87	126	Vestas V126-3.3	2016
	295 - 15	5000	80	108	Siemens 2.3 MW	2015

3.0 Instrumentation

Wind velocity, temperature, relative humidity, and turbulence are key parameters that are measured during a wind resource assessment campaign. At the SGP site, a variety of instruments provide a long-term archive of atmospheric parameters from the subsurface to the stratosphere. Measurements from 2010 to 2020 are analyzed in this report, depending on instrument and data availability. Below, a short summary of instruments used in this report, their expected accuracy and availability, is also provided.

3.1 Surface Instruments

A variety of surface-based instruments are deployed at the SGP C1 site to measure standard meteorological parameters (such as wind speed, temperature, relative humidity, etc.). The surface meteorological observation system (SMOS) provides 10-minute average surface wind speed, wind direction, air temperature, relative humidity (RH), barometric pressure, and precipitation at approximately 3 m above ground level (AGL) (Ritsche 2008). Surface measurements from the energy balance Bowen ratio (EBBR) and eddy-correlation (ECOR) flux stations were used to calculate atmospheric fluxes (Cook and Sullivan 2020). Fluxes were calculated using a 30-minute average mean estimate of winds and temperature from the EBBR sensors. Precipitation measurements from either laser disdrometer data or an optical rain gauge, depending on their availability, was used in the analysis below (Bartholomew 2016). Soil moisture and temperature were measured from flux plates at 5, 10, 20, 50, and 75 cm below the surface (Cook 2016). The 60 m meteorological tower at the SGP site has measurements of winds (from sonic anemometers) at 4 m and 25 m AGL, and temperature and RH at 4 m, 25 m, and 60 m AGL. A collage of all surface instruments at the SGP site is shown in Figure 3.

Because of the lack of instrument availability during certain years and time periods, multiple instruments were used to create the multi-year database. Appropriate quality control procedures, as recommended by the instrument mentors, was performed prior to merging multiple data sources. Although the uncertainty of each instrument is different, this is not expected to make a difference to the overall assessment of the climatology at the SGP site. All the data sets were averaged or interpolated to 15-minute increments. The data availability varied each year, and only instruments with data availability greater than 70 percent for any given year were chosen for use in this analysis. Limited surface-based instruments are deployed at the extended facilities.

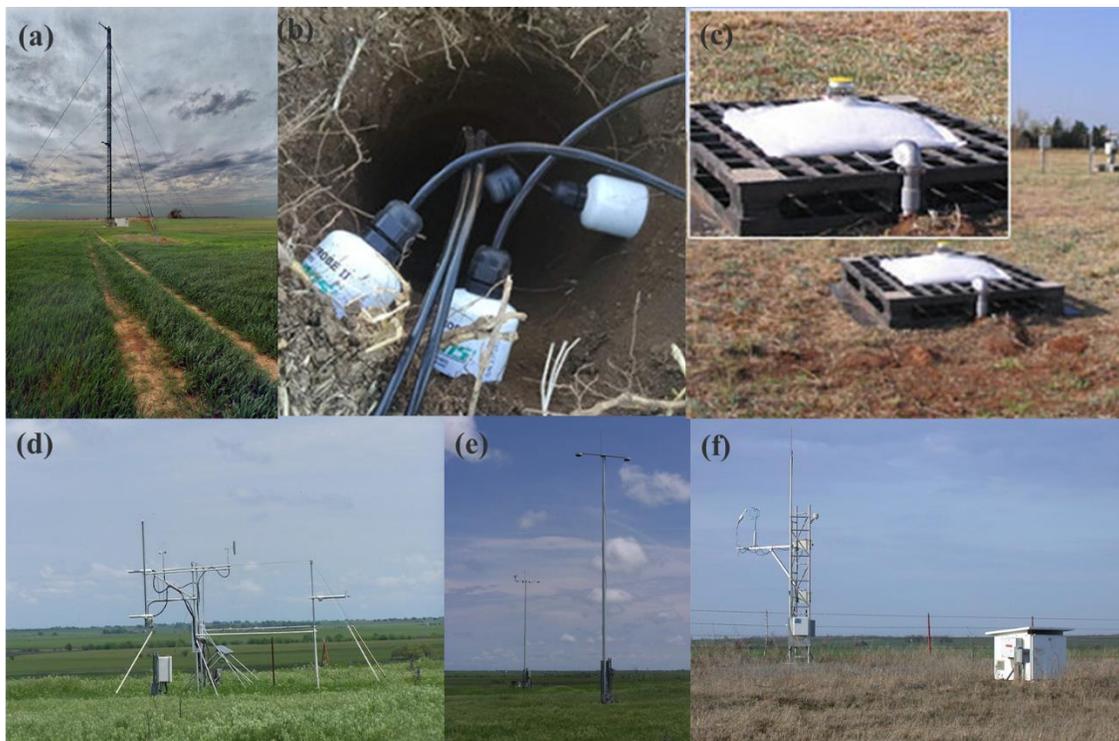


Figure 3. Suite of subsurface and surface measurements at the SGP central facility. (a) 60 m meteorological tower, (b) soil temperature and moisture flux measurements, (c) disdrometer, (d) energy balance Bowen ratio (EBBR) station, (e) surface meteorological observation system (SMOS), and (f) eddy-correlation (ECOR) flux station. (Images courtesy of the U.S. Department of Energy ARM user facility.)

3.2 Doppler Lidars

Doppler lidars provide range-resolved measurements of attenuated backscatter, signal-to-noise ratio, and radial velocity (Pearson et al. 2010; Newsom and Krishnamurthy 2020). At the SGP C1, the Doppler lidar is placed on top of a shipping container, as shown in Figure 4. The ARM Doppler lidars are operated on a fixed scan schedule consisting of vertical stares and plan position indicator (PPI) scans. The lidars predominantly stare vertically up and perform occasional PPI scans. The PPI scans are performed once every 10 or 15 minutes (depending on the site) and take approximately 40 seconds to complete one cycle. The vertical stares provide measurements of clear-air vertical velocity profiles in the lower troposphere with a temporal resolution of about 1 second and a height resolution of 30 m. Several VAPs are provided as a standard output from the Doppler lidars (Newsom et al. 2015; Newsom et al. 2019)— horizontal wind speed, wind direction, vertical velocity variance, vertical velocity skewness, vertical velocity kurtosis, cloud base height, cloud fraction, and cloud base vertical velocity (to name a few). Raw lidar measurements and associated VAPs from the SGP C1 lidar for 2010 to 2020 are currently available. At the satellite sites (E32, E37, E39 and E41), lidar measurements from 2016 to 2020 are currently available.

Doppler lidars are frequently used in the wind energy industry for wind resource assessment, wind turbine control, and continued site assessment. The accuracy of average lidar-derived wind speed and direction measurements has been thoroughly validated in various field campaigns and consultant reports, with an uncertainty less than 2 percent (Krishnamurthy et al.

2013; Lundquist et al. 2015; Lundquist et al. 2017; Vassallo et al. 2021). Lidars are also included in the latest edition of International Electrotechnical Commission (IEC) Standard IEC 61400-12-1 (2017) as one of the standard instruments for power curve assessment and can be used as a stand-alone device. Because lidars provide vertical profiles of winds through the rotor-swept area, they are an attractive and cheaper alternative for deployment in challenging locations (offshore/complex terrain).



Figure 4. ARM Doppler lidar (inset) installed at the SGP central facility. (Image courtesy of the U.S. Department of Energy ARM user facility.)

Figure 5 shows the yearly average data availability from 2012 to 2019 from the Doppler lidar wind (DLWIND) VAP at the SGP central facility. The data availability is the percentage of measurements with valid wind speed and direction retrieval and signal-to-noise ratio (SNR) > 0.008 (Newsom et al. 2015). As shown in Figure 5, lidar availability is variable over the course of the period (2012 to 2019), and the lidars observe ~1 km at least 50 percent of the time at the SGP C1 site.

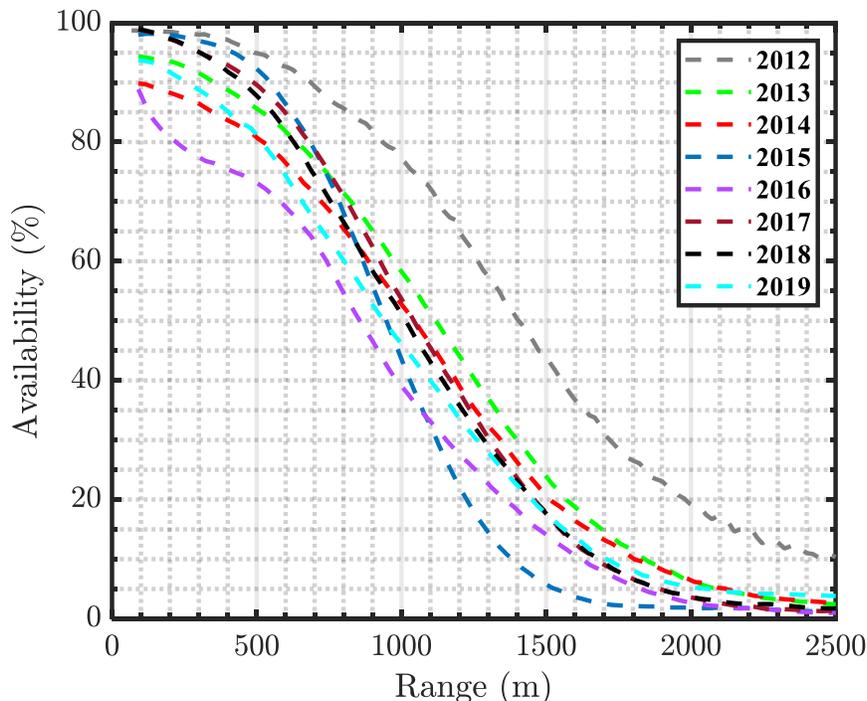


Figure 5. Yearly Doppler lidar range availability from 2012 to 2019 from filtered wind profile VAP estimates at the SGP C1.

3.3 Raman Lidars

The ARM program has operated a Raman lidar (RL) at the SGP central facility since 1996 (Goldsmith et al. 1998; Turner and Goldsmith 1999; Newsom et al. 2009; Newsom et al. 2013; Turner et al. 2016). This system operates at a wavelength of 355 nm, with a pulse energy of 300 mJ, and a pulse repetition frequency of 30 Hz. The design incorporates several detection channels for sensing Raman backscattered radiation from atmospheric N_2 , O_2 , and H_2O , as well as elastic backscatter from aerosol. Figure 6 shows the location of the RL at the SGP C1, where it is installed on top of a shipping container.

Processing the signals from the various detection channels enables retrieval of several quantities, including water vapor mixing ratio (q), temperature, and various cloud and aerosol optical properties, such as extinction and depolarization ratio. For this study we are interested in the RL water vapor measurement, which is derived from the ratio of the water vapor channel backscatter at 408 nm to the nitrogen channel backscatter at 387 nm. The raw ratio is then corrected for range-dependent effects, such as two-way molecular attenuation and geometrical overlap, as described by Newsom et al. (2020). The corrected ratio profile is then calibrated using simultaneous radiosonde measurements.

The water vapor channel at 408 nm is quite sensitive to solar radiation. The result is that measurement error increases and the range performance degrades somewhat during the daytime. At night the RL can make valid water vapor mixing ratio measurements up to a distance of, and in some cases exceeding, 10 km. During the day, the maximum range reduces to about 4 to 6 km, and sometime lower depending on conditions.



Figure 6. Raman lidar installed on top of a shipping container at the SGP central facility in 2007. (Image courtesy of the U.S. Department of Energy ARM user facility.)

3.4 Radiosondes

Radiosondes provide in situ vertical profiles of wind speed, direction, temperature, RH, and pressure of the atmosphere. At the SGP C1 site, radiosondes are launched four times daily, nominally at 0530, 1130, 1730, and 2330 UTC each day. During select field campaigns, additional radiosonde data sets are made available (generally once every 3 hours). In addition to standard parameters, radiosonde VAPs also provide three different estimates of boundary layer height. An evaluation of the various boundary layer height estimates is provided by Sivaraman et al. (2013). Figure 7 shows the radiosonde launch pad at the SGP C1 site. In this report, the radiosonde data are used to train the boundary layer height random forest model (see Section 4.4).



Figure 7. New generation radiosondes released on October 12, 2013, at the SGP site. (Image courtesy of the U.S. Department of Energy ARM user facility.)

3.5 Satellite Products

Moderate Resolution Imaging Spectroradiometer (MODIS) satellite data products have been widely used to understand the spatial variability of cloud fraction, cloud top and base height, and vegetation index fraction estimates. Compared to Geostationary Satellite (GOES), sun-synchronous satellites measurements (e.g., MODIS) have several additional spectral bands that are highly sensitive to cloud, air, and surface properties, thereby enabling comprehensive surface and atmospheric parameters at a global scale. In contrast to GOES, MODIS satellites do not provide a continuous time series of data at a given location. A comparison of MODIS cloud fraction estimates with in situ ground sensors at the SGP site at around noon revealed that monthly averaged MODIS satellite estimates exhibit a very good correlation (97 percent) with a root mean square error (RMSE) of less than 3 percent (Wang and Zhao 2016). Therefore, a monthly averaged estimate of cloud properties can represent the SGP site reasonably well. High-resolution vegetative index fraction (horizontal grid resolution of 500 m) estimates provide insight into the spatial variability of surface roughness during various seasons at the SGP site. Satellite measurements are key for understanding spatial variability at a given site. In this report, MODIS data for 20 years (from 1999 to 2019) are evaluated.

4.0 Mathematical Preliminaries and Algorithms

Here, we provide some mathematical relations related to turbulence and other boundary layer parameters, which are used in following sections.

4.1 Surface Fluxes and Atmospheric Stability

We define the wind vectors as $U = (u, v, w)$, such that u (longitudinal component) is in the mean wind direction, v (transverse component) is perpendicular to the mean wind direction, and w is the vertical direction (positive upwards). Wind velocity fluctuations (u', v', w') are estimated by subtracting the instantaneous velocity by a 30-minute average of individual velocity components.

The turbulence kinetic energy (TKE) is then defined as

$$TKE = \frac{1}{2}(\langle u'^2 \rangle + \langle v'^2 \rangle + \langle w'^2 \rangle) \quad (1)$$

In wind energy, a more frequently used statistic is turbulence intensity (TI), which is defined as

$$TI = \frac{\sigma_u}{u} \quad (2)$$

where $\sigma_u = \sqrt{\langle u'^2 \rangle}$ is the standard deviation of the horizontal wind speed u .

In similarity theory, one of the key scaling parameters of the atmospheric boundary layer is friction velocity. Here, surface friction velocity is given by

$$u_* = [(\overline{u'w'})^2 + (\overline{v'w'})^2]^{\frac{1}{4}} \quad (3)$$

Vertical kinematic heat flux ($\overline{w'\theta'}$) is used to scale turbulence due to buoyant production. With these parameters—vertical heat flux and surface friction velocity—a length scale (L) can be defined (Obukhov 1946), and is given by

$$L = \frac{\overline{\theta}_v u_*^3}{\kappa g (\overline{w'\theta'_v})} \quad (4)$$

where

- $\overline{\theta}_v$ = mean virtual potential temperature,
- u_* = surface friction velocity,
- g = acceleration of gravity,
- $\kappa \approx 0.4$ = the von Kármán constant, and
- $(\overline{w'\theta'_v})$ = kinematic virtual heat flux evaluated at the surface.

The combination of $(g/\overline{\theta}_v)$ is sometimes defined as the buoyancy parameter, and the length scale is generally referred to as the Obukhov length scale. The Obukhov length scale provides the height above the surface where the buoyancy production dominates over shear production. This length scale, along with the height of measurement z , provides a measure of atmospheric stability in flat terrain conditions, (z/L) . The various stability classes were defined as ranges of L , as shown in Table 2 below.

Table 2. Stability classification based on L thresholds.

Stratification	L
Very stable	$10 < L < 50$
Stable	$50 < L < 200$
Near-neutral stable	$200 < L < 500$
Neutral	$ L > 500$
Near-neutral unstable	$-500 < L < -200$
Unstable	$-200 < L < -100$
Very unstable	$-100 < L < -50$

Within the surface layer, u_* and $(\overline{w'\theta'_v})$ are assumed to be constant, suggesting that fluxes can be estimated from measurements at a given height. Within the surface layer one can estimate the Obukhov length from sonic anemometers without much sacrifice to the accuracy of the estimate (Cheynet et al. 2018).

The latent heat flux (Q), sensible heat flux (H), and momentum flux (τ) are defined as

$$H = \rho C_p \overline{w'\theta'} \tag{5}$$

$$Q = \rho L_v \overline{w'q'} \tag{6}$$

and

$$\tau = -\rho \overline{u'w'} = \rho u_*^2 \tag{7}$$

where

- ρ = density of air,
- C_p = specific heat of air at constant pressure,
- L_v = latent heat of vaporization of water,
- θ = potential temperature, and
- q = specific humidity.

τ assumes that the wind is along the longitudinal axis (x-axis). Another important scaling parameter is the mixed layer or boundary layer height which is defined in Section 4.4.

4.2 Similarity Functions

Accurate characterization of the near-surface wind speed profile is key for wind power applications. Surface measurements are generally extrapolated using Monin–Obukhov similarity theory (MOST) to either hub height or the top of the rotor layer. Various similarity function formulations have been proposed based on empirical fits to available data (Businger et al. 1971; Dyer 1974 Nieuwstadt 1984; Höglström 1988; Beljaars and Holtstag 1991; Foken 2006; Kelly and Gryning 2010). These similarity relationships have been mostly valid for flat terrain conditions, while in complex terrain sites, the applicability of MOST can be questionable (Fernando et al. 2015). The widely used flux-profile relations for wind shear and the temperature gradient (Businger et al. 1971) are based on observations from the 1968 KANSAS experiment (Izumi 1971). The universal function of momentum (φ_m or nondimensional wind shear) is given

as a function of the dimensionless stability parameter (z/L , or ζ), based on Businger et al. (1971) and Dyer (1974), hereafter BD74:

$$\varphi_m(\zeta) = (1 - 15\zeta)^{-1/4}, \quad -2 < \zeta < 0 \quad (8)$$

$$\varphi_m(\zeta) = (1 + 4.7\zeta), \quad 0 < \zeta < 1 \quad (9)$$

There have been many variations of these equations, under stable atmospheric conditions (Högström 1988; Beljaars and Holtslag 1991; Foken 2006). Flux observations in the very stable regime are subject to issues related to sampling and exhibit large variability (Nieuwstadt 1984). Therefore, it is difficult to empirically establish which function is most suitable at a given site. Many similarity theory formulations in the literature have been observed to collapse under unstable conditions (Högström 1988).

The Beljaars and Holtslag 1991 (hereafter BH91) universal form of φ_m under stable conditions, which were derived from the Cabauw data in The Netherlands, is given as

$$\varphi_m(\zeta) = -a\zeta - b\left(\zeta - \frac{c}{d}\right)e^{-d\zeta} + \frac{bc}{d} \quad (10)$$

with $a = 1$, $b = 2/3$, $c = 5$, and $d = 0.35$. For unstable conditions, BH91 and BD74 formulations are similar. Hence, BD74 equations are used as a standard for all unstable conditions. For stable conditions, formulations shown in Equations (8) and (10) are used as estimates from BH91 and BD74, respectively.

The validity of above surface layer similarity functions at the SGP central facility is shown in Section 5.1. The dimensionless wind shear from the tower measurements can be calculated using the following expression:

$$\varphi_m(\zeta) = \frac{kz}{u^*} \frac{\partial \bar{U}}{\partial z} \quad (11)$$

Here we analyze the dimensionless wind shear using surface-layer values for u^* and L at 25 m AGL. The horizontal wind speed vertical gradient ($\partial \bar{U} / \partial z$) is calculated using the 4 m and 25 m level measurements from the 60 m meteorological tower. Surface roughness height (z_o) was calculated, under neutral conditions (i.e., $|L| > 500$), as follows:

$$z_o = z e^{-(kU(z)/u^*)} \quad (12)$$

where, k is the Von Kármán constant ($= 0.4$), z is the measurement height, and $U(z)$ is the mean wind speed at height z .

4.3 Remote-Sensing Turbulence

Surface heating generates thermal plumes and turbulence induced by wind shear generates horizontal convective rolls (in the absence of surface heat fluxes) in a continental convective boundary layer (CBL). These plumes lead to mixing of heat, momentum, aerosols, and moisture through the top of the boundary layer (Lenschow and Stephens 1980). Vertical velocity variance profiles are key to understanding the updraft and downdraft mechanisms. Also, as shown by Taylor (1922, 1935), the vertical size of a plume growing due to homogeneous turbulent motion

is proportional to σ_w^2 . We can therefore also use σ_w^2 as an indicator of vertical mixing (except for gravity waves during stable stratification).

The observed radial velocity from the ARM Doppler lidars can be expressed as

$$\tilde{u}_r = u_r + u_n \quad (13)$$

where u_r is the true atmospheric radial velocity, and u_n is the instrumental noise in the measurements. We assume that the noise is zero-mean Gaussian random variable. If the noise is uncorrelated with radial velocity, then the observed variance is given by

$$\sigma_{\tilde{u}_r}^2 = \sigma_{u_r}^2 + \sigma_{u_n}^2 \quad (14)$$

where $\sigma_{u_r}^2$ is the true atmospheric variance, and σ_n^2 is the variance of the noise. The noise generally increases with increasing range and decreasing SNR. Further details about calculating the vertical velocity variance accounting for the instrumental noise are provided by Newsom and Krishnamurthy (2020). The ARM Doppler lidars are typically staring vertically upward for prolonged durations since 2012 at the SGP; therefore, range- and time-resolved vertical velocity statistics (such as variance, skewness, and kurtosis) can be estimated (Newsom et al. 2019). Currently a VAP—Doppler Lidar W velocity Statistics (DLWSTATS) — provides these estimates from ARM Doppler lidars at the SGP.

Turbulence within the boundary layer can also be characterized by estimating the eddy dissipation rate. Within the boundary layer, as turbulence decays with altitude, the height at which either TKE or eddy dissipation rate or vertical velocity variance are near zero can be approximated as the boundary layer height (Vakkari et al. 2015). Traditional techniques that rely on vertical velocity variance profiles to estimate the top of the boundary layer height (Tucker et al. 2009) fail during nocturnal conditions because vertical velocity variance estimates are very low through the boundary layer. Therefore, eddy dissipation rate estimates could be useful in accurately predicting nocturnal boundary layer heights. In case of homogeneous and isotropic turbulence, the eddy dissipation rate estimates from vertical velocity stares can be calculated by detecting the inertial subrange from the lidar velocity spectra (Champagne et al. 1978; O'Connor et al. 2010):

$$\varepsilon = \frac{\alpha}{\kappa} (2\pi/U)^{2/3} \langle f^{5/3} [\varphi(f_{cut}) - \varphi(f_{noise})] \rangle \quad (15)$$

where,

- α = $3/4$,
- κ = 0.54,
- U = mean wind speed,
- φ = energy spectrum within the inertial subrange,
- φ_{noise} = Mean white noise spectrum (Hildebrand and Sekhon 1974),
- f_{cut} = high-frequency cutoff spectral index,
- f_{noise} = noise frequency index, and
- $\langle \rangle$ = denotes the mean.

4.4 Boundary Layer Height Using Machine Learning

Measuring the growth of the planetary boundary layer (PBL) height is crucial for understanding the turbulent transfer of air mass between the boundary layer and free atmosphere, which in

turn strongly influences the winds, temperature, and moisture within the atmospheric boundary layer. During daytime, the air within the PBL is well mixed because of convection, due to surface heating and the weakening of turbulence at the top of the PBL (entrainment zone). Modern remote-sensing instruments can provide continuous estimates of boundary layer dynamics. Doppler lidars, when staring vertically upward, measure the vertical velocities through the PBL. Generally, the convective PBL height from a Doppler lidar can then be estimated by applying a cutoff threshold to the vertical velocity variance estimates (Lenschow et al. 1980; Lenschow et al. 2000; Tucker et al. 2009; Lenschow et al. 2012; Berg et al. 2017). A vertical velocity variance threshold of $0.04 \text{ m}^2\text{s}^{-2}$ is often used in the literature (Tucker et al. 2009). Doppler lidar estimates of PBL height, using a constant threshold for vertical velocity variance, have shown considerable error when compared to radiosonde estimates of PBL height (Emeis et al. 2008; Lenschow et al. 2012; Schween et al. 2014; Krishnamurthy et al. 2021).

Near-surface fluxes, such as latent heat flux, and other surface and subsurface properties, such as soil temperature, soil moisture, surface potential temperature, surface humidity, etc., have been shown to influence the PBL height (Santanello Jr et al. 2005). A multi-year analysis using radiosonde data over Europe, showed that the boundary layer height was also closely correlated with surface RH and air temperature (Zhang et al. 2013). Machine learning models, such as Random Forest, Artificial Neural Networks, have gained significant attention for understanding and classifying various atmospheric phenomena (McGovern et al. 2017; Gagne II et al. 2019; Vassallo et al. 2019, Vassallo et al. 2020a, Vassallo et al. 2020b). In this study, a Random Forest (RF) algorithm was used to estimate planetary boundary layer heights (PBLHs).

RF regression (Breiman 2001) is an ensemble method that is made up of a population of decision or decorrelated trees. Figure 8 provides a graphical illustration of the RF bootstrapping process. Bootstrap aggregation (bagging) is used so that each tree can randomly sample from the data set, while only a random subset of the list of input parameters/data is given to each individual tree. The premise behind RF is to improve the variance reduction of bagging by reducing the correlation between the trees, without increasing the variance. The trees can be truncated to add further diversification. After construction, the population's individual predictions are averaged to give a final prediction of the target variable. Ideally, this process results in a diversified and decorrelated set of trees whose predictive errors cancel out, producing a more robust final prediction. An advantage of RFs is their ability to determine the importance of all input features for the predictive process. This is done by calculating the mean decrease in variance that is achieved during a given split in each decision tree. The decrease in variance for each input feature (or input variable) can be averaged over the entire forest, providing an approximation of the feature's importance for the prediction (feature importance estimates sum to 100 percent to ease interpretability). However, if two input variables are highly correlated (as is expected when testing atmospheric forcing), it is highly unlikely that the reported importance values will accurately represent each variable's significance (Breiman 2001).

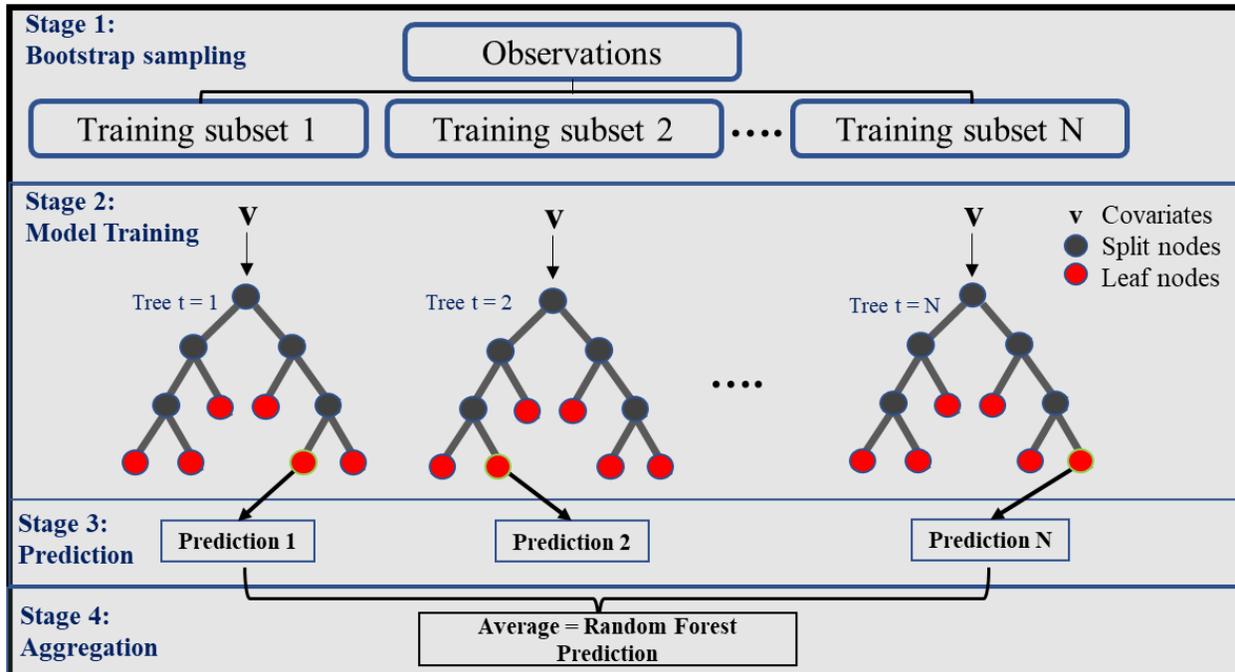


Figure 8. Graphical representation of a standard random forest algorithm (recreated from online sources and observed sample trees from current RF model).

Here, we use a series of 15 surface parameters (from mean winds to fluxes) and several other derived parameters from a Doppler lidar (such as SNR, backscatter, eddy dissipation rate, low-level jet height, etc.) as input for the RF model. Boundary layer height estimates from intermittent radiosondes released every 6 hours at the SGP site were used to train the RF algorithm. In this report, when estimating the PBL height of any given year (say 2016), data from at least 3 years (say 2015, 2017, and 2018) were used in training. Because Doppler lidar data were only available from 2012, we used the RF algorithm to estimate the PBL height from 2012 to 2019. Hyperparameters (i.e., parameters whose values are used to control the learning process in a machine learning model) for the RF algorithm, such as learn rate, number of trees, and tree splits, were calculated based on a Bayesian optimization approach, which minimizes the model mean square error by varying the chosen hyperparameters (Snoek et al. 2012). Further details are provided by Krishnamurthy et al. (2021). A sample time series of boundary layer height estimate from the RF algorithm is provided in Figure 9 below. Compared to estimates from the lidar vertical velocity variance threshold, the machine learning estimates of boundary layer height closely match the radiosonde observations.

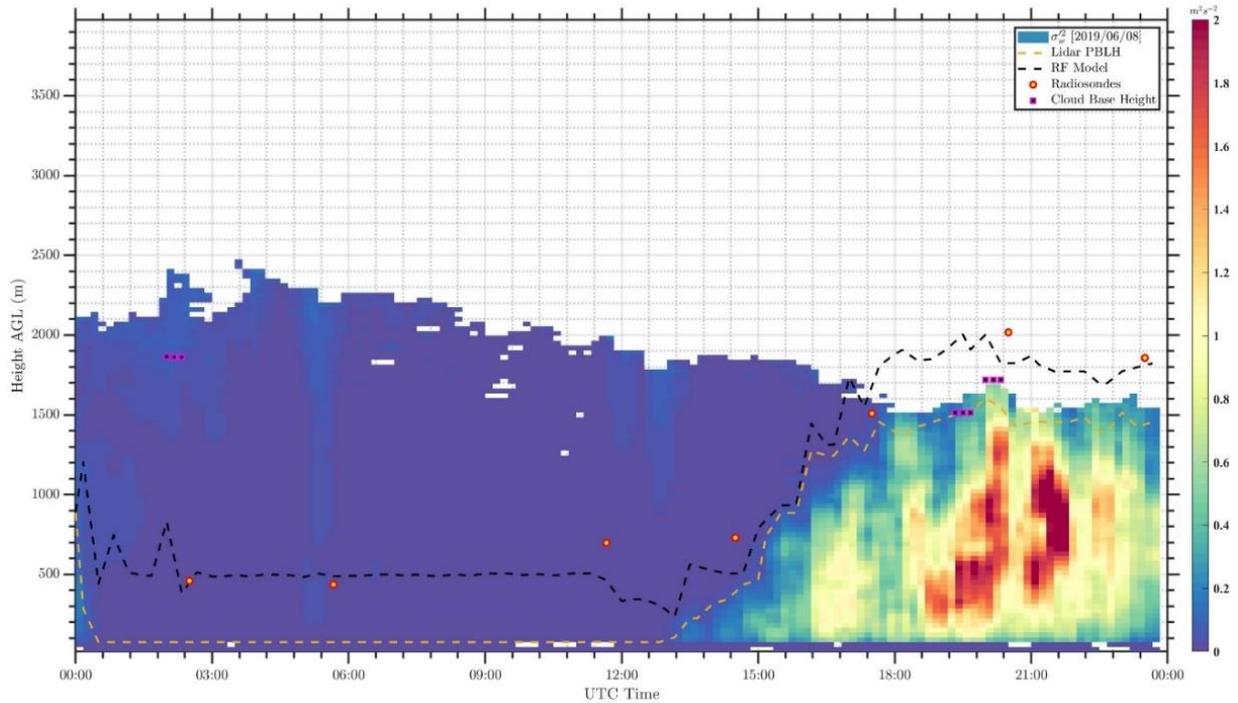


Figure 9. Vertical velocity variance on June 8, 2019, at the SGP C1 site with boundary layer height estimates from the lidar thresholding technique (orange dotted line, Tucker et al. 2009), random forest model (black dotted line), and radiosondes (filled circles, Sivaraman et al. 2013). Cloud base height estimates from the lidar are also shown (filled squares).

5.0 Atmospheric Boundary Layer at the SGP

In this section, a detailed analysis of the data collected at the SGP site from January 2010 to May 2020 is provided. Specifically, we analyze data from surface measurements, Doppler lidars, RLs, and satellite products over the SGP site. The long-term variability of atmospheric mean and turbulent parameters (such as winds, temperature, TKE, TI, etc.) from ground surface to the top of the boundary layer are discussed. They provide a holistic view of various atmospheric conditions at the SGP site, which would aid in model development and future campaigns to be conducted around the SGP site.

5.1 Subsurface and Surface Layer

The surface and subsurface soil layer measurements are key for understanding land–atmosphere interactions. Land–atmosphere interactions drive Earth’s surface water and energy budgets. They can alter clouds and precipitation in a region, affect the growth of the PBLH, and influence the persistence of extremes such as droughts. Therefore, a thorough knowledge of the land–atmosphere properties is key in predicting the local climate.

A study of the local climate using almost a decade of data surface-layer variables from 2010 to 2019 shows that there is a strong diurnal behavior at the SGP site. To develop this climatology, all mean measurements were averaged or interpolated to 15 minutes. Surface fluxes were calculated using a 30-minute mean averaging interval. Figure 10 shows hourly averaged estimates of air temperature, friction velocity, TKE, moisture flux, and heat flux at 10 m AGL from 2010 to 2019. At the SGP site, solar noon during central standard time occurs approximately at 18:00 hours UTC. The trends in the surface fluxes are similar, with a peak observed at approximately 20:00 hours UTC (~15:00 hours local time). A temporal delay of an hour is observed between the peak heat flux and TKE estimates, indicating that a gradual increase in vertical mixing results in higher turbulence within the CBL. Hourly averaged temperature measurements indicate the daytime peak is observed at 22:00 hours UTC (~17:00 hours local time) with a gradual reduction over night and a minimum just before sunrise at 12:00 hours UTC (~05:00 hours local time). Nocturnal conditions show that the stationarity assumption can be valid within the surface layer because of near constant friction velocity and other fluxes. Friction velocity is a key velocity scaling parameter in similarity theory. At the SGP site a constant mean u_* is observed during nocturnal conditions (Figure 10b).

Using a similarity theory derived wind profile (Dyer 1974 and others) as an input to theoretical atmospheric boundary layer models commonly used in wind energy (Jackson and Hunt 1975; Hunt et al. 1988) can cause issues due to the improper treatment of gravity wave propagation outside the stable boundary layer and large velocity perturbation errors near the surface. Modifications to this theory by Weng (1997), assume a constant friction velocity independent of atmospheric stability, which also has been shown to cause large errors in estimating certain atmospheric conditions. Recent studies have also suggested that u_* reduces with increasing atmospheric stability, at least in complex terrain (Argaín et al. 2009). Figure 11 shows dependence of friction velocity with normalized Obukhov length scale in stably stratified atmospheric conditions. Significant scatter is observed for any given stability value, for which the friction velocity varies by two orders of magnitude. Overall, the friction velocity is observed to reduce with increasing atmospheric stability, suggesting that using a constant u_* for modeling studies in flat terrain conditions can lead to erroneous results, and supporting the conclusions of Argaín et al. (2009). Therefore, for future boundary layer modeling studies, varying the friction velocity as function of stability is recommended even in flat terrain conditions, such as those of

the Southern Great Plains, for accurate wind resource assessment. It is particularly important for idealized simulations over a longer time frame, for example a diurnal case.

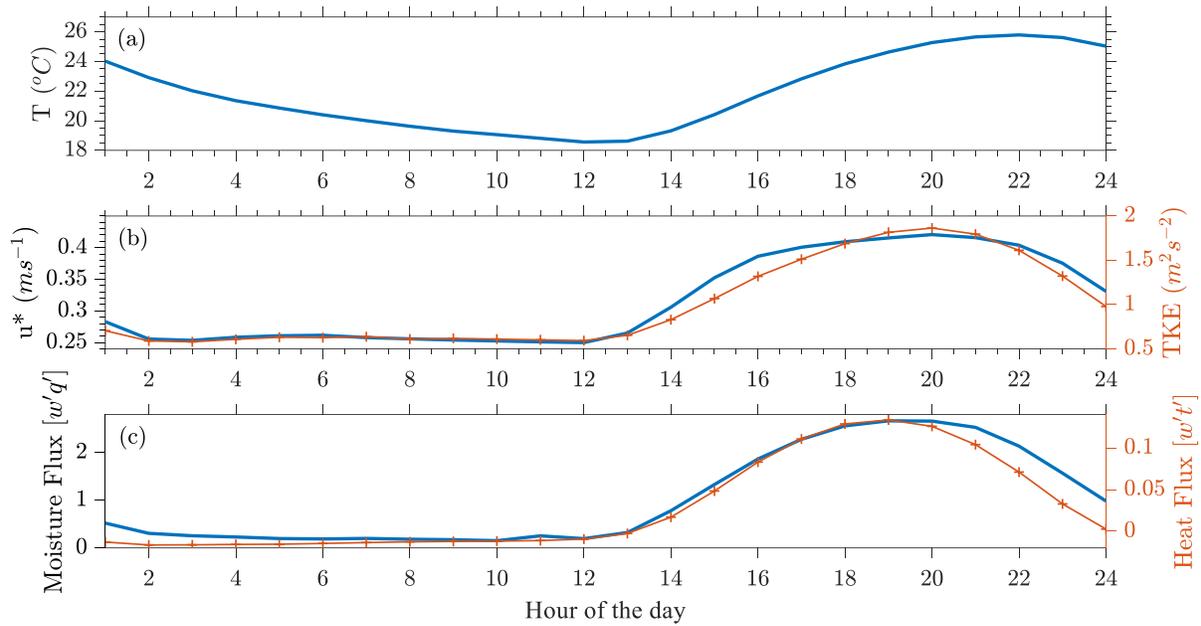


Figure 10. Hourly averaged (UTC time) surface (a) air temperature, (b) friction velocity (u^*) and turbulence kinetic energy (TKE), and (c) moisture and heat flux at 10 m AGL from 2010 to 2019 at the SGP C1 site.

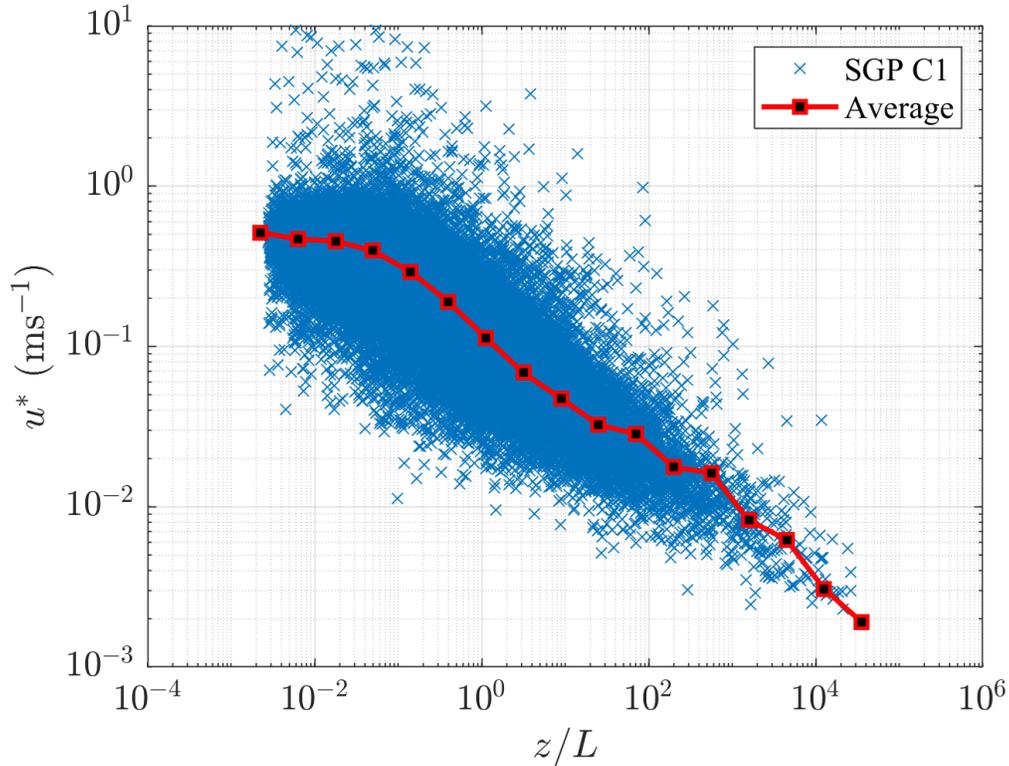


Figure 11. Friction velocity vs atmospheric stability (z/L) at the SGP C1 site for stable stratification.

The SGP site experiences significant seasonal variability with regard to winds, temperature, and turbulence. We have classified the SGP data into four distinct seasons: DJF for winter, MAM for spring, JJA for summer, and SON for fall conditions. Figure 12 shows hourly averaged near-surface air temperature and TKE observed during the four seasons from 2010 to 2019. The average daily maximum air temperature in summer is $\sim 34^{\circ}\text{C}$ and in winter $\sim 16^{\circ}\text{C}$. Spring and fall have similar temperature variability with a daytime peak near 25°C . A strong diurnal variability is observed during all seasons, with approximately a 6°C difference between the daytime peak and nighttime low. Interesting seasonal TKE trends are observed at the SGP site. Summer months show primarily high convective turbulence due to increased solar irradiance, as expected, but lower nocturnal turbulence compared to other seasons. During spring, daytime highs in TKE could be due to multiple sources (e.g., surface connected convective turbulence, cloud-driven top-down turbulence, and intermittent turbulence). During winter and spring, higher nocturnal turbulence is observed than during other months. One reason for low TKE during summer months could be the presence of a low-level jet (LLJ), which increases the stratification below the LLJ, thereby resulting in differential warm air advection, stratified turbulence, and development of very stable atmospheric conditions (Bonin et al. 2015). This is observed, especially in the residual layer, where our measurements are taken.

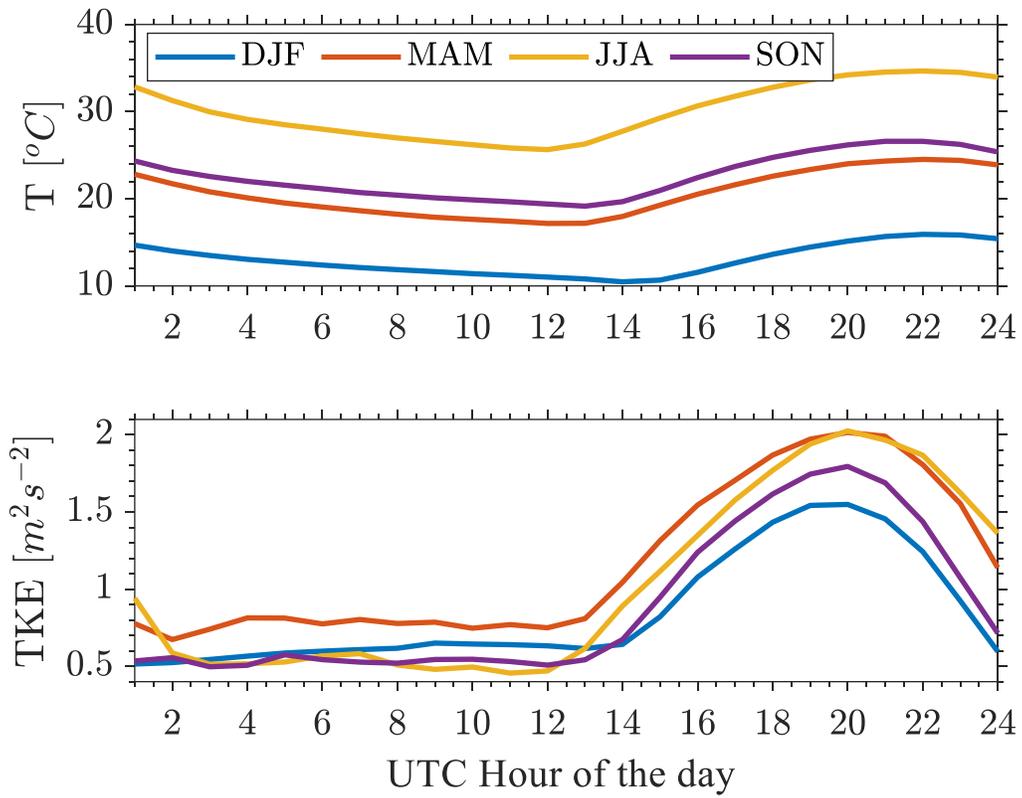


Figure 12. Seasonal hourly averaged surface air temperature and, TKE for four seasons—DJF, MAM, JJA and SON—from 2010 to 2019 at the SGP C1 site.

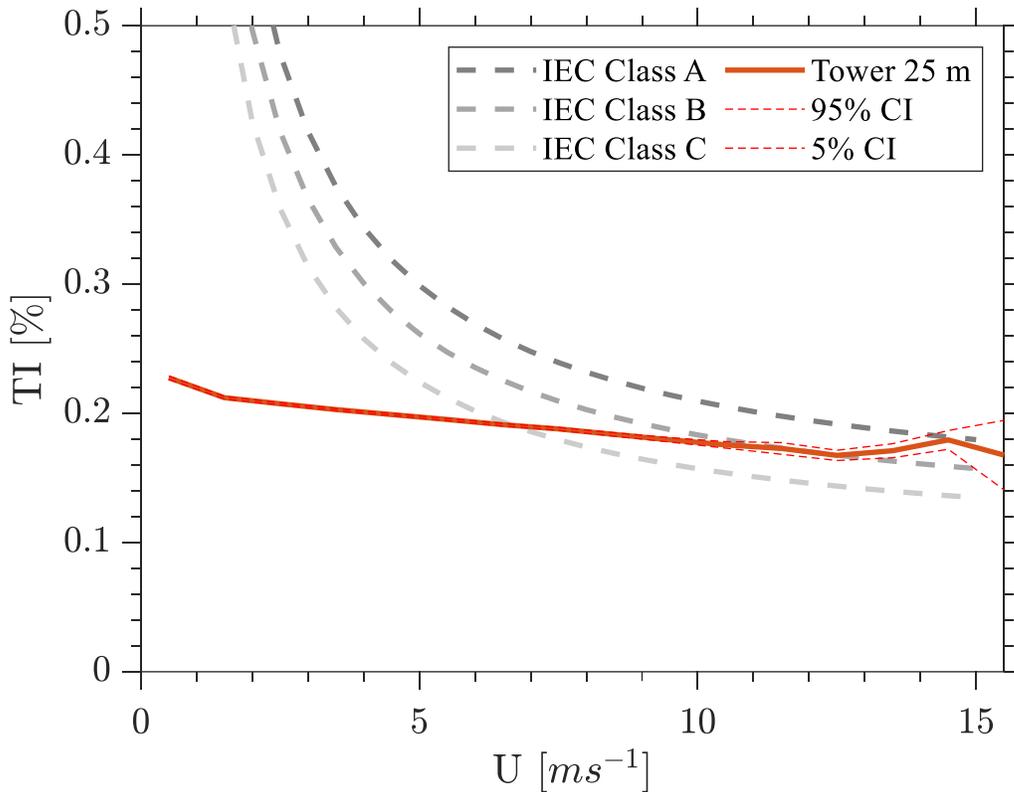


Figure 13. Turbulence intensity versus average wind speed at 25 m AGL from 2012 to 2019 at the SGP C1 site. The dashed lines show IEC curves for various input turbulence intensity estimates defined in IEC 61400-1.

TI is a key parameter for assessing wind turbine loads. Figure 13 shows the average TI at 25 m AGL from 2012 to 2019. The three classes of IEC 61400-1 TI curves are also shown, and are defined as

$$TI = TI_{IEC} [0.75U(z) + b]/U(z) \quad (16)$$

where TI_{IEC} is a reference TI (= 0.16, 0.14, and 0.12 for turbine classes A, B, and C, respectively) and b is a constant (= 0.56 ms^{-1}). In accordance with the IEC standard, the 90 percent confidence interval from the observed TI should be below the curve of the selected turbine within the operating range of the wind turbine. As illustrated, the TI at the SGP site in the range of interest (0 to 15 ms^{-1}) is observed to be consistently higher than two wind turbine classes (B & C) and lower than Class A wind turbines. Because the mean TI estimates were made at 25 m AGL, it can be assumed that TI estimates at typical hub heights (80 to 90 m) would be lower because of the wind shear observed at the SGP site during both daytime and nighttime conditions (see Figure 23). Generally, higher TI results in higher uncertainty in wind turbine power curves.

The atmospheric stability at the SGP site was classified using the Obukhov length scale values listed in Table 2. Figure 14 shows atmospheric stability as a function of months, time of the day, average wind speed, and southerly/northerly wind direction regimes. Higher percentages of unstable conditions are observed during summer months (45 percent) compared to winter

months (25 percent) (Figure 14a). Stable atmospheric conditions are observed in near equal proportions during all seasons, varying between 45 percent, during spring and fall, and 35 percent, during summer and winter months. It is interesting to note that neutral atmospheric conditions are seldom observed at the SGP site and occur less than 10 percent of the time. Diurnal variability in atmospheric stability observations is distinctly observed at the SGP site. Daytime unstable conditions are observed more than 65 percent of the time, and nighttime stable conditions are observed ~80 percent of the time (Figure 14b). Nocturnal stable conditions are predominantly observed during southerly winds, because of the presence of a LLJ creating a stable atmospheric boundary layer below the nose of the LLJ. A marginally larger percentage of northerly winds is observed to be unstable during daytime conditions compared to similar time periods during southerly winds (Figure 14d and e). Neutral and near-neutral unstable conditions increase in percentage with increasing wind speeds (Figure 14c); at $\approx 15 \text{ m s}^{-1}$ more than 60 percent of the data are near-neutral. Figure 15 shows distributions of atmospheric stability as a function of surface wind direction from 2012 to 2019. Southwesterly wind directions (200 to 250 degrees) are observed to be more unstable than other wind directions. One of the reasons for this could be the presence of a wind farm, since 2011, southwest of SGP C1 (see Table 1). The increased vertical mixing due to wind farm wakes could create more unstable conditions downwind of the wind farm. The predictions made for wind farm wake recovery distances range between 2 and 14 km (Barthelmie et al. 2004). At the SGP site, the distance from the wind farm to C1 is approximately 30 rotor diameters (or 4.5 km). Therefore, the plausibility of the wind farm affecting the stability around the SGP site in certain wind directions cannot be ignored, but further research on this subject is a topic of future work. Since 2016, several other wind farms have been built around the SGP site and a thorough analysis of every wind farm's effect on measurements at the SGP site should be evaluated.

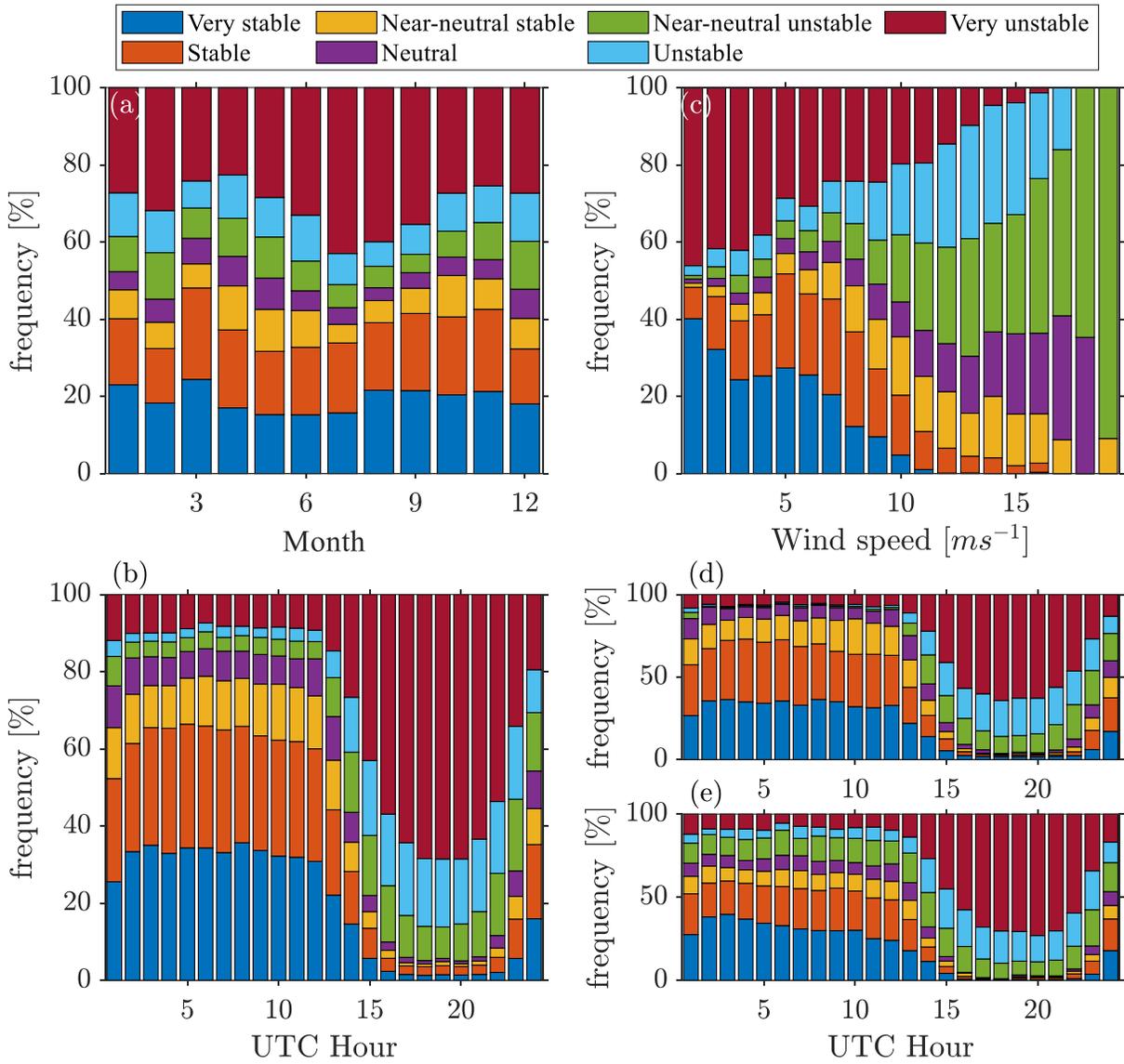


Figure 14. Atmospheric stability classification from 2012 to 2019 (a) monthly averaged, (b) hourly averaged, (c) wind speed bins, and (d) and (e) hourly averaged for southerly and northerly wind direction sectors, respectively at the SGP C1 site.

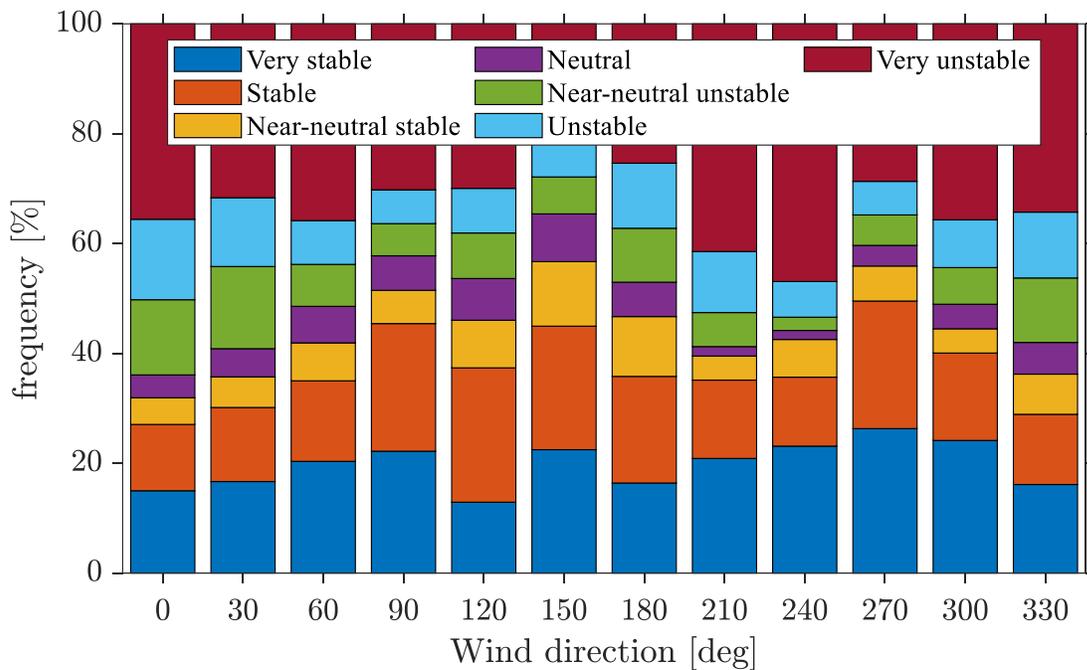


Figure 15. Atmospheric stability as a function of wind direction from 2012 to 2019 at the SGP C1 site.

The MOST (Obukhov 1946; Monin and Obukhov 1954) is the most widely used theory of atmospheric turbulence within the surface layer. MOST is also used in large-eddy simulations (LESs) as a surface boundary condition to estimate scalar fluxes and shear stress based on the gradients between the first grid point above surface and surface. Many of the similarity functions were developed using data from a flat terrain site (KANSAS, Cabau, Riso, etc.). Therefore, it behooves us to test the two well-known similarity theories, BD74 and BH91, relative to the data set at the SGP site. The SGP site features flat terrain and a predominant southerly wind direction. Figure 16 shows the nondimensional wind shear (ϕ_m) as a function of atmospheric stability (ζ). Wind speed gradient ($\partial U/\partial z$) was estimated using the measurements at 25 m and 4 m AGL and atmospheric stability estimated at 25 m AGL was used for the analysis. Estimates with friction velocity less than 0.1 ms^{-1} and surface wind speeds less than 3 ms^{-1} were removed from the analysis, because they considerably reduce the scatter in ϕ_m estimates. Both the BD74 and BH91 nondimensional wind shear estimates are also shown. BD74 formulations are observed to overestimate under stable conditions, while BH91 similarity functions tend toward the observational mean. In near-neutral conditions ($\zeta \sim 0$) and unstable conditions ($\zeta < 0$), there is no difference between the formulations, and both are observed to underestimate the observed nondimensional shear by approximately 0.5.

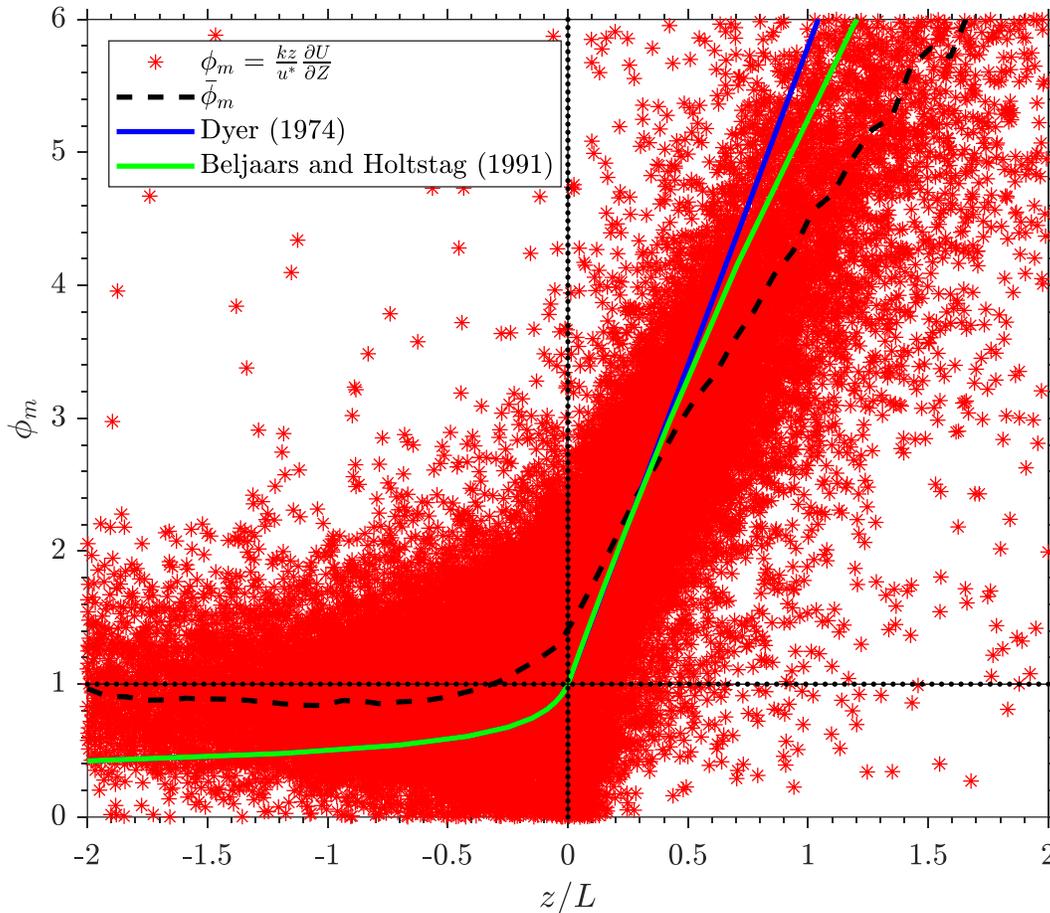


Figure 16. Dimensionless wind shear ϕ_m as a function of dimensionless stability z/L from 2010 to 2014. Data with $u^* < 0.1$ and $U < 3$ m/s were filtered. The average of all data points is given by $\bar{\phi}_m$.

Based on our observations, we found that at the near-neutral limit the nondimensional shear (ϕ_m) is 1.29. This neutral limit value as per MOST should be unity. Possible explanations for higher ϕ_m could be the uncertainty in the tower measurements and the violation of the constant flux layer assumption at the SGP site (not shown). Moreover, we used $k = 0.40$ in our analyses, but values of $k = 0.4 \pm 0.04$ have been reported in the literature (Högström 1996; Foken 2006). Businger et al. (1971) obtained a von Kármán constant of 0.35, based on the finding that $\phi_m = 1.15$ when using $k = 0.4$. Figure 17 shows the average nondimensional wind shear estimates for southerly and northerly wind directions. In stable atmospheric conditions, wind shear in southerly wind directions is larger than that in northerly wind directions. There is not much difference in ϕ_m during near-neutral and unstable conditions. The BH91 formulations are observed to better fit northerly wind directions during stable conditions. Therefore, the effect of either the nocturnal LLJ and/or wind farm wakes suggest that MOST assumptions may not be always valid under stable conditions at the SGP site. A modest mismatch between similarity models and observations in near-neutral and unstable conditions could allude to prior arguments about instrument uncertainty, violation of the constant flux layer at the SGP site, and the appropriate von Kármán constant.

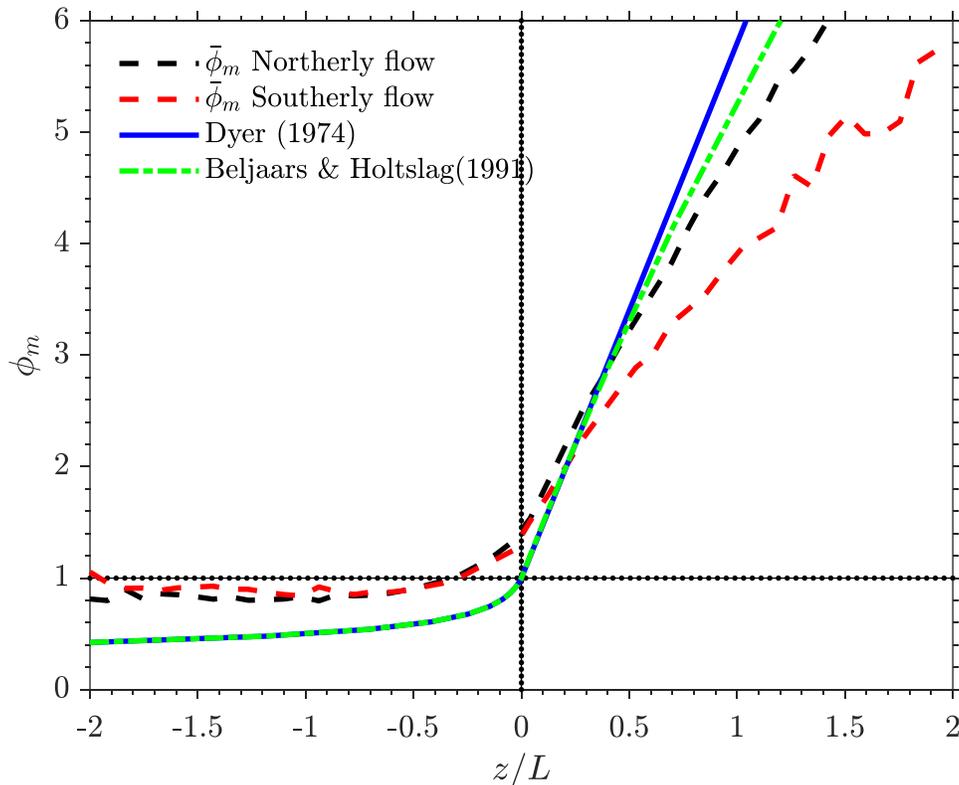


Figure 17. Same as Figure 16, but for the southerly and northerly wind direction regimes.

Because of surface heterogeneity at the SGP site, a wind profile continually encounters changes in surface roughness. A step change in surface roughness can lead to the formation of an internal boundary layer, changing shear stresses and turbulent response to such a change. Surface roughness length (z_0) is therefore an important parameter in boundary layer meteorology that is highly correlated to surface properties, and it varies as a function of space and time. At the SGP site, because of the presence of farmland, surface roughness varies every season and as a function of wind direction. Figure 18a shows surface roughness as a function of wind direction at the SGP 60 m meteorological tower during near-neutral conditions. Measurements show an increase in surface roughness east to southeast ($\sim 80 - 120$ degrees), due to either local terrain variability (a dip of ~ 20 m observed ~ 2 km away) or maintenance buildings at SGP C1. Other small peaks are observed at ~ 30 degrees and ~ 270 degrees, which could correspond to variability in the type of crops or other structures at those locations. The average surface roughness in other wind directions is generally below 0.02 m (typical flat terrain site z_0 is 0.03). As mentioned earlier, the surface roughness is expected to change monthly due to the crop production in the neighboring regions. Figure 18b shows surface roughness averaged over each month from 2003 to 2014. Surface roughness is observed to be at least two to three times greater in spring and summer months compared to winter months. At the SGP C1 the effects of two types of crops, grassland/pasture and winter wheat, are observed. Although the type of crop grown each season at the SGP site is not available, a generic trend related to the effect of crops on surface roughness could be analyzed. Winter wheat accounts for more than 70 percent of total U.S. wheat production and is a prominent crop in the Great Plains region. Winter wheat is planted in the fall and harvested around May; a very different growing cycle than summer crops or grasses. Near-surface air is cooler and moister over Oklahoma's winter wheat belt than over nearby grassland. Double-cropping occurs intermittently, when

winter wheat is harvested in spring and, another crop, such as soy or wheat, is planted afterwards and harvested in fall (Bagley et al. 2017). But from the cumulative surface roughness estimates, a distinctive seasonal effect is observed.

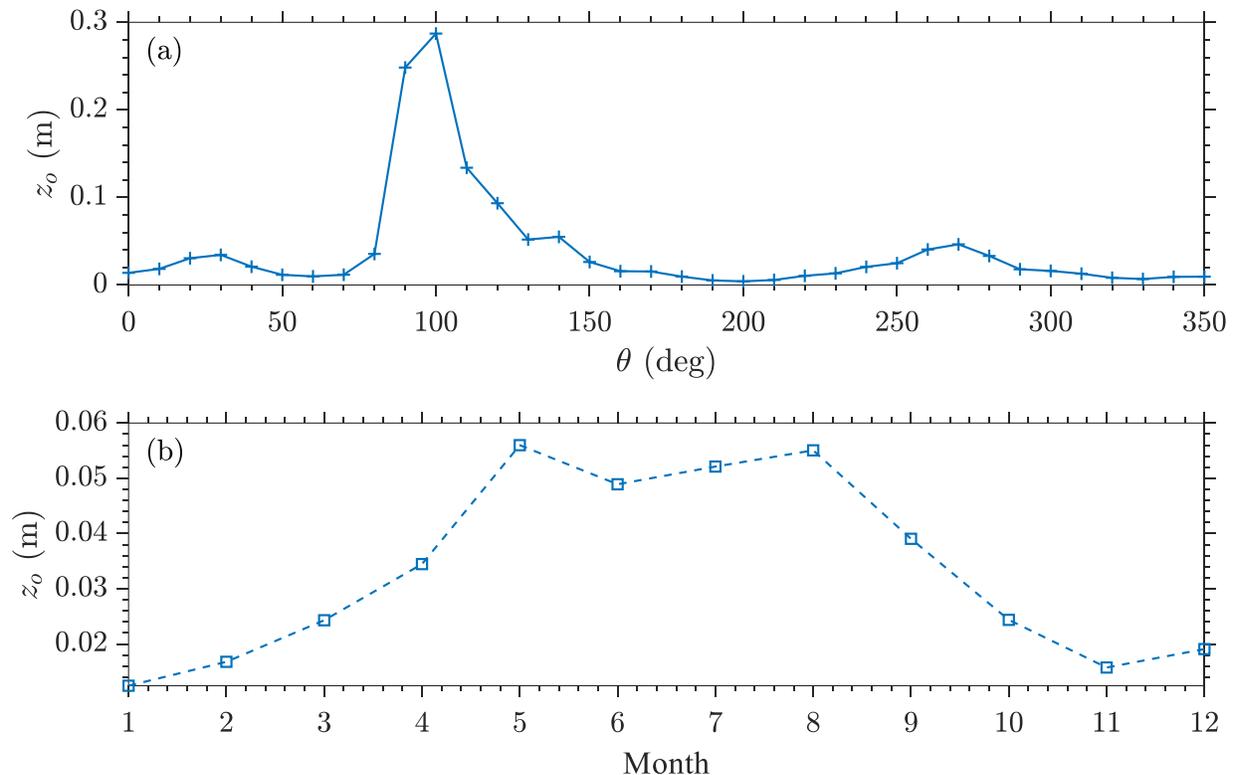


Figure 18. Average surface roughness (z_0) in near-neutral conditions ($|L| > 500$) at the SGP C1 site versus (a) wind direction and (b) monthly at 25 m AGL from 2003 to 2014.

Because of the presence of a variety of crops at the SGP site, soil moisture and soil temperature can play a crucial role in altering the surrounding boundary layer, such as boundary layer height, latent heat fluxes, etc. Figure 19 shows the diurnal and seasonal variability of soil temperature and soil moisture profiles from 5 cm to 75 cm below ground level at the SGP site. Strong diurnal variations are observed in soil temperature near the surface, with peaks observed between 23:00 and 01:00 UTC hours (~18:00 – 20:00 hours local time) and troughs after sunrise at 15:00 hours UTC (~10:00 hours local time). A similar trend is observed at depth below the surface, but with somewhat smaller variability. On a seasonal basis (Figure 19b), soil temperature during summer months is the highest at approximately 30°C. Soil temperature at the SGP site is generally lower than the peak air temperature during summer time, but well correlated (Figure 12). Near-uniform heating is observed through the depth of measurements in various seasons. On the other hand, no soil moisture (or volumetric content) diurnal trends are observed at any given surface depth. Although soil moisture might not vary diurnally, there is strong seasonal variability near the surface. Soil integrates past rainfall events and other weather anomalies (e.g., heat waves). It has been shown that soil memory extends the impact of climate and weather events and can aid in improving the predictability of weather systems (Santanello Jr et al. 2018). Near-surface peak soil moisture content at the SGP site is observed during the months of January, February, and March. The soil moisture content at depth below 20 cm is observed to be less variable, and exhibits higher volumetric content.

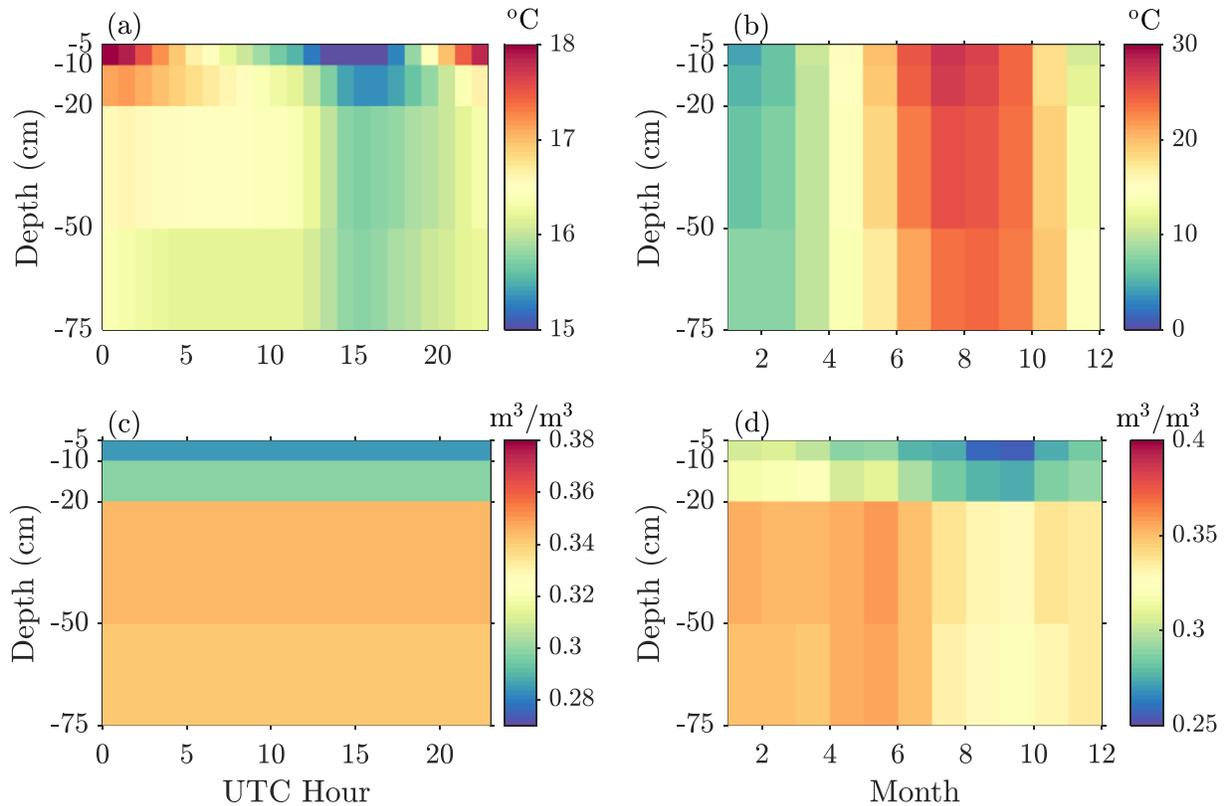


Figure 19. (a), (c) Hourly and (b), (d) monthly averaged soil temperature from 2012 to 2019 and soil specific water content profiles from 2012 to 2015 at the SGP C1 site.

Figure 20 shows the average temperature difference between surface soil temperature and air temperature at 4 m AGL from 2015 to 2019. Positive nighttime temperature difference is observed (i.e., the soil is warmer than the surrounding winds) and maximum difference of approximately 4°C is observed at 10:00 hours UTC (~05:00 hours local time). Daytime the surrounding winds are warmer than the surface and the difference reaches up to 3°C around 17:00 hours UTC (~12:00 hours local time). During sunrise transition time periods, a drastic change in the temperature difference is observed, as the soil layer warms up rapidly because of incoming solar radiance. On a monthly time scale, the lowest soil and air temperature difference is observed during summer and the maximum temperature difference during winter months. But overall, a net positive difference is observed; i.e., the soil temperature is warmer than the air temperature at the SGP site. The effect of various land-surface parameters on the near-surface and boundary layer winds is a topic of research at the SGP site (Santanello Jr et al. 2005; Fast et al. 2019).

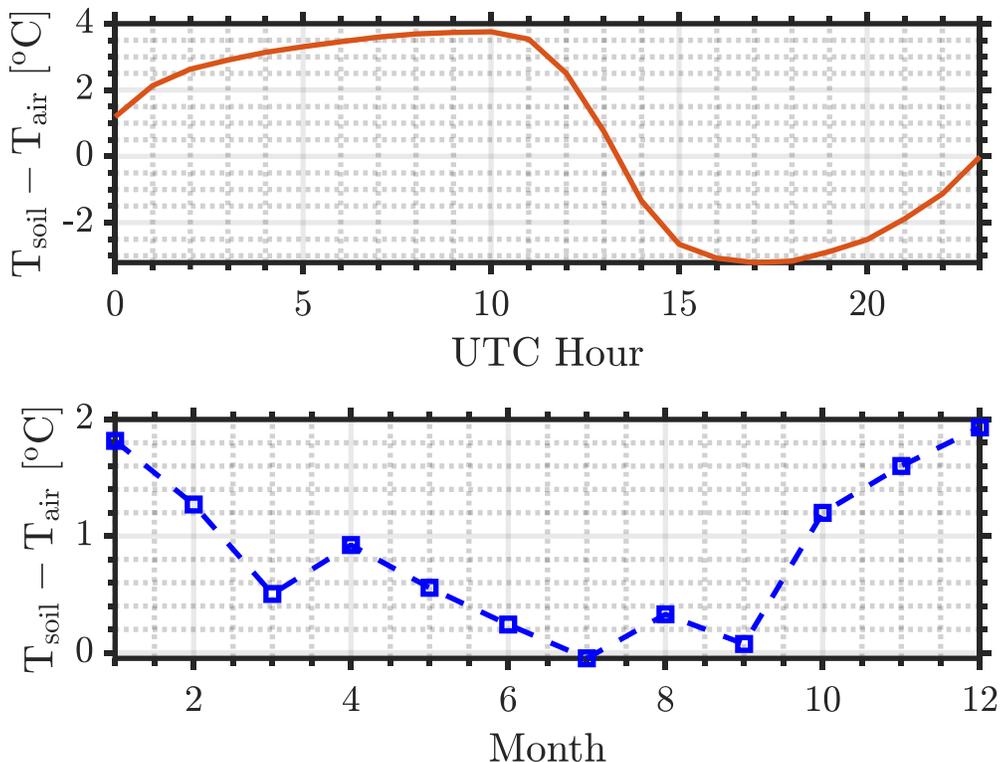


Figure 20. Difference between soil temperature and air temperature (a) hourly and (b) monthly from 2015 to 2019 at the SGP C1 site.

Figure 21 shows monthly and hourly averaged precipitation rate measurements at the SGP C1 from 2012 to 2019. The monthly peak precipitation rate occurs during spring, mostly in the month of May, with a median rain rate of 0.12 mm/hr. The maximum rain rate observed between 2012 to 2019 was ~91 mm/hr (3.6 in./hour) on 24 June 2018 at 07:00 hours UTC (~02:00 hours local time). Nighttime showers are more frequent at the SGP site, as shown in Figure 21b. One of the reasons for the increased frequency of showers at nighttime is the mesoscale convective systems (MCSs). A field experiment—the Plains Elevated Convection at Night (PECAN)—was designed at the SGP site to study such nocturnal precipitation maximums (Geerts et al. 2017). The PECAN study results support the conclusion that MCSs cause nocturnal precipitation at the SGP site. The primary sources of rainfall were divided into four different sources: (a) mountain-initiated precipitation, (b) plains-initiated precipitation, (c) precipitation advecting over the border of the radar domain, and (d) episodes in which different initiation categories were merged together (Weckwerth and Romatschke 2019).

Figure 22 shows the hourly averaged longwave (L up and L down), shortwave (K up and K down), and net radiation observed from the EBBR data at the SGP site from 2012 to 2019. Net radiation (Q) is defined as

$$Q = (K \downarrow - K \uparrow) + (L \downarrow - L \uparrow) \tag{17}$$

Incoming shortwave radiation is generally a combination of direct solar energy and diffused or scattered solar radiation, while outgoing shortwave radiation is a function of reflected, albedo, and downwelling shortwave radiation. Incoming longwave terrestrial radiation is the radiation emitted by the atmosphere and is generally determined by air temperature, while outgoing

longwave radiation is the radiation emitted by the Earth and is determined by surface temperature. The net radiation is also a function of the sensible heat flux, latent heat flux, and ground heat flux (i.e., the energy used to change temperature of subsurface). Therefore, the temperature patterns at a location are affected by spatial and temporal imbalances in radiation and energy fluxes, the heat capacity and reflectivity of Earth's surface, and type of work the energy supports (either sensible or latent heat flux). At the SGP site, peak outgoing shortwave radiation is observed at 18:00 hours UTC (~12:00 hours local time). A temporal delay is observed between the peak convective activity (at 20:00 hours UTC in Figure 12) at the SGP site and shortwave radiation (at 18:00 hours UTC in Figure 22). Peak net radiation is observed during summertime and lower net radiation is observed during winter months (not shown).

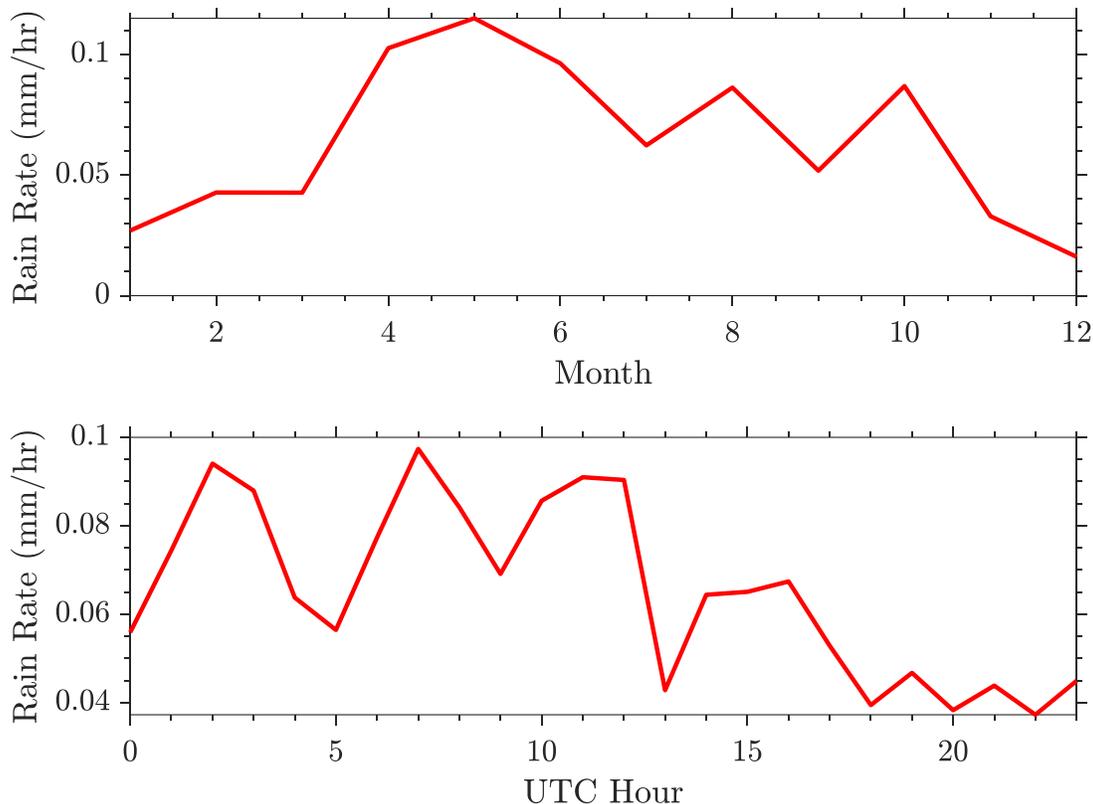


Figure 21. (a) Monthly and (b) hourly median precipitation rate from 2012 to 2019 at the SGP C1 site.

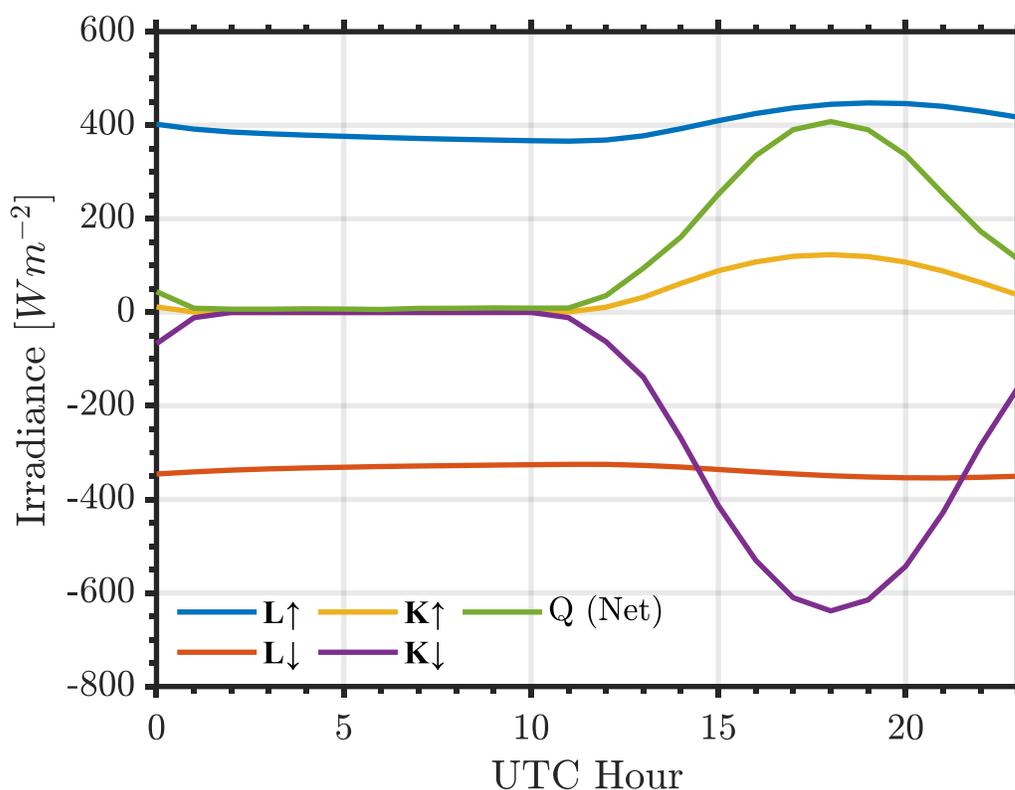


Figure 22. Hourly averaged longwave (L), shortwave (K) and Net (Q) radiation from 2012 to 2019 at the SGP C1 site.

5.2 Boundary Layer Horizontal Wind Profiles

Winds within the atmospheric boundary layer are generally variable, gusty, and are influenced by the surrounding topography, land-surface features, friction, local pressure difference, and other large-scale synoptic flows. Current state-of-the-art mesoscale models show the highest error within the first few hundred meters of the boundary layer in both flat and complex terrain sites (Zhang et al. 2013; Jiménez and Dudhia 2013; Krishnamurthy et al. 2013; Wharton et al. 2013). Wind profilers (such as Doppler lidars) have been instrumental in characterizing the atmospheric boundary layer and can continuously monitor the wind variability. In this section, the yearly, diurnal, and seasonal variability of the atmospheric boundary layer, as measured by a Doppler lidar at the SGP C1, are presented. Figure 23 shows hourly averaged horizontal wind speed contours and wind direction vectors for each year from 2014 to 2019. As mentioned earlier, Doppler lidars at the SGP site are programmed to measure a wind profile once every 15 minutes (Newsom et al. 2019; Newsom and Krishnamurthy 2020). These data sets are averaged to create an hourly vertical profile. The data availability during all the years was similar, with the exception of 2016, during which 15 percent less data were measured compared to other years (see Figure 5). Nocturnal wind direction at the SGP is predominantly southerly at lower heights, and veers westerly at ~ 1 km AGL. A nocturnal LLJ is observed during all years, but the intensity of the jet varies considerably each year. The timing and the height of the peak jet velocity is also observed to shift by a few hours and few hundred meters each year, respectively. The nocturnal LLJ is observed to span from approximately 03:00 to 14:00 hours UTC ($\sim 22:00$ to 09:00 hours local time). The height of the LLJ nose is observed to be generally

below 600 m AGL. In Figure 23, winds during 2014 are observed to be of highest intensity with peak hourly averaged winds greater than 15 ms^{-1} at approximately 08:00 hours UTC (~02:00 hours local time). Daytime wind directions are predominantly south to southwesterly and winds are less variable with height.

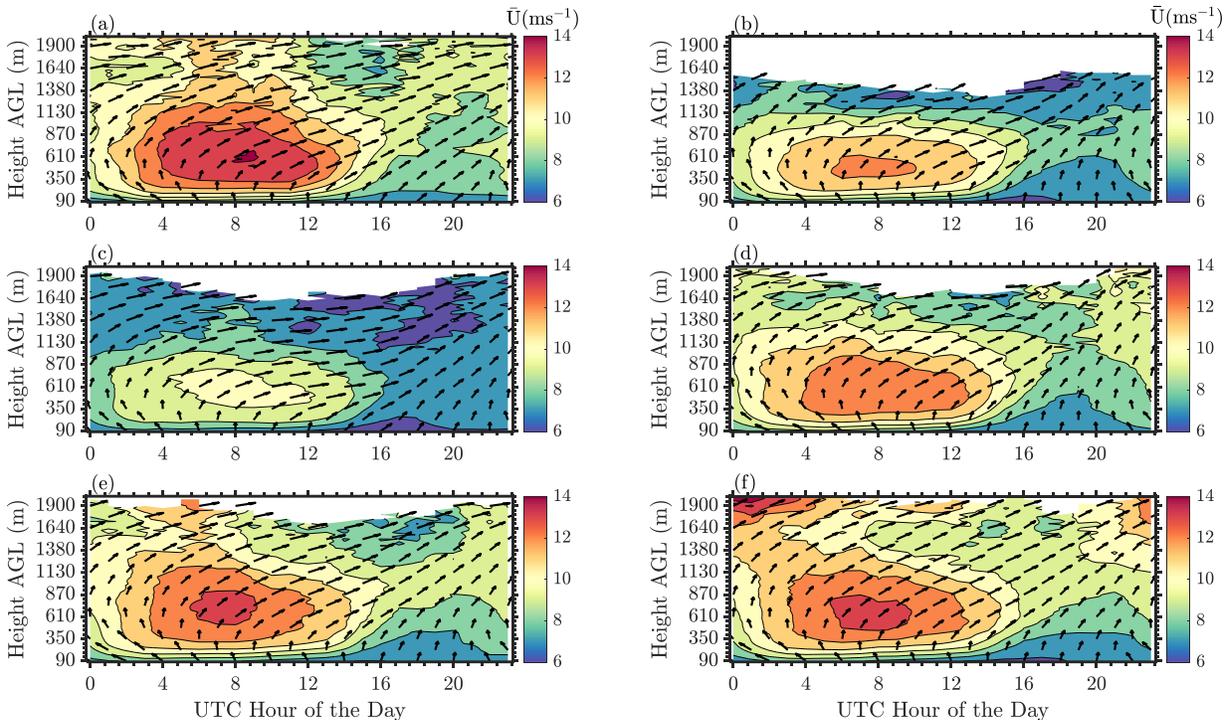


Figure 23. Hourly averaged wind speed (\bar{U}) and direction from (a) 2014, (b) 2015, (c) 2016, (d) 2017, (e) 2018, and (f) 2019 at the SGP central facility. Colored contours of \bar{U} for every 1 ms^{-1} are displayed and the arrows depict mean wind direction. Data availability below 5 percent is not shown. Apart from year 2016 (subplot c), which had 15 percent less data availability compared to other years, all other years had a minimum data availability greater than 95 percent at the lowest range-gate of 90 m.

Figure 24 shows hourly averaged seasonal winds and direction for winter (DJF), spring (MAM), summer (JJA), and fall (SON) from 2013 to 2019. The mean structure of the wind profile for each season is different, but a nocturnal LLJ is observed in every season. Winds are highest during spring and winter months, with peak of the nocturnal jet ($>14 \text{ ms}^{-1}$) between 600 m to 800 m AGL. Lowest winds are observed during summer, but a well-defined LLJ peak is observed. Daytime winds are stratified during winter months, while summer months featured near constant winds through the boundary layer.

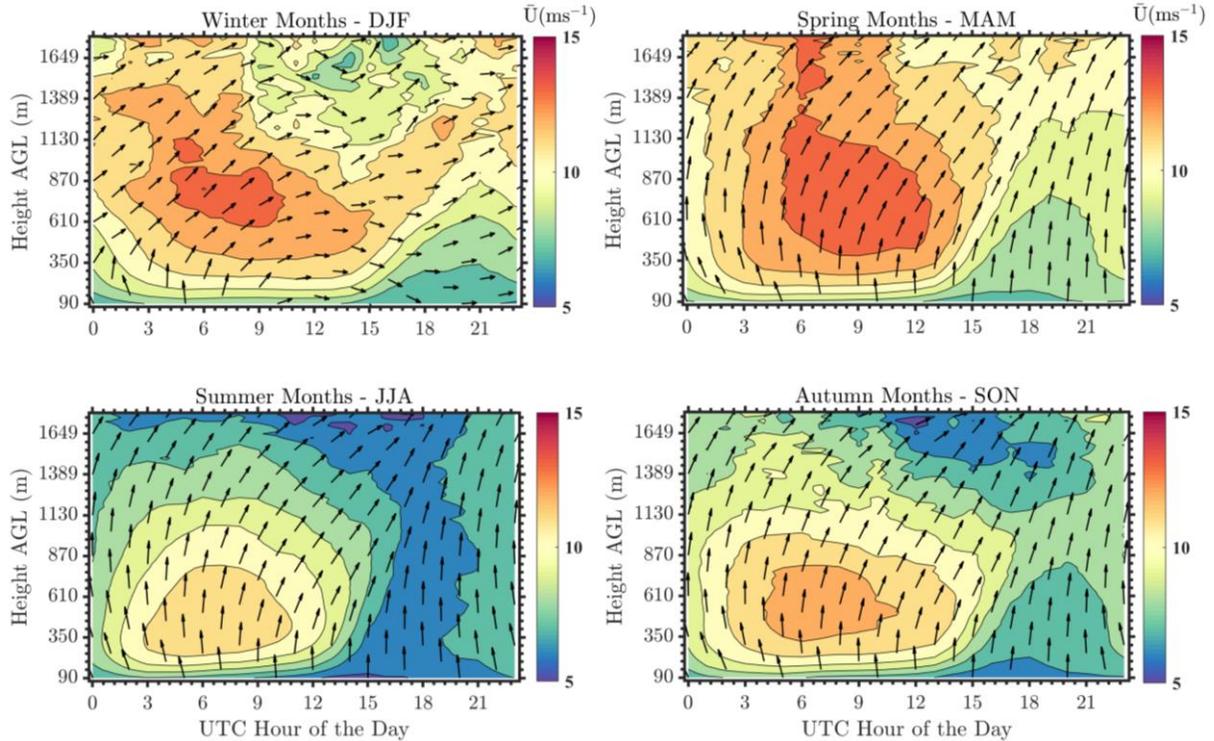


Figure 24. Hourly averaged wind speed (\bar{U}) and direction seasonal variability from 2013 to 2019 at the SGP central facility. Colored contours of \bar{U} for every 1 ms^{-1} are displayed and the arrows represent mean wind direction. Measurements at higher altitudes are irregular due to reduced data availability.

The hourly, daily, monthly, and annual variability of wind speed is shown in Figure 25. The interannual variability is low; 90 m hourly wind speeds exceed 20 m s^{-1} on occasions. While wind speeds tend to have a relatively stable multi-year mean and variance, the effects of interannual variability (IAV) may skew any single year's wind speed distribution. IAV is typically calculated as

$$IAV = 100 \frac{\sigma_{\bar{U}}}{\bar{U}} \tag{18}$$

where \bar{U} is the mean of a representative period's annual wind speeds and $\sigma_{\bar{U}}$ is the standard deviation of annual wind speeds over the entire representative period, thus making IAV a percentage of the period's overall mean wind speed (Pryor et al. 2018). IAV values of 6 percent are used as a representative estimate by the wind energy industry (Brower 2012). The IAV at the SGP site from 2012 and 2019 is approximately 3.37 percent. Therefore, future annual mean winds at the SGP can be reasonably expected to be similar. An important note about the frequency of data used for this analysis: the Doppler lidar measures a wind profile once every 15 minutes, therefore the gustiness of the winds at the SGP site are not accounted for in the IAV analysis. But wind gustiness is not expected to have a large impact on the IAV estimates provided above (a shorter analysis of IAV using tower data is provided in Section 6.3).

The abundant use of machine learning algorithms for forecasting in the wind energy sector is primarily due to the gap in the applicability of commonly used numerical models or techniques for short-term forecasting. Numerical models (such as weather research forecasting [WRF]) do not accurately predict small-scale atmospheric variability in surface wind speeds and are

generally used for 6+ hour forecasts because of their high computational costs. A simple persistence model (where $U_{t+n} = U_t$; U is the streamwise wind speed, t is the timestep, and n is the number of timesteps ahead to be forecasted) is generally used for ultra-short-term forecasting (<30 minutes), because its accuracy reduces for larger time steps. Therefore, machine learning models can fill the gap (between 30 minutes to 6 hours ahead) between inexpensive persistence models and numerical models. Machine learning algorithms can be applied to historical data to learn various trends at the SGP site and test the efficacy of short-term forecasting (Vassallo et al. 2021). Further research on the applicability of machine learning algorithms for ultra/short-term wind forecasting at the SGP site is warranted (mainly because of the low IAV).

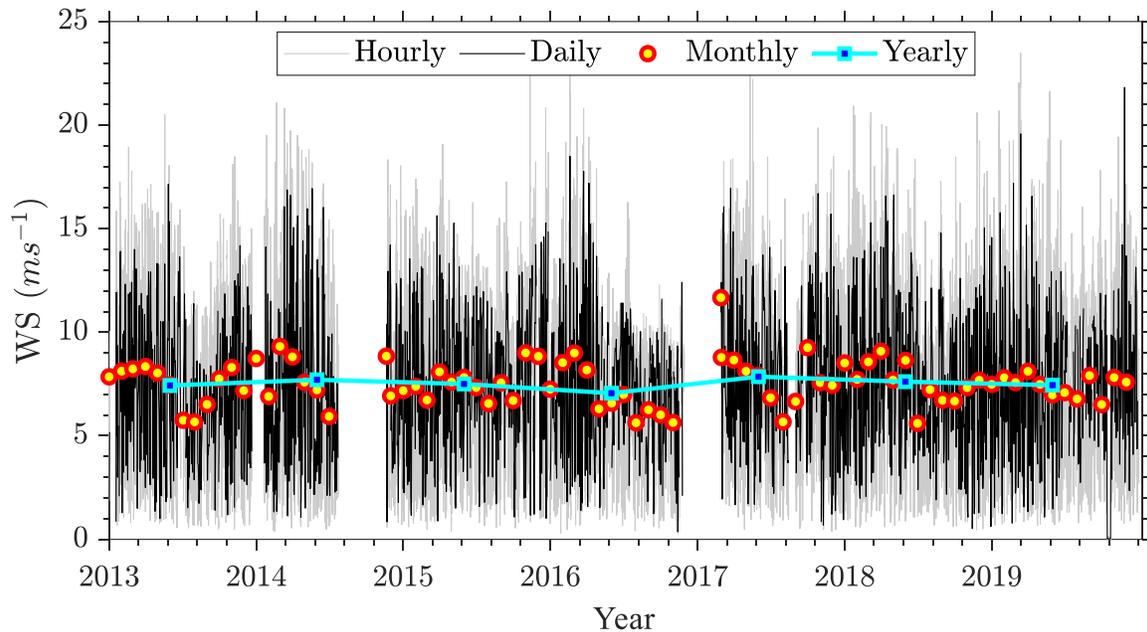


Figure 25. Time series of hourly, daily, monthly, and yearly averaged windspeeds (\bar{U}) at 90 m AGL from 2013 to 2019 at the SGP central facility.

Figure 26 shows the wind rose at 90 m (typical hub height at the SGP) and 500 m AGL. Because the Doppler lidar measures a wind profile once every 15 minutes, the gustiness of the winds is not observed in this wind rose. Wind and direction measurements every 15 minutes are hourly averaged. Each spoke of the wind rose provides a percentage of winds in the sector displayed, and cumulatively all spokes add up to 100 percent. Bimodal distribution of wind directions is observed at the SGP site, predominantly from southerly and northerly wind directions. The predominant wind direction at 90 m is ~ 170 degrees (southerly) and at 500 m it is ~ 210 degrees (south southwesterly). Overall, southerly winds are observed ~ 36 percent of the time at the SGP site. Northerly winds are less frequent, approximately 25 percent of the time, and the speed of the winds is also observed to be low. At 500 m AGL, the winds are stronger and have a higher frequency of average winds greater than 25 ms^{-1} from southerly wind directions.

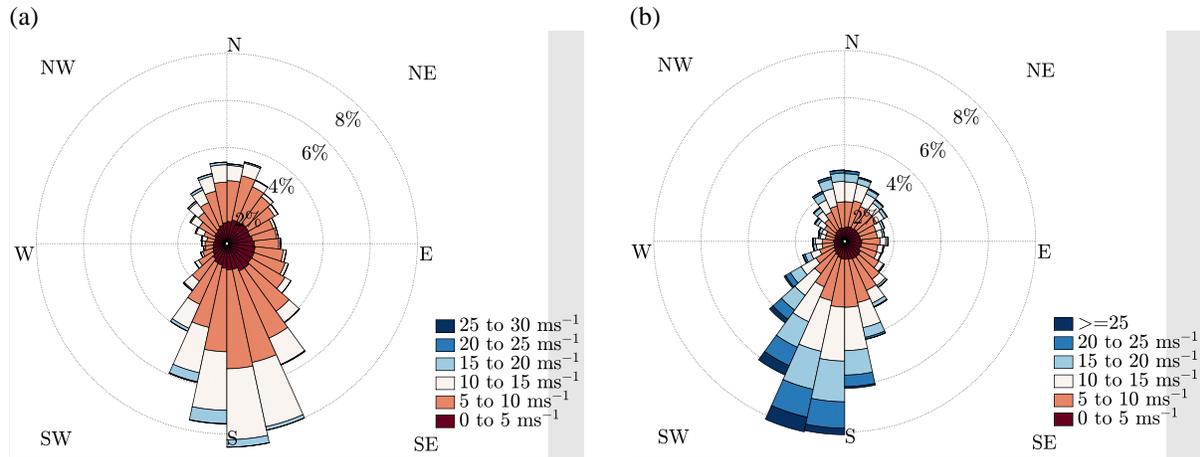


Figure 26. Average wind rose at (a) 90 m and (b) 500 m AGL from 2013 to 2019 at the SGP central facility.

For field campaign planning and certain modeling studies, it is important to understand the orientation of the winds each month and/or season. Figure 27 shows monthly averaged wind roses from 2013 to 2019 at 90 m AGL. Summer months (JJA) predominantly feature southerly wind speeds that are less intense than those during other seasons. July and August also show an increased distribution of easterly winds. Southerly and northerly wind directions are predominant during winter months (DJF), but are skewed toward northerly wind directions. Westerly winds are seldom observed at the SGP site—less than 10 percent of the time for any given month. Typically, westerly, or easterly wind directions are transient in nature. Therefore, when designing a field campaign to analyze wind farm wakes it is advisable to orient the instruments in either northerly or southerly wind directions to accurately capture the downwind wakes. For modeling studies, it would be beneficial to see the impact of southerly wind directions and the impact of LLJ on wind turbine loads and annual estimated production.

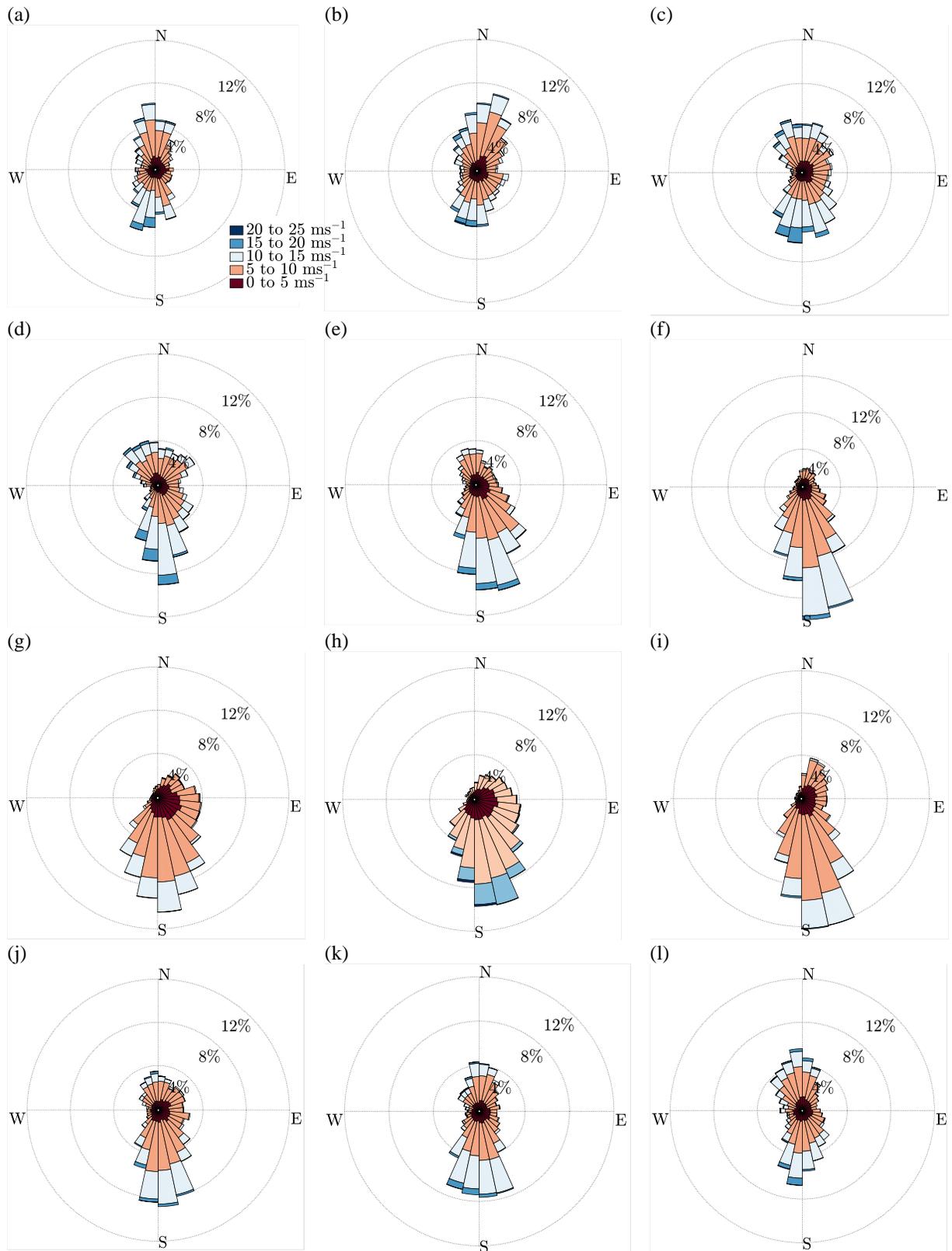


Figure 27. Monthly averaged (January to December [a to l]) wind rose at 90 m AGL from 2013 to 2019 at the SGP central facility.

The Weibull distribution is commonly used in wind energy to characterize mean wind speed frequencies per sector. It is defined as

$$f(x) = \frac{\alpha}{\beta} \left(\frac{x}{\beta}\right)^{\alpha-1} \exp[-(x/\beta)^\alpha] \tag{19}$$

where $f(x)$ is the frequency of a variable x (in our case wind speed), and α and β are the shape and scale parameters of the distribution, respectively. For wind energy studies, power density ($1/2 \rho \overline{U^3}$, where ρ is the air density and U is the mean wind speed) can be expressed using Weibull parameters by $1/2 \rho \beta^3 \Gamma(1 + 3/\alpha)$, where Γ is the Euler gamma function. Figure 28a shows the all-sector mean wind speed histogram with Weibull fit parameters at 90 m AGL from 2013 to 2019. The Weibull parameters, calculated using a maximum likelihood estimator, represent the frequency distribution of the time-series data accurately. On a yearly basis, only small variations in α and β parameters are observed. Therefore, the average power density of a wind turbine at the SGP central facility would be approximately 460.5 Wm^{-2} . Figure 28b shows a bimodal wind direction distribution observed at the SGP site from 2013 to 2019. As discussed earlier, the winds are predominantly southerly within a secondary peak along northerly wind directions. Easterly winds are also observed for a short time period. Westerly wind directions are seldom observed at the SGP site, as indicated by the clear dip in the histogram from approximately 250 degrees to 280 degrees. Therefore, wind farms to the west-southwest of the SGP C1 are expected to marginally impact the long-term measurements at the SGP site in these wind directions.

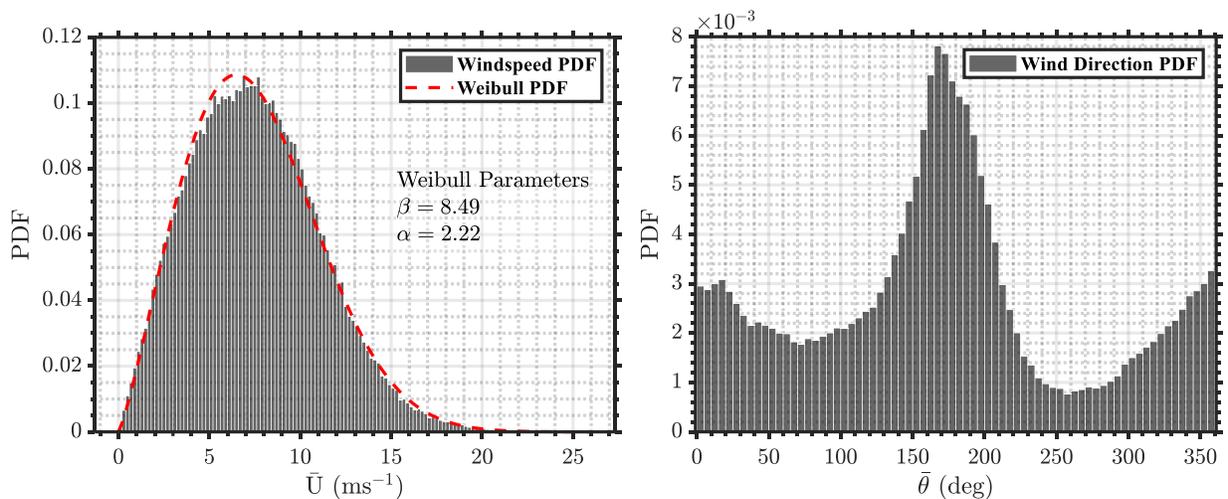


Figure 28. Wind speed and direction distribution at 90 m AGL from 2013 to 2019 at the SGP central facility.

5.3 Boundary Layer Turbulence

Turbulence is a key process within the atmospheric boundary layer, which controls the vertical exchange of momentum, heat, and moisture. In a CBL, vertical velocity variance has been shown to accurately depict the extent of the turbulent atmospheric boundary layer (Berg et al. 2017). Height-resolved measurements of clear-air vertical velocity variance, skewness, and kurtosis are useful in understanding turbulent mixing in the atmospheric boundary layer, convective initiation, and cloud life cycles. Figure 29 shows the hourly averaged vertical velocity

variance from 2014 to 2019. Peak convective activity is observed at approximately 20:00 hours UTC (~14:00 hours local time) and the maximum variance is observed below approximately 500 m AGL. The vertical velocity variance profile is observed to first increase with height and then decrease to the top of the boundary layer. At the SGP site, the scaled vertical velocity variance profile (w/w_* , where w is the vertical velocity and w_* is the convective velocity scale) is expected to peak at $\sim 1/3 z/z_i$, (where z_i is the boundary layer height and z is the height AGL) (Berg et al. 2017; Krishnamurthy et al. 2021). The strength of the vertical velocity variance varies each year but the height of peak variance is generally between 300 and 400 m AGL.

The top of the boundary layer can be classified when the turbulence is near zero. The Doppler lidar signal strength is a function of aerosol loading in the atmosphere, which generally reduces with height. Therefore, when the aerosol loading above the boundary layer height is acceptable for a good return, the lidar provides accurate measurements above the boundary layer height (for example during transition time periods). Nocturnal vertical velocity variance estimates are extremely low to detect any turbulence levels at the SGP site; therefore, TKE or dissipation rate profiles can be used to detect the level of turbulence.

A few system issues have been reported previously (Berg et al. 2017). In data prior to 2015, the lidar system at the SGP C1 observed meandering of vertical velocity variance with height (see Figure 29a, year 2014). The true reason for this meandering was not provided by the vendor (Halo Photonics), and after maintenance in 2015 the system showed realistic velocity variance profiles. Several techniques were used to filter out the meandering (such as local meander min/max fits, interpolation techniques etc.), but because of the lack of knowledge about the true vertical velocity variance behavior during these times, the lidar mentors are unsure of the accuracy of these various techniques. Although the lidar shows meandering with height, the structure of the vertical velocity variance profiles is observed to be representative of other years. These vertical velocity variance measurements can still be used to reasonably estimate the turbulence and boundary layer heights (using machine learning algorithms, discussed later). The lowest vertical velocity variance was observed in year 2019, compared to the rest of the years (i.e., from 2012 to 2020). The reason for this behavior is still under investigation, because no indication of improper system behavior or a large-scale atmospheric phenomenon at the SGP site has been identified that might explain the low variance estimates.

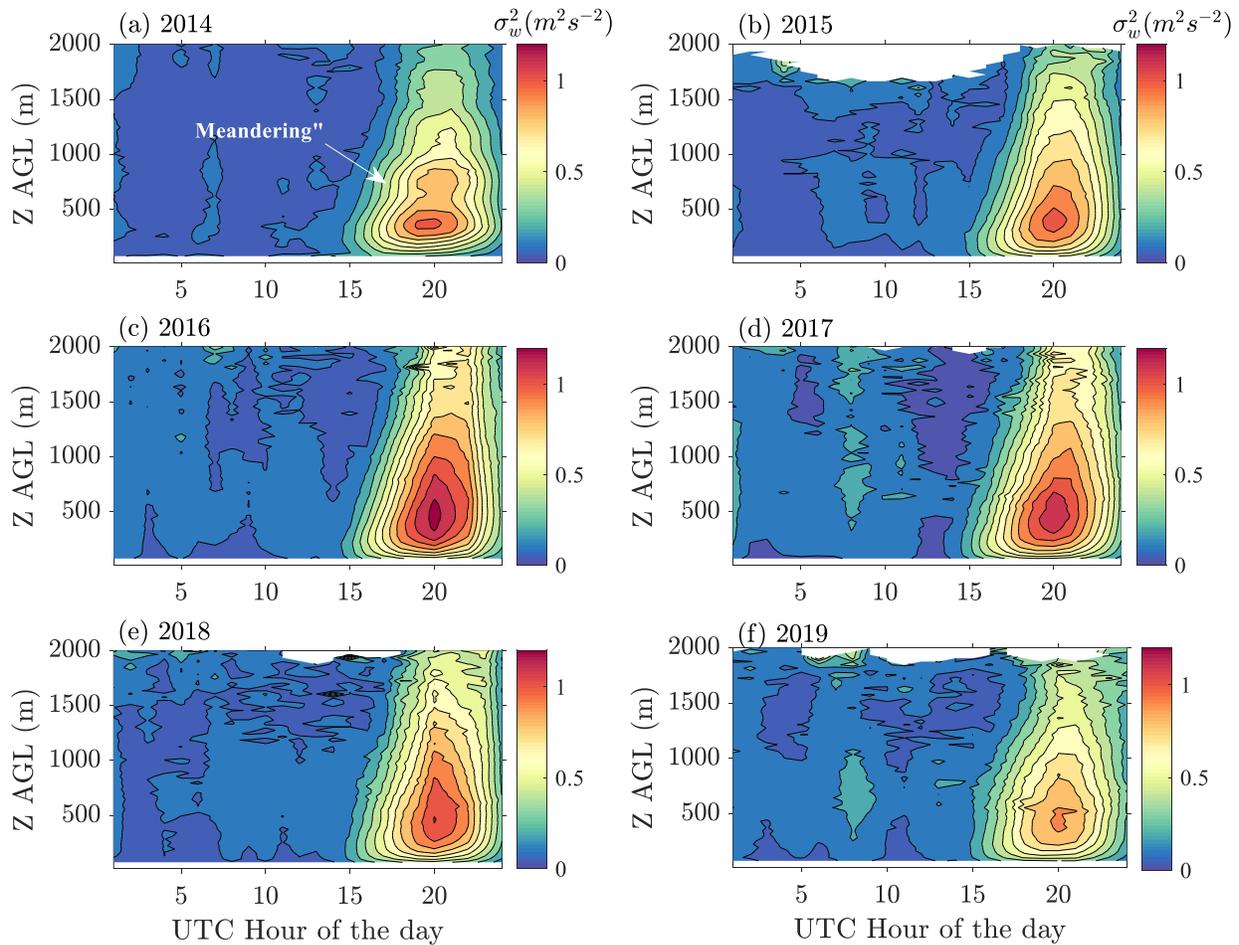


Figure 29. Hourly averaged vertical velocity variance ($\bar{\sigma}_w^2$) from 2014 to 2019 (a-f) at the SGP central facility. Colored contours of $\bar{\sigma}_w^2$ for every $0.1 \text{ m}^2\text{s}^{-2}$ are displayed. Data availability below 5 percent are not displayed. Apart from year 2016 (subplot c), which had 15 percent less data availability compared to other years, all other years had a minimum data availability greater than 95 percent at the lowest range-gate of 90 m.

Figure 30 shows hourly averaged vertical velocity variance estimates for the four seasons observed at the SGP site (i.e., DJF, MAM, JJA, and SON) from January 2016 to May 2020. Data sets prior to 2015 were not used because of the meandering artifact mentioned above. The largest vertical velocity variance or turbulence is observed during summer times when peak vertical velocity variances of greater than $2 \text{ m}^2\text{s}^{-2}$ at approximately 500 m AGL are observed. The lowest daytime vertical velocity variance is observed during winter convective time periods when peak vertical velocity variances below $1 \text{ m}^2\text{s}^{-2}$ at approximately 300 m AGL are observed. Similar convective turbulence levels are observed in fall and spring, with moderately higher vertical velocity intensities during spring seasons. During spring and summer months, nocturnal convection initiation events (Reif and Bluestein 2017) occur, which increase the vertical velocity variance during certain time periods. Generally, increased storm activity is observed during spring and summer months. As shown previously, increase precipitation is also observed during spring/summer months; therefore, higher vertical velocity variance during nighttime conditions could be due to nocturnal elevated convection events at the SGP site (Weckwerth and Romatschke 2019). During summer months, the LLJ is also observed to substantially increase

mixing below the LLJ height. Figure 31 shows the average horizontal wind speed and vertical velocity variance profiles from 20 to 27 June 2018. A period of consistent LLJ was observed from 20 to 27 June 2018, which shows higher mixing (aka higher vertical velocity variance) near the surface and up to ~400 m AGL. When the LLJ is well-defined and consistent, enhanced mixing is observed near the surface. Overall, nocturnal turbulence at the SGP site is complex and is a topic of future research. The lidar data availability is also observed to be a function of seasons; winter seasons show low aerosol loading and summer months show high aerosol concentration at the SGP site. Figure 32 shows vertical velocity variance wind roses at 90 m and 500 m AGL from 2012 to 2019. One notable observation at 500 m is that several instances of high vertical velocity variance are observed from southerly wind directions. Predominantly lower vertical velocity variance estimates are associated with nocturnal conditions.

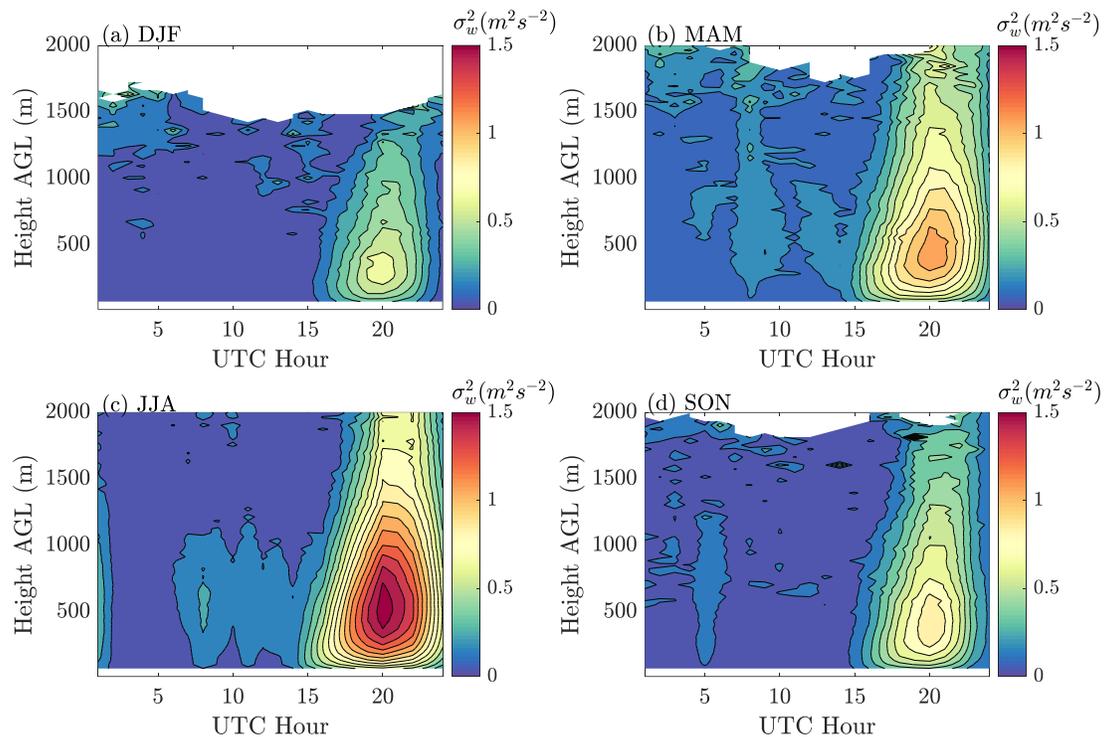


Figure 30. Hourly averaged vertical velocity variance ($\bar{\sigma}_w^2$) estimates for (a) DJF, (b) MAM, (c) JJA, and (d) SON seasonal variability from January 2016 to May 2020 at the SGP central facility. Colored contours of $\bar{\sigma}_w^2$, for every 0.1 ms^{-1} are displayed. Data availability below 5 percent are not displayed.

Figure 33 shows hourly averaged vertical eddy dissipation rate estimates at 100 m AGL from January 2013 to May 2020 for each of the four seasons (DJF, MAM, JJA, and SON). The dissipation rate estimated using the spectral slope algorithm (see section 4.3) generally underpredicts the dissipation rate estimated from a sonic anemometer during nocturnal conditions (Wilzack et al. 2019). Structure function methods accounting for the volume-averaging effects of the lidar tend to provide better dissipation rate estimates than sonic anemometers (Frehlich et al. 2006; Krishnamurthy et al. 2011; Bodini et al. 2019). At the SGP site, we see a maximum average dissipation rate of $5 \times 10^{-4} \text{ m}^2 \text{ s}^{-3}$ during peak convective periods at 100 m AGL. During nocturnal conditions, a near constant dissipation rate is observed at 100 m AGL, and it is generally below $1 \times 10^{-4} \text{ m}^2 \text{ s}^{-3}$. On a seasonal basis, higher dissipation rates are observed during summer and fall than during spring and winter. Yearly dissipation rate

variability is observed but overall similar trends are observed all years, except for years 2018 and 2019 when significantly lower dissipation rates are observed. This is consistent with prior observations of vertical velocity variance in 2019, which also were lower compared to other years. Further investigation is required to assess whether the cause is either lidar system degradation (since summer 2018) or any large-scale atmospheric phenomenon causing lower turbulence levels at the SGP site.

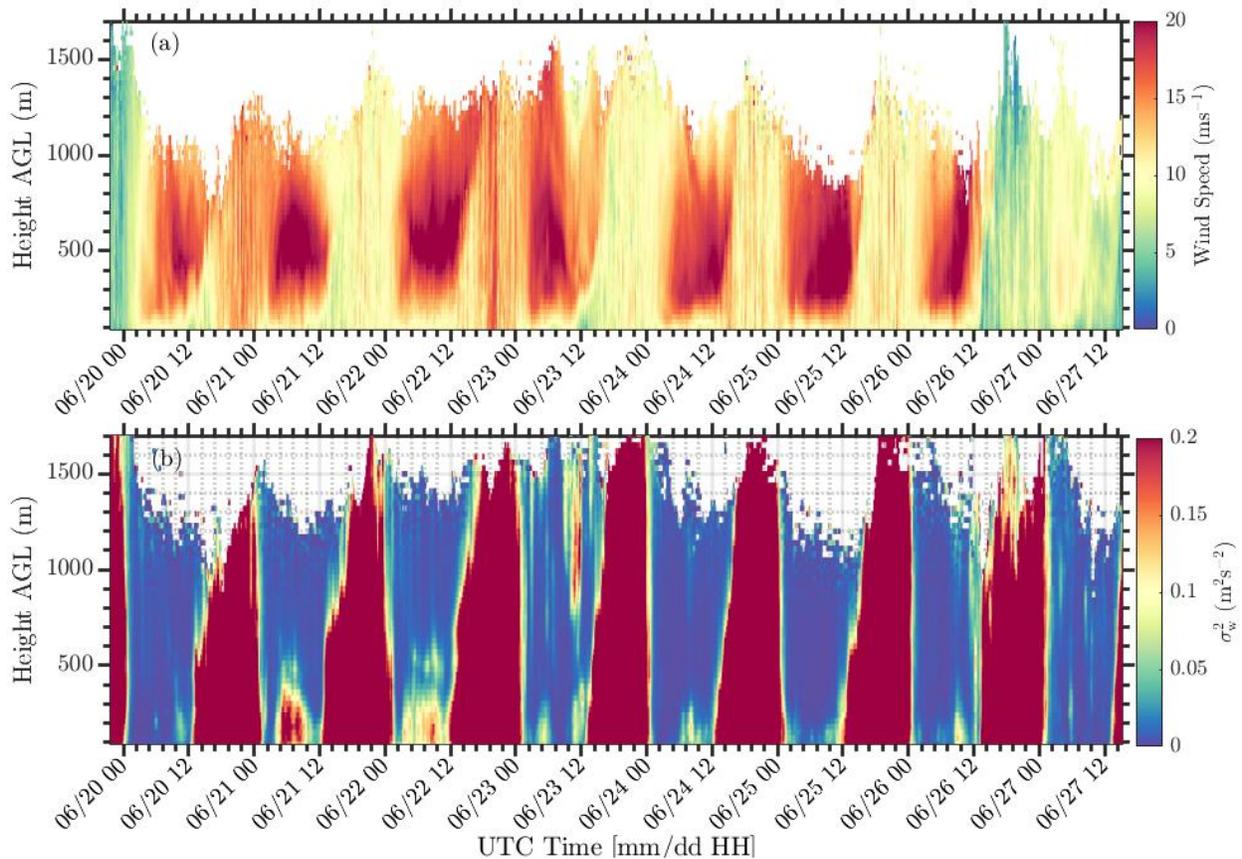


Figure 31. (a) Average horizontal wind speed and (b) average vertical velocity variance from 20 June to 27 June, 2018 at the SGP central facility.

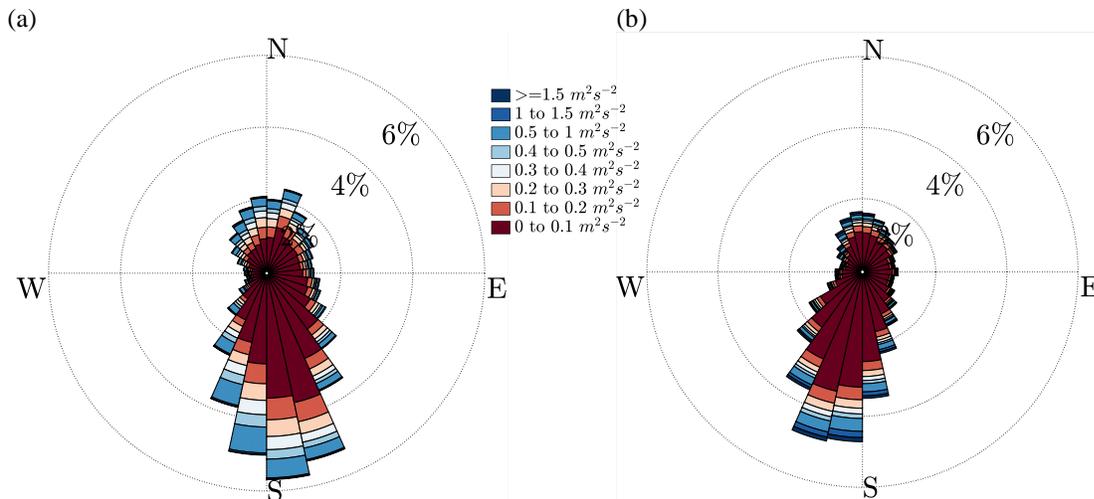


Figure 32. Average vertical velocity variance rose at (a) 105 m and (b) 500 m AGL from 2012 to 2019 at the SGP central facility.

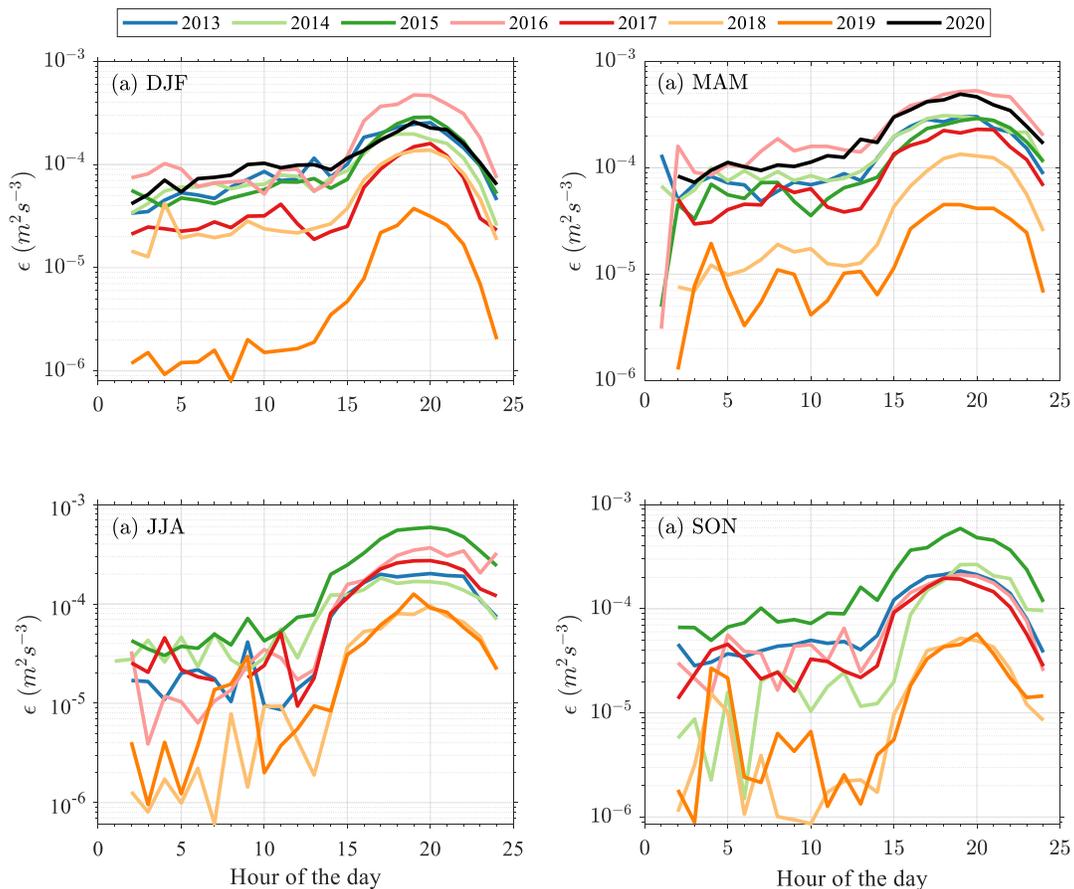


Figure 33. Hourly averaged eddy dissipation rate from vertical staves from January 2013 to May 2020 at 100 m AGL for seasons (a) DJF, (b) MAM, (c) JJA, and (d) SON at the SGP central facility.

5.4 Boundary Layer Water Vapor and Relative Humidity

In this section we examine the climatology of the water vapor column at the SGP central facility as observed by the RL from January 2016 through June 2020. This includes an analysis of the water vapor mixing ratio, RH, and horizontal water vapor mass flux.

Figure 34 through Figure 37 display the mean diurnal variation in water vapor mixing ratio (q) and RH from the RL for each season. Also shown is the mean water vapor mixing ratio and RH at 10 m AGL from a nearly collocated surface meteorological station (approximately 200 m east-southeast of the RL). Figure 34 shows that there is very little diurnal variation in q during the winter months (DJF); typical values are in the range of 3 to 4 g kg⁻¹ with a slight increase during the daytime. The diurnal variability and the median q increase as the weather warms. The daytime PBL is clearly identifiable in spring (Figure 35), summer (Figure 36), and fall (Figure 37), and shallower in winter (Figure 34). Surface q measurements generally indicate less variation than that observed by the RL in the boundary layer above the surface. Figure 36 suggests that in summer q gradually decreases with the development of the nocturnal boundary layer. After sunrise, q increases with the buildup of the CBL. This results in a local minimum in the surface q at or shortly after sunrise (this minimum is seen in all seasons). The RL measurements clearly show a similar minimum lagging the one seen at the surface.

We note that the mixing ratio is a direct measurement of the RL, which is determined by the calibration of the water vapor-to-nitrogen SNR, as described in Section 3.3. The radiosonde provides direct measurements of RH, pressure, and temperature. As mentioned in Section 3.3, radiosondes are typically launched four times daily at the SGP central facility (nominally at 05:30, 11:30, 17:30 and 23:30 UTC daily). Temperature and pressure profiles from the radiosonde are linearly interpolated to the sampling intervals of the RL q profiles, and then used to convert q to RH. Because the frequency of radiosonde launches is not sufficient to capture rapidly changing conditions, it is likely that use of the interpolated radiosonde data may introduce artifacts, particularly during the morning and evening transition periods.

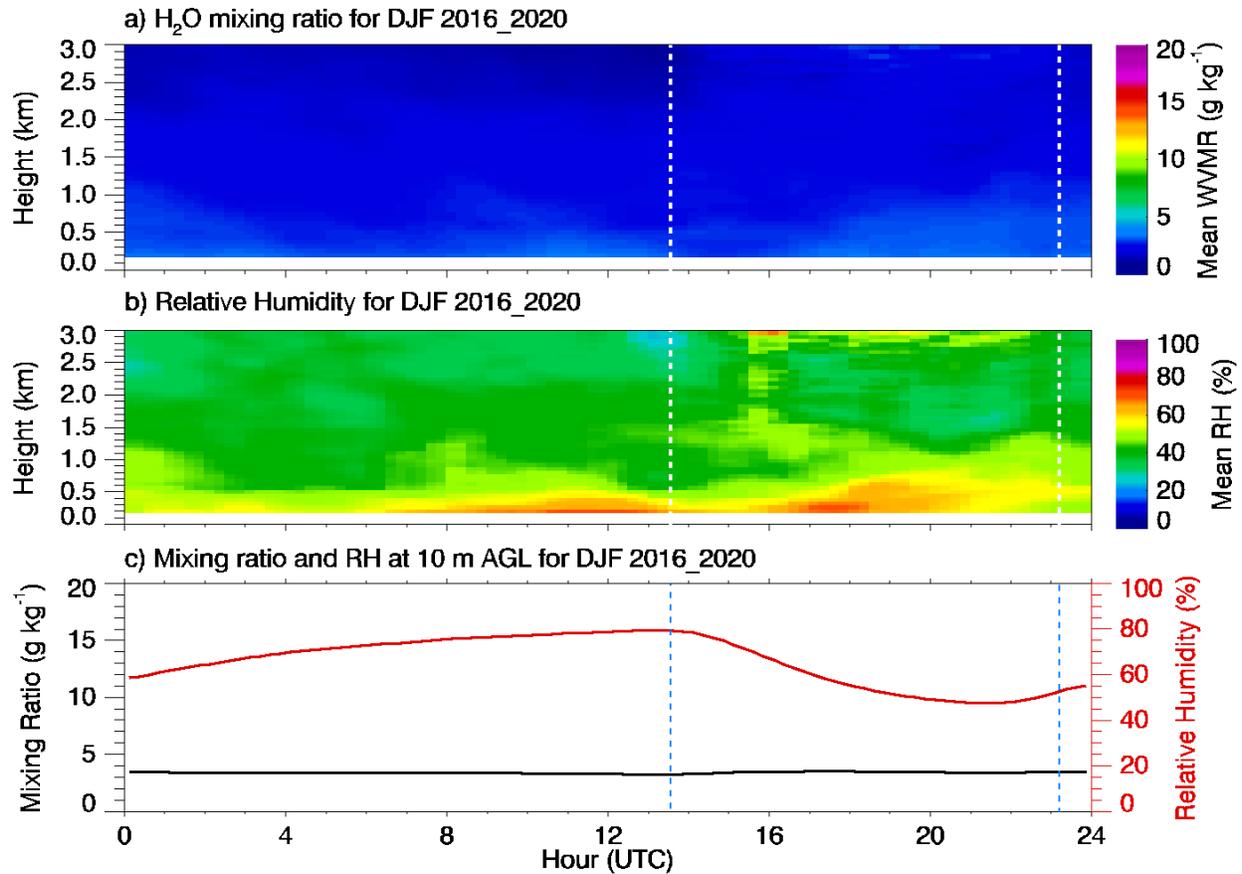


Figure 34. Diurnal mean (a) water vapor mixing ratio and (b) RH as observed by the Raman lidar at the SGP central facility for December, January, and February (DJF), 2016 through 2020. Panel (c) shows the diurnal mean mixing ratio (black) and RH (red) from the met station at 10 m AGL. Sunrise (~13:30 UTC) and sunset (~23:10 UTC) times are indicated by the dashed vertical lines in each panel.

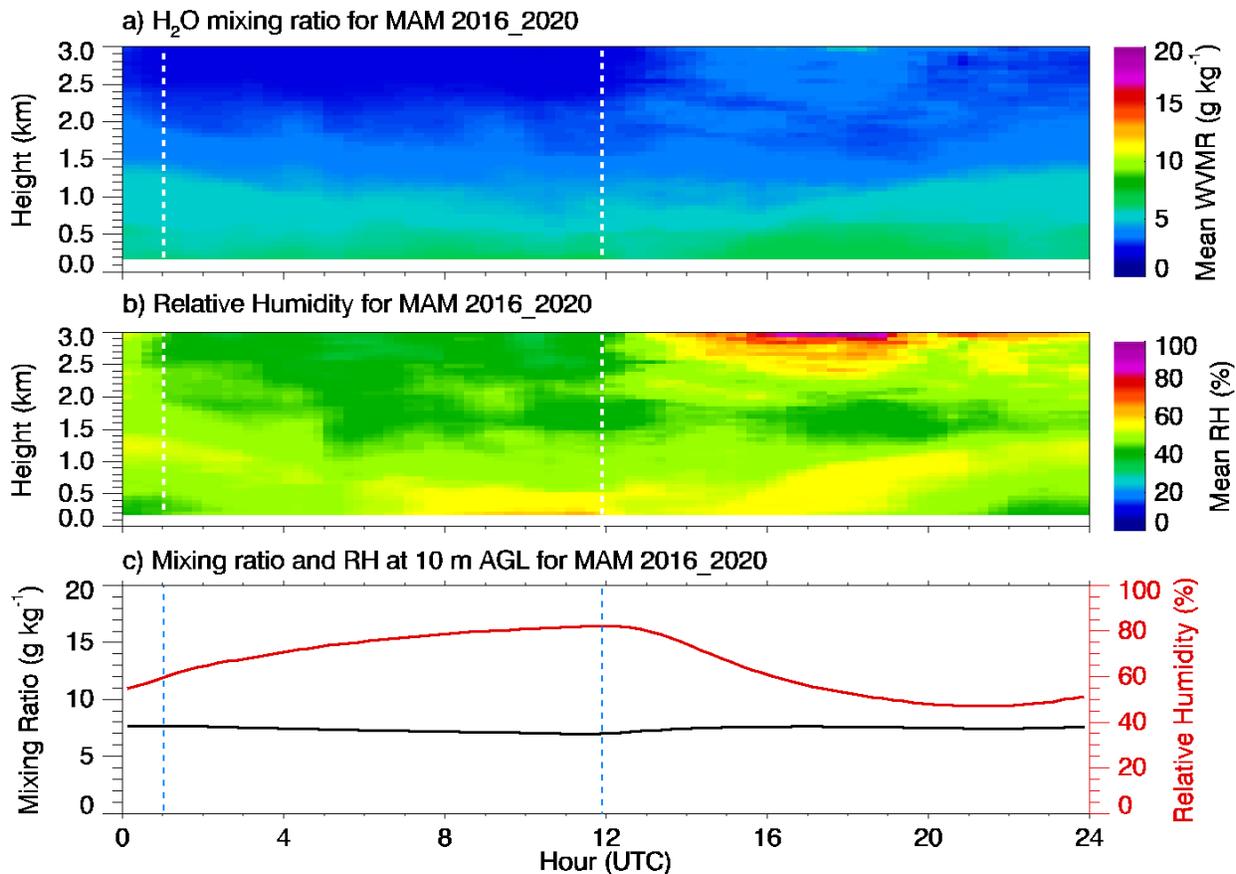


Figure 35. Diurnal mean (a) water vapor mixing ratio and (b) RH as observed by the Raman lidar at the SGP central facility for March, April, and May (MAM), 2016 through 2020. Panel (c) shows the diurnal mean mixing ratio (black) and RH (red) from the met station at 10 m AGL. Sunrise (~11:50 UTC) and sunset (~01:00 UTC) times are indicated by the dashed vertical lines in each panel.

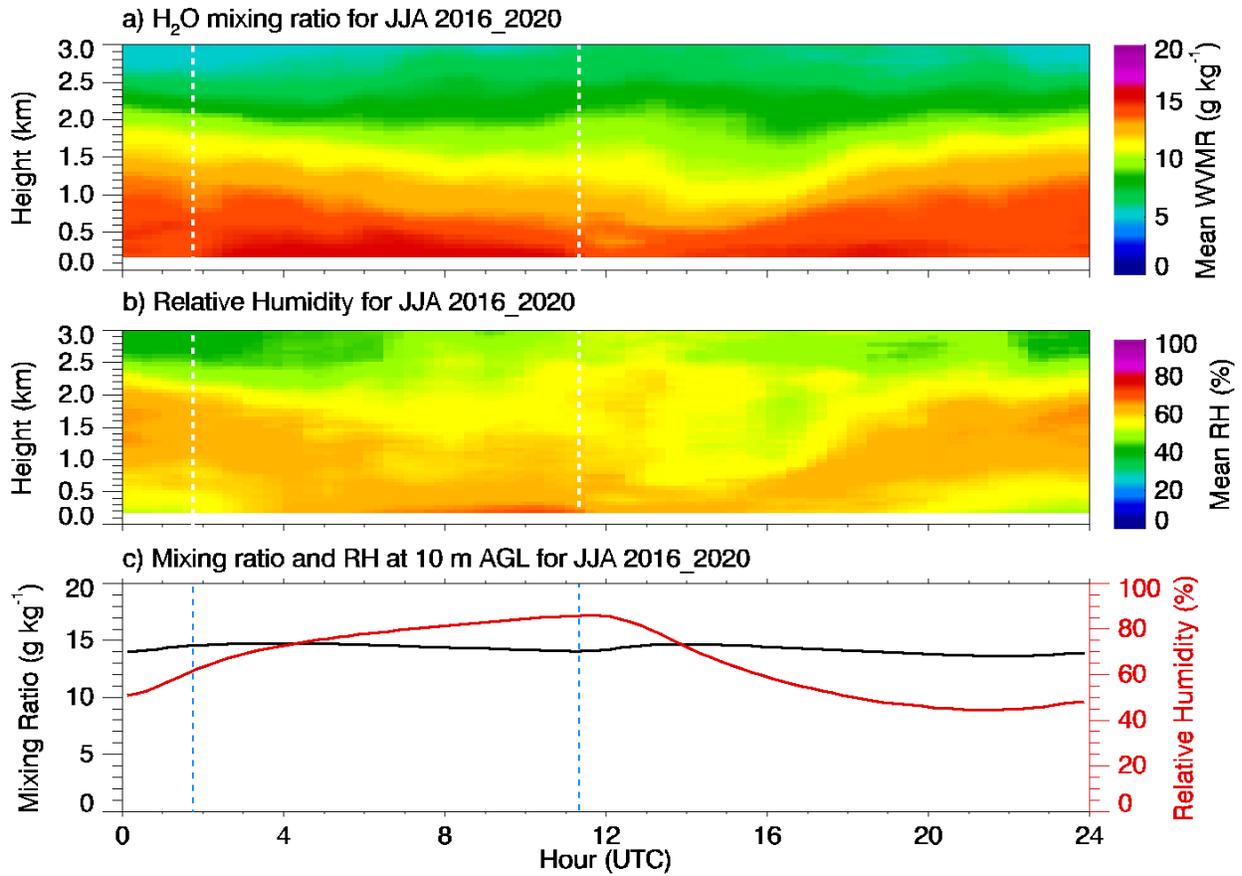


Figure 36. Diurnal mean (a) water vapor mixing ratio and (b) RH as observed by the Raman lidar at the SGP central facility for June, July, and August (JJA), 2016 through 2020. Panel (c) shows the diurnal mean mixing ratio (black) and RH (red) from the met station at 10 m AGL. Sunrise (~11:20 UTC) and sunset (~01:45 UTC) times are indicated by the dashed vertical lines in each panel.

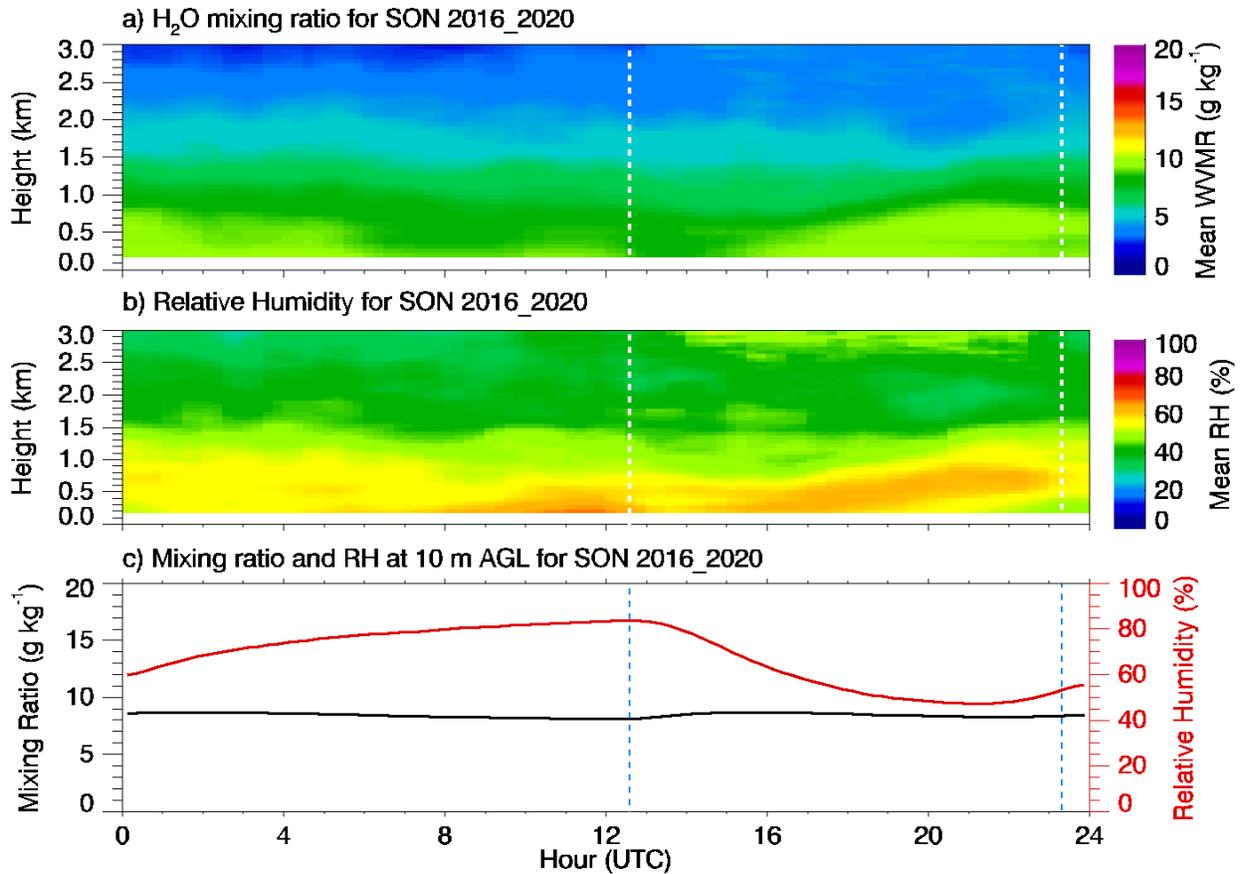


Figure 37. Diurnal mean (a) water vapor mixing ratio and (b) RH as observed by the Raman lidar at the SGP central facility for September, October, and November (SON), 2016 through 2020. Panel (c) shows the diurnal mean mixing ratio (black) and RH (red) from the met station at 10 m AGL. Sunrise (~12:30 UTC) and sunset (~23:10 UTC) times are indicated by the dashed vertical lines in each panel.

In addition to water vapor, it is also important to understand the directional source of water vapor. For this study, we examined horizontal moisture mass flux as computed from the product of the water vapor density from the RL with the horizontal wind speed measurement from a collocated Doppler lidar. This quantity represents the mass of water vapor passing through a unit area per unit time (i.e., it has units of mass per area per time).

Figure 38 through Figure 41 show the median water vapor mass flux as a function of time of day during each of the four seasons. Also shown is the median wind direction over the same averaging period. During winter, Figure 38 indicates the most significant source of water vapor is from the west and northwest (mostly above 1 km AGL) during nighttime. In spring the moisture transport increases overall, with higher values above 1 km AGL coming mostly from the southwest quadrant. This is also similar to the fall periods. Figure 40 clearly shows a large spike in the water vapor flux associated with the nocturnal LLJ during summer. The largest values appear to occur between the surface and the nose of the LLJ. In fall, the most significant source of water vapor is observed from the east.

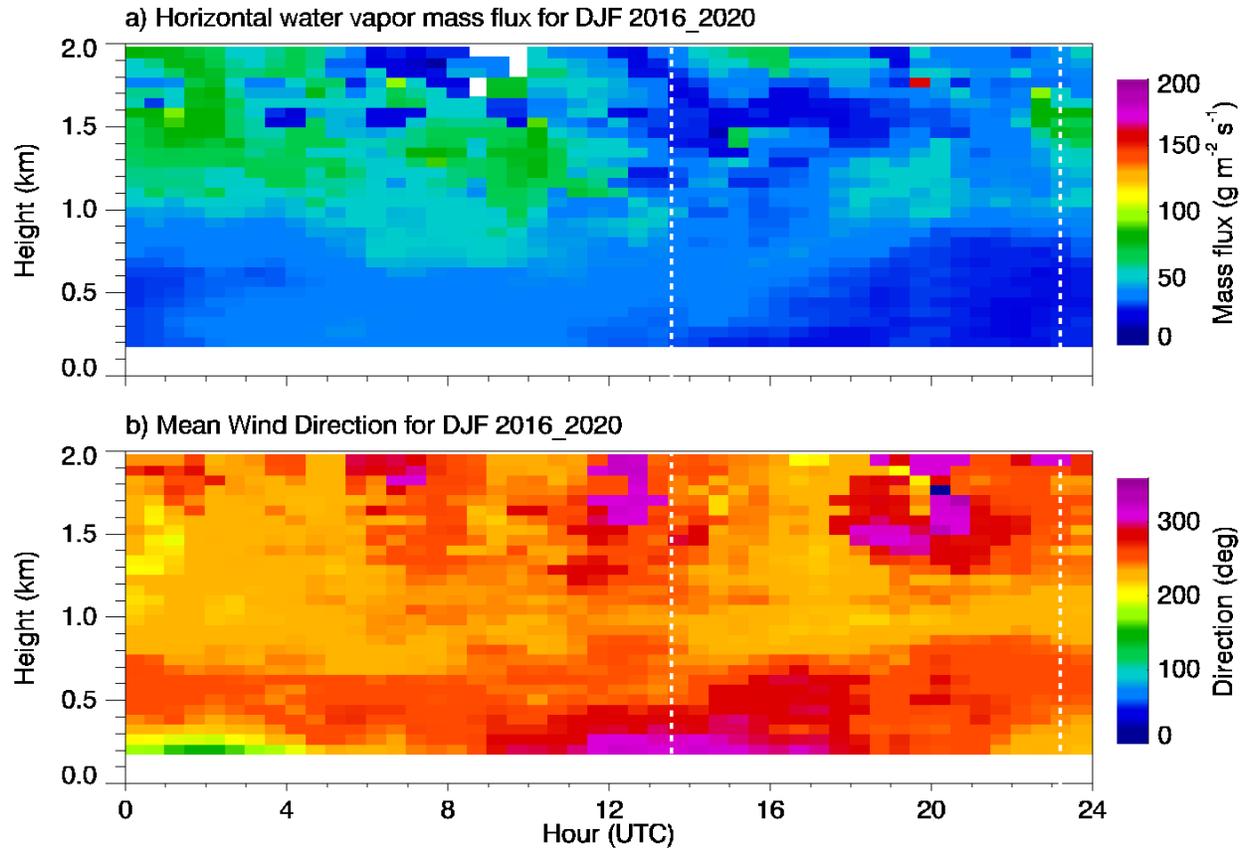


Figure 38. Diurnal mean (a) water vapor mass flux and (b) wind direction as observed by the Raman and Doppler lidars at the SGP central facility for DJF, 2016 through 2020. The dotted lines indicate sunset (~23:10 UTC) and sunrise (~13:30 UTC).

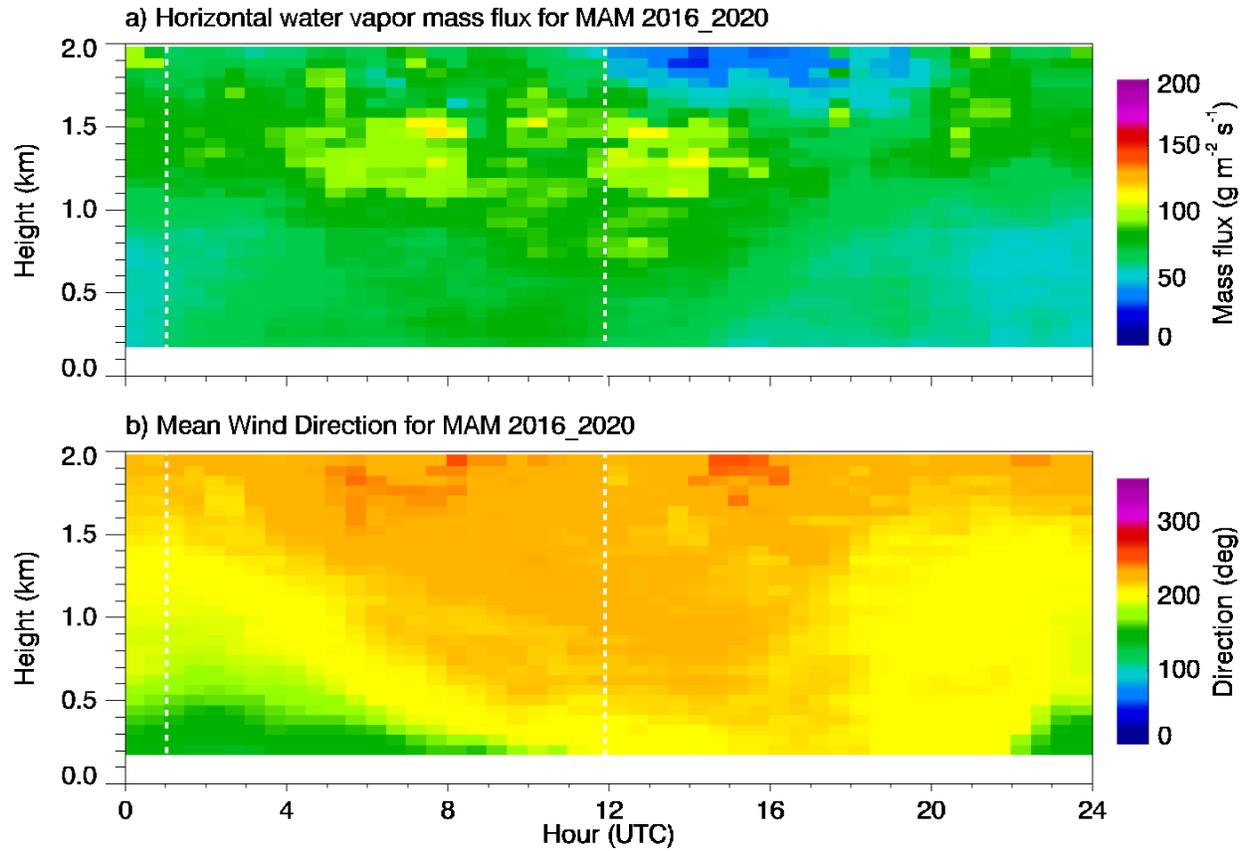


Figure 39. Diurnal mean (a) water vapor mass flux and (b) wind direction as observed by the Raman and Doppler lidars at the SGP central facility for MAM, 2016 through 2020. The dotted lines indicate sunrise (~11:50 UTC) and sunset (~01:00 UTC).

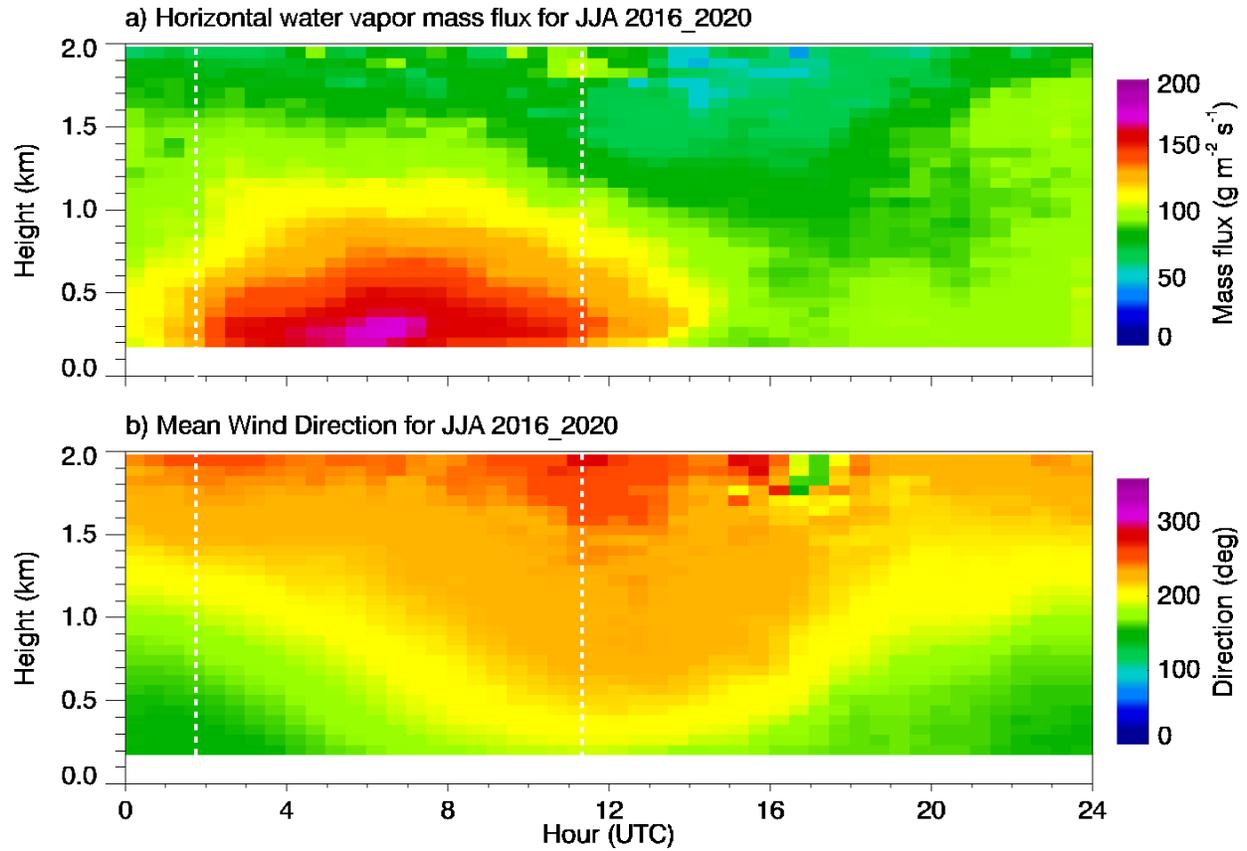


Figure 40. Diurnal mean (a) water vapor mass flux and (b) wind direction as observed by the Raman and Doppler lidars at the SGP central facility for JJA, 2016 through 2020. The dotted lines indicate sunrise (~11:20 UTC) and sunset (~01:45 UTC).

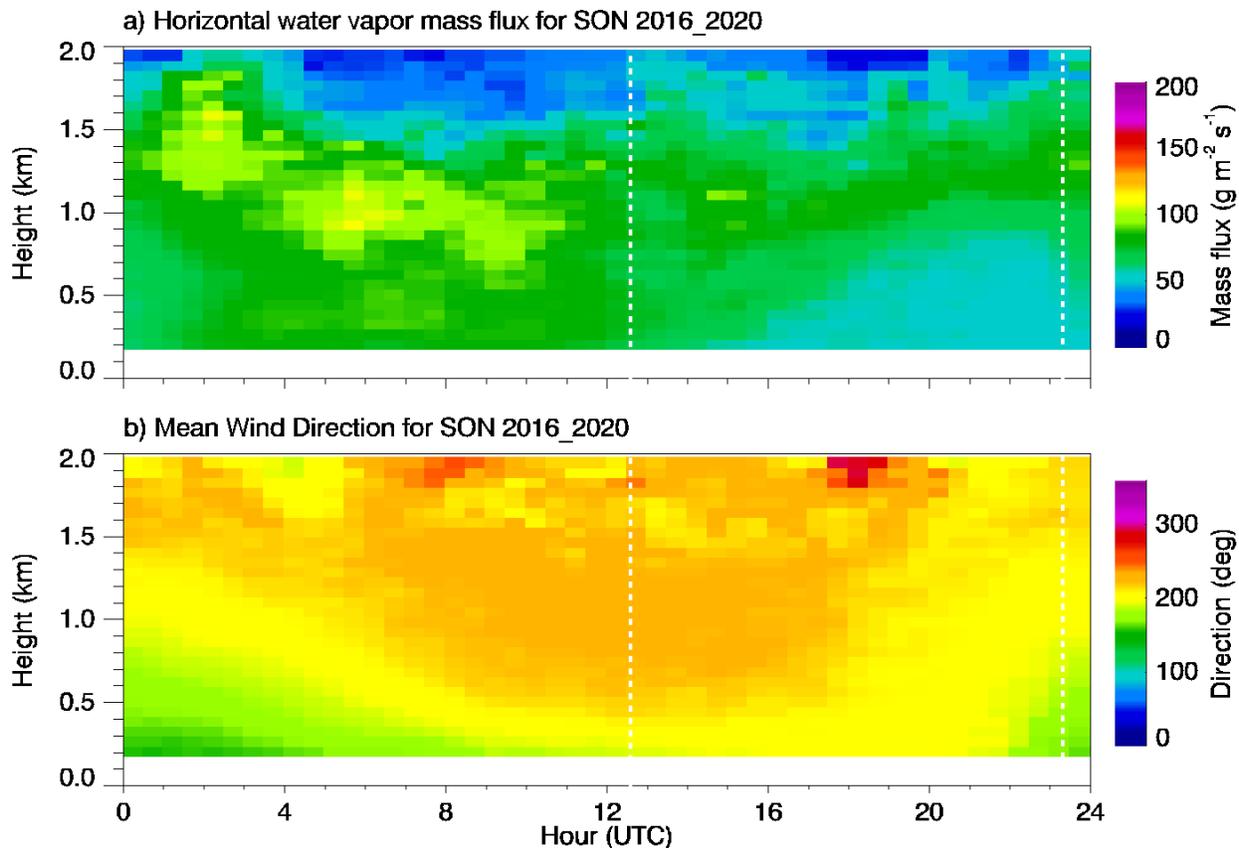


Figure 41. Diurnal mean (a) water vapor mass flux and (b) wind direction as observed by the Raman and Doppler lidars at the SGP central facility for SON, 2016 through 2020. The dotted lines indicate sunrise (~12:30 UTC) and sunset (~23:10 UTC).

Figure 42 shows the median surface water vapor flux as a function of wind direction. As indicated, all seasons show a maximum for wind directions from the south. There is some indication of a bimodal structure during the fall, with a weak maximum occurring in the northerly direction. In these plots we observed that the percentage of mass flux coming from the southern quadrant (centered on south) was 47 percent in the winter, 52 percent in the spring, 67 percent in the summer, and 66 percent in the fall. This compares to 25 percent if the distributions were uniform. As shown in Figure 24, the nocturnal LLJ at the SGP site is generally observed from the south during all seasons and is the primary source of water vapor coming into the great plains (Berg et al. 2015).

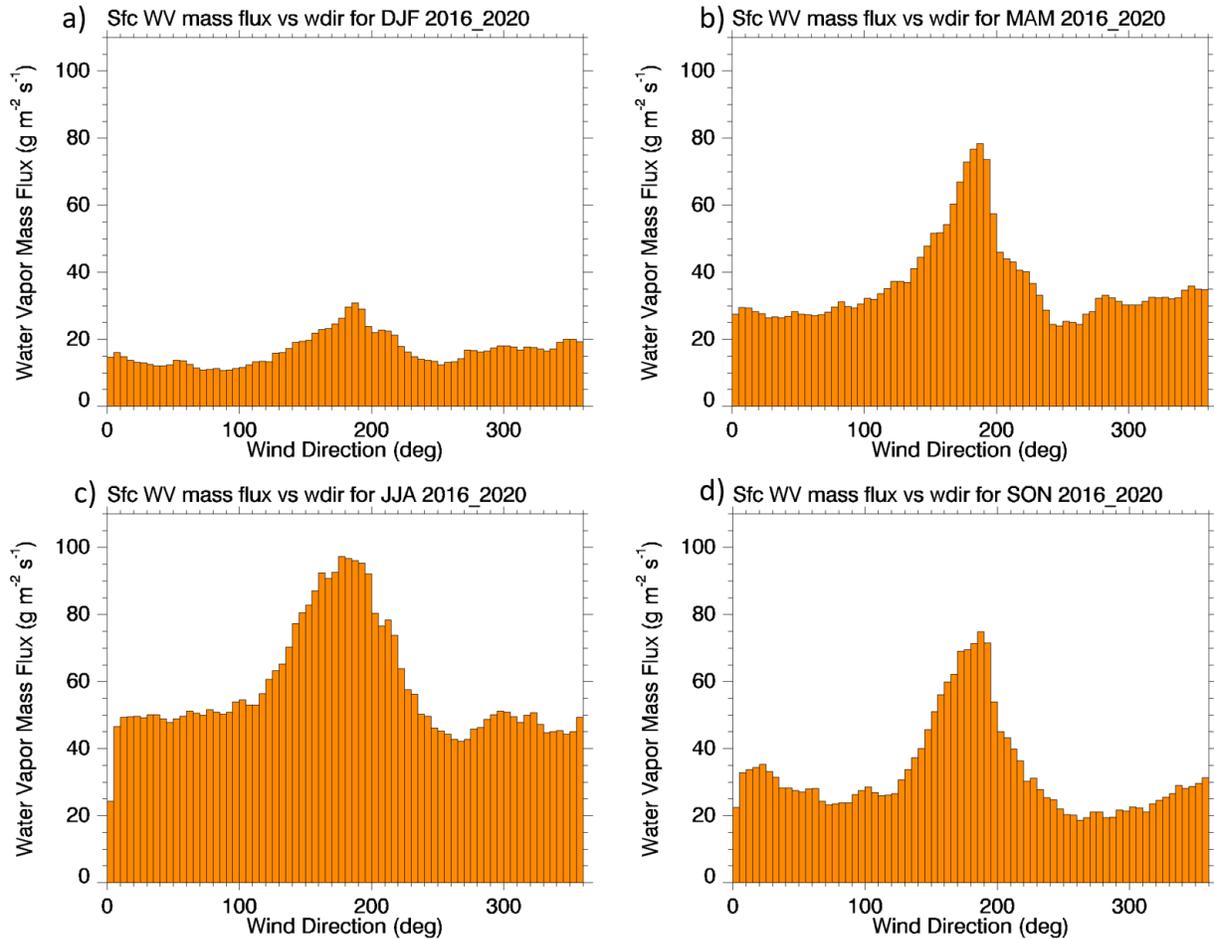


Figure 42. Median water vapor mass flux as a function of wind direction at 10 m AGL for (a) DJF, (b) MAM, (c) JJA, and (d) SON at the SGP central facility. Statistics were computed using data from January 2016 to June 2020.

With regard to field campaign planning and modeling studies at the SGP site, boundary layer profiles from Doppler lidars and RLs provide insight into some key aspects:

1. LLJ is observed to increase the water vapor flux into the SGP site. LLJ variability from year to year can increase (decrease) the moisture intake from the Gulf of Mexico to the Southern Great Plains, thereby affecting the local climate significantly (see Section 5.4).
2. With regard to nearby and future wind farms, the impact of LLJ on wind turbines should be assessed. Because the mean LLJ height is below 500 m AGL, increased wind shear below a LLJ can increase fatigue and dynamic equivalent loads. Further analysis of LLJ height seasonal variability and distribution is provided in Section 6.1.
3. Low-altitude wind directions are predominantly southerly during most of the years, with little variability from year to year. Winds are observed to veer southwesterly at higher altitudes. Accounting for this is essential during campaign designing, especially when designing lidar and radar scan patterns.
4. Daytime winds are observed to be uniform with minimal wind shear, except during winter.

5. Seasonal variability is high at the SGP site; winter and spring feature high winds through and above the boundary layer and summer features the lowest winds. Although summer winds are lower, they are observed to be within Region 2 (operational mode of the wind turbine during which the turbine captures as much power as possible from the wind) of a wind turbine power curve, where the turbine controller maximizes power output. The effect of seasonal winds on wind turbine loads, wakes, and annual energy production needs to be further studied near the SGP site.

5.5 Cloud Base and Boundary Layer Height

Cloud base height can be retrieved from a Doppler lidars when staring vertically up (Newsom et al. 2015). A thorough validation of the algorithm with measurements from a ceilometer was performed by Newsom et al. (2015). Figure 43 shows distribution of cloud base height from 2012 to 2019 at the SGP C1. Predominantly low-level clouds below 500 m AGL are observed at the SGP site. Sunny Lim et al.'s (2019) climatological analysis showed that the SGP site encounters a high frequency of low-level clouds and cirrus clouds. Formally, a cloud forms when the humid air becomes saturated and water vapor condensation occurs by either cooling air to its dew-point (temperature decreases) or by adding enough moisture to reach saturation (dew-point increases). Fog in general can be considered to be a stratus cloud at or near the ground or ocean surface ("a cloud when it is low or touching the surface"). Several instances of fog or mist or haze are observed periodically at the SGP site, especially in the early morning. Fog, mist, and haze are the terms generally used to describe low visibility caused by water droplets or/and dry particles suspended in the air. In terms of visibility, fog reduces it to less than 1 km (the international definition of fog), mist reduces visibility to between 1 km and 2 km, while haze can reduce visibility to between 2 km and 5 km (Fernando et al. 2021). The Doppler lidar signal attenuates rapidly through a fog layer due to increased scattering. The droplets in the fog or the rain absorb or scatter the near infrared laser. The severity depends on the water content percentage, or droplet size distribution. Visibility is inversely proportional to the lidar range and extinction coefficient (Stoelinga and Warner 1999). The radial velocity measurements within the fog/cloud layer are generally not considered reliable from a Doppler lidar because of the effects of hygroscopic growth on the laser signal, and a possible double peak in Doppler spectra of wind and droplet concentration sometimes observed during conditions with higher precipitable water. Therefore, measurements affected by fog/within a cloud were removed from this analysis. Although Doppler lidars at the SGP site can measure up to 10 km, we have limited our analysis to clouds below 4 km AGL.

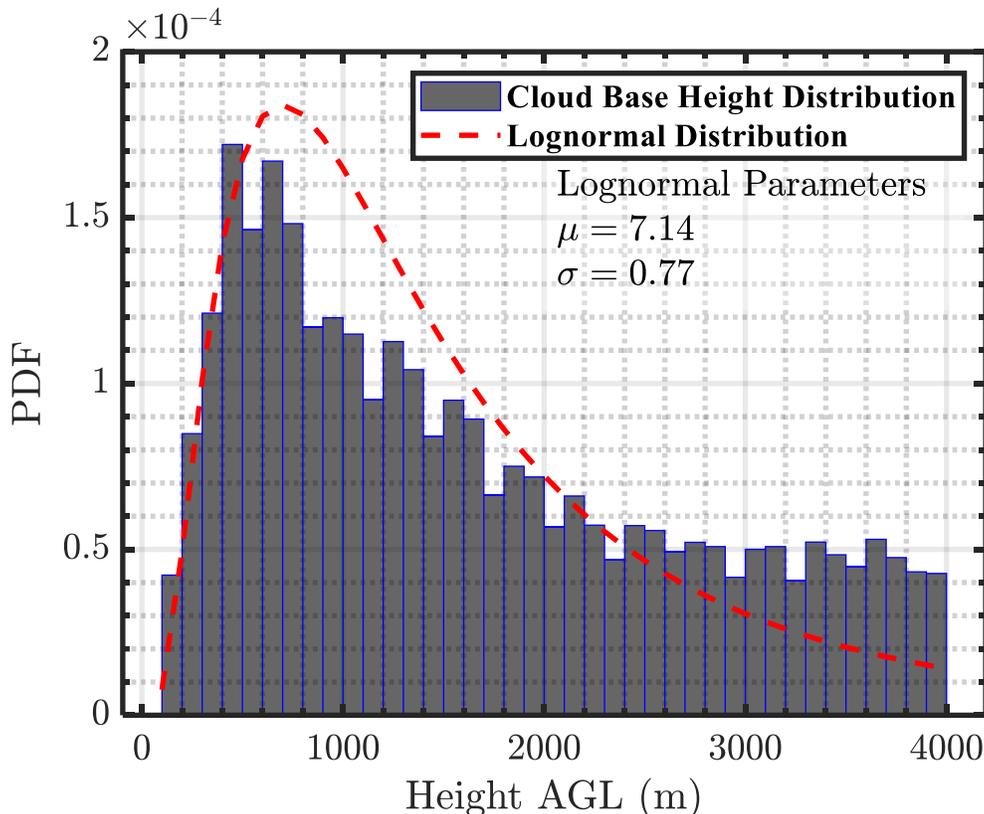


Figure 43. Cloud base height distributions as observed from a Doppler lidar from 2012 to 2019 at the SGP central facility.

A seasonal distribution of cloud base height is provided in Figure 44. A lognormal distribution was fit to the cloud base height data to represent a skewed distribution. Winter (DJF) and spring (MAM) months show the highest low-level cloud probability at the SGP site. A temperature inversion layer often occurs immediately above a cloud layer that cools radiatively and thereby strengthens the capping temperature inversion. Therefore, the cloud height can be considered to be a proxy for inversion height, but not always (Shin and Park 2020). During summer months (JJA), clear sky conditions are generally observed at the SGP site and the cloud distribution peaks at 1500 m AGL. Similar cloud height distributions are observed during spring and fall seasons, with highest distributions above 500 m and below 1000 m AGL.

Cloud height measurements from lidars (either Raman, Doppler, or ceilometer) at the SGP site are an instantaneous snapshot in space, as we observe a passing cloud at a given location. Measurements at a single location do not accurately depict the spatial distribution of clouds within the area. Satellites provide a holistic view of cloud layers over a given region. Even though most satellites are not geostationary, they can still provide information about the spatial variability of cloud layers within a domain, which is extremely helpful when evaluating model results. MODIS is a key instrument deployed onboard the Terra and Aqua satellites. Both satellites pass any given location on Earth's surface at least once every 2 days. In this report, 20 years of monthly cloud top minimum and cloud fractions from MODIS satellites were analyzed. MODIS data sets are available at high spatial resolution (1 km²), these data sets are used in cloud fraction calculations. At the SGP site, 1-degree resolution pixels are statistically grouped into four seasons (DJF, MAM, JJA, SON).

Figure 45 shows the seasonal distributions of the monthly averaged cloud fraction and cloud top height from a 1° pixel over the SGP site. As mentioned earlier, a long-term comparison of monthly averaged MODIS and cloud base height estimates from a fish-eye camera showed very good correlations (>98 percent) and RMSE below 3 percent (Wang and Zhao 2017). Satellite measurements also confirm that low-level clouds (average height below 500 m) are predominantly observed at the SGP site in all the seasons, except summer when the mean cloud level height is above 500 m. The cloud fraction estimates at the SGP site indicate lesser cloud fractions during summer and fall and higher during winter and spring. On an average, clouds cool the surface by reflecting the shortwave radiation. Therefore, seasons that have a larger cloud fraction, i.e., winter and spring, will experience relatively low surface-driven convection at the SGP site. This is confirmed in the seasonally averaged Doppler lidar vertical profiles of vertical velocity variance, shown in Figure 30, where low vertical velocity variance during these seasons is seen.

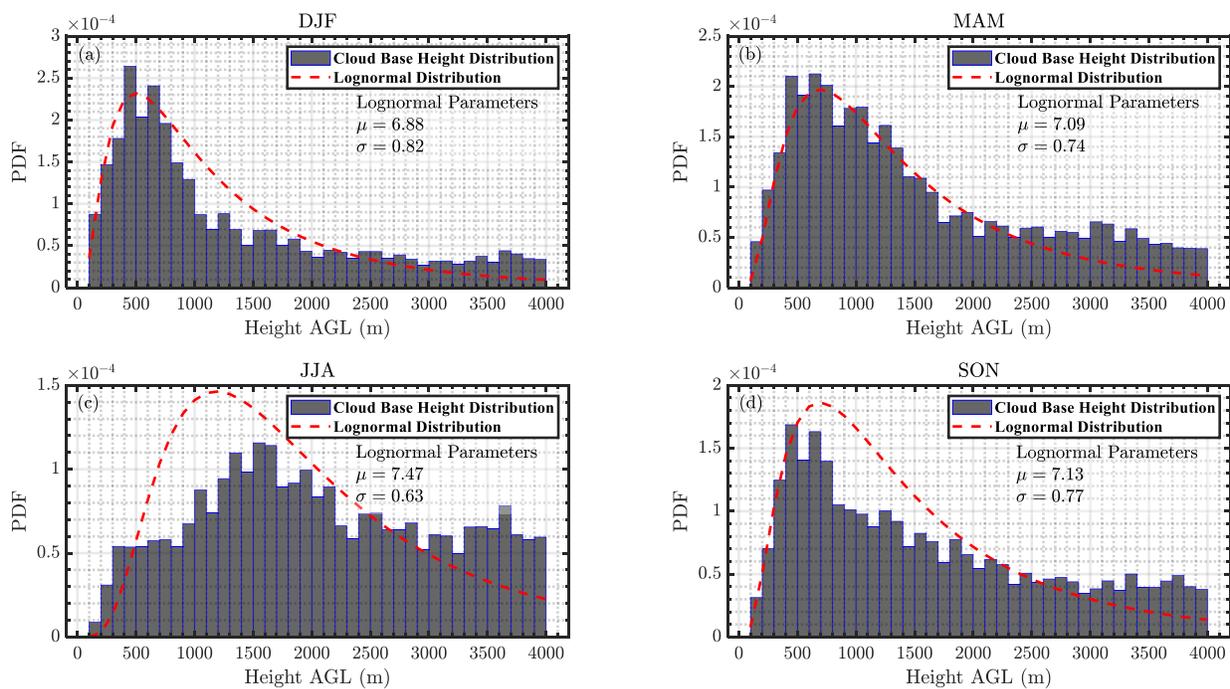


Figure 44. Hourly averaged cloud base height distributions over four seasons: (a) DJF, (b) MAM, (c) JJA, and (d) SON from 2012 to 2019 at the SGP central facility.

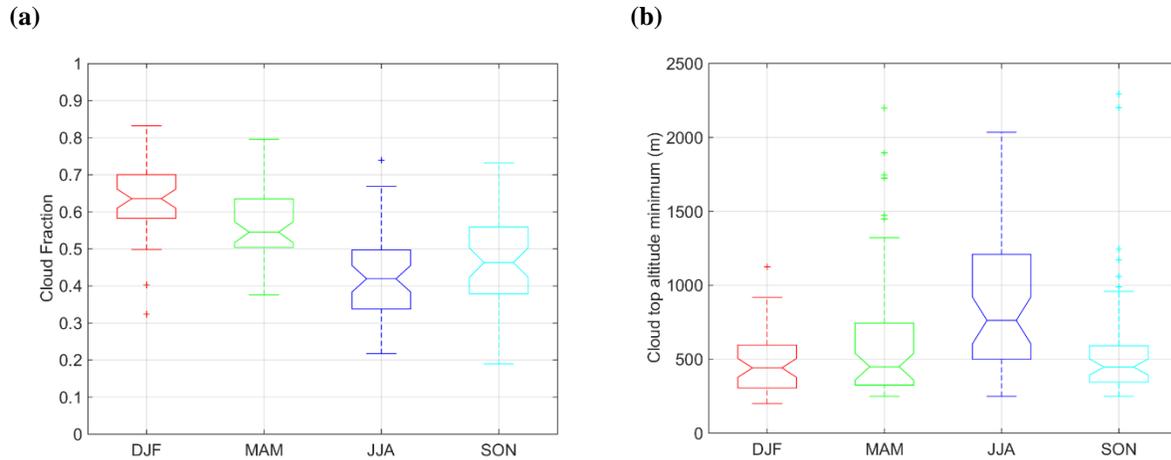


Figure 45. Box plot seasonal distribution of (a) cloud fraction and (b) cloud top minimum altitude from 1° resolution pixel over the ARM SGP site. The statistics are developed using 20 years of MODIS monthly averaged data. The center line within the box indicates the median value of the distribution, the box represents the first and third quartile of the distribution, the error bars represent the minimum and maximum values in the distribution, and the markers indicate outliers not used in the calculation of the median value.

We also examined the statistics of 10 sec cloud base height (CBH) estimates from the RL during the period from January 2016 to June 2020. Figure 46 shows the median CBH as a function of the time of day for the four seasons. Generally, CBHs are lowest in the winter and spring seasons. There is very little diurnal variability during the winter, when CBHs are generally between 500 to 900 m AGL. During spring, CBHs are smallest during the pre- and post-dawn hours. However, an odd local maximum occurs right at sunrise. During the late morning and afternoon, the CBHs increase with time as a result of cumulus forming near the top of the developing CBL. The summertime sees the largest variation in CBH throughout the course of a day, with values ranging from about 1.7 to 4.2 km. As in the spring, the largest CBHs occur near sunset, and the smallest values occur just before solar noon.

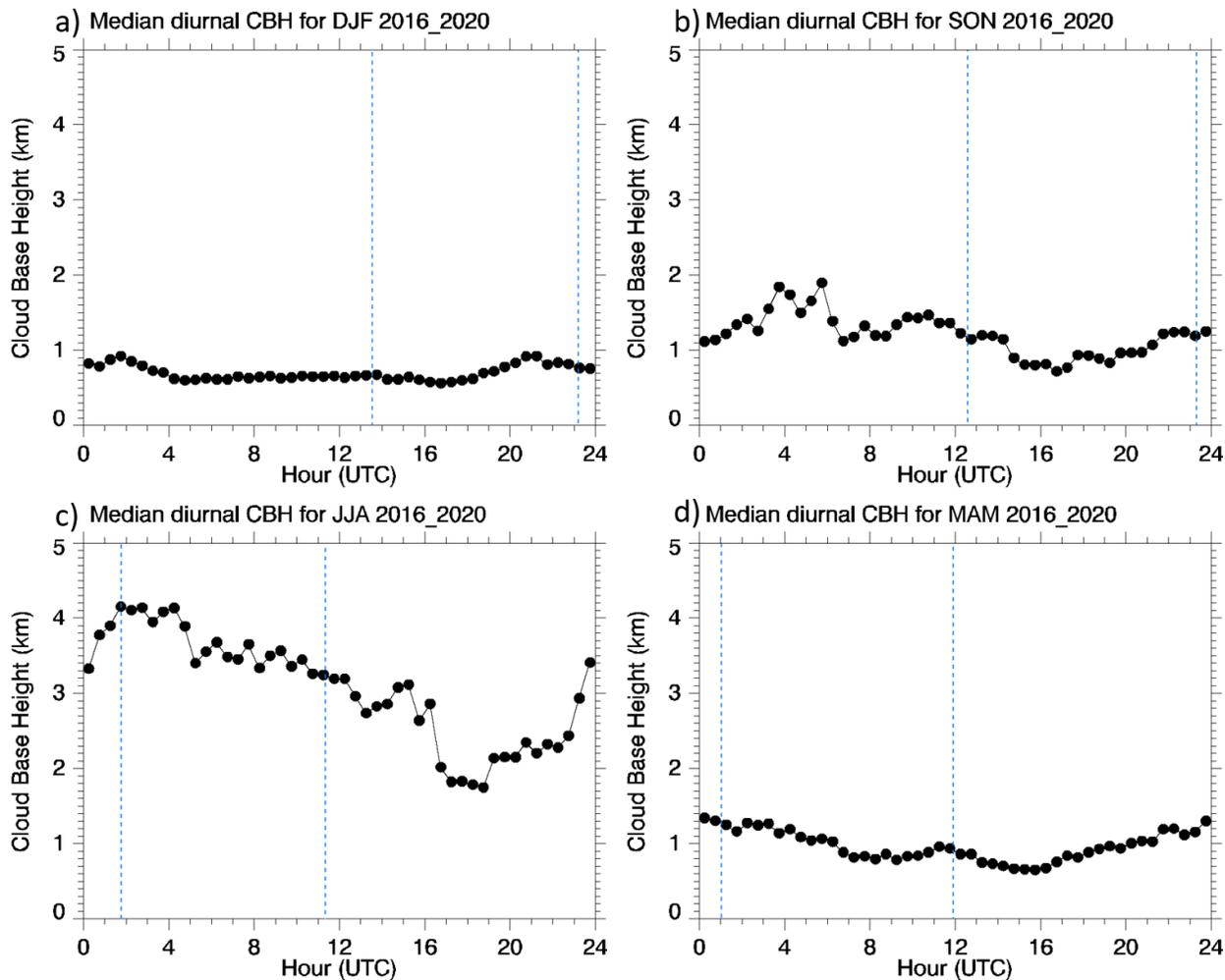


Figure 46. Median diurnal CBH from RL for four seasons (a) DJF, (b) SON, (c) JJA, and (d) MAM from 2016 to 2020 at the SGP central facility. The dotted lines indicate sunset and sunrise for each season.

Figure 47 shows distributions of CBH for each of the four seasons derived from RL. The winter distribution is strongly peaked below 1 km AGL. In contrast, the summer distribution is much broader, with a slight dip around 3 km AGL that suggests a bimodal distribution. In a convective boundary layer at the SGP, cloud bases are near the lifting condensation level. The lowest mode is due in large part to shallow cumulus forming near the top of the convective boundary layer (Sengupta et al., 2004, Zhang and Klein 2010). The spring and fall distributions are similar to one another. Both exhibit a strong fall-off at about 1 km AGL, as does the winter distribution. These observations correlate well with both satellite and Doppler lidar estimates shown earlier.

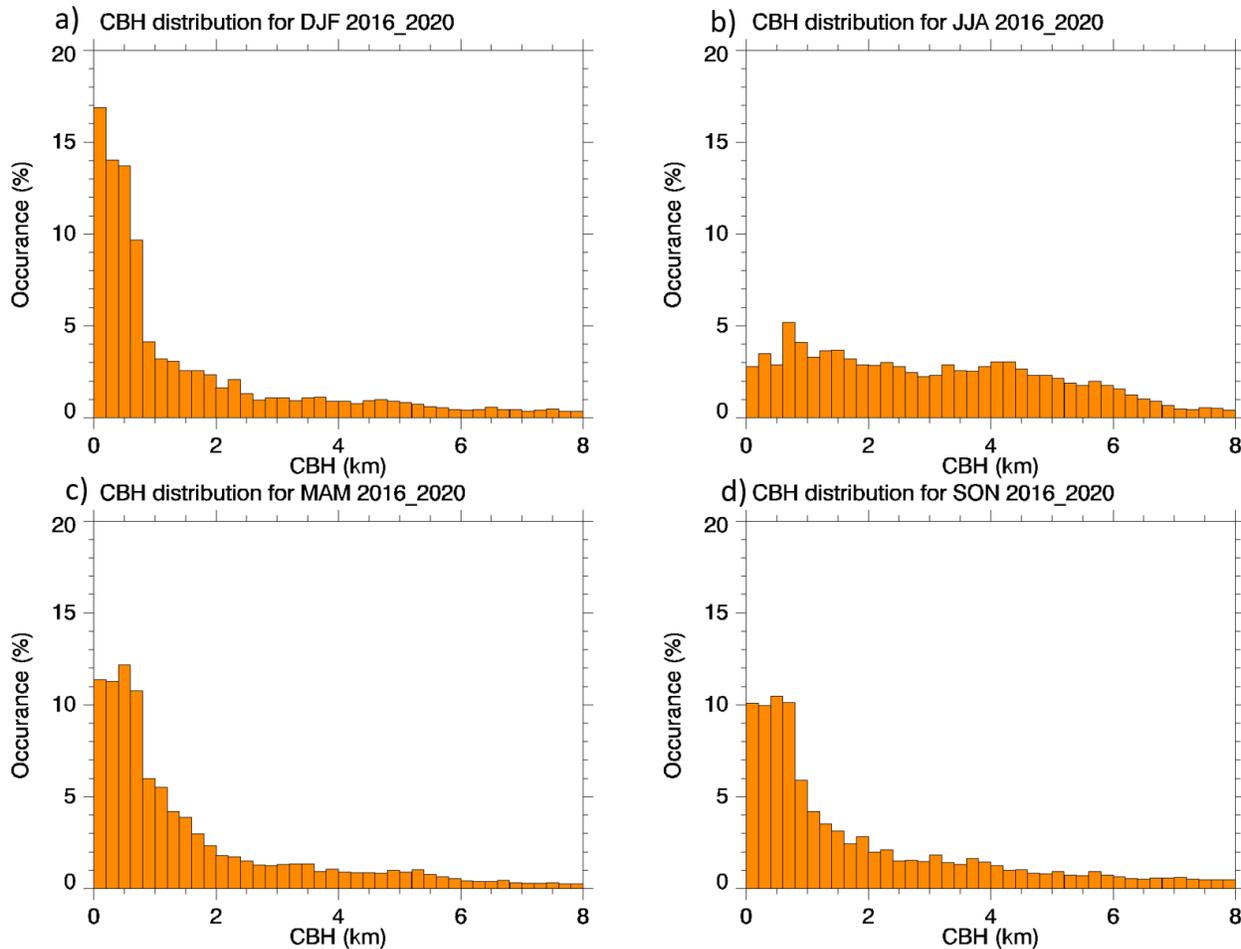


Figure 47. Distributions of cloud base height estimates from the Raman lidar during the period from January 2016 to June 2020 at the SGP central facility.

Boundary layer height estimates are essential for improving the skill of climate, weather, and air quality models. They play a crucial role in our understanding of near-surface turbulent mixing, entrainment, and genesis of shallow cumulus clouds. Although Doppler lidars provide reasonable estimates of connective boundary layer heights, a low bias is generally observed compared to radiosonde estimates. During nocturnal conditions, lidars do not provide reliable boundary layer height estimates. The top of the boundary layer is generally determined when the turbulent fluxes are near zero (Deardorff 1979; Fedorovich and Mironov 1995). For a Doppler lidar, the top of the CBL height coincides with the top of the boundary layer when the vertical velocity variance is near zero. As is observed in Figure 48, the top of the CBL height estimated by the RF model coincides with the minimum vertical velocity variance from the Doppler lidars.

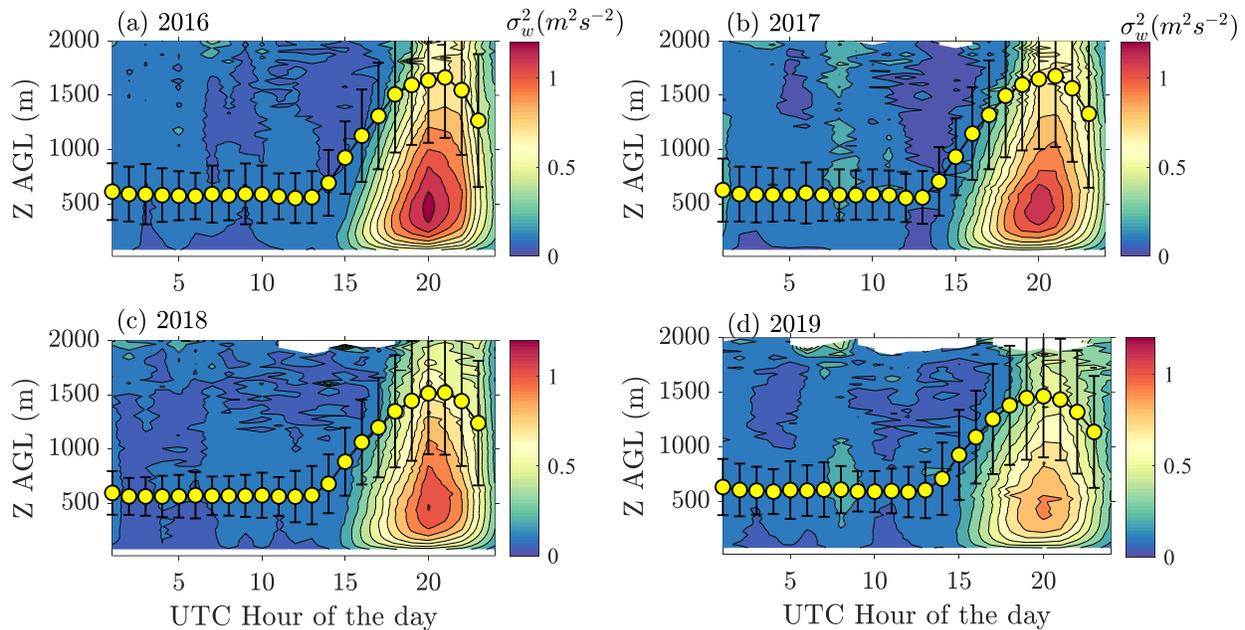


Figure 48. Hourly averaged RF-estimated boundary layer heights from 2016 to 2019, overlaid on vertical velocity variance estimates (shown in Figure 29) at the SGP central facility.

Figure 49 provides a multi-year comparison of boundary layer height at the SGP C1 from 2016 to 2019. Over the years, an approximately 15 percent variability in peak CBL height is observed. Compared to the rest of the years, the lowest boundary layer heights were observed in year 2019, which follows with the lowest turbulence levels observed during 2019. The highest boundary layer heights were observed in year 2017, followed by 2016. High vertical velocity estimates are also observed for years 2016 and 2017 (shown in Figure 29 and Figure 48). The variability in boundary layer height is also linked to surface properties, such as soil temperature, soil, and land-air interaction effects. At the SGP site, lower soil temperatures result in lower boundary layer heights (Krishnamurthy et al. 2021). High correlations are observed between boundary layer height and other surface parameters such as RH, surface air temperature, Obukhov length, and TKE.

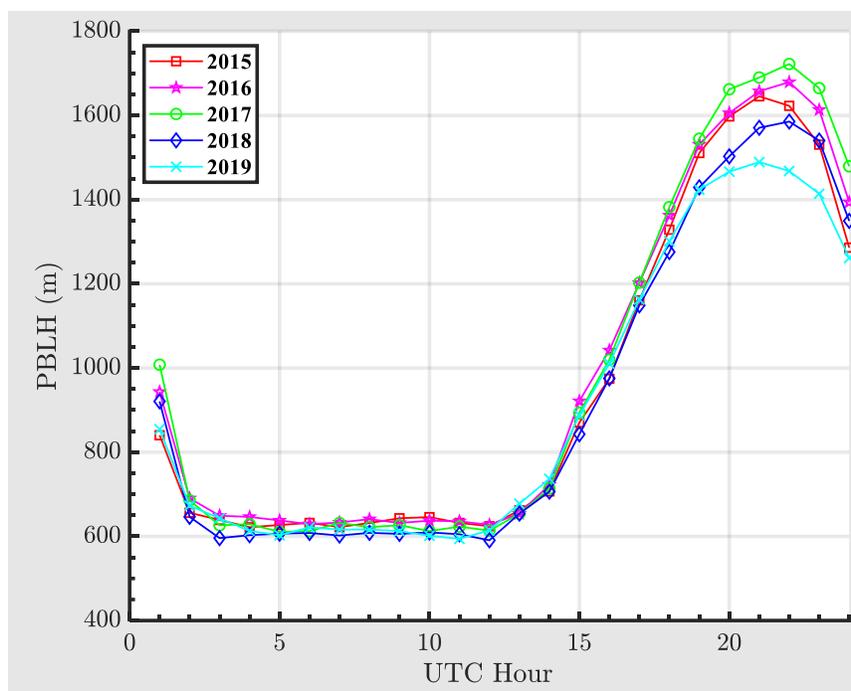


Figure 49. Hourly averaged PBLH from RF algorithm from 2015 to 2019 at the SGP central facility.

On a seasonal basis, the boundary layer heights vary significantly. Figure 50 shows the diurnal variability of boundary layer heights estimated from an RF algorithm for each season (DJF, MAM, JJA, and SON). Peak average CBL heights vary from 1200 m AGL in winter to approximately 2000 m AGL during summer seasons. During summertime, a deep boundary layer is observed, primarily due to high solar irradiance, increase in soil temperature, and related surface fluxes. The transitional boundary layer is seldom picked up by atmospheric models (WRF or LES), so the steep transitions during summertime are generally not modeled accurately (Krishnamurthy et al. 2021). Average nocturnal boundary layer heights at the SGP site are observed to be near constant across multiple seasons. An average nocturnal boundary layer depth of approximately 600 m is normally capped by the presence of the nocturnal LLJ at the SGP site. During winter and spring months, the LLJ is observed to be higher in altitude compared to other seasons (see Figure 24). Figure 51 shows maximum daily boundary layer heights for each season at the SGP site. As expected, the maximum daily boundary layer heights are observed during summertime with a median boundary layer height of ~ 2200 m AGL. During spring and fall, the median daily maximum boundary layer height is ~ 2000 m AGL. The lowest daily maximum PBLHs were observed during winter periods, which featured with a median value of approximately 1500 m AGL. The seasonal cycle of the PBLH provides insight into the turbulence exchange of heat, momentum, and moisture. The cloud type and coverage that affect the radiation budget are also determined by the depth of the PBL (Wood 2012). This can be clearly observed when looking at seasonal averaged cloud top altitude from satellite measurements (seen in Figure 45b) and maximum daily boundary layer heights for the same seasons observed in Figure 51.

As mentioned earlier, the RF algorithm is built using a multitude of surface and lidar-derived meteorological parameters. The marginal effect of each input parameter on the estimated PBLH can be analyzed using the RF framework (Friedman et al. 2001). The partial dependence of PBLH on RH and soil temperature is shown in Figure 52. Large values of partial dependence

indicate strong dependence of PBLH on the input parameter, whereas small ranges show weaker dependence. The higher dependence of PBLH on extreme soil temperatures and low RH correlates well with previous observations during summer and winter time conditions. Further analysis of PBLH partial dependence on various surface parameters is provided by Krishnamurthy et al. (2020).

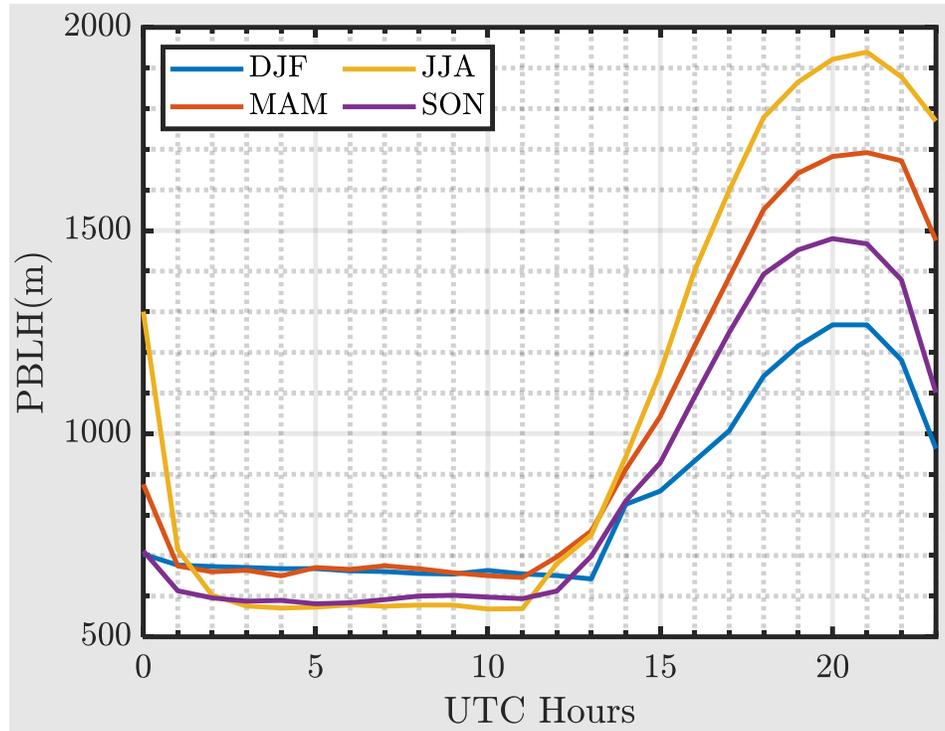


Figure 50. Hourly averaged boundary layer height over four seasons at the SGP central facility from 2015 to 2019 at the central facility.

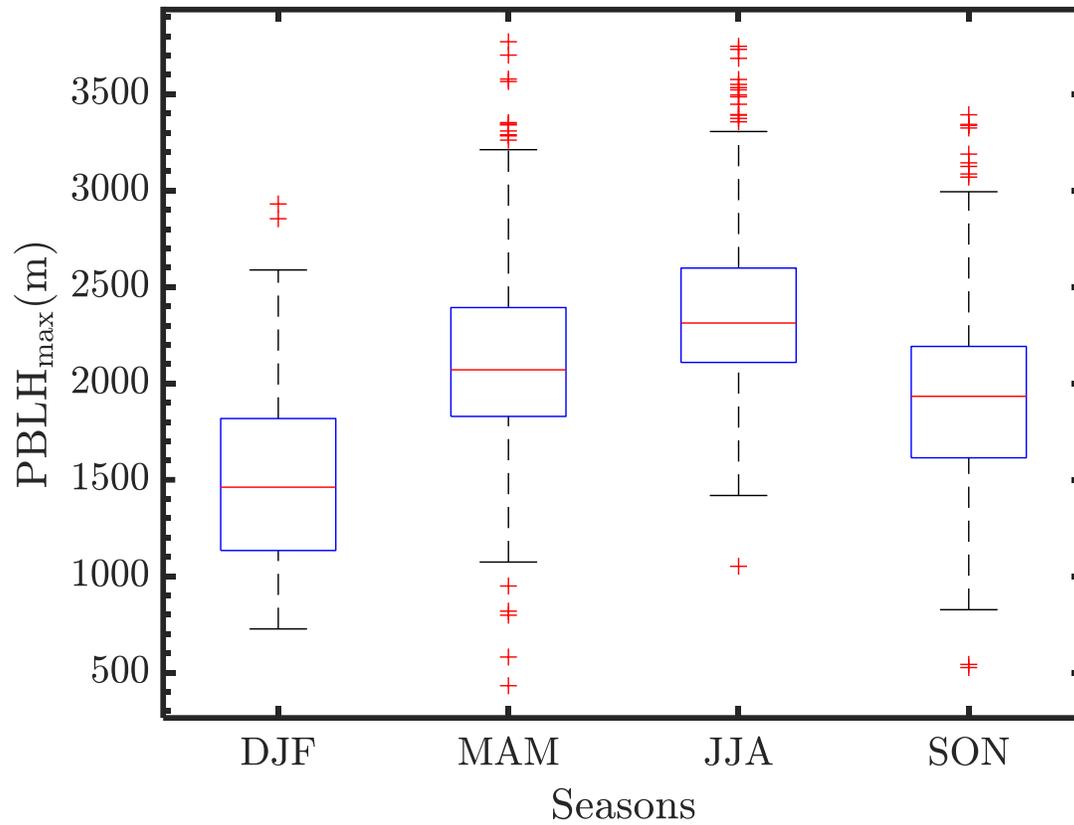


Figure 51. Daily maximum boundary layer height averaged from 2015 to 2019 for each season at the SGP central facility. The red horizontal line indicates the median value, the blue box represents the 25th and 75th percentile levels, and the + signs indicate possible outliers not used to calculate the median value.

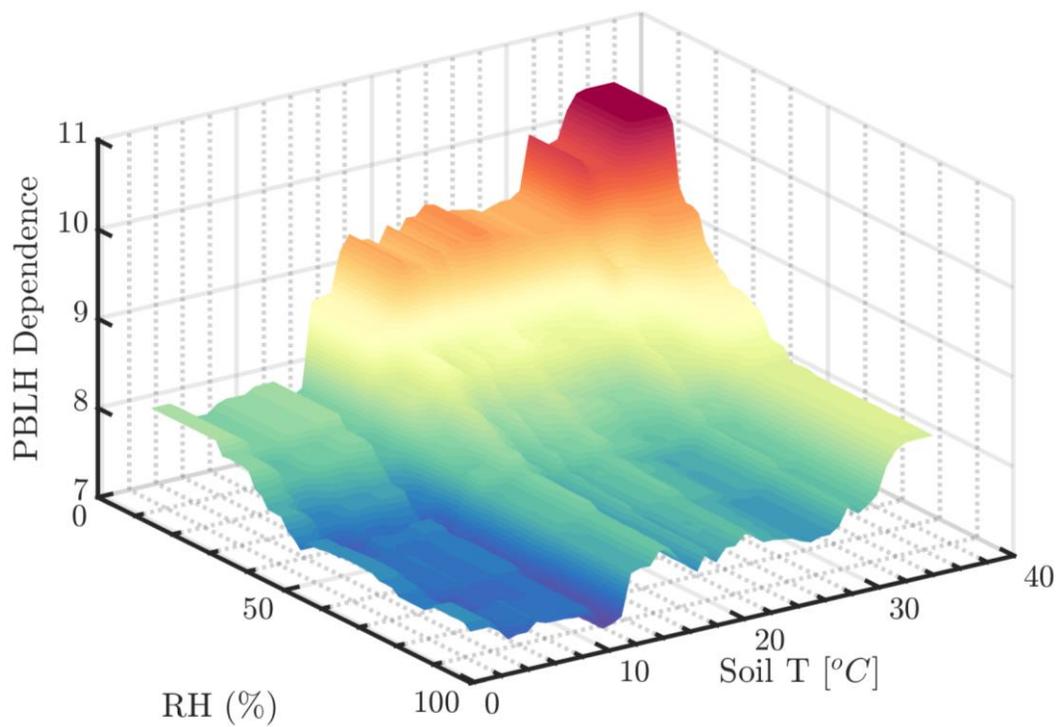


Figure 52. Partial dependence between PBL height, RH, and soil temperature using RF regression.

6.0 Site-specific Atmospheric Conditions

6.1 Nocturnal Low-level Jet

Nocturnal LLJs are commonly observed in the Southern Great Plains. Nocturnal LLJs are mostly generated by inertial oscillations, which are initiated by a drop in eddy viscosity at sunset due to development of a stable boundary layer and radiative cooling of the ground surface (e.g., Parish et al. 1988; Zhong et al. 1996; Parish and Oolman 2010). A Great Plains nocturnal LLJ climatology by Bonner (1968) indicated that the average altitude of LLJ wind maxima was approximately 800 m AGL, using approximately a 2-year radiosonde data set around the Great Plains. Using better time- and height-resolved data, Whiteman et al. (1997) estimated that half of the SGP site's nocturnal LLJ wind maxima are located below 500 m. Berg et al. (2015) also provided an assessment of the SGP LLJs using radiosonde data and provided model results showing that LLJs are key for effective moisture intake into the Great Plains from Gulf of Mexico leading to deep convection. Bonin et al. (2015) classified the stability of the nocturnal LLJ using vertical velocity variance estimates within the LLJ. Several modeling studies were conducted to analyze the impact of the LLJ on the Great Plains (some recent studies: Berg et al. 2015, Smith et al. 2018). In studies thus far, sophisticated models (such as the WRF model) have consistently under-predicted the magnitude and height of the LLJ near the Great Plains (Berg et al. 2015; Wharton et al. 2013; Smith et al. 2018). There are several reasons for this under-prediction; for example, improper land-surface coupling in the models could result in significant differences. Zhong et al. (1996) showed that soil moisture changes in the Great Plains affected jet amplitudes; drier soils lead to stronger nocturnal LLJs. Drier soils also lead to higher boundary layer height estimates at the SGP site (Krishnamurthy et al. 2021). Therefore, a thorough study of the evolution, sustenance, and dissipation of nocturnal LLJs at the SGP site is warranted.

To determine the height and strength of the LLJ, several criteria are available in the literature (Blackadar 1957; Bonner 1968; Whiteman et al. 1997). In this work, we determine the height of the LLJ when the wind speed maximum reduces by at least 2 ms^{-1} on either side of the maximum (Blackadar 1957). Doppler lidar wind profile data from May 2011 to May 2020 (approximately 9 years' worth) were used in this analysis. Because lidar wind profiles are available every hour, a daily median LLJ height per day is chosen and is limited to nighttime conditions (i.e., 5 to 13 hours UTC or ~00:00 to ~08:00 hours local time). Any observation below 200 m, above 3 km, and at jet wind speeds greater than 40 ms^{-1} was omitted, because of possible errors in lidar processing. Ignoring measurements below 200 m could potentially bias the LLJ height estimates high during certain seasons (JJA), but because no LLJ observations have been documented below 200 m at the SGP site, any potential bias is expected to be low. Omitting periods when the lidar data were unavailable or filtered, a total of 2550 observations of LLJ were identified from May 2011 to May 2020. Figure 53 shows the daily median LLJ height distribution for four seasons (DJF, MAM, JJA, SON) from May 2011 to May 2020. In summer, the median height of LLJ is observed to be predominantly below 600 m AGL and very few LLJs above 1 km are observed. In winter, the lowest percentage of LLJs were observed at the SGP site; the majority of daily median LLJ heights were below 500 m AGL and up to a maximum height of 2800 m. In spring and summer, the highest percentages of LLJ were observed at the SGP site, both totaling to approximately 60 percent of the LLJ observations. Spring, summer and fall show similar distributions, except for larger distribution of daily median LLJ heights observed in spring. In winter, fewer LLJs was observed, predominantly below 500 m AGL. Figure 54 shows the daily maximum LLJ wind speed for all four seasons from May 2011 to May 2020. Winter and spring observe higher maximum LLJ wind speed distributions than other

seasons. Several instances of LLJ wind speed maximums exceeding 30 ms^{-1} are observed during both spring and winter seasons. In summer, lower maximum wind speeds are observed, which is consistent with seasonal average wind speeds observed by Doppler lidar (Figure 24).

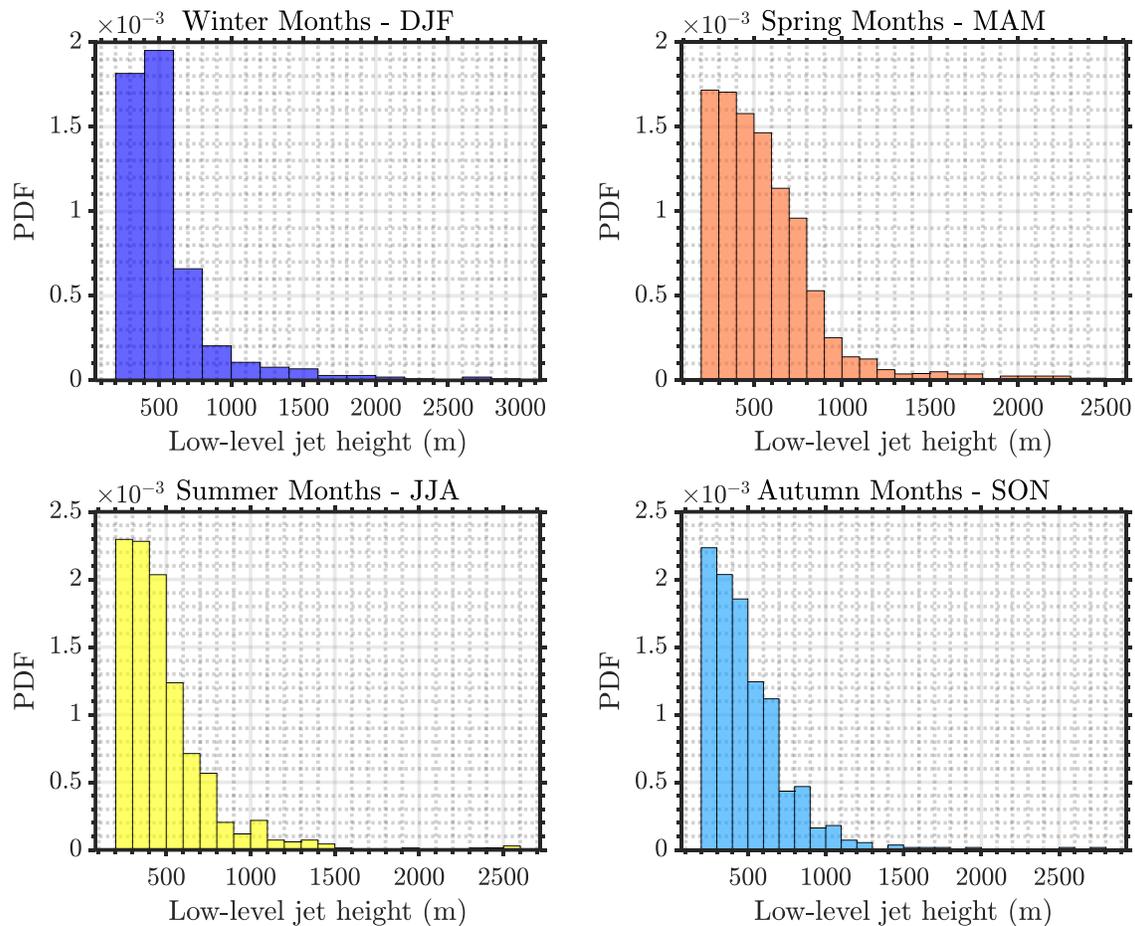


Figure 53. Daily median low-level jet height distributions at the SGP central facility from May 2011 to May 2020 for (a) DJF, (b) MAM, (c) JJA, and (d) SON.

Because the nocturnal LLJ at the SGP site is generally below 600 m AGL, it significantly influences the winds near the surface. The atmosphere is generally weakly stable during an LLJ, which is expected to cause the most damaging fatigue loads on a wind turbine. Gutierrez et al. (2016) evaluated the effect of LLJ on wind turbine power and loads and observed that when the wind turbine experiences negative wind shear (i.e., the jet height is at the bottom edge of the wind turbine) the effect of LLJ is minimal. When the wind shear is positive (i.e., the jet height is above the tip of the turbine), the impact on the turbines (with respect of loads) is expected to be higher. From simulation studies, it is observed that when the LLJ is located at hub height, the root mean square rotor aerodynamic loads increase by two times, when the winds increase from 8 to 16 m/s (Zhang et al. 2019). Even though the wind turbines are pitched at higher winds, the wind shear across the rotor-swept area is expected to cause significant loading. Most damaging fatigue loads occur within the weakly stable range, which results in intense vertical mixing, with the gradient Richardson number between + 0.01 and +0.05.

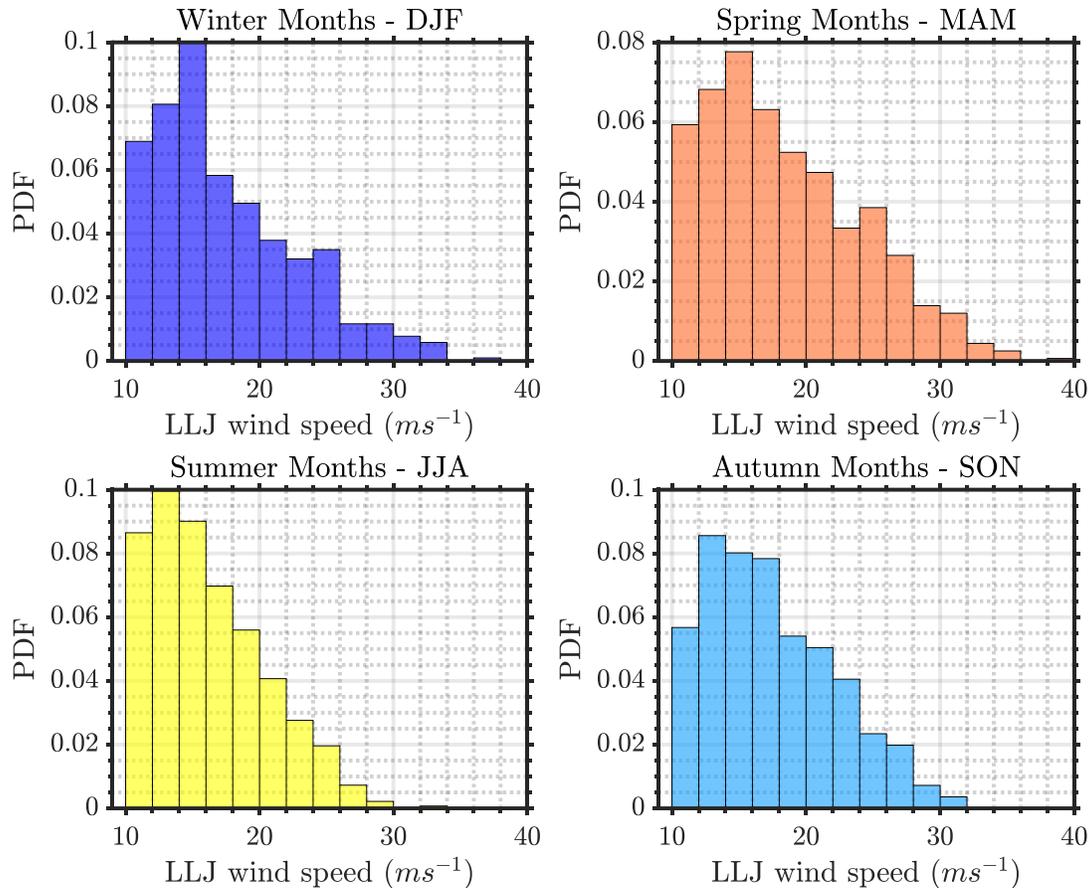


Figure 54. Daily maximum low-level jet wind speed distributions at the SGP central facility from May 2011 to May 2020 for (a) DJF, (b) MAM, (c) JJA, and (d) SON.

6.2 Spatial Variability

Analogous to cloud properties, surface green vegetation fractions in the form of fraction of photosynthetically active radiation at 500 m horizontal grid resolution are key to providing guidance for crop management and modeling the exchange of momentum, energy, and mass between the PBL and land surface. Vegetation cover (VC) defines the percentage of soil that is covered by green vegetation. Figure 55 shows the temporal variability in VC at the 1° domain for 2018 (the SGP site is at the center of the domain). At the SGP site, it can be observed that the VC in winter is a factor of 2 lower than it is in summer. The primary reason for this variability is the different crops grown in each season around the SGP site. Winter wheat is grown in small sections around the SGP site, but primarily summer crops are more frequently observed. Increase in pastureland during summertime also results in higher VC around the SGP site. A steep gradient in VC is observed in spring and fall over the SGP site. This gradient in VC also results in a gradient in surface roughness length (see Figure 18b). Therefore, it is indeed interesting to observe a near-linear relationship between VC from satellites and surface roughness length measurements at the SGP site (Yu et al. 2018). This change in roughness causes changes in wind profiles within the surface layer and is key to understanding the surface energy budget at the SGP site. The changes in VC also are sensitive to surface albedo (Berg et al. 2020). A spatial analysis of the VC using satellite data would provide further details about the land–atmosphere feedback processes. Figure 56 shows the spatial variability in VC over the SGP site at a resolution of 1 degree for a given day in January (winter) and June (summer) of

2018. The VC maps on both winter and summer days show clear east-west gradients. The VC gradients are a result of variations in agriculture practices around the SGP site (Berg et al. 2020).

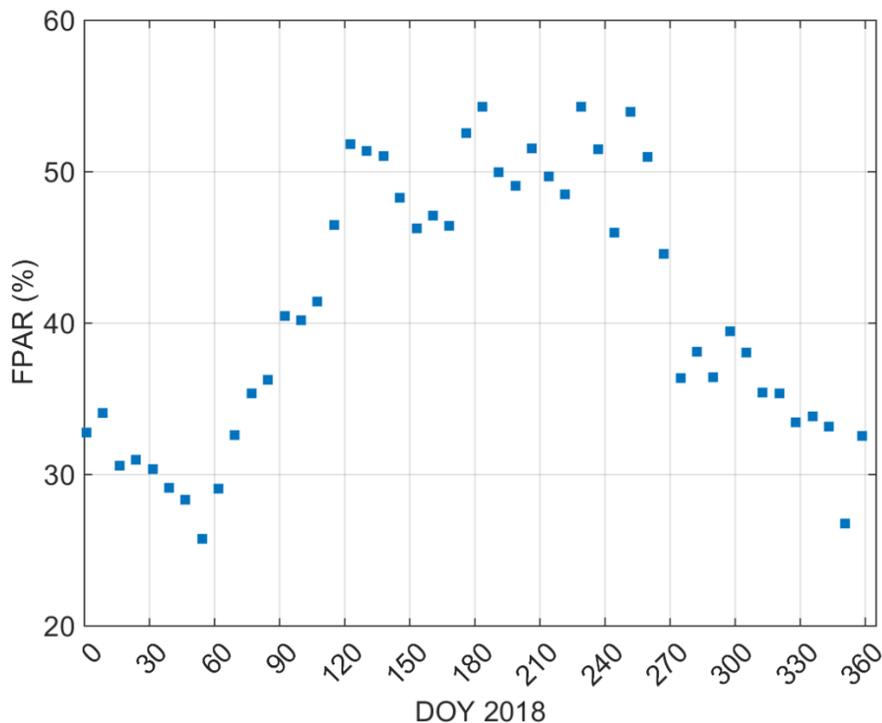


Figure 55. Temporal variation in vegetation cover (%) from Fraction of Photosynthetically Active Radiation using MODIS satellite observations for the year 2018. The observations at an interval of ~8 days are compiled over 1x1 deg domain with the SGP C1 site at the center.

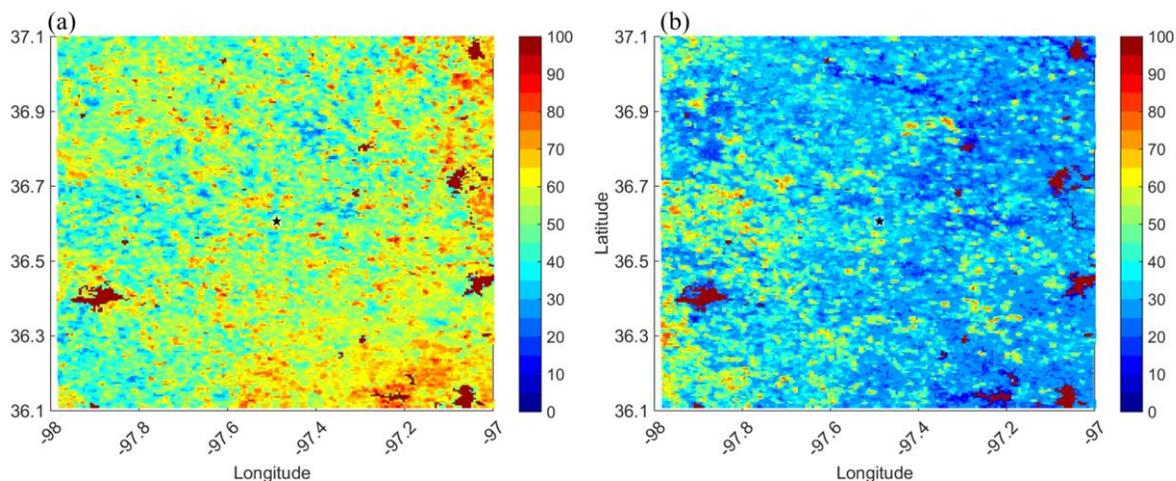


Figure 56. Spatial distribution of vegetation cover (%) from Fraction of Photosynthetically Active Radiation using MODIS satellite observation for (a) a day in the first week of January and (b) a day in last week of June 2018 (right panel). The VC>100 filled values are from water, rocks, or other non-vegetated reflected pixels. The SGP site is shown in the middle of map by a black star enclosed in a square.

As observed in Figure 56, west of the SGP site, a larger percentage of the area is observed to be covered in vegetation compared to the area east of the SGP C1. Land-air interaction effects are prominent at the SGP site because of the heterogeneous surface. Soil moisture, soil temperature and air temperature variability across the SGP site have been shown to affect the height of the LLJ (Zhong et al. 1996), the boundary layer height (Krishnamurthy et al. 2021), and other boundary layer processes (Berg et al. 2020). Figure 57 shows monthly averaged soil moisture, soil temperature, and soil-air temperature differences at the four satellite sites shown in Figure 1 from 2014 to 2019. These are roughly 40 to 50 km from the SGP C1 and are located in corners of the VC satellite image shown in Figure 56. Significant variability is observed among the sites, especially in air-land temperature difference. At Site E39 (southeast of the SGP C1), the highest monthly averaged soil temperature and highest air-soil temperature gradient are observed. The effect of soil temperature is also known to affect boundary layer height estimates. Figure 58 shows the hourly averaged RF-estimated PBLH for 2019 at three satellite sites (E32, E37, and E39) and the central facility. It should be noted that the RF model built at SGP C1 was used at these extended facilities. Typically, RF models are site specific, therefore the accuracy of the PBLH estimates at other satellite sites needs to be evaluated using a reference (such as radiosondes). Local surface and lidar measurements from the given site were used as input features in the model estimation. The central facility shows the lowest convective PBLH compared to the rest of the satellite sites, and the E39 shows the highest PBLH. This corresponds to previous findings that the increase in soil temperature results in higher PBLH. Although multiple drivers affect the growth and development of PBLH, local surface parameters such as RH, soil temperature, and TKE seem to provide the bulk of the effect on convective PBLH (Krishnamurthy et al. 2021). The terrain height difference between all the three sites is within 50 m. Nocturnal conditions on the other hand do not seem to vary considerably across multiple sites and this lack of variation is probably due to the presence of the LLJ at the SGP site. Site E41 was not used in this analysis, because lidar data were not available for more than 6 months due to a system failure.

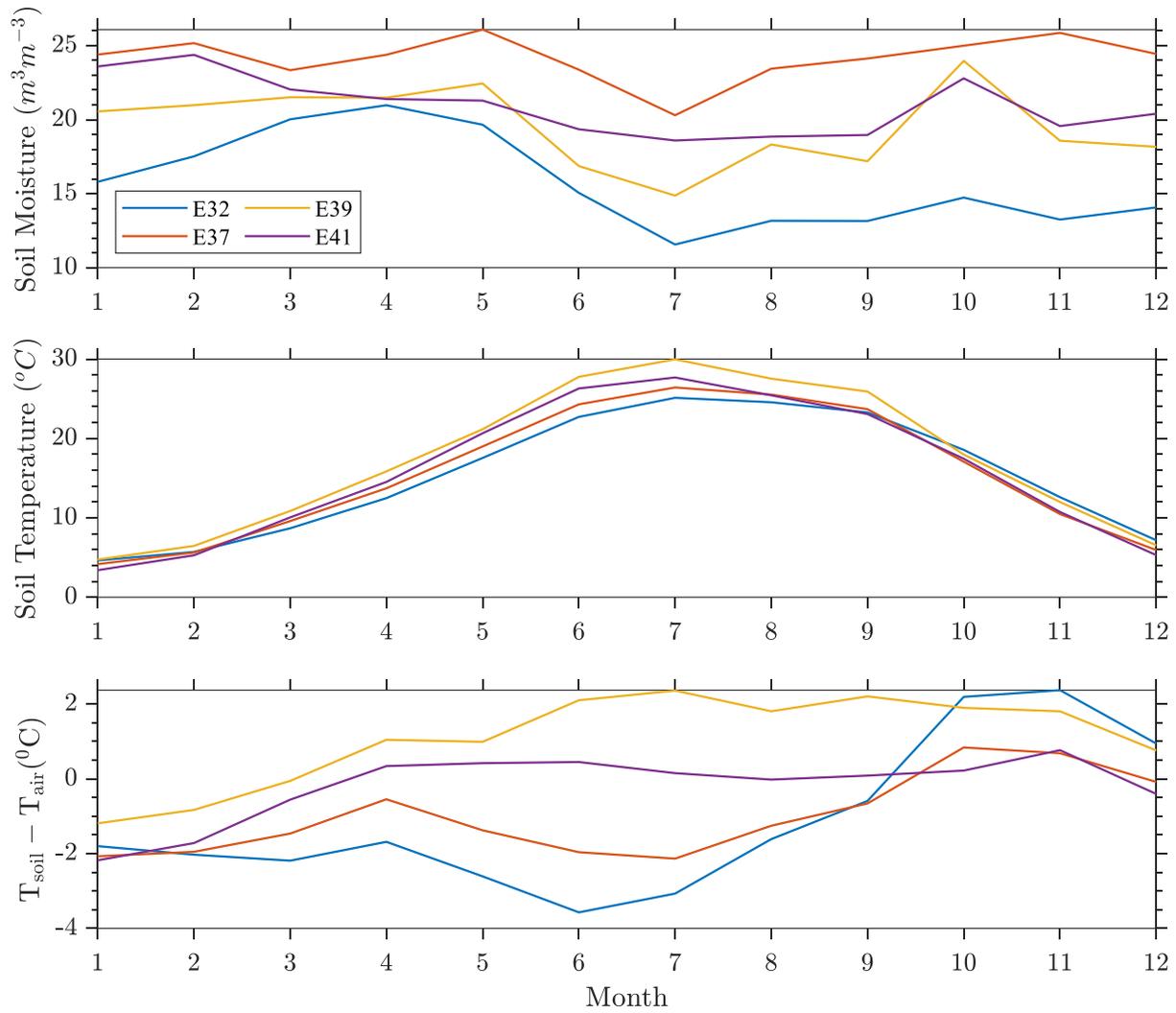


Figure 57. Spatial variability at four satellite ARM sites (E32, E37, E39, and E41) showing monthly averaged (a) soil moisture, (b) soil temperature, and (c) soil-air temperature difference from 2014 to 2019.

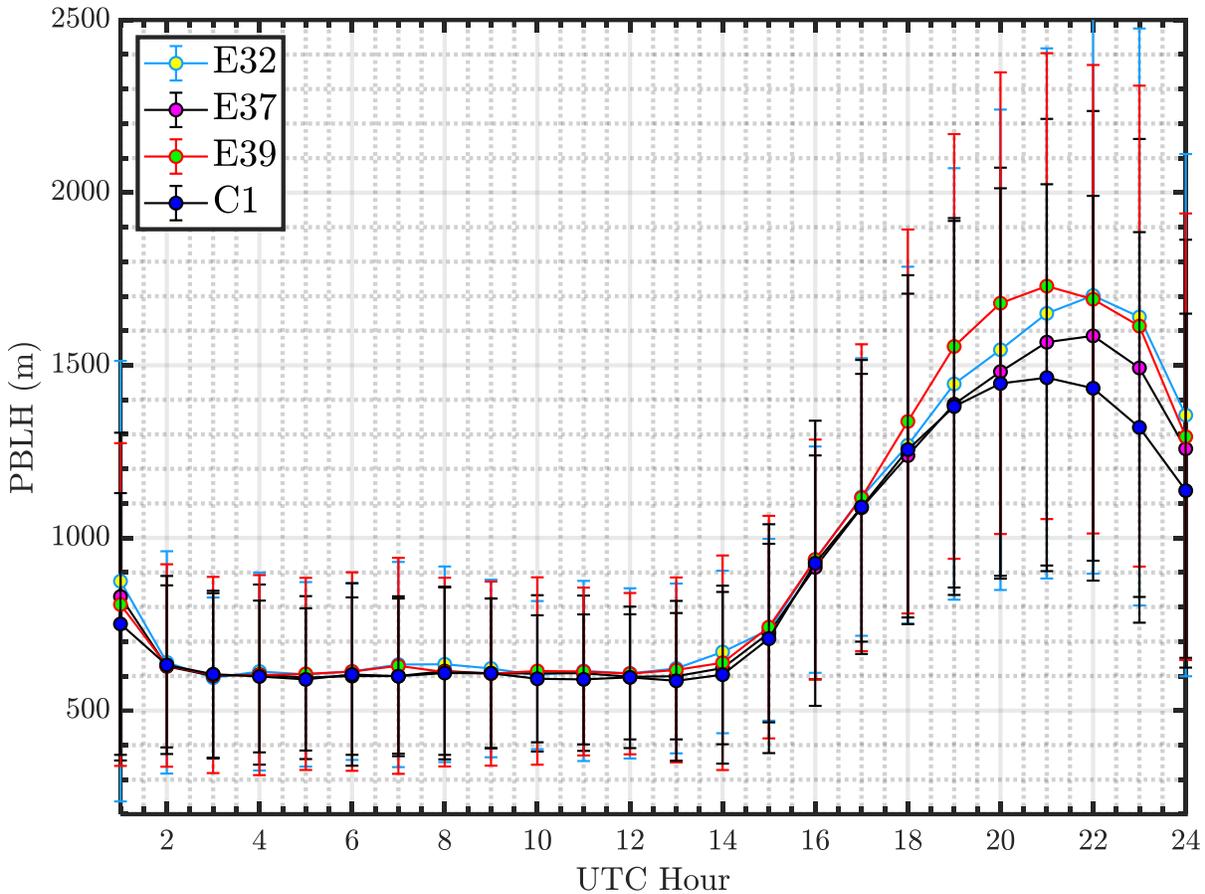


Figure 58. Hourly averaged random forest estimated PBLH in 2019 at E32, E37, E39, and C1.

6.3 Possible Effect of Wind Farms at the SGP Site

The growing presence of wind farms around the SGP site, mean the central facility site can no longer be considered devoid of wind farm wake effects. However, no concrete evidence of these effects is observed based on the data available in the literature, mainly due to the distance of wind farms from the central facility. A more in-depth analysis of wind data related to pre- and post-wind farm construction is needed to isolate effects due to climate change, IAV, and wind farms, etc. Some preliminary thoughts about how the IAV of wind speed at the SGP site has varied over the years, the effects of the internal boundary layer developed by a wind farm, and how wind farms can affect the local microclimate are presented below.

Based on the tower measurements from 2003 to 2014 at 25 m AGL the IAV of wind speed ($IAV = 100(\sigma_U/U)$ where U is the annual mean wind speed and σ_U is the standard deviation of annual wind speed) at the SGP C1 site was 3.19 percent. Regions with IAV less than 6 percent are preferred for wind farm development, because of the reduced interannual wind uncertainty (Brower 2012), i.e., the winds are reasonably predictable from year to year. Estimates at hub height are generally used for IAV calculations. Therefore, further analysis was performed using the available data from 2012 to 2019 measured by a Doppler lidar at 90 m AGL. Based on wind farm development at the SGP site, shown in Table 1, time periods prior to 2016 were considered “pre-wind farm” and post 2016 was considered “post-wind farm”. Pre-wind farm IAV at the SGP site was 1.85 percent, while post-wind farm IAV was 4.48 percent. This increase in

IAV could be attributed to the presence of wind farms built south of the SGP site but could also be due to the small sample size used in estimating the IAV. Generally, IAV is calculated based on a larger sample, say 10 to 20 years of data. Therefore, a longer data set at hub height is needed to evaluate the true impact of wind farms on the Southern Great Plain's climatology. In addition, an analysis at SGP C1 using the Doppler lidar wind speed profiles showed that under stable atmospheric conditions, hub-height wind speeds greater than 6 ms^{-1} to 10 ms^{-1} (range based on the operating power curve of the GE turbines) and select wind directions (primarily from the three turbines south of SGP C1 within 4 km), the maximum reduction in wind speed at hub-height and above was less than 2%. More analysis is indeed needed and would benefit the research community in understanding long-term impacts of wind farms at SGP C1 data.

Although wind speed deficits are theorized to recover over 20–25 km downwind of the wind farms, the temperature and RH of the waked atmosphere for a large wind farm are expected to be prolonged further when the inversion height is below the hub height and the atmosphere is stably stratified (Siedersleben et al. 2018). The cumulative wake of the entire wind farm can reach up to 50 km downstream (Lundquist et al. 2018). Immersed boundary layer LES models of large wind farms show that the growth internal boundary layer by a wind farm affects the formation of stratocumulus clouds downwind of the wind farm (Sharma et al. 2018). Increased vertical entrainment during periods of stratocumulus-topped boundary layers tends to destabilize the cloud layer, resulting in the destruction of the cloud layer. Based on a quick study of cloud base heights and winds from Doppler lidars (at C1 and E39), some evidence alluded to stratocumulus cloud variability at these sites. Bands of stratocumulus clouds do exist at the SGP site, but no direct evidence of either destruction of stratocumulus clouds or wind farm wake deficits were observed at the SGP site using the Doppler lidar CBH data. This could be due to the size and orientation of the wind farms south of the SGP site (single row of wind turbines and not a deep array), but further investigation is needed. Other satellite sites, such as E32 and E41 can expect higher influences due to the proximity, orientation, and size of surrounding wind farms.

The phenomenon of wind turbine microclimate occurs as a result of changes—in local temperature, moisture, and CO_2 levels—due to vertical mixing, turbulence, and wakes created by the wind turbines. Because an increasing number of wind turbines are being built on farmlands surrounding the SGP site, it is imperative to assess the long-term effects of wind farms on crops. Given the correlation between climate change and crop yield, as recent studies have found (Schlenker et al. 2005; Schlenker et al. 2009), it is important to understand the implications of wind turbine microclimate on local farming communities. The potential crop yield variability, due to the microclimate, could have a vast impact on the micro and macroeconomics of local counties and the country throughout the lifetime of a wind farm. Therefore, interdisciplinary research, delving into both the physical and social aspects of crop yield variability, is necessary. This would aid in developing optimal strategies for better application of our knowledge for the benefit of wind and agricultural communities, locally and internationally. Ultimately, the goal of energy independence requires a balanced ecosystem, *sine qua non* of science serving society.

Wind turbine land lease agreements can pay landowners up to \$6,000 per year per megawatt installed. The capacities of modern generation wind turbines range from 2 MW to 5 MW, so a farmer can earn up to \$30,000 per turbine per year. Thus, wind farming has quickly become quite valuable, especially for larger farms that can host several hundred wind turbines. As of today, U.S. wind farms are paying a total of more than \$250 million dollars to property owners across the country, and this number is poised to grow over the coming years (AWEA 2017 annual report). Most wind farm projects require at least 60 acres of land per megawatt

produced, but only a small percentage (about less than 5 percent) of this land will be used for the placement of wind turbines and other supporting infrastructure (such as a substation, maintenance facilities, and storage offices, etc.). The total acreage depends on the wind turbine spacing to account for wind turbine wakes, which can span up to 15 times the rotor diameter for a single turbine. The cumulative wake of the entire wind farm can reach up to 50 km downstream (Lundquist et al. 2018). This creation of atmospheric instability can have a positive or a negative impact on crop production.

Results show that wind farms significantly affect near-surface air temperature and humidity as well as surface sensible and latent heat fluxes (Roy et al. 2004; Rajewski et al. 2013). Given that climate affects crop efficiency (Schlenker et al. 2005; Deschênes et al. 2007; Schlenker et al. 2009; Deschênes et al. 2012; Fisher et al. 2012) and wind farms create microclimate around the crops, evidently wind farms in turn may affect crop yield. Recent field studies looked at local small-scale impacts (Rajewski et al. 2013) of increased vertical mixing and CO₂ concentration levels on crops, but no conclusions were drawn about the crop yield. A recent study conducted by the National Aeronautics and Space Administration (Deryng et al. 2016), shows that increased CO₂ concentration levels can both help and hurt crops. The increased concentration levels boost crop yield by increasing the rate of photosynthesis, but they reduce the amount of water crops lose through transpiration (reduced water efficiency). Therefore, microclimates created by wind farms could possibly improve the vertical mixing close to the surface and, in turn, improve crop yield. This could have a different impact depending on the type of crop downwind of the wind turbine/farm and geographic location. For example, wine makers in New York and California currently use large fans to reduce the impact of frost on grape yield by entraining the warm air from above. But the inherent uncertainty in wind production from year to year could create variability in crop yield and could have a compounding effect over the lifetime of the wind farm (~25 years). This could affect the local revenue earned by farmers, trade, food supply and prices, state or federal policies on tax credits, etc. There is also a risk of farmers reducing their yield because of revenue gained from leasing their sites to wind farms, resulting in serious implications to climate change (albedo effects). Therefore, in addition to understanding the variability in annual yield due to microclimates, it is very important to address the interplay between science and policy due to crop yield variability in wind turbines adjacent to farms. A more detailed study of the possible positive/negative effects of wind farms on crops at the SGP site is necessary.

7.0 Perspectives and Future Work

The SGP site provides long-term, high-quality meteorological data for over two decades (since 1992) in southern Oklahoma, USA. Atmospheric scientists have used the data effectively to unravel several mysteries related to complex land-air interaction effects, mesoscale processes, convective systems, and aerosol and cloud dynamics. However, the data are far from being exploited in various other research domains. Wind energy studies have seldom compared microscale model results and field observations. The diurnal, seasonal, and annual variability can aptly be evaluated using the long-term data. Microscale models are frequently used in wind energy studies to assess the wind resource and annual energy production at a given site. In the same vein, very few studies have been conducted to evaluate reanalysis products such as Modern-Era Retrospective analysis for Research and Applications (MERRA) and the North American Regional Reanalysis (NARR) (Kennedy et al. 2011). These reanalysis data are generally used as input conditions for many mesoscale and microscale models and would affect the outcome of the numerical simulations in the region being studied. The array of Doppler radar and lidar data at the SGP site can characterize both the boundary layer and mesoscale processes, such as MCSs and nocturnal convection initiation. The transfer of energy from such large mesoscale systems to the surface can be quantified using the SGP in situ and remote-sensing data.

Since early 2015, an LES study (LES ARM Symbiotic Simulation and Observation workflow, LASSO) is being performed for select case studies around the SGP site. Currently, the focus is only on using LES to improve our understanding of shallow clouds. A more thorough analysis of using LES using ARM observations for wind energy applications is warranted. This would provide context and self-consistent representation of the atmosphere surrounding the SGP observatory. Because the SGP data are being used for validating other models, the effect of nearby wind farms on the SGP data has still not been explored. As a part of future work, the authors request the modeling teams to possibly study the effect of turbines on the SGP data. Especially on remote-sensing instruments, which provide spatially averaged measurements, wind turbine wake signatures are averaged out and sometimes not clearly detectable in the processed signal at SGP C1. High-fidelity LESs can provide an uncertainty map for remote-sensing instruments based on downwind distances from the wind farm at the SGP site. Particularly, how far does the internal boundary layer developed by the wind farm affect remote-sensing measurements? In addition, analysis of using LES data to characterize vertical mixing within a wind farm and the transport of soil moisture/RH downwind of a wind farm can be especially interesting to understand crop yield variability around wind farms. Targeted field campaigns, such as AWAKEN, can help provide answers to some of the above questions at the SGP site.

8.0 References

Argaín, J. L., Miranda, P. M., & Teixeira, M. A. (2009). Estimation of the friction velocity in stably stratified boundary-layer flows over hills. *Boundary-layer meteorology*, 130(1), 15-28.

AWEA. 2017. U.S. Wind Industry Annual Market Report: Year Ending 2016. Washington, D.C. <http://www.awea.org/market-reports>.

AWEA. 2019. U.S. Wind Industry Annual Market Report: Year Ending 2018. Washington, D.C. <http://www.awea.org/market-reports>.

Bagley, J. E., Kueppers, L. M., Billesbach, D. P., Williams, I. N., Biraud, S. C., & Torn, M. S. (2017). The influence of land cover on surface energy partitioning and evaporative fraction regimes in the US Southern Great Plains. *Journal of Geophysical Research: Atmospheres*, 122(11), 5793-5807.

Baidya Roy, S., Pacala, S. W., & Walko, R. L. (2004). Can large wind farms affect local meteorology?. *Journal of Geophysical Research: Atmospheres*, 109(D19).

Bartholomew, M. J. (2016). *Impact Disdrometers Instrument Handbook* (No. DOE/SC-ARM-TR-111). DOE Office of Science Atmospheric Radiation Measurement (ARM) Program (United States).

Beljaars, A. C. M., & Holtslag, A. A. M. (1991). Flux parameterization over land surfaces for atmospheric models. *Journal of Applied Meteorology*, 30(3), 327-341.

Berg, L. K., Riihimaki, L. D., Qian, Y., Yan, H., & Huang, M. (2015). The low-level jet over the southern Great Plains determined from observations and reanalyses and its impact on moisture transport. *Journal of Climate*, 28(17), 6682-6706.

Berg, L. K., & Lamb, P. J. (2016). Surface properties and interactions: Coupling the land and atmosphere within the ARM program. *Meteorological Monographs*, 57, 23-1.

Berg, L. K., Newsom, R. K., & Turner, D. D. (2017). Year-long vertical velocity statistics derived from Doppler lidar data for the continental convective boundary layer. *Journal of Applied Meteorology and Climatology*, 56(9), 2441-2454.

Berg, L. K., Long, C. N., Kassianov, E. I., Chand, D., Tai, S. L., Yang, Z., ... & Mendoza, A. (2020). Fine-Scale Variability of Observed and Simulated Surface Albedo Over the Southern Great Plains. *Journal of Geophysical Research: Atmospheres*, 125(7), e2019JD030559.

Blackadar, A. K. (1957). Boundary layer wind maxima and their significance for the growth of nocturnal inversions. *Bulletin of the American Meteorological Society*, 38(5), 283-290.

Bodini, N., Lundquist, J. K., Krishnamurthy, R., Pekour, M., Berg, L. K., & Choukulkar, A. (2019). Spatial and temporal variability of turbulence dissipation rate in complex terrain. *Atmospheric Chemistry and Physics*, 19, 4367-4382.

Bonin, T. A., Blumberg, W. G., Klein, P. M., & Chilson, P. B. (2015). Thermodynamic and turbulence characteristics of the southern great plains nocturnal boundary layer under differing turbulent regimes. *Boundary-Layer Meteorology*, 157(3), 401-420.

- Bonner, W. D. (1968). Climatology of the low level jet. *Mon. Wea. Rev*, 96(12), 833-850.
- Breiman, L. (2001). Random forests. *Machine learning*, 45(1), 5-32.
- Brower, M. (2012). *Wind resource assessment: a practical guide to developing a wind project*. John Wiley & Sons.
- Businger, J. A., Wyngaard, J. C., Izumi, Y., & Bradley, E. F. (1971). Flux-profile relationships in the atmospheric surface layer. *Journal of the Atmospheric Sciences*, 28(2), 181-189.
- Champagne, F. H. (1978). The fine-scale structure of the turbulent velocity field. *Journal of Fluid Mechanics*, 86(1), 67-108.
- Cheyne, E., Jakobsen, J. B., & Reuder, J. (2018). Velocity spectra and coherence estimates in the marine atmospheric boundary layer. *Boundary-layer meteorology*, 169(3), 429-460.
- Cook, D. R. (2016). *Soil temperature and moisture profile (STAMP) system handbook* (No. DOE/SC-ARM-TR-186). DOE Office of Science Atmospheric Radiation Measurement (ARM) Program (United States).
- Cook DR and RC Sullivan. (2020). *Eddy Correlation Flux Measurement System (ECOR) Instrument Handbook*. (No. DOE/SC-ARM/TR-052). DOE Office of Science Atmospheric Radiation Measurement (ARM) Program (United States).
- Deardorff, J. W. (1979). Prediction of convective mixed-layer entrainment for realistic capping inversion structure. *Journal of the Atmospheric Sciences*, 36(3), 424-436.
- Deryng, D., Elliott, J., Folberth, C., Müller, C., Pugh, T. A., Boote, K. J., ... & Khabarov, N. (2016). Regional disparities in the beneficial effects of rising CO₂ concentrations on crop water productivity. *Nature Climate Change*, 6(8), 786-790.
- Deschênes, O., & Greenstone, M. (2012). The economic impacts of climate change: evidence from agricultural output and random fluctuations in weather: reply. *American Economic Review*, 102(7), 3761-73.
- Deschênes, O., & Greenstone, M. (2007). The economic impacts of climate change: evidence from agricultural output and random fluctuations in weather. *American Economic Review*, 97(1), 354-385.
- Dyer, A. (1974). A review of flux-profile relationships. *Boundary-Layer Meteorology*, 7(3), 363-372.
- Emeis, S., Schäfer, K., & Münkel, C. (2008). Surface-based remote sensing of the mixing-layer height—a review. *Meteorologische Zeitschrift*, 17(5), 621-630.
- Fast, J. D., Berg, L. K., Feng, Z., Mei, F., Newsom, R., Sakaguchi, K., & Xiao, H. (2019). The Impact of Variable Land-Atmosphere Coupling on Convective Cloud Populations Observed During the 2016 HI-SCALE Field Campaign. *Journal of Advances in Modeling Earth Systems*, 11(8), 2629-2654.

- Fedorovich, E. E., & Mironov, D. V. (1995). A model for a shear-free convective boundary layer with parameterized capping inversion structure. *Journal of the atmospheric sciences*, 52(1), 83-96.
- Fernando, H. J. S., Pardyjak, E. R., Di Sabatino, S., Chow, F. K., De Wekker, S. F. J., Hoch, S. W., ... & Steenburgh, W. J. (2015). The MATERHORN: Unraveling the intricacies of mountain weather. *Bulletin of the American Meteorological Society*, 96(11), 1945-1967.
- Fernando, H. J., Gultepe, I., Dorman, C., Pardyjak, E., Wang, Q., Hoch, S. W., ... & Wauer, B. (2021). C-FOG: Life of coastal fog. *Bulletin of the American Meteorological Society*, 102(2), E244-E272.
- Fisher, M. C., Henk, D. A., Briggs, C. J., Brownstein, J. S., Madoff, L. C., McCraw, S. L., & Gurr, S. J. (2012). Emerging fungal threats to animal, plant and ecosystem health. *Nature*, 484(7393), 186-194.
- Foken, T. (2006). 50 years of the Monin–Obukhov similarity theory. *Boundary-Layer Meteorology*, 119(3), 431-447.
- Frehlich, R., Meillier, Y., Jensen, M. L., Balsley, B., & Sharman, R. (2006). Measurements of boundary layer profiles in an urban environment. *Journal of applied meteorology and climatology*, 45(6), 821-837.
- Friedman, J., Hastie, T., & Tibshirani, R. (2001). *The elements of statistical learning* (Vol. 1, No. 10). New York: Springer series in statistics.
- Gagne II, D. J., Haupt, S. E., Nychka, D. W., & Thompson, G. (2019). Interpretable deep learning for spatial analysis of severe hailstorms. *Monthly Weather Review*, 147(8), 2827-2845.
- Geerts, B., Parsons, D., Ziegler, C. L., Weckwerth, T. M., Biggerstaff, M. I., Clark, R. D., ... & Haghi, K. (2017). The 2015 plains elevated convection at night field project. *Bulletin of the American Meteorological Society*, 98(4), 767-786.
- Gibbs, J. A., Fedorovich, E., & Van Eijk, A. M. (2011). Evaluating Weather Research and Forecasting (WRF) model predictions of turbulent flow parameters in a dry convective boundary layer. *Journal of applied meteorology and climatology*, 50(12), 2429-2444.
- Goldsmith, J. E. M., Blair, F. H., Bisson, S. E., & Turner, D. D. (1998). Turn-key Raman lidar for profiling atmospheric water vapor, clouds, and aerosols. *Applied Optics*, 37(21), 4979-4990.
- Gutierrez, W., Araya, G., Kiliyanpilakkil, P., Ruiz-Columbie, A., Tutkun, M., & Castillo, L. (2016). Structural impact assessment of low level jets over wind turbines. *Journal of Renewable and Sustainable Energy*, 8(2), 023308.
- Hildebrand, P. H., & Sekhon, R. S. (1974). Objective determination of the noise level in Doppler spectra. *Journal of Applied Meteorology*, 13(7), 808-811.
- Högström, U. L. F. (1988). Non-dimensional wind and temperature profiles in the atmospheric surface layer: A re-evaluation. In *Topics in Micrometeorology. A Festschrift for Arch Dyer* (pp. 55-78). Springer, Dordrecht.

- Hunt, J. C. R., & Richards, K. J. (1984). Stratified airflow over one or two hills. In *Boundary Layer Structure* (pp. 223-259). Springer, Dordrecht.
- International Electrotechnical Commission. (2017). Wind turbines-Part 12-1: Power performance measurements of electricity producing wind turbines. *IEC 61400-12-1*.
- Jackson, P. S., & Hunt, J. C. R. (1975). Turbulent wind flow over a low hill. *Quarterly Journal of the Royal Meteorological Society*, *101*(430), 929-955.
- Jiménez, P. A., & Dudhia, J. (2013). On the ability of the WRF model to reproduce the surface wind direction over complex terrain. *Journal of Applied Meteorology and Climatology*, *52*(7), 1610-1617.
- Kennedy, A. D., Dong, X., Xi, B., Xie, S., Zhang, Y., & Chen, J. (2011). A comparison of MERRA and NARR reanalyses with the DOE ARM SGP data. *Journal of Climate*, *24*(17), 4541-4557.
- Kelly, M., & Gryning, S. E. (2010). Long-term mean wind profiles based on similarity theory. *Boundary-layer meteorology*, *136*(3), 377-390.
- Krishnamurthy, R., Calhoun, R., Billings, B., & Doyle, J. D. (2013). Mesoscale model evaluation with coherent Doppler lidar for wind farm assessment. *Remote sensing letters*, *4*(6), 579-588.
- Krishnamurthy, R., Calhoun, R., Billings, B., & Doyle, J. (2011). Wind turbulence estimates in a valley by coherent Doppler lidar. *Meteorological Applications*, *18*(3), 361-371.
- Krishnamurthy, R., Newsom, R. K., Berg, L. K., Xiao, H., Ma, P. L., & Turner, D. D. (2021). On the estimation of boundary layer heights: A machine learning approach. *Atmospheric Measurement Techniques Discussions*, 1-34.
- Emeis, S., Schäfer, K., & Munkel, C. (2008). Surface-based remote sensing of the mixing-layer height—a review. *Meteorologische Zeitschrift*, *17*(5), 621-630.
- Lareau, N. P. (2020). Subcloud and Cloud-Base Latent Heat Fluxes during Shallow Cumulus Convection. *Journal of the Atmospheric Sciences*, *77*(3), 1081-1100.
- Lareau, N. P., Zhang, Y., & Klein, S. A. (2018). Observed boundary layer controls on shallow cumulus at the ARM Southern Great Plains site. *Journal of the Atmospheric Sciences*, *75*(7), 2235-2255.
- Lenschow, D. H., Wyngaard, J. C., & Pennell, W. T. (1980). Mean-field and second-moment budgets in a baroclinic, convective boundary layer. *Journal of the Atmospheric Sciences*, *37*(6), 1313-1326.
- Lenschow, D. H., & Stephens, P. L. (1980). The role of thermals in the convective boundary layer. *Boundary-Layer Meteorology*, *19*(4), 509-532.
- Lenschow, D. H., Wulfmeyer, V., & Senff, C. (2000). Measuring second-through fourth-order moments in noisy data. *Journal of Atmospheric and Oceanic technology*, *17*(10), 1330-1347.

Lenschow, D. H., Lothon, M., Mayor, S. D., Sullivan, P. P., & Canut, G. (2012). A comparison of higher-order vertical velocity moments in the convective boundary layer from lidar with in situ measurements and large-eddy simulation. *Boundary-layer meteorology*, 143(1), 107-123.

Liu, S., & Liang, X. Z. (2010). Observed diurnal cycle climatology of planetary boundary layer height. *Journal of Climate*, 23(21), 5790-5809.

Liu, Y. (2019). Introduction to the Special Section on Fast Physics in Climate Models: Parameterization, Evaluation, and Observation. *Journal of Geophysical Research: Atmospheres*, 124(15), 8631-8644.

Lundquist, J. K., Churchfield, M. J., Lee, S., & Clifton, A. (2015). Quantifying error of lidar and sodar Doppler beam swinging measurements of wind turbine wakes using computational fluid dynamics. *Atmospheric Measurement Techniques*, 8, 907-920.

Lundquist, J. K., Wilczak, J. M., Ashton, R., Bianco, L., Brewer, W. A., Choukulkar, A., ... & Gunter, S. (2017). Assessing state-of-the-art capabilities for probing the atmospheric boundary layer: the XPIA field campaign. *Bulletin of the American Meteorological Society*, 98(2), 289-314.

Lundquist, J. K., DuVivier, K. K., Kaffine, D., & Tomaszewski, J. M. (2019). Costs and consequences of wind turbine wake effects arising from uncoordinated wind energy development. *Nature Energy*, 4(1), 26-34.

Mather, J. H., & Voyles, J. W. (2013). The ARM Climate Research Facility: A review of structure and capabilities. *Bulletin of the American Meteorological Society*, 94(3), 377-392.

McGovern, A., Elmore, K. L., Gagne, D. J., Haupt, S. E., Karstens, C. D., Lagerquist, R., ... & Williams, J. K. (2017). Using artificial intelligence to improve real-time decision-making for high-impact weather. *Bulletin of the American Meteorological Society*, 98(10), 2073-2090.

Monin, A. S., & Obukhov, A. M. (1954). Basic laws of turbulent mixing in the surface layer of the atmosphere. *Contrib. Geophys. Inst. Acad. Sci. USSR*, 151(163), e187.

Newsom, R. K., Turner, D. D., Mielke, B., Clayton, M., Ferrare, R., & Sivaraman, C. (2009). Simultaneous analog and photon counting detection for Raman lidar. *Applied optics*, 48(20), 3903-3914.

Newsom, R. K., Turner, D. D., & Goldsmith, J. E. (2013). Long-term evaluation of temperature profiles measured by an operational Raman lidar. *Journal of Atmospheric and Oceanic Technology*, 30(8), 1616-1634.

Newsom, R. K., Sivaraman, C., Shippert, T. R., & Riihimaki, L. D. (2015). *Doppler lidar vertical velocity statistics value-added product* (No. DOE/SC-ARM/TR-149). DOE ARM Climate Research Facility, Washington, DC (United States).

Newsom, R. K., Sivaraman, C., Shippert, T. R., & Riihimaki, L. D. (2015). *Doppler Lidar WIND Value-Added Product* (No. DOE/SC-ARM/TR-148). DOE ARM Climate Research Facility, Washington, DC (United States).

Newsom, R. K., Turner, D. D., Lehtinen, R., Münkel, C., Kallio, J., & Roininen, R. (2020). Evaluation of a Compact Broadband Differential Absorption Lidar for Routine Water Vapor

Profiling in the Atmospheric Boundary Layer. *Journal of Atmospheric and Oceanic Technology*, 37(1), 47-65.

Newsom, R. K. & Krishnamurthy, R. (2020). *Doppler lidar (DL) handbook* (No. DOE/SC-ARM/TR-101). DOE Office of Science Atmospheric Radiation Measurement (ARM) Program (United States). Available online: http://www.arm.gov/publications/tech_reports/handbooks/dl_handbook.pdf.

Nieuwstadt, F. T. (1984). The turbulent structure of the stable, nocturnal boundary layer. *Journal of the atmospheric sciences*, 41(14), 2202-2216.

Obukhov, A. (1946). Turbulence in thermally inhomogeneous atmosphere. *Trudy Inst. Teor. Geofiz. Akad. Nauk SSSR*, 1, 95-115.

O'Connor, E. J., Illingworth, A. J., Brooks, I. M., Westbrook, C. D., Hogan, R. J., Davies, F., & Brooks, B. J. (2010). A method for estimating the turbulent kinetic energy dissipation rate from a vertically pointing Doppler lidar, and independent evaluation from balloon-borne in situ measurements. *Journal of atmospheric and oceanic technology*, 27(10), 1652-1664.

Parish, T. R., Rodi, A. R., & Clark, R. D. (1988). A case study of the summertime Great Plains low level jet. *Monthly weather review*, 116(1), 94-105.

Parish, T. R., & Oolman, L. D. (2010). On the role of sloping terrain in the forcing of the Great Plains low-level jet. *Journal of the atmospheric sciences*, 67(8), 2690-2699.

Pearson, G., Davies, F., & Collier, C. (2010). Remote sensing of the tropical rain forest boundary layer using pulsed Doppler lidar. *Atmospheric Chemistry & Physics Discussions*, 10(2).

Pryor, S. C., Shepherd, T. J., & Barthelmie, R. J. (2018). Interannual variability of wind climates and wind turbine annual energy production. *Wind Energy Science*, 3(2), 651.

Rajewski, D. A., Takle, E. S., Lundquist, J. K., Oncley, S., Prueger, J. H., Horst, T. W., ... & Doorenbos, R. K. (2013). Crop wind energy experiment (CWEX): observations of surface-layer, boundary layer, and mesoscale interactions with a wind farm. *Bulletin of the American Meteorological Society*, 94(5), 655-672.

Reif, D. W., & Bluestein, H. B. (2017). A 20-year climatology of nocturnal convection initiation over the central and southern Great Plains during the warm season. *Monthly Weather Review*, 145(5), 1615-1639.

Ritsche, M. T. (2008). *Surface meteorological observation system (smos) handbook* (No. DOE/SC-ARM/TR-031). DOE Office of Science Atmospheric Radiation Measurement (ARM) Program (United States).

Santanello Jr, J. A., Friedl, M. A., & Kustas, W. P. (2005). An empirical investigation of convective planetary boundary layer evolution and its relationship with the land surface. *Journal of Applied Meteorology*, 44(6), 917-932.

- Santanello Jr, J. A., Dirmeyer, P. A., Ferguson, C. R., Findell, K. L., Tawfik, A. B., Berg, A., ... & Roundy, J. (2018). Land-atmosphere interactions: The LoCo perspective. *Bulletin of the American Meteorological Society*, 99(6), 1253-1272.
- Schlenker, W., Hanemann, W. M., & Fisher, A. C. (2005). Will US agriculture really benefit from global warming? Accounting for irrigation in the hedonic approach. *American Economic Review*, 95(1), 395-406.
- Schlenker, W., & Roberts, M. J. (2009). Nonlinear temperature effects indicate severe damages to US crop yields under climate change. *Proceedings of the National Academy of sciences*, 106(37), 15594-15598.
- Schween, J. H., Hirsikko, A., Löhnert, U., & Crewell, S. (2014). Mixing-layer height retrieval with ceilometer and Doppler lidar: from case studies to long-term assessment. *Atmospheric Measurement Techniques*, 7(11).
- Sengupta, M., Clothiaux, E. E., & Ackerman, T. P. (2004). Climatology of warm boundary layer clouds at the ARM SGP site and their comparison to models. *Journal of climate*, 17(24), 4760-4782.
- Sharma, V., Pressel, K., Parlange, M. B., Lehning, M., & Calaf, M. (2018). Large wind farms can destroy stratocumulus clouds-evidence from large eddy simulations. *AGUFM*, 2018, A31J-3006.
- Shin, J., & Park, S. (2020). The relationship between low-level cloud amount and its proxies over the globe by cloud type. *Atmospheric Chemistry and Physics*, 20(5), 3041-3060.
- Siedersleben, S. K., Lundquist, J. K., Platis, A., Bange, J., Bärfuss, K., Lampert, A., ... & Emeis, S. (2018). Micrometeorological impacts of offshore wind farms as seen in observations and simulations. *Environmental Research Letters*, 13(12), 124012.
- Sivaraman, C., McFarlane, S., Chapman, E., Jensen, M., Toto, T., Liu, S., & Fischer, M. (2013). Planetary Boundary Layer (PBL) Height Value Added Product (VAP): Radiosonde Retrievals. *Department of Energy Office of Science Atmospheric Radiation Measurement (ARM) Program (United States)*.
- Smith, E. N., Gibbs, J. A., Fedorovich, E., & Klein, P. M. (2018). WRF Model study of the Great Plains low-level jet: Effects of grid spacing and boundary layer parameterization. *Journal of Applied Meteorology and Climatology*, 57(10), 2375-2397.
- Snoek, J., Larochelle, H., & Adams, R. P. (2012). Practical bayesian optimization of machine learning algorithms. *Advances in neural information processing systems*, 25, 2951-2959.
- Stoelinga, M. T., & Warner, T. T. (1999). Nonhydrostatic, mesobeta-scale model simulations of cloud ceiling and visibility for an East Coast winter precipitation event. *Journal of applied meteorology*, 38(4), 385-404.
- Sunny Lim, K. S., Riihimaki, L. D., Shi, Y., Flynn, D., Kleiss, J. M., Berg, L. K., ... & Johnson, K. L. (2019). Long-term retrievals of cloud type and fair-weather shallow cumulus events at the ARM SGP site. *Journal of Atmospheric and Oceanic Technology*, 36(10), 2031-2043.

- Taylor, G. I. (1922). Diffusion by continuous movements. *Proceedings of the London mathematical society*, 2(1), 196-212.
- Taylor, G. I. (1935). Statistical theory of turbulence. *Proceedings Royal Society, A*, 156, 307-317.
- Taylor, G. I. (1935). Statistical theory of turbulence IV-diffusion in a turbulent air stream. *Proceedings of the Royal Society of London. Series A-Mathematical and Physical Sciences*, 151(873), 465-478.
- Tucker, S. C., Senff, C. J., Weickmann, A. M., Brewer, W. A., Banta, R. M., Sandberg, S. P., ... & Hardesty, R. M. (2009). Doppler lidar estimation of mixing height using turbulence, shear, and aerosol profiles. *Journal of Atmospheric and Oceanic Technology*, 26(4), 673-688.
- Turner, D. D., & Goldsmith, J. E. M. (1999). Twenty-four-hour Raman lidar water vapor measurements during the Atmospheric Radiation Measurement Program's 1996 and 1997 water vapor intensive observation periods. *Journal of Atmospheric and Oceanic Technology*, 16(8), 1062-1076.
- Turner, D. D., Goldsmith, J. E. M., & Ferrare, R. A. (2016). Development and applications of the ARM Raman lidar. *Meteorological Monographs*, 57, 18-1.
- Vakkari, V., O'Connor, E. J., Nisantzi, A., Mamouri, R. E., & Hadjimitsis, D. G. (2015). Low-level mixing height detection in coastal locations with a scanning Doppler lidar. *Atmospheric Measurement Techniques*, 8(4), 1875-1885.
- Vassallo, D., Krishnamurthy, R., & Fernando, H. J. (2020a). Decreasing wind speed extrapolation error via domain-specific feature extraction and selection. *Wind Energy Science*, 5(3), 959-975.
- Vassallo, D., Krishnamurthy, R., Sherman, T., & Fernando, H. J. (2020b). Analysis of Random Forest Modeling Strategies for Multi-Step Wind Speed Forecasting. *Energies*, 13(20), 5488.
- Vassallo, D., Krishnamurthy, R., & Fernando, H. J. (2021). Utilizing physics-based input features within a machine learning model to predict wind speed forecasting error. *Wind Energy Science*, 6(1), 295-309.
- Vassallo, D., Krishnamurthy, R., Menke, R., & Fernando, H. J. (2021). Observations of stably stratified flow through a microscale gap. *Journal of the Atmospheric Sciences*, 78(1), 189-208.
- Wang, Y., & Zhao, C. (2017). Can MODIS cloud fraction fully represent the diurnal and seasonal variations at DOE ARM SGP and Manus sites?. *Journal of Geophysical Research: Atmospheres*, 122(1), 329-343.
- Weckwerth, T. M., & Romatschke, U. (2019). Where, when, and why did it rain during PECAN?. *Monthly Weather Review*, 147(10), 3557-3573.
- Weng, W. (1997). Stably stratified boundary-layer flow over low hills: a comparison of model results and field data. *Boundary-layer meteorology*, 85(2), 223-241.

- Wharton, S., Simpson, M., Osuna, J., Newman, J., & Biraud, S. (2013). *Assessment of Land Surface Model Performance in WRF for Simulating Wind at Heights Relevant to the Wind Energy Community* (No. LLNL-TR-643914). Lawrence Livermore National Lab.(LLNL), Livermore, CA (United States).
- Whiteman, C. D., Bian, X., & Zhong, S. (1997). Low-level jet climatology from enhanced rawinsonde observations at a site in the southern Great Plains. *Journal of Applied Meteorology*, 36(10), 1363-1376.
- Wilczak, J. M., Stoelinga, M., Berg, L. K., Sharp, J., Draxl, C., McCaffrey, K., ... & Muradyan, P. (2019). The Second Wind Forecast Improvement Project (WFIP2): Observational Field Campaign. *Bulletin of the American Meteorological Society*, 100(9), 1701-1723.
- Wood, R. (2012). Stratocumulus clouds. *Monthly Weather Review*, 140(8), 2373-2423.
- Yang, F., Pan, H. L., Krueger, S. K., Moorthi, S., & Lord, S. J. (2006). Evaluation of the NCEP Global Forecast System at the ARM SGP site. *Monthly weather review*, 134(12), 3668-3690.
- Yu, M., Wu, B., Yan, N., Xing, Q., & Zhu, W. (2017). A method for estimating the aerodynamic roughness length with NDVI and BRDF signatures using multi-temporal Proba-V data. *Remote Sensing*, 9(1), 6.
- Zhang, Y., & Klein, S. A. (2010). Mechanisms affecting the transition from shallow to deep convection over land: Inferences from observations of the diurnal cycle collected at the ARM Southern Great Plains site. *Journal of the Atmospheric Sciences*, 67(9), 2943-2959.
- Zhang, Y., Seidel, D. J., & Zhang, S. (2013). Trends in planetary boundary layer height over Europe. *Journal of climate*, 26(24), 10071-10076.
- Zhang, H., Pu, Z., & Zhang, X. (2013). Examination of errors in near-surface temperature and wind from WRF numerical simulations in regions of complex terrain. *Weather and Forecasting*, 28(3), 893-914.
- Zhang, X., Yang, C., & Li, S. (2019). Influence of Low-Level Jet intensity on aerodynamic loads of horizontal axis wind turbine rotor. *Engineering Applications of Computational Fluid Mechanics*, 13(1), 300-308.
- Zhong, S., Fast, J. D., & Bian, X. (1996). A case study of the Great Plains low-level jet using wind profiler network data and a high-resolution mesoscale model. *Monthly Weather Review*, 124(5), 785-806.

Appendix A – Instrument List

Table A.1 presents a detailed list of the instruments, atmospheric radiation measurement data stream identification numbers needed to easily redo the analysis, the parameters, and the time period used in the analysis. Because some data streams had limited data, either due to instrument availability or failure, similar data available from other instruments were used in the analysis. Therefore, a near continuous data set from 2000 to 2020 was created, depending on the parameter and site.

Table A.1. SGP Instrument list, data stream, and parameters used for the analysis.

Instrument	ARM Data Stream	Parameters	Time Period Used in the Analysis	Comments
60-m Met Tower	spgco2flx4mC1.b1	Surface wind speed, wind direction, air temperature, relative humidity	2000 – 2019	The data were clean and had small data gaps.
60-m Met Tower	spgco2flx25mC1.b1	Turbulence kinetic energy, latent heat flux, sensible heat flux, Obukhov length, friction velocity, vertical velocity variance	2000 – 2019	The data were clean and had small data gaps. Depending on data availability, either the EBBR heat flux or the flux tower heat flux estimates were picked. Preference was given to the met tower.
EBBR	sgpebbrE13.00	Latent heat flux and sensible heat flux	2000 – 2019	Depending on data availability, either the EBBR heat flux or the flux tower heat flux estimates were picked. Preference was given to the met tower.
SEBS/SIRS	sgpbeflux1longC1.c1	Longwave radiation, shortwave radiation, normal radiation	2000 – 2019	Best estimate surface radiative flux from multiple instruments.
SWATS/STAMP	sgpstampe13.b1 sgpswatsE13.b1	Soil temperature, soil moisture	2012 – 2019	STAMP replaced SWATS in 2015. The data since 2015 are observed to have

Instrument	ARM Data Stream	Parameters	Time Period Used in the Analysis	Comments
				higher diurnal variations in soil moisture than prior to 2015. Therefore, soil moisture data past 2015 are not used in the analysis.
Doppler lidar	sgpdfptC1.b1 sgpdfptE32.b1 sgpdfptE37.b1 sgpdfptE39.b1 sgpdfptE41.b1	Signal-to-noise ratio, attenuated backscatter, eddy dissipation rate	2011 – 2020	Some gaps in data exist, due to system maintenance or failure.
Doppler lidar	sgpdprofwstats4newsC1.c1 sgpdprofwstats4newsE32.c1 sgpdprofwstats4newsE37.c1 sgpdprofwstats4newsE39.c1 sgpdprofwstats4newsE41.c1	Vertical velocity variance profile, cloud base height	2011 – 2020	Some gaps in data exist, due to system maintenance or failure.
Doppler lidar	sgpdprofwind4newsC1.c1	Horizontal wind speed profile, horizontal wind direction profile	2011 – 2020	Some gaps in data exist, due to system maintenance or failure.
Raman lidar	sgprlproffex1thorC1.c0 sgp10rlproftemp2newsC1.c0	Relative humidity, water vapor mixing ratio, back-scattering coefficient, lidar ratio, scattering coefficient	2016 – 2020	Some gaps in data exist, due to system maintenance or failure.
Radiosondes	sgppblhtsonde1mcfarIC1.c1	Boundary layer height estimates, inversion height, wind profiles, temperature, and relative humidity profiles	2011 – 2019	Boundary layer height estimates from Liu and Liang (2010) are used in the Random Forest algorithm.

EEBR = Energy Balance Bowen Ratio; SEBS/SIRS = Surface Energy Balance System/Surface Infrared Radiation Station; SWATS/STAMP = Soil Water and Temperature System/Soil Temperature and Moisture Profile.

Appendix B – Code and Plot Repository

A commercial software, MATLAB, was used for a major portion of this analysis. For the ease of redoing some of the analysis in this report, a repository of MATLAB codes used to read the raw/processed data from the instruments listed in Table A.1 is provided in GitHub. The repository of codes can be found in <https://github.com/rkpnnl/ARM.git>.

The lidar mentors also maintain a webpage for near real-time observations from the lidar (both Raman and Doppler lidar) at all the SGP measurement sites. Near real-time data visualization from vertical velocity variance, skewness, kurtosis, wind speed, and wind direction for the Doppler lidar are provided. Retrievals of aerosol backscatter, extinction, linear depolarization ratio, scattering ratio, and lidar ratio from the Raman lidar are provided. The quick looks can be accessed here: <https://engineering.arm.gov/~newsom/>.

Time-height cross-section plots from multi-year Doppler lidar data are also available and can be accessed here: <https://engineering.arm.gov/~raghuvaidhya/>. This includes data availability, wind distributions, wind profiles, and vertical velocity statistics at sites C1, E32, E37, E39, and E41 from 2012 to 2019. In the future, near real-time estimates of boundary layer height using the Random Forest algorithm described in the report will also be provided.

Pacific Northwest National Laboratory

902 Battelle Boulevard
P.O. Box 999
Richland, WA 99354
1-888-375-PNNL (7665)

www.pnnl.gov