MINISTERE DE L'ENSEIGNEMENT SUPERIEUR
ET DE LA RECHERCHE SCIENTIFIQUE

Felix houphouet-boigny university

UNITE DE FORMATION ET DE
RECHERCHE SCIENCES DES
STRUCTURES DE LA MATIERE ET DE
TECHNOLOGIE

RÉPUBLIQUE DE CÔTE D'IVOIRE
UNION - DISCIPLINE - TRAVAIL
Institute of Bio- and Geosciences (IBG-3)

Nº: 661

SPONSORED BY THE
Federal Ministry
of Education
and Research

MASTER IN RENEWABLE ENERGY AND CLIMATE CHANGE

SPECIALITY: GREEN HYDROGEN / GEORESOURCES

MASTER THESIS:

Subject/Topic:

Evaluation of Wind and Solar Power Generation over South Africa According to Different Data Sources

Presented on the 26th of September 2023 by

Houssoukri Zounogo Wahabou YODA

JURY:

Mr. Olivier OBROU	**President**	Full professor UFHB
Mr. Alexandre N'GUESSAN	**Examinor**	Associate professor UFHB
Mr. Wanignon Ferdinand FASSINOU	**Director**	Associate professor UFHB
Mr. Harrie-Jan Hendricks Franssen	**Co-Director**	Full professor FZJ

Academic year 2022-2023

DEDICACE

"Dedicated to the unwavering support and boundless love of my cherished family members. To my guiding light, my mother, whose strength and compassion have shaped me into who I am today. And to my late father, whose memory continues to inspire and motivate me to reach for the stars. Your love and guidance have been the pillars of my journey, and this achievement is a testament to the values you instilled in me. With heartfelt gratitude, this is for you."

ACKNOWLEDGEMENT

Foremost, I extend my profound gratitude to the Almighty for endowing me with the strength, knowledge, abilities, and opportunities to embark on this enriching research journey and ultimately bring it to fruition. This accomplishment stands as a testament to His blessings without which this endeavour would not have been achievable.

My heartfelt appreciation goes to the German Federal Ministry of Education and Research (BMBF) for their invaluable financial support through the West African Science Service Centre for Climate Change and Adapted Land Use (WASCAL) project. This support has been the cornerstone enabling the realization of this research endeavour.

I am deeply indebted to the entire WASCAL project team and the collective at the International Master Programme in Energy and Green Hydrogen Production (IMP-EGH). Their unswerving dedication to training, support, opportunities, and guidance has been a constant beacon throughout my research journey from the roots in Africa to my academic journey in Germany.

Being a part of the exceptional research team at the Institute of Bio- and Geosciences (IBG-3), Forschungszentrum Jülich GmbH, Juelich, Germany, has been a true privilege. I extend my sincere gratitude to my primary supervisor, Prof. Dr. Harrie-Jan Hendricks Franssen. His expert guidance, continuous support, and valuable insights have been the driving force. His mentorship transcended research to fruitful discussions and constructive suggestions, enriching the work. His willingness to share wisdom has been transformative in shaping my research direction.

I am equally appreciative of my local supervisor, Dr. Wanignon Ferdinand FASSINOU, at the Université Felix Houphouet Boigny (UFHB) in Cote d'Ivoire. His prompt responses and both administrative and scientific counsel have been indispensable.

My gratitude extends to my co-supervisor, Shuying Chen. Her unwavering support since the inception of my thesis has been instrumental. Her technical expertise and scientific advices has been paramount in shaping my research.

Their generosity in sharing knowledge enabled me to grasp concepts and express them creatively. Their patient explanations of coding processes were pivotal, and their support in addressing challenges has been cherished. Despite the tight timeline of four months, they were

consistently available for guidance, be it in research setup, analysis, or result interpretation. Their logical thinking, constructive feedback, and attentive mentoring have shaped my research outcomes.

I would like also to express my deepest gratitude to the members of the jury for their valuable time, expertise, and contributions to the evaluation of my master thesis. Their guidance and feedback have been instrumental in shaping the final outcome of this work.

I am thankful to all my IMP-EGH promotion mates, classmates, and office mates for their fellowship and support. To my family, your unwavering support has been my foundation.

In conclusion, this research would not have been possible without the steadfast support and encouragement of those mentioned above. Their contributions have played an indispensable role, and I am truly grateful for their assistance and guidance.

ABSTRACT

Reanalysis data and regional downscaling using atmospheric models have become integral tools for assessing wind and solar energy potential in renewable energy simulations. This study leverages ERA5 reanalysis and the high-resolution ICON in Limited Area Mode (ICON-LAM) simulated dataset, along with the Renewable Energy Simulation toolkit (RESKit), to evaluate wind speed and solar radiation, followed by the evaluation of their conversion to power using the RESKit model. South Africa is chosen as our study domain for its current national wide poor electricity supply and the availability of observation data. The focus time period spans a duration of three years, ranging from 2017 to 2019. In the first step, observation wind speeds collected from weather masts and observed solar radiation collected from Southern African Science Service Centre for Climate Change and Adaptive Land Management (SASSCAL) and Trans-African Hydro-Meteorological Observatory (TAHMO) stations are compared to modelled wind speeds and solar radiation. In the second step, observation wind and solar power generation collected from the Renewable Energy Data Information Service (REDIS) are compared to modelled wind and solar power generation. This comparison is conducted through diverse indicators, including Pearson correlation (r), root mean square error (RMSE), mean error (ME), coefficient of determination (R-squared or R^2), mean absolute error (MAE), and Perkins skill score (PSS). Our findings underscore the superiority of ICON-LAM over ERA5 in terms of ME, MAE, R-squared, RMSE, r and PSS for wind speed assessment. For solar radiation analysis, ICON-LAM outperforms ERA5 in terms of PSS, MAE, and RMSE. Shifting focus to wind power estimation, ERA5 shows better performance than ICON-LAM in ME, R-squared, RMSE, Pearson correlation, and PSS. Similarly, in solar power estimation, ERA5 excels over ICON-LAM in MAE, ME, RMSE, r, and PSS. It's crucial to highlight that the power generation comparison has been carried out using observation data that is significantly aggregated, resulting in notable uncertainty that may constrain the performance evaluation of the high-resolution ICON-LAM. This study not only underscores the significance of advanced datasets and tools but also sheds light on their nuanced performance in assessing renewable energy potential across various metrics.

Keywords: dynamical downscaling; ICON-LAM; reanalysis; ERA5; renewable energy; RESKit; evaluation; Africa.

RESUMÉ

Les données de réanalyse et la réduction d'échelle régionale à l'aide de modèles atmosphériques sont devenues des outils essentiels pour évaluer le potentiel éolien et solaire dans les simulations d'énergie renouvelable. Cette étude utilise la réanalyse ERA5 et l'ensemble de données simulées ICON-LAM à haute résolution, ainsi que l'outil de simulation d'énergie renouvelable (RESKit), pour évaluer la vitesse du vent et le rayonnement solaire, suivis de l'évaluation de leur conversion en puissance à l'aide du modèle RESKit. L'Afrique du Sud est choisie comme domaine d'étude en raison de son actuelle pénurie d'électricité à l'échelle nationale et de la disponibilité de données d'observation. La période d'étude s'étend sur une durée de trois ans, de 2017 à 2019. Dans la première étape, les vitesses du vent observées collectées à partir de mâts météorologiques et le rayonnement solaire observé collecté auprès du Centre de Services Scientifiques de l'Afrique Australe pour le Changement Climatique et la Gestion Adaptative des Terres (SASSCAL) et des stations de l'Observatoire Hydrométéorologique Trans-Africain (TAHMO) sont comparés aux vitesses du vent et au rayonnement solaire modélisés. Dans la deuxième étape, la production d'énergie éolienne et solaire observée collectée auprès du Service d'Information sur les Données d'Energie Renouvelable (REDIS) est comparée à la production d'énergie éolienne et solaire modélisée. Cette comparaison est réalisée à l'aide de divers indicateurs, notamment la corrélation de Pearson (r), l'erreur quadratique moyenne (RMSE), l'erreur moyenne (ME), le coefficient de détermination (R-carré ou R^2), l'erreur absolue moyenne (MAE) et le score de compétence de Perkins (PSS). Nos résultats soulignent la supériorité d'ICON-LAM par rapport à ERA5 en termes de ME, MAE, R-carré, RMSE, r et PSS pour l'évaluation de la vitesse du vent. Pour l'analyse du rayonnement solaire, ICON-LAM surpasse ERA5 en termes de PSS, MAE et RMSE. En ce qui concerne l'estimation de la puissance éolienne, ERA5 présente de meilleures performances que ICON-LAM en termes de ME, R-carré, RMSE, corrélation de Pearson et PSS. De même, dans l'estimation de la puissance solaire, ERA5 se distingue par rapport à ICON-LAM en termes de MAE, ME, RMSE, r et PSS. Il est important de souligner que la comparaison de la production d'énergie a été réalisée à l'aide de données d'observation significativement agrégées, ce qui entraîne une incertitude notable pouvant limiter l'évaluation des performances de l'ICON-LAM à haute résolution. Cette étude met en évidence non seulement l'importance des ensembles de données et des outils avancés, mais aussi leur performance nuancée dans l'évaluation du potentiel de l'énergie renouvelable selon diverses mesures.

Mots clés : affinage dynamique ; ICON-LAM ; réanalyse ; ERA5 ; énergie renouvelable ; RESKit ; évaluation ; Afrique.

ACRONYMS AND ABBREVIATIONS

AC : Alternative Current

CDS : Climate Data Store

CdTe : Cadmium Telluride

CF : Capacity factor

CIGS : Copper Indium Gallium Selenide

CLLJs : Coastal low-level wind jets

CORDEX : Coordinated Regional Climate Downscaling Experiment

CSP : Concentrated Solar Power

DC : Direct Current

DWD : Deutscher Wetterdienst

EBCS : Eastern Boundary Current Syste

ECMWF : European Centre for Medium-Range Weather Forecasts

EU : European Union

GaAs : Gallium Arsenide

GCM : Global Climate Model

GHI : Global Horizontal Irradiance

GIS : Geographic Information System

GSR : Global Solar Radiation

ICON-LAM : ICOsahedral Nonhydrostatic in Limited Area Mode

IEA : International Energy Agency

IFC : International Finance Corporation

MAE : Mean Absolute Error

MARS : Meteorological Archival and Retrieval System

ME : Mean Error

MPI-M : Max-Planck-Institute for Meteorology

NWP : Numerical Weather Prediction

Obs : Observation

PDF : Probability Density Function

PSS : Perkins Skill Score

PV : Photovoltaic

r : Pearson correlation

R^2 : Coefficient of determination

RCM : Regional Climate Model

RE : Renewable Energy

REDIS : Renewable Energy Data Information Service

RESKit : Renewable Energy Simulation toolkit

RMSE : Root Mean Square Error

RSDS : Surface Solar Irradiance Downwards

SASSCAL : Southern African Science Service Centre for Climate Change and Adaptive Land Management

TAHMO : Trans-African Hydro-Meteorological Observatory

UTC : Coordinated Universal Time

VRE : Variable Renewable Energy

LIST OF TABLES

LIST OF FIGURES

TABLES OF CONTENTS

GENERAL INTRODUCTION

GENERAL INTRODUCTION

Energy takes on the role of instigating change and sustaining ways of life in the context of society. Throughout history, civilizations have thrived due to the sun's energy harnessed by plants. However, the profound societal shifts spurred by the industrial revolution across the past two centuries demanded substantial energy, a need fulfilled primarily by coal and later petroleum. Presently, societal advancement endures, and the predominant source of energy remains non-renewable fossil fuels (coal, petroleum, and natural gas). These resources fuel various aspects of life, including transportation (mainly powered by petroleum), heating (utilizing natural gas), and electricity generation (involving coal, gas, nuclear, hydroelectric, and other renewable sources). The combustion of fossil fuels comes hand in hand with carbon emissions. As carbon dioxide accumulates in the atmosphere, apprehensions escalate regarding potential drastic alterations in climate patterns (Santamarina, 2006). Over the span of the last seven decades, the increase in atmospheric CO_2 levels has occurred at a rate approximately 100 times faster than observed at the conclusion of the previous glacial era (Lenton et al., 2019). Furthermore, the average global temperature now stands 1.1℃ higher than the pre-industrialization period (Lenton et al., 2019; Rhodes, 2017).

Moreover, we find ourselves ensnared within an unparalleled energy crisis. The illusion of boundless fossil fuel resources has been shattered, compounded by the exacerbating effects of the Russia-Ukraine conflict, which has further intensified this crisis (Turiel, 2022). However, the contemporary worldwide energy crisis transcends previous instances in its breadth and intricacy. The oil shocks of the 1970s predominantly concerned oil, prompting policymakers to focus on a relatively clear objective: diminishing reliance on oil, particularly oil imports. In contrast, the present energy crisis encompasses numerous facets. It encompasses not only natural gas and oil but also coal, electricity, food security, and the imperative issue of climate change (Birol, 2022). A crucial energy transition becomes imperative, necessitating a comprehensive consideration of both climate and economic factors. Several initiatives are already in progress much like the European Green Deal (Fetting, 2020). The European Union (EU) through this initiative aimed at addressing climate change, fostering sustainable growth, and achieving carbon neutrality by 2050. This ambitious agenda encompasses a wide array of sectors and policies, ranging from energy and transportation to agriculture and biodiversity (Fetting, 2020). Following a similar trajectory, the aspiration to attain carbon peak by 2030 and carbon neutrality by 2060 heralds the advent of a climate-conscious economy in China (Birol, 2022). This shift is crucial given that China has held the mantle of the world's foremost carbon

emitter since 2006 (Ritchie et al., 2020). The nation's CO_2 emissions surged to a staggering 10.67×10^9 tons in 2020, constituting a substantial 30.65% of the global aggregate (Ritchie et al., 2020).

The African continent is already facing more severe climate change than most other parts of the world, despite bearing the least responsibility for the problem. With nearly one-fifth of the world's population today, Africa accounts for less than 3% of the world's energy-related carbon dioxide (CO_2) emissions. Besides this climate crisis, 600 million people, or 43% of the total population, lack access to electricity, most of them in sub-Saharan Africa (IEA, 2022). These drivers paved the way for a large integration of renewable energies such as solar and wind in Southern Africa's energy mix seeing their huge potential (Roehrkasten et al., 2016).

The power output profiles of most renewable energy sources are highly dependent on weather conditions, resulting in a power source of a stochastic rather than deterministic nature (Njiri & Söffker, 2016). Solar PV and wind power, classified as variable renewable electricity (VRE) resources, exhibit weather-related variability on all timescales from sub-hourly to interannual, and their yield is site-specific (Sterl, 2021). Changes in weather conditions, particularly in terms of sunlight and wind speed, are increasingly influencing energy production and the reliability of energy supply. An accurate characterization of the variability in the irradiance and wind speed is important to quantify the impact on the power production, and to identify the best locations for installing PV power plants and wind power plants (Gruber et al., 2022).

Thus, nowadays, reanalysis climate datasets, regional dynamical downscaling, Numerical Weather Prediction (NWP) models are involved in several studies to generate time series of power generation from wind and solar to assess the viability of future electricity systems with high shares of renewables. This context brings us to the use of data sources as the reanalysis dataset ERA5 (Hersbach et al., 2020), the regional dynamical downscaling model of ICOsahedral Nonhydrostatic numerical weather prediction model in its Limited Area Mode (ICON-LAM) (Zängl et al., 2015) and the renewable energy simulation model of Renewable Energy Simulation toolkit (RESKit) (Ryberg et al., 2019a) for wind and solar assessment over South Africa. This backdrop gives rise to a series of research questions, namely: How accurate are the wind speed and solar irradiance simulations obtained from ERA5 and ICON-LAM models when compared to in-situ observations? What are the key performance indicators that can be used to evaluate the results of wind speed and solar irradiance simulations? How does the simulated electricity generation from RESKit models driven by the output of ERA5 and ICON-LAM compare with the true province-level generation time series data for wind and solar energy in South Africa? Following these research questions, the subsequent research

hypotheses come into focus: the wind speed and solar irradiance simulations obtained from ERA5 and ICON-LAM models will show a high degree of correlation with in-situ observations, indicating their accuracy; the key performance indicators used to evaluate the simulation results will demonstrate the reliability and suitability of the ERA5 and ICON-LAM models for renewable energy generation assessment in South Africa; the power generated from the RESKit models will closely align with the true province-level generation time series data for wind and solar energy in South Africa, confirming the driven models' effectiveness. It is important to note that ICON-LAM is a model with higher spatial resolution, which means it can provide more detailed information for specific geographic areas. On the other hand, ERA5 covers the entire globe but at a coarser spatial resolution. This suggests that ICON-LAM has the potential to outperform ERA5 in terms of accuracy and precision. Aligned with the preceding hypotheses, the main objective of this study is to facilitate the adoption of strategies for switching the energy sector to renewable energy in South Africa by evaluating renewable generation based on multiples inputs. This consideration gives rise to the following specific objectives: assess the accuracy of wind speed and solar irradiance simulations through the comparison with on-site observations, employing key performance indicators to gauge the reliability of the models; evaluate the accuracy of wind power and solar power output simulations through the comparison with on-site observations, employing key performance indicators to measure the effectiveness of the models . Transitioning to renewable energy (RE) is crucial for increasing access to electricity and combating climate change. However, this transition should be accompanied by sensitization on the advantages of RE and educational activities in this field.

This study is structured into three principal sections. Chapter 1 initiates with a comprehensive literature review, followed by concise definitions of key concepts. Additionally, an introduction to the underlying physics of wind and solar energy is provided, along with an exploration of the renewable energy potential within South Africa. Chapter 2 is devoted to delving into the study area and outlining the methodology adopted for the research. Lastly, Chapter 3 is dedicated to the thorough discussion of the results obtained, presenting a comprehensive analysis of the findings.

CHAPTER I: MAIN CONCEPTS AND LITERATURE REVIEW

CHAPTER I: MAIN CONCEPTS AND LITERATURE REVIEW

Introduction

This chapter begins by providing clear definitions of some key concepts. It then delves into the underlying physics behind wind and solar energy, explores their potential within the context of South Africa, and concludes by presenting a literature review.

I.1 Definition of main concepts

I.1.1 Reanalysis datasets

A reanalysis dataset, often referred to as atmospheric or climate reanalysis, is a comprehensive set of climate or weather data that assimilates a wide range of observations from various sources such as satellites, weather stations, and ocean buoys, into a numerical model to generate a consistent and coherent depiction of past weather and climate conditions. Reanalysis datasets are used to create historical records of the Earth's atmospheric conditions, which are crucial for climate research, weather forecasting, and studying long-term trends (Fujiwara et al., 2017).

I.1.2 Atmospheric models

Atmospheric models are mathematical representations of the Earth's climate system that use physical, chemical, and biological principles to simulate the behaviour of the atmosphere, oceans, land, and ice. These models are essential tools for understanding past climate variations, projecting future climate changes, and studying the impacts of human activities on the climate. There are several types of atmospheric models used for different purposes (Schneider & Dickinson, 1974):

- **Global Climate Models (GCMs)**

Global Climate Models (GCMs), also known as General Circulation Models, are the most comprehensive type of climate models. They simulate the Earth's climate system, including the atmosphere, ocean, land surface, and sea ice. GCMs are used to study long-term climate trends, natural climate variability, and the response of the climate system to changes in external forcing, such as greenhouse gas concentrations and solar radiation (Walsh et al., 2008).

- **Regional Climate Models (RCMs)**

RCMs are usually nested within GCMs and focus on specific regions or smaller areas. They use the output from GCMs as boundary conditions and then provide higher-resolution simulations for a particular region. RCMs are useful for studying local climate processes, extreme events, and assessing the impacts of climate change on a more regional scale (Laprise, 2008).

- **Dynamical Downscaling**

Dynamical downscaling is a technique that involves using high-resolution regional models, like RCMs, to provide more detailed information for specific areas. The downscaling process takes the coarse output from GCMs and refines it to capture local topography and land-use patterns, which can influence climate at smaller scales. This approach is particularly valuable when fine-scale information is needed for climate impact assessments or adaptation planning (Giorgi & Gutowski, 2015).

Climate models are continuously improved through a combination of advances in scientific understanding and increased computational power.

I.2 Wind and solar energy basics

I.2.1 Physics behind wind energy

(Wind is the result of the expansive motion of air particles on a significant scale. This motion is instigated by variations in atmospheric pressure caused by variations in temperature. Essentially, it originates from the sun. When the sun heats the Earth's surface, it also warms the atmosphere. This disparity in solar energy, in conjunction with the Earth's rotation, gives rise to our worldwide circulation patterns and prevailing surface winds. At the equator, hot air ascends, then travels north towards the poles, descends, and eventually flows southward to replenish the air at the equator)(Schneider & Sobel, 2022).

I.2.1.1 Wind turbine physics

A wind turbine is a machine that converts kinetic energy from the wind into electricity. The blades of a wind turbine turn between 13 and 20 revolutions per minute, depending on their technology, at a constant or variable velocity, where the velocity of the rotor varies in relation to the velocity of the wind in order to reach a greater efficiency (Tong, 2010a). The kinetic energy in Joules (J) of a mass m (kg) with the velocity v (m/s) is (Tong, 2010b):

$$E_k = \frac{1}{2}mv^2 \tag{1}$$

The air mass m can be determined from the air density ρ (kg/m^3) and the air volume V (m^3) according to

$$m = \rho V \tag{2}$$

Then,

$$E_{k,wind} = \frac{1}{2}\rho V v^2 \tag{3}$$

Power is energy divided by time. We consider a small time, Δt (s), in which the air particles travel a distance s=vΔt (m). We multiply the distance with the rotor area of the wind turbine, A (m^2), resulting in a volume of

$$\Delta V = Av\Delta t \tag{4}$$

which drives the wind turbine for the small period of time. Then the wind power in Watt (W) is given as:

$$P_{wind} = \frac{E_{k,wind}}{\Delta t} = \frac{\Delta \rho V v^2}{2\Delta t} = \frac{\rho A v^3}{2} \tag{5}$$

The wind power increases with the cube of the wind speed. In other words: doubling the wind speed gives eight times the wind power. Therefore, the selection of a "windy" location is very important for a wind turbine. The effective usable wind power is less than indicated by the above equation. The wind speed behind the wind turbine cannot be zero, since no air could follow. Therefore, only a part of the kinetic energy can be extracted considering the Figure 1:

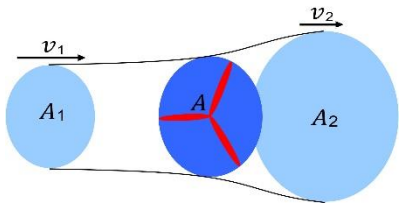


Figure 1: Speed variation in an ideal model of wind turbine

Source: (*Physics of Wind Turbines | Energy Fundamentals*)

The wind speed before the wind turbine is larger than after. Because the mass flow must be continuous, Av=constant, the area A_2 after the wind turbine is bigger than the area A_1 before. The effective power is the difference between the two wind powers:

$$C_p = \frac{P_{eff}}{P_{wind}} = \frac{(v_1 + v_2)(v_1^2 - v_2^2)}{2v_1^3} \quad (6)$$

A maximum drawing power is then obtained for $v_2 = v_1 / 3$ after deriving the above equation, and the ideal power coefficient is given by:

$$C_p = \frac{P_{eff}}{P_{wind}} = \frac{16}{27} \approx 59\% \quad (7)$$

This is referred to as the Betz Criterion or the Betz Limit and it was first formulated in 1919, and applies to all wind turbine designs. It is the theoretical power fraction that can be extracted from an ideal wind stream. Modern wind machines operate at a slightly lower practical non-ideal performance coefficient. It is generally reported to be in the range of 35% - 45% (Huleihil et al., 2012).

I.2.1.2 Turbine electricity generation

When wind strikes the rotor blades, they start rotating, and the gearbox transforms the rotor rotation from low to high speed. The high-speed shaft from the gearbox is coupled with the rotor of the generator, and the electrical generator runs at a higher speed. An exciter is needed to give the required excitation to the magnetic coil of the generator field system so that it can generate the required electricity. The generated voltage at output terminals of the alternator is proportional to both the speed and field flux of the alternator. The speed is governed by wind power, which is out of control. Hence, to maintain uniformity of the output power from the alternator, excitation must be controlled according to the availability of natural wind power. The exciter current is controlled by a turbine controller that senses the wind speed. Then the output voltage of the electrical generator is given to a rectifier, which converts it into direct current (DC). This rectified DC output is given to a line converter unit to convert it into stabilized AC output, which is ultimately fed to either the electrical transmission network or transmission grid with the help of a step-up transformer (https://www.facebook.com/electrical4u, 2020).

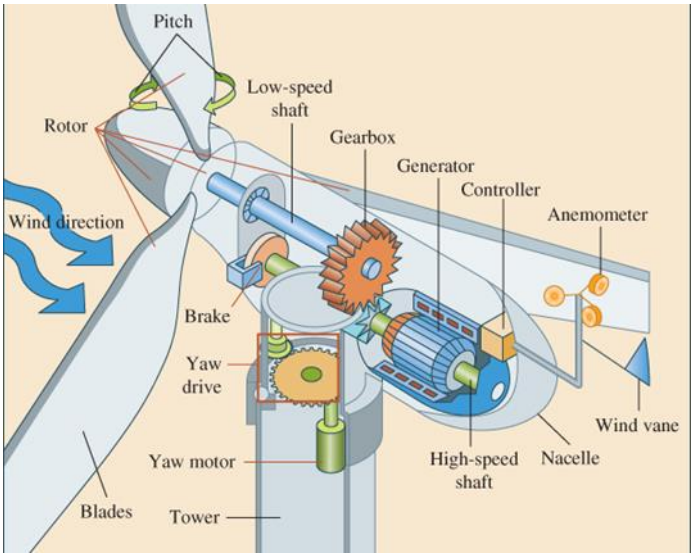


Figure 2: Components of wind turbine

Source: (https://www.facebook.com/linquip, 2021)

I.2.1.3 Wind power curve

The power curve of a wind turbine shows the relationship between the turbine's power output and the mean wind speed. The curve is determined from field measurements and shows that the turbine starts producing usable power at the cut-in speed. The power output increases until it reaches the rated power output at the rated speed, after which further increases in wind speed will not increase power output due to power control. The cut-out speed is the wind speed at which the turbine must shut down to avoid damage as increased mechanical stress on the rotor blades, heightened vibration levels in the tower structure, and accelerated wear and tear on various components that lead to the efficiency decrease. The cut-in and cut-out speeds define the operating limits of the wind turbine (Tong, 2010a).

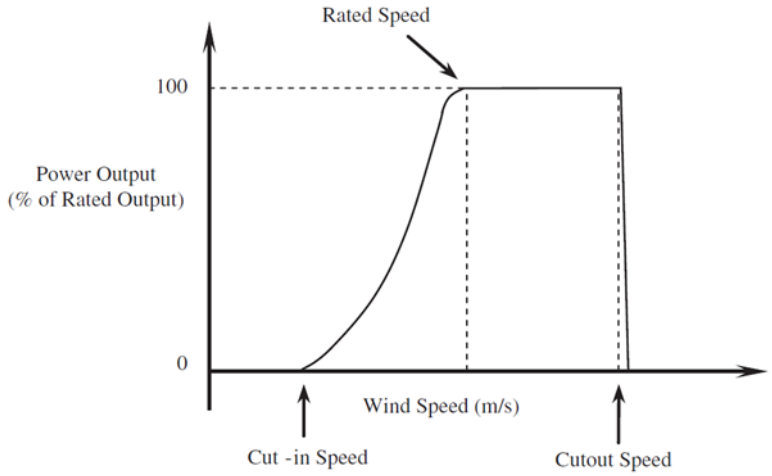


Figure 3: Typical wind turbine power curve (Tong, 2010a)

I.2.2 Physics behind photovoltaic (PV) energy

Photovoltaics is the process of directly converting sunlight into electrical energy using thin layers of semiconducting materials, which possess characteristics that lie between those of metals and insulators (Green, 2000).

I.2.2.1 The photovoltaic effect

The phenomenon of converting solar radiation into electrical energy is achieved through the photovoltaic effect, initially discovered by Becquerel (Copeland et al., 1941). This effect is commonly defined as the generation of an electric voltage when light is applied to a solid or liquid system with two electrodes. In practical terms, most photovoltaic devices utilize a semiconductor's pn junction, where the photovoltage is created (Goetzberger et al., 2002).

I.2.2.2 Photovoltaic technologies

Different technologies of photovoltaic are already in the market and research is still ongoing to improve their efficiency. Here we mention some PV technologies:

- **Crystalline silicon PV cells**

These cells are referred to as the first-generation PV cells because they were developed as early as the 1950s. They are produced from 100 to 200 µm thick wafers sliced from bulks of solar grade silicon (Goodrich et al., 2013). These solar cells are alternatively referred to as conventional, traditional, or wafer-based solar cells. Their production methods determine whether they are mono-crystalline (mono-Si) or polycrystalline (multi-Si) in nature.

- **Thin film PV cells**

The second-generation PV cell, known for its flexibility and light weight, is produced by applying one or more thin film layers of photovoltaic material onto substrates like glass, plastic, ceramic, or metal. These thin film layers can range from nanometers (nm) to micrometers (µm) in thickness. Unlike crystalline silicon PV cells that rely on costly and intricate ingot-growth techniques, thin film cells are more affordable (Águas et al., 2015). Thin film PV cells encompass several types, with the market-dominant ones being amorphous silicon, Copper Indium Gallium Selenide (CIGS), Cadmium Telluride (CdTe), and Gallium Arsenide (GaAs).

- **State-of-the-art PV cells**

These PV cells belong to the third generation and are often referred to as emerging technologies due to their experimental stage and limited market significance. They utilize a wide variety of substances, predominantly organic and often organo-metallic compounds. Perovskite PV cells, Dye PV cells, Concentrated PV (CPV) cells, Silicon germanium (SiGe) PV cells, Nanofibre

PV cells, and Multi-junction/Tandem/Cascaded PV cells are all diverse types of photovoltaic cells belong to this third generation of solar cells (Ogbomo et al., 2017).

I.2.2.3 The current-voltage characteristic

The I-V characteristic curves of a solar cell, module, or array depending on solar radiation provide insights into its current-voltage relationship and serve as a valuable tool for evaluating solar energy conversion efficiency. These curves offer a detailed description of the device's ability to convert solar energy and its overall efficiency. Understanding the electrical I-V characteristics, particularly the maximum power point (Pmax), is crucial in assessing the performance and solar efficiency of a solar cell or panel. Solar cells generate direct current (DC) electricity, and the relationship between current, voltage, and power is defined by the equation P = I × V. Therefore, it is possible to construct I-V curves for solar cells, which illustrate the current versus voltage characteristics of a photovoltaic device. The diagram below displays the current-voltage (I-V) behaviour of a standard silicon PV cell operating under normal conditions. The power output of an individual solar cell or panel is determined by multiplying its current and voltage values (I x V). By performing this multiplication at various voltage points, ranging from short-circuit to open-circuit conditions, the resulting power curve is obtained for a specific level of radiation (Morgan et al., 1994).

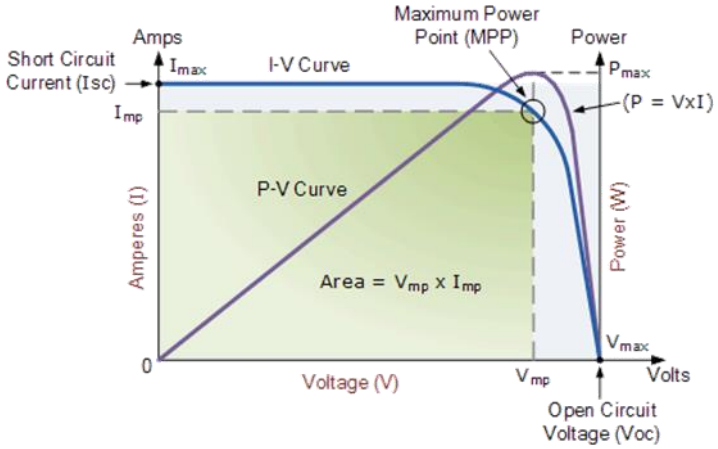


Figure 4: Solar cell current-voltage characteristic

Source: (*Solar Cell I-V Characteristic and the Solar Cell I-V Curve*, 2021)

I.2.2.4 Electricity generation

In terms of solar power generation integrated with grid power, there are essentially two types: concentrated solar power (CSP) and photovoltaic (PV) power. This study focuses on PV power. CSP generation, also referred to as solar thermal power generation, operates similarly to

traditional thermal power generation, where thermal energy (steam) is converted into electricity. PV solar panels differ from solar thermal systems as they do not rely on the sun's heat to generate thermal power. Instead, they utilize sunlight through the "Photovoltaic effect" to produce direct electric current (DC). Subsequently, this direct current is transformed into alternating current, typically with the help of inverters and other components. This conversion enables the electricity to be distributed across the power grid network or used locally as a stand-alone system (Khare & Rangnekar, 2014).

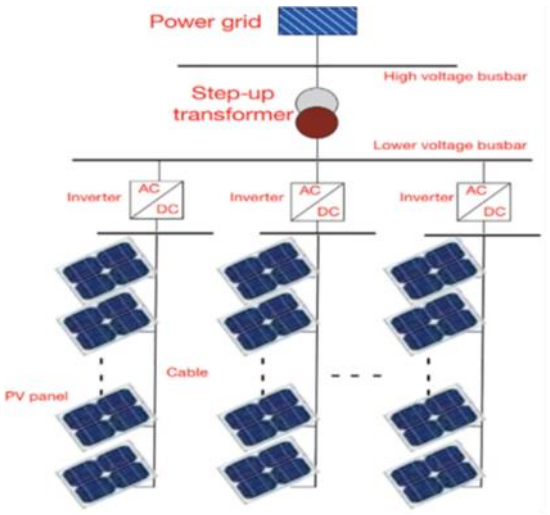


Figure 5: On grid power station (Nwaigwe et al., 2019)

I.3 Wind and solar potential in South Africa

I.3.1 Wind energy potential

Wind has been used as an energy source in South Africa for many decades, with early settlers using windmills to pump water for agricultural purposes (Van Der Linde, 1996). Akinbami et al., (2021) estimated the wind energy potential in South Africa based on meteorological data from about 170 stations at varying heights (2–10 m). The atlas categorized the nation's wind energy potential into regions of good, moderate, and low wind potential areas. The coastal areas of Western and Eastern Cape regions were classified as good wind energy potential regions with an average annual speed of over 4 m/s at 10 m above the ground level, while the Cape Middleveld and the Bushveld basin were categorized as low wind potential areas. Diab (1995) in her study concluded that over 7 TWh/year of energy can be potentially generated in South Africa. The wind atlas created in 2008 modelled wind speeds throughout the country at heights above 10 m, revealing that the best regions for harnessing wind power are the Western, Northern, and Eastern Cape regions. The study showed that about 6 GW of wind energy can be potentially harnessed in South Africa, and under optimistic assumptions, this value can be well over 50 GW (Szewczuk & Prinsloo, 2010).

I.3.2 Solar energy potential

Solar energy is one of the renewable energy resources with highest potential in South Africa. In South Africa, there is a total area of approximately 194,000 km^2 of high solar radiation potential including Northern Cape which is one of the best solar resource areas in the world (Aliyu et al., 2018). The highest DNI observed in South Africa is in the Northern Cape region with the DNI peaking at over 3200 kWh/m^2 while the KwaZulu-Natal region has a modest annual DNI of over 1400 kWh/m^2 (Akinbami et al., 2021).

I.4 Literature review

The literature review provided focuses on works related to evaluation studies of wind speed and solar irradiance that are served for the downstream associated power estimate application. The evaluation of wind and solar power is crucial for the planning and development of renewable energy systems. Various scientists have conducted studies globally, in Africa, and in specific African countries, using different methods to estimate solar and wind potential.

- **Wind Energy**

In a study conducted by Ohunakin (2011), statistical analysis was employed to explore the wind behavior across five stations in Nigeria. Through this analysis, the study determined that the mean annual wind speeds ranged from 2.7 m/s to 4.7 m/s, while the yearly wind power densities spanned from 16 W/m^2 to 95 W/m^2.

Thus, a recent study by Libanda & Paeth, (2023) used reanalysis datasets and state-of-the-art regional climate models from the Coordinated Regional Climate Downscaling Experiment (CORDEX Africa) to study the current and future wind energy potential in Zambia. The study found that winds in Zambia are dominated by southeasterlies and are rarely strong, with an average speed of 2.8 m/s. The study also found that although wind speed is increasing, it is still generally too weak to support large-scale wind power generation. The highest wind power densities were projected in the northern and central parts of the country, while the lowest were expected along the Luangwa valley. The study concluded that small-scale wind turbines that accommodate cut-in wind speeds of 3.8 m/s or lower would be suitable for Zambia.

Moreover, Zheng et al., (2018) delved into the global offshore wind potential utilizing Era-interim datasets. Through their research, they introduced a novel categorization for the world's offshore wind resource. Their findings unveiled regions of immense energy potential situated within mid and high latitudes. Notably, these areas encompass the Southern Hemisphere easterlies and the North Atlantic. Intriguingly, certain lower latitude regions, such as the Somalian coast, emerged as high-wind potential areas. The study further showcased that

elevated offshore wind resources are characteristic of the Northwestern African region. This prominence is attributed to the Eastern Boundary Current System (EBCS), which defines the circulation patterns. Coastal low-level wind jets (CLLJs) play a significant role in this phenomenon. Similarly, like to the Findlater jet in Madagascar, some other African regions exhibit promising prospects for wind energy development.

Elsner (2019) conducted an evaluation of the geographical diversity of overlooked wind resources across the African continent. Their research shed light on specific African countries, including Mozambique, South Africa, Somalia, Madagascar, and Morocco, which possess exceptional wind energy potential. Elsner's study emphasized certain limitations in the field, such as the absence of robust quantitative analyses that comprehensively assess wind energy potential and the relative immaturity of technology. These factors stand as obstacles in the advancement of wind energy development within the continent. Also, in their study titled "Development of Wind Energy in Africa", Mukasa et al., (2015) elaborated on the progression of Africa's wind energy sector. Their study delved into the complexities of the region's wind energy projects, examining the diverse factors that impede their growth. The authors underscored the relatively modest scale of Africa's wind market, despite the continent's abundant offshore and onshore wind energy potential. As per the comprehensive reports delivered by the International Finance Corporation (IFC) (Newell & Bulkeley, 2017), it has been determined that Africa boasts ample onshore wind resources with the potential to consistently power its population. Surprisingly, wind energy's current contribution to Africa's existing power grid remains below one percent. Within this context, a notable 27 African nations have been identified as holding significant wind energy potential. Among these states, Algeria emerges as a leader, boasting a substantial untapped total resource estimated at a staggering 7,700 gigawatts of electrical power. To put this in perspective, this amount equates to a remarkable 11 times the entirety of the world's installed wind capacity.

Mentis (2013), assessed wind energy potential in Africa using GIS (Geographic Information System) analysis with a 1km×1km grid size resolution and the Inverse Distance Weighted method. They found that Sudan, South Africa, and Algeria had the highest potential. Thus, Akor (2021) evaluated wind energy potential in Africa using ERA5 data, providing insights into spatial and temporal variations and identifying areas with high wind resources. The study found good representation of temperature but biases in precipitation in certain regions.

- **Solar Energy**

Diabaté et al., (2004) proposed a map of the solar radiation climate in Africa for preliminary assessment and modelling of solar energy systems. The study assembled a dataset of monthly

means of the daily clearness index for 62 sites in Africa and created a solar radiation climate map comprising 20 climates based on cluster analysis. The study utilized known features and other atlases to map the solar radiation climate in Africa. The map and its companion tables provided a basis for performing elementary energy calculations. The study acknowledged the need for further improvements in the mapping process. Moreover, Sawadogo et al., (2023) evaluated the performance of reanalysis and satellite-based products for estimating hourly GHI in Burkina Faso and Ghana. The findings indicated that satellite data outperformed reanalysis data, and the study introduced a new measure of overall performance. The results suggested that CAMS and SARAH-2 data could be used as alternative data sources for assessing solar energy in different climatic zones of West Africa. Similarly, Amillo et al., (2018) conducted a study to evaluate various satellite-based solar radiation methodologies aimed at estimating the solar resource potential in South Africa. Through their research, they identified the most effective satellite method, which exhibited a notably low relative mean bias. This method was subsequently integrated into a geographical information system module that accounted for the influence of surrounding topography. The derived estimates from this methodology are valuable for subsequent analyses and applications, such as the creation of maps depicting the potential output anticipated from photovoltaic systems within South Africa.

Aliyu et al., (2018) in their study reviewed the prospects of renewable energy sources for South Africa, Egypt, and Nigeria. The findings indicated that these countries encourage energy efficiency and have unique potential for hydro, solar, wind, and biomass energy. The study emphasized the importance of providing the necessary technology and skills to harness these resources effectively. It also highlighted the need to address energy challenges and consider factors such as energy efficiency, grid extension, energy storage, and seasonal variation in the implementation of renewable energy systems.

Wright et al., (2019) quantified the significance linked with enhanced accuracy in forecasting variable renewable energy (VRE) and examines the potential implications of improved VRE forecasting on both the South African power system and comparable power system domains within the United States. The study's findings highlight that the value associated with refined VRE forecasts demonstrates an upward trajectory in tandem with the escalation of VRE penetration. Interestingly, this value range widens as the penetration increases. Nevertheless, the study underscores that this value ultimately reaches a saturation point at high levels of improved VRE forecasts, reflecting a diminished incremental impact on further improvement.

Additionally, another study conducted by Govindasamy & Chetty, (2018) assessed the quantity of Global Solar Radiation (GSR) received in Pietermaritzburg, South Africa. It rigorously

validated the efficiency and precision of forecasting models specifically designed for areas in close proximity. The outcomes of this investigation accentuated the potential benefits of employing comprehensive prediction and forecasting models. Such models are envisaged to facilitate strategic placements of solar technologies for effective harnessing of GSR, thereby optimizing energy utilization.

Fant & UNU-WIDER, (2016) developed hourly estimates for onshore wind and solar photovoltaic (PV) electricity generation for the period from 1979 to 2010 across Africa. This was achieved through the utilization of publicly available data and established methodologies. The primary intention behind creating this dataset was to facilitate its integration into an energy-expansion-planning model designed for the African continent, forming the basis for a forthcoming study. The study's findings brought to light distinctive patterns in wind and solar resources across Africa. Wind resources exhibited more substantial variations both temporally and spatially compared to solar resources. It was discerned that these variations in wind resources significantly influenced the outcomes. Notably, the East African Power Pool emerged as an area with noteworthy potential for both wind and solar energy generation, attributed in large part to the observed variations in wind resources. On the other hand, the Central African Power Pool exhibited comparatively lower potential for both wind and solar energy due to more constrained wind resources.

In conclusion, the literature review offers valuable insights into the untapped potential of wind and solar energy. It also underscores the existing gap stemming from the scarcity of datasets across the continent, essential for the assessment of wind and solar resources. This scarcity of data compelled us to turn to alternative data sources such as ERA5 and ICON-LAM for the estimation of wind and solar power generation in the context of South Africa.

Conclusion

To summarize, this chapter has covered key concepts in climate science and renewable energy. The discussion began by defining reanalysis datasets, Global Climate Models, Regional Climate Models, and Dynamical Downscaling. The fundamental principles of wind turbine physics, electricity generation in turbines, the wind power curve, and the photovoltaic effect were also explored. Photovoltaic technologies were examined, including crystalline silicon PV cells, thin film PV cells, and state-of-the-art PV cells, with insights into the current-voltage characteristic. The discussion then delved into solar power generation. Finally, the wind and solar potential in South Africa was assessed, followed by an overview on the subject matter.

CHAPTER II: MATERIALS AND METHODOLOGY

CHAPTER II: MATERIALS AND METHODOLOGY

Introduction

In this section, we provide an overview of the study area, offer detailed descriptions of the materials used, explain the process of data acquisition, delineate the methods employed, and outline the key analysis parameters for our study.

II.1 Study area

In this research, an assessment of wind and solar energy potential is conducted in the vast and diverse region of South Africa. South Africa is a country located at the southernmost tip of the African continent. It is bordered by the Atlantic Ocean to the west and the Indian Ocean to the east. The country shares land borders with Namibia, Botswana, Zimbabwe, Mozambique, Eswatini (formerly known as Swaziland), and Lesotho. South Africa's geographic coordinates range from approximately 22.5 degrees to 35 degrees south latitude and 16.5 degrees to 33 degrees east longitude. Its location at the convergence of two major oceans and its diverse landscapes contribute to the country's unique climate and rich natural resources. Due to the substantial population, thriving industries, ports, and cities, there is a significant surge in power demand, necessitating the planning of feasible renewable energy projects to achieve sustainable development goals. Among the available renewable energy sources in South Africa, this is focused on wind and solar energy. Wind power stands out as a promising eco-friendly option, with abundant availability in both onshore and offshore locations. Onshore wind is affected by surface roughness and undulating topography.

In addition to onshore wind, an assessing of solar energy potential is conducted in the study area. South Africa benefits from ample sunshine throughout the year, making solar energy a valuable and readily available resource for sustainable power generation.

II.2 Materials description

This section provides an overview of the datasets used for assessing solar and wind resources in this study. The wind speed analysis as well as solar irradiance relies on data from ERA5 and ICON-LAM with a focus time period of 2017 to 2019. Additionally, wind power output evaluation and solar power output will also be estimated using RESKit model.

II.2.1 Reanalysis dataset ERA5

In this study, ERA5 is used to conduct a comprehensive assessment of wind and solar resources. ERA5 is a reanalysis product developed by the European Centre for Medium-Range Weather Forecasts (ECMWF) as part of the Copernicus Climate Change Service.

It represents a significant improvement over its predecessor, ERA-Interim and offering higher temporal resolution with hourly data. Moreover, ERA5 boasts a more refined horizontal grid spacing of approximately 30 km, enabling a more accurate representation of regional wind and solar characteristics. Furthermore, ERA5 incorporates enhanced assimilation techniques, resulting in a better global balance of precipitation and evaporation data, which significantly improves the accuracy and reliability of the reanalysis. As a result, it has already been utilized successfully for energy-related studies, demonstrating remarkable improvements over previous reanalysis datasets (Ramon et al., 2019; Silva et al., 2021). As for its storage volume, ERA5 is a massive dataset due to its high spatiotemporal resolution and long-time span. The comprehensive dataset is stored and organized in the ECMWF's MARS (Meteorological Archival and Retrieval System). Starting from 1950 to present, the dataset's overall size amounts to approximately 5 petabytes. To facilitate quick retrieval of ERA5 data, a refined version (approximately 1 petabyte) is also accessible on the CDS (Climate Data Store) cloud server (Hersbach et al., 2020).

II.2.2 Regional Dynamical Downscaling model

The ICOsahedral Nonhydrostatic (ICON) model is an atmospheric model mainly developed by the German Weather Service (Deutscher Wetterdienst, DWD) and Max-Planck-Institute for Meteorology (MPI-M). The ICON model utilizes a non-hydrostatic dynamical core and a hexagonal grid, known as the icosahedral-triangular Arakawa C grid, which allows for more accurate representation of terrain surface and its derived atmospheric processes at various scales. Distinguished by its sophisticated design, the ICON model boasts a significantly enhanced spatial resolution when compared to ERA5. Data from ICON in Limited Area Mode (ICON-LAM) simulations with grid spacing of 3.3km over South Africa, which is produced by dynamical downscaling of the ICON global weather forecast, is utilized in this study. ICON-LAM has been used for short-range weather forecasts, typically covering areas as small as a few hundred kilometers and it provides more detailed and accurate forecasts for mesoscale weather phenomena, such as local winds, and other small-scale atmospheric features (Pham et al., 2021).

II.2.3 The RESkit model

The Renewable Energy Simulation toolkit (RESKit) stands as an all-encompassing solution meticulously crafted to streamline large-scale simulations of renewable energy systems. Its primary objective revolves around the seamless generation of input data for Energy System

Design Models. Covering an extensive spectrum, this toolkit boasts simulation capabilities spanning onshore and offshore wind turbines, as well as solar PV systems. In addition, it features a set of tools that empower users to manipulate general weather data. Of paramount significance, RESKit executes simulations at the granularity of individual units while maintaining an exceptional level of computational efficiency. This remarkable trait empowers RESKit to swiftly simulate considerable quantities of individual turbines and PV systems within mere minutes, contingent upon the utilization of suitable hardware. It's worth highlighting that this performance benchmark is achieved even as RESKit operates at the intricate level of unit-level wind turbines and PV modules. Prominent attributes of RESKit encompass the capacity to generate synthetic wind turbine power curves, seamless access to the latest PV module databases sourced from esteemed institutions such as Sandia and the CEC, and the flexibility to configure settings to harness climate model datasets (*RESKit - Renewable Energy Simulation Toolkit for Python*, 2020/2023; Ryberg et al., 2019b).

II.3 Data acquisition

In this section, we outline the process of data acquisition for both wind and solar resources, encompassing the collection and preparation of various datasets, including in-situ wind-speed and solar irradiance measurements from meteorological weather masts and weather nets in different locations over South Africa, as well as regional-level solar and wind power observations from the Renewable Energy Data and Information Service (REDIS). Additionally, we incorporate information about existing wind farms and solar installations in the study area, which will be instrumental in simulating the energy production of these renewable energy facilities.

II.3.1 Wind observation data

II.3.1.1 Wind speed

Wind-speed measurements at levels close to turbine hub heights are challenging to access due to confidentiality reasons. In this study, we collected wind-speed data from 19 meteorological weather masts (http://wasadata.csir.co.za/wasa1/WASAData) installed at various locations across South Africa. Wind speeds at 10m and 60m were collected. These weather masts provided valuable measurements for our analysis, allowing us to assess the wind resources and their potential for wind energy production in the study area. The 10-minute wind-speed series underwent a meticulous manual filtering process to eliminate erroneous measurements, including extended periods of zeros and abnormal behaviour of individual anemometers compared to the others. Fortunately, the number of anomalies detected was minimal, indicating

that the collected measurements were of high quality and reliability. Subsequently, the data were aggregated by averaging the 10-minute readings to hourly values, providing a more manageable dataset for our analysis.

For each location, we have access to wind-speed measurements spanning one to three years, ensuring a sufficiently representative timeframe for our study. Moreover, we only considered data with an availability rate of at least 50%, guaranteeing that we work with substantial and reliable data sets.

II.3.1.2 Wind power production

For wind power observation, we collected hourly time series data from the Renewable Energy Data and Information Service (http://redis.energy.gov.za/electricity-production-details/). These datasets represent the wind energy produced at the regional level, corresponding to different provinces in South Africa, including the Eastern Cape, Western Cape, and Northern Cape. To ensure data quality and reliability, a careful manual filtering process was employed to remove any erroneous or inconsistent measurements. Fortunately, the number of identified anomalies was minimal, affirming the overall good quality of the collected data. The filtered time series data have an impressive availability rate of at least 73%, indicating a substantial and robust dataset for our analysis.

II.3.2 Solar observation data

II.3.2.1 Solar irradiance

The in-situ observation of surface solar irradiance downwards (RSDS) consists of two local networks: Trans-African Hydro-Meteorological Observatory (TAHMO) and Southern African Science Service Centre for Climate Change and Adaptive Land Management (SASSCAL) Weather Net. From these two networks, a total of 22 stations has been selected for data collection. To ensure data quality and consistency, the RSDS observations underwent a thorough data cleaning and filtering process. Several rules were applied to standardize the data and remove any erroneous or unreliable values. These rules included unifying the recorded time zone to Coordinated Universal Time (UTC) to maintain consistency across all stations. Additionally, implausible values were excluded from the dataset. To further enhance the reliability of the data, stations with missing data over 70% of the considered hourly data points were skipped. This decision was made to ensure that only stations with a significant amount of reliable data are considered for analysis.

II.3.2.2 Solar power

For solar power observation, hourly time series data were collected from the Renewable Energy Data and Information Service (http://redis.energy.gov.za/electricity-production-details/). These time series data represent the solar energy produced at the regional level, corresponding to two provinces in South Africa, including the Western Cape, and Northern Cape. To ensure data quality and reliability, a meticulous manual filtering process was employed to remove any erroneous and missing data respect to the simulation time series. The cleaned time series data demonstrates a commendable availability rate of approximately 60%.

II.3.3 Solar and wind farms

II.3.3.1 Wind farms

In this section, we detail the data acquisition process for the simulation of wind power generation in South Africa. We utilized turbine-specific information, including location, installed capacity, hub height, and rotor diameter, to conduct the analysis. Data on 39 wind parks in South Africa were sourced from the Renewable Energy Data and Information Service (REDIS) (http://redis.energy.gov.za/electricity-production-details/). To enhance the dataset, additional details such as rotor diameters, hub heights, and capacities were gathered from the Wind Power (https://www.thewindpower.net/store_country_en.php?id_zone=58). However, to address certain erroneous or missing data points, we conducted online searches and completed the database. The resulting comprehensive dataset of South African wind parks is available upon request. Generation data is provided by REDIS (http://redis.energy.gov.za/electricity-production-details/) as capacity factors. Table 1 gives an overview on the wind parks.

Table 1: Wind Parks data sets used in the simulation.

Provinces	***Eastern Cape***	***Northern Cape***	***Western Cape***	***Overall***
Location information on level of	Wind parks	Wind parks	Wind parks	Wind parks
Turbines	547	541	359	1447
Parks	16	10	13	39
Total capacity (GW)	1.435	1.2076	0.914	3.557
Avg. Park capacity (MW)	89.740	120.76	70.327	91.22

Avg. turbine capacity (MW)	2.62	2.232	2.546	2.446
Avg. rotor diameter (m)	107.92	107.1	97.958	105
Avg. hub height (m)	90	98.45	92.654	93.7

II.3.3.2 Solar farms

In the simulation of solar power generation, we gathered data from 44 solar farms, including their installed capacity and corresponding technology. The information was obtained from REDIS (http://redis.energy.gov.za/electricity-production-details/), where we discovered that the most commonly installed solar panel technology is crystalline. Table 2 gives an insight about the locality and the installed capacity.

Table 2: Solar Parks data sets used in the simulation.

Provinces	***Northern Cape***	***Western Cape***	***Overall***
Location information on level of	Solar parks	Solar parks	Solar parks
Parks	39	5	44
Total capacity (MW)	1550.7	133.82	1684.52
Avg. Park capacity (MW)	39.76	26.64	66.4

Figure 6 illustrates the spatial distribution of the solar and wind farms, which have been chosen for the purpose of simulation.

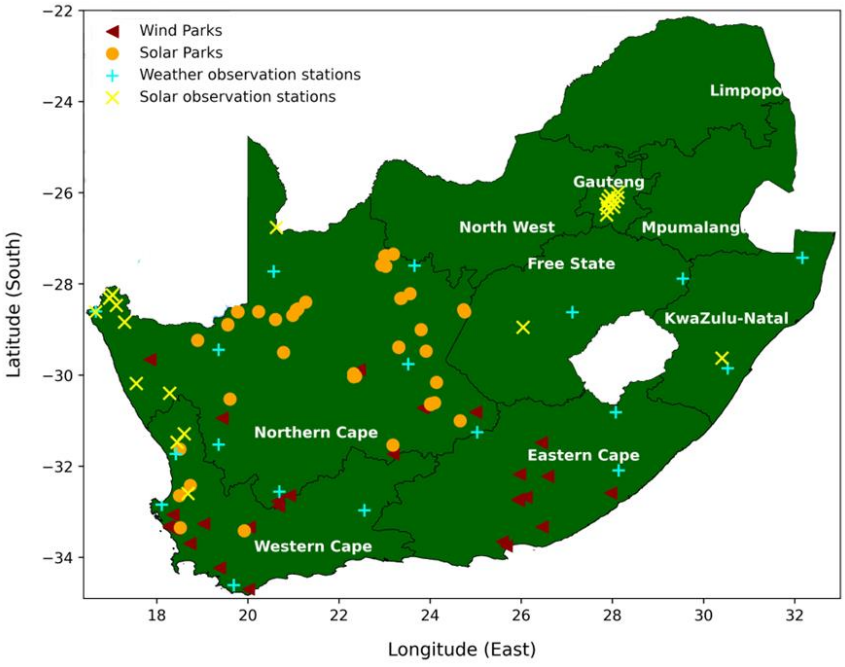


Figure 6: Spatial distribution of solar and wind parks over South Africa. Weather masts and solar observation stations are also indicated.

II.4 Methods

In this section, we outline the methodologies employed for extracting hourly wind-speed and solar irradiance time series from the datasets. The primary objective is to compare the observed and modelled wind speeds, as well as the observed and modelled solar irradiance, thereby assessing the accuracy of the models. Subsequently, we convert the wind speeds and solar radiation data into power series, enabling a comprehensive comparison between the observed and simulated power outputs.

II.4.1 Extraction of modelled data

II.4.1.1 Wind speed extraction

In a preparatory step, effective wind speeds were derived from the eastward (u) and northward (v) wind speed components in the reanalysis data using the Pythagorean theorem. This approach combines the U and V components to calculate the resultant wind speed at each grid cell. The effective wind speed represents the magnitude of the wind vector. Next, using the geographical location information of wind turbines or weather masts, the reanalysis and the ICON-LAM wind speeds were interpolated to the nearest neighbour. This interpolation process ensures that the wind speed values are assigned to the specific locations of the wind turbines or weather masts based on their latitude and longitude coordinates. Then an extrapolation to the hub height was made using:

1*) Linear interpolation*: It can be performed based on the distance proportion of the target height located between two closest heights. The general formula for linear interpolation can be used to interpolate wind speed values between known data points. The formula is given by (Pownuk & Kreinovich, 2017):

$$y = y_1 + \left|\frac{x - x_1}{x_1 - x_2}\right| (y_2 - y_1) \tag{8}$$

where:

x (m) is the input height for which we want to interpolate the wind speed

y is the estimated wind speed (m/s) at x level

y_1 and y_2 are the corresponding wind speeds (m/s) at x_1 and x_2

x_1 and x_2 are the known heights (m) that bracket the input height

2) *Logarithmic law*: Based on the Monin–Obukhov similarity theory, the logarithmic law is a physical model stating that the wind velocity at a determined height h_2 can be calculated by (Gualtieri, 2015; Tizgui et al., 2017):

$$v_2 = v_1 \frac{ln\left(\frac{h_2}{z_0}\right) - \psi_m\left(\frac{h_2}{L}\right)}{ln\left(\frac{h_1}{z_0}\right) - \psi_m\left(\frac{h_1}{L}\right)} \quad (9)$$

where:

v_2 (m/s) is the speed of the wind to be calculated at the height h_2 (m)

v_1 (m/s) is the speed of the wind measured at anemometer height h_1 (m)

ψ_m is the Monine Obukhov stability function.

L is the Monine Obukhov length (m).

z_0 is the roughness length (m).

When atmospheric stability is considered neutral (ψ_m=0), the logarithmic wind profile can thus be simplified. This simplified equation depends solely on the roughness length (z_0) (Gualtieri, 2015; Tizgui et al., 2017):

$$v_2 = v_1 \frac{ln\left(\frac{h_2}{z_0}\right)}{ln\left(\frac{h_1}{z_0}\right)} \quad (10)$$

3) *Power law:* The wind speed at the anemometer height or turbine hub height is determined by vertically interpolating between the wind speeds at the two closest levels (v_2 and v_1) using a power law fit. This computation is done based on equation (10) which enables the estimation of wind speed at height h_2 from the wind speeds measured at heights h_2 and h_1.

$$v_2 = v_1 \left(\frac{h_2}{h_1}\right)^{\alpha} \quad (11)$$

The power law equation (10) (Bailey et al., 1997) is commonly employed to extrapolate wind speeds from a single height, usually 10 meters above the ground. This method assumes a fixed value for the wind shear exponent known as the Hellmann (or friction) exponent, which allows estimating wind speeds at other heights. However, this fixed value may introduce inaccuracies since the wind shear varies with atmospheric stability. To address this issue and improve accuracy, the wind shear exponent is calculated dynamically at each time step by solving the power law for heights h_1 and h_2, allowing the computation of the wind shear exponent based on the specific wind speeds measured at those heights. By considering the dynamically

computed shear exponent, the extrapolation of wind speeds to different heights takes into account the varying atmospheric stability, leading to more precise results compared to using a fixed average value.

II.4.1.2 Solar irradiance extraction

For solar radiation assessment, ERA5 also serves as the primary data source. It provides higher spatial resolution of 0.25° and hourly temporal resolution, making it a valuable resource for solar energy analysis. The variable of interest is "surface solar radiation downwards" (SSRD), which represents Global Horizontal Irradiance (GHI) in units of J/m². To convert SSRD values into the standard unit of solar radiation (W/m²), a simple conversion factor of 1/3600 is applied, considering the hourly data. The ERA5 data, including the SSRD variable, is accessible through the Climate Data Store (CDS) (Yang & Bright, 2020).

II.4.2 Comparing modelled and observed data

At each weather mast location and weather station, the observed hourly wind speed series and solar radiation data are combined with the wind speed and solar radiation series extracted from the two datasets. To ensure data consistency, any time step where observation data are missing is removed from the comparison. These consolidated time series, encompassing both wind speed and solar radiation data, are then utilized to compute several important metrics. The results presented in chapter 3 highlight the most compelling findings, including, mean error (ME), coefficient of determination (R^2 or R-squared), root mean square error (RMSE), the correlation coefficient (r), the mean absolute error (MAE) and the Perkins skill score (PSS). These metrics offer valuable insights into the data analysis.

II.4.3 Converting wind speed and solar irradiance to power output

II.4.3.1 Wind speed conversion

To estimate wind power generation, the power curve model introduced by Ryberg et al. (2019b) was employed. This model employs an empirical approach to derive power curves, leveraging the concept of specific power, denoting the installed capacity per rotor swept area of wind turbines. Through an intricate analysis of variations in power output relative to specific power, this model delivers a refined and precise estimation of wind power for individual wind turbines. However, due to the lack of technical parameters for individual wind turbines at each location, a simplifying assumption was made. The analysis assumed that each wind farm utilizes only one type of wind turbine, irrespective of the number of turbines present, and the installed

capacity of the wind farm is evenly distributed among all the wind turbines. RESKit wind workflow driven by ERA5 and ICON-LAM was then utilized to calculate the generated electricity for each wind farm location, assuming it operates as if it has only one wind turbine installed. Consequently, the calculated capacity factor for each wind farm represents the performance of a single wind turbine. To obtain the overall wind power generation for the entire wind farm, upscaling is necessary. The process of upscaling depends on the reference data available, particularly the generation data collected from REDIS (http://redis.energy.gov.za/electricity-production-details/).

II.4.3.2 Solar irradiance conversion

The solar power simulation involves utilizing data from various solar parks across South Africa. The simulation encompasses three pivotal technologies: 'Photovoltaic Crystalline Fixed,' 'Photovoltaic Crystalline Tracking Single Axis,' and 'Photovoltaic Thin Film Fixed.' While detailed specifics regarding the on-site solar module were not available, the simulation aims for utmost accuracy by factoring in solar module industry viability and affordability. Ultimately, the choice gravitates towards **'JA_Solar_JAP6_72_335_4BB_RE'** for the crystalline silicon solar module, and **'NuvoSun_FL1132_430'** for the thin film solar module. These chosen modules' particulars are then integrated into the RESKit model solar workflow, facilitating the subsequent power simulation process.

II.4.4 Comparing simulated and observed power outputs

The power-output time series, representing the wind and solar power generation, are aggregated spatially over the provinces. This means that the individual power-output data from different wind farms and solar farms within each province are combined to create a single time series representing the total wind and solar power generation in terms of capacity factor for each corresponding province. Once the provincial wind and solar power generation time series are obtained, they are compared to the hourly capacity factor from the Renewable Energy Data Information Service data (REDIS). The provincial aggregation of the power-output time series is done because the REDIS database provides only provincial hourly time series data for electricity generation in term of capacity factor. Since the REDIS database does not provide specific wind and solar power generation data at a fine-grained level (such as individual wind and solar farm location), the analysis or comparison can only be done at the provincial level.

II.5 Analysis parameters

II.5.1 Metrics of error analysis

To enhance our comprehension of error patterns, we undertake a comprehensive evaluation by juxtaposing our observation data against simulation data from ERA5 and ICON-LAM.

II.5.1.1 The standard deviation

The standard deviation is a statistical measure that quantifies the amount of variation or dispersion in a set of values. The standard deviation is the square root of the average of the squared differences between each data point and the mean. It gives an idea of how much the data points are spread out around the mean. A higher standard deviation indicates greater variability, while a lower standard deviation indicates less variability. The standard deviation is calculated using the following formula (Norquist & Meeks, 2010):

$$\sigma = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(x_i - \bar{x})^2} \tag{12}$$

Where: N is the number of data points in the dataset, x_i represents each individual data point and $\bar{x}$ is the mean (average) of all data points.

II.5.1.2 Root Mean Square Error (RMSE)

Root Mean Squared Error (RMSE) provides insight into how well a predictive model's predictions match the observation data. RMSE provides an understanding of the average magnitude of errors between predicted and observed values. Larger RMSE values indicate that the model's predictions deviate more from the actual values on average.

$$RMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(y_i - x_i)^2} \tag{13}$$

Where:

N is the number of observations.

y_i represents the predicted value for observation i.

x_i represents the actual observed value for observation i.

II.5.1.3 Mean Error (ME)

Mean Error (ME) is a statistical metric that quantifies the average difference between the predicted values and the observed values in a dataset. It provides insight into the direction and magnitude of the errors in a predictive model or estimation method. The sign of the Mean Error indicates the overall direction of the errors. A positive ME implies that, on average, the predictions are overestimating the observed values, while a negative ME implies underestimation.

$$ME = \frac{1}{N}\sum_{i=1}^{N}(y_i - x_i) \tag{14}$$

II.5.1.4 Coefficient of determination (R^2)

The coefficient of determination, often denoted as R^2 or R-squared, is a statistical measure that provides insights into the proportion of the variance in the dependent (prediction) variable that is explained by the independent (observation) variables in a regression model. In other words, it quantifies the goodness of fit of the regression model to the observed data.

A higher R^2 indicates that a larger proportion of the variance in the dependent variable is explained by the regression model's independent variables. An R^2 of 0 implies that the model does not explain any variability, while an R^2 of 1 indicates a perfect fit.

$$R^2 = 1 - \frac{\sum_{i=1}^{N}(y_i - x_i)^2}{\sum_{i=1}^{N}(x_i - \bar{x_i})^2} \tag{15}$$

Where:

- N is the number of observations.
- y_i represents the predicted value for observation i.
- x_i represents the actual observed value for observation i.
- $\bar{x}_i$ is the mean of the observation.

II.5.1.5 Pearson correlation coefficient (r)

The Pearson correlation coefficient is a statistical measure of the strength of a linear relationship between two variables. It was first proposed by statistician Carl Pearson and is used to analyze and study the degree of linear correlation between variables. There are several different correlation coefficients for different research subjects, but the most commonly used

correlation coefficient is the Pearson correlation coefficient. The Pearson correlation coefficient ranges from -1 to 1, where -1 indicates a perfect negative correlation, 0 indicates no correlation, and 1 indicates a perfect positive correlation. The Pearson correlation coefficient can be calculated using statistical software such as Excel or a calculator, or by hand using the formula:

$$r_{xy} = \frac{\sum_{i-1}^{m}(x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i-1}^{m}(x_i - \bar{x})^2}\sqrt{\sum_{i-1}^{m}(y_i - \bar{y})^2}} \quad (16)$$

Where:

- r_{xy} is the degree of correlation between different variables.
- m is the number of data in the sequence.
- $\bar{x}$ represents the mean of the elements in vector 1.
- $\bar{y}$ represents the mean of the elements in vector 2.
- x_i represents individual data points of variable x
- y_i represents individual data points of variable y

II.5.1.6 Mean Absolute Error (MAE)

Mean Absolute Error is a metric used to measure the accuracy of a prediction model by calculating the average absolute differences between the predicted values and the actual observed values. A lower MAE indicates that the model's predictions are, on average, closer to the observed values, signifying better accuracy. The formula for calculating MAE is (Ramirez Camargo et al., 2019):

$$\mathrm{MAE} = \frac{1}{\mathrm{N}}\sum_{\mathrm{i}=1}^{\mathrm{N}}|y_i - x_i| \quad (17)$$

II.5.2 Weibull distribution function

The effectiveness of wind energy systems at a particular location and time hinges significantly on the distribution of wind speeds. To comprehensively evaluate wind energy potential, it is essential to have a clear understanding of the wind speed probability distribution. This distribution is pivotal in calculating the wind energy distribution associated with the location (Kantar & Usta, 2008). Among the various distributions used for wind data analysis, the Weibull distribution stands out as a potent choice. It provides a robust framework for analysing

wind data, offering valuable insights into wind speed probabilities. Despite its limitations in precisely capturing the probabilities of low wind speeds, the Weibull distribution remains a valuable tool in assessing wind energy resources and optimizing wind energy systems (Daoudi et al., 2022; Michael et al., 2021; Shu & Jesson, 2021). The Weibull distribution has two parameters: the shape parameter (k) and the scale parameter (c). The probability density function f(v) is used to determine the probability of a given wind speed v when analyzing the average wind speed of the data. The probability density function (PDF) of the Weibull distribution is given by (Albani & Ibrahim, 2017; Ay & Kisi, 2015):

$$f(v) = \left(\frac{k}{c}\right)^{k-1} exp\left(-\left(\frac{v}{c}\right)^{k}\right) \tag{18}$$

Where k and c symbolize the shape and scale parameters, respectively. The cumulative distribution is derived by integrating the Weibull probability density function, which is denoted as:

$$f(v) = 1 - exp\left(-\left(\frac{v}{c}\right)^{k}\right) \tag{19}$$

II.5.3 Perkins skill score (PSS)

This metric evaluates the shared region between two probability density functions (PDFs) by calculating the cumulative minimum value for each corresponding bin. In this way, it quantifies the extent to which the two distributions align. When a model replicates observed conditions accurately, the skill score (S_{score}) will attain a value of one. This value represents the total probability accumulated at the canter of each bin within a given PDF (Perkins et al., 2007). Mathematically, it is expressed as:

$$S_{score} = \sum_{1}^{n} minimum(Z_m - Z_o) \tag{20}$$

Where: n is the number of bins used to calculate the probability density function (PDF) for a specific region, Z_m represents the frequency of values in the i-th bin from the model, Z_o represents the frequency of values in the i-th bin from the observed data. If the model poorly simulates the observed PDF, the skill score approaches zero, indicating minimal overlap between the two distributions.

II.5.4 Capacity factor

The capacity factor is a crucial metric in assessing the efficiency and utilization of a power generation system. It represents the ratio of actual energy output from a power plant or system over a given period to the maximum possible energy output if the plant operated at its full capacity during the same period. In other words, it measures how much of the potential power generation capacity is actually realized (Boccard, 2009; Miller & Keith, 2018). Mathematically, the capacity factor (CF) is calculated using the formula:

$$CF = \frac{E_{actual}}{E_{max}} \times 100\% \quad (21)$$

Where:

- E_{actual} is the total energy generated by the power plant or system during the specified period.
- E_{max} is the theoretical maximum energy output that the system could generate if it operated at full capacity continuously during the same time period.

The capacity factor is usually expressed as a percentage.

II.5.5 Percentage error

A percentage error, also known as a percent error, is used to quantify the accuracy of a measured value compared to a true or observed value. It's often used in scientific and experimental contexts to assess the degree of deviation between the measured value and the expected value. Percent error is expressed as a percentage and indicates how far off the measured value is from the true value (De Myttenaere et al., 2016).

$$Percent\ error = \left|\frac{y - x}{x}\right| \times 100\% \quad (22)$$

Where:

- The absolute value $|\cdot|$ ensures that the result is positive, regardless of whether the measured value is greater or smaller than the true value.
- y is the value obtained through measurement or simulation.
- x is the known or observed value

Conclusion

The chapter provides an overview of the study area and describes the reanalysis dataset ERA5, the regional dynamical downscaling model ICON in Limited Area Mode (ICON-LAM), and the ResKit model. It also explains the wind and solar data acquisition process, including wind speed, wind power production, solar irradiance, and solar power, and provides information on the solar and wind farms. The discussion further explores the wind speed extraction, solar irradiance extraction, and the comparison between modeled and observed data, as well as the conversion of wind speed and solar irradiance to power. Additionally, the chapter develops several analysis parameters, such as standard deviation, root mean square error, mean error, coefficient of determination, Pearson correlation coefficient, mean absolute error, Weibull distribution function, Perkins skill score, percentage of error, and capacity factor.

CHAPTER III: RESULTS AND ANALYSIS

CHAPTER III: RESULTS AND ANALYSIS

Introduction

In this section, we embark on the crucial phases of our study, beginning with simulation validation. We delve into an analysis of wind speeds at both 10 meters and 60 meters above ground level, followed by an examination of wind power. Additionally, we turn our attention to the analysis of solar radiation and solar power.

III.1 Simulation validation

The validation dataset encompasses a span of three years, from 2017 to 2019, and comprises observed data related to wind speed, solar radiation, wind power, and solar power. The data points were aggregated on an hourly basis. In the context of wind speed, measurements were taken at both 10 and 60 meters above ground level. These measurements were collected from 19 weather masts strategically positioned across diverse locations within South Africa. For solar radiation, the dataset drew upon information from 22 stations. Among these, 11 were affiliated with Southern African Science Service Centre for Climate Change and Adaptive Land Management (SASSCAL) and the remaining 11 were part of the Trans-African Hydro-Meteorological Observatory (TAHMO). In the case of wind power, data were sourced from three distinct provinces: Northern Cape, Western Cape, and Eastern Cape. For solar power, the dataset exclusively covered two provinces: Northern Cape and Eastern Cape. Validation of power output relied on the utilization of capacity factors. These factors play a crucial role in gauging the performance of power generation systems. The evaluation of the simulated time series involved a comprehensive assessment of their quality through the application of six statistical parameters including Pearson correlation (r), root mean square error (RMSE), mean error (ME), mean absolute error (MAE), coefficient of determination (R-squared or R^2) and Perkins skill score (PSS). In the case of wind power losses attributed to wake effects, downtimes resulting from maintenance activities, interruptions caused by avian or bat interactions, and instances of curtailment were not factored into the analysis.

III.2 Results and Analysis

III.2.1 Wind speeds

III.2.1.1 Wind speed at 10 m

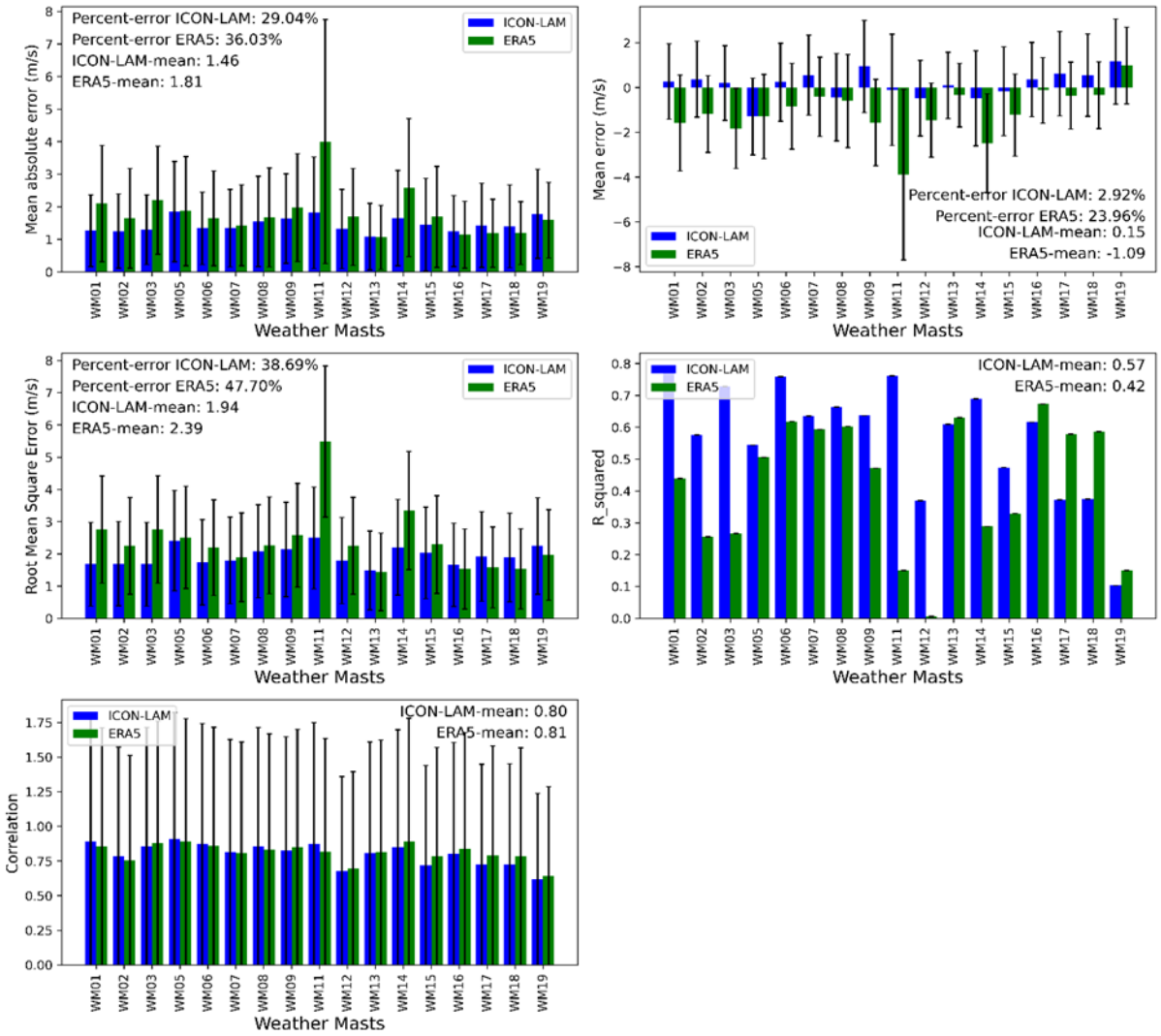


Figure 7: Comparison of statistical metrics of observed and simulated wind speeds based on ERA5 and ICON-LAM (Colour bars represent metrics values and solid black lines represent standard deviation).

In this section, the study explores the variation in simulation quality resulting from the utilization of distinct datasets, specifically ICON-LAM and ERA5. The assessment is depicted in

Figure **7** which offers a comprehensive comparison of statistical metrics across various locations. In term of mean absolute error (MAE), the percent-error for ERA5 is 36.0%, indicating that the simulation, on average, deviates from observed values by this percentage. Conversely, ICON-LAM demonstrates a lower percent-error of 29.0%, suggesting a comparatively better agreement with observed data. ICON-LAM achieves a mean MAE of 1.46 m/s, while ERA5 records a slightly higher mean MAE of 1.81 m/s. This indicates that ICON-

LAM's simulations are closer to the observed values in terms of mean absolute error. ERA5 exhibits a percent-error of 30.0% for ME, signifying its average deviation from the observed values. ICON-LAM, on the other hand, displays a notably lower percent-error of 2.9%, indicating a better alignment with the observed data in terms of mean error. The mean ME values further confirm this trend, with ICON-LAM recording a mean ME of 0.15 m/s and ERA5 registering a mean ME of -1.09 m/s. ICON-LAM's ME values are much closer to zero, implying a smaller average discrepancy. ERA5's percent-error for RMSE is 47.7%, signifying its average deviation from observed values. ICON-LAM displays a comparatively lower percent-error of 38.7%, implying a relatively better agreement with observed data in terms of the root mean square error. The mean RMSE values are 1.94 m/s for ICON-LAM and 2.39 m/s for ERA5, indicating that ICON-LAM's simulations have a slightly smaller overall error magnitude. Both ICON-LAM and ERA5 demonstrate similar Pearson correlation values, with ICON-LAM at 0.8 and ERA5 at 0.81. This suggests a strong linear relationship between the simulated and observed values in both cases. However, there is a small difference in R-squared values, with ICON-LAM's mean at 0.57 and ERA5's mean at 0.42. This indicates that ICON-LAM's simulations explain a larger proportion of the observed variability compared to ERA5. In summary, while both ICON-LAM and ERA5 exhibit relatively strong correlations with the observed data, ICON-LAM consistently demonstrates better performance in terms of MAE, ME, and RMSE metrics. This suggests that ICON-LAM's simulations are generally closer to the observed values and have a smaller overall error magnitude compared to ERA5. Additionally, ICON-LAM's simulations seem to capture a larger proportion of the observed variability as indicated by the higher R-squared values.

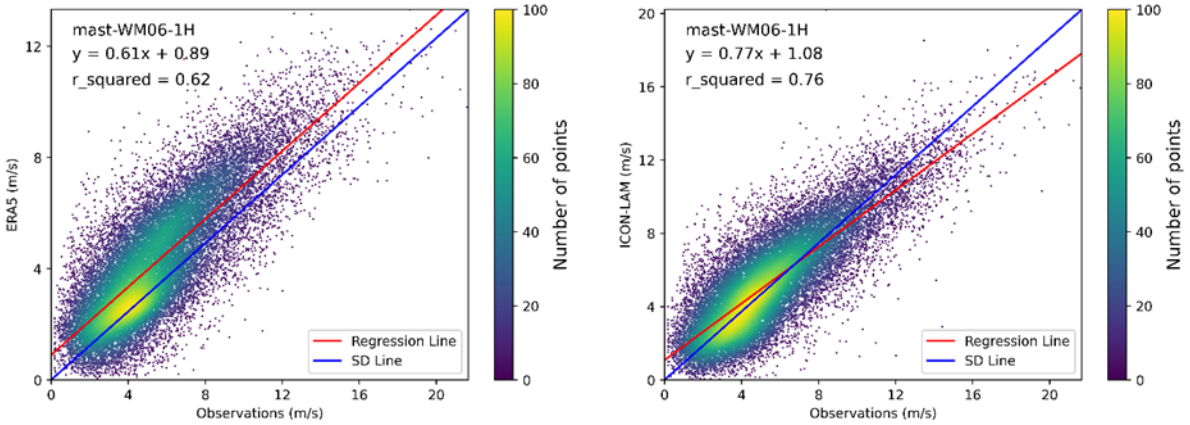


Figure 8: Comparison between observed wind speeds and simulated wind speeds (left: ERA5, right: ICON-LAM) at one location (weather mast 06) over South Africa.

The scatter plot, Figure 8 serves as a visual representation of the comparative performance between the two distinct datasets. Weather mast 6 has been selected due to its ability to effectively represent both the simulated and observed data. By plotting the observed values against the simulated ones, the scatter plot enables us to discern patterns, trends, and deviations between the datasets. The higher R-squared value of 0.76 for ICON-LAM indicates that its simulations account for a larger share of the observed variability compared to ERA5's R-squared of 0.62.

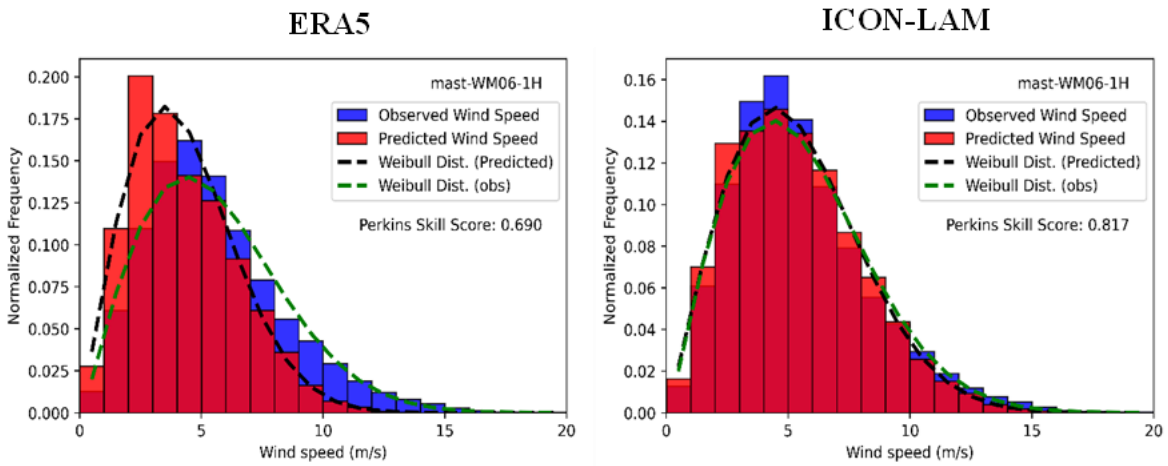


Figure 9: Comparison of probability density function (PDF) of observed and simulated wind speeds based on ERA5 (left) and ICON-LAM (right) (Obs stands for observation). The fitted PDF´s according the Weibull distribution are also indicated.

In Figure 9 the PDF plots for both ERA5 and ICON-LAM give insights into the wind speed distribution. For ERA5, the peak of the observed data's PDF is for wind speeds between 4 and 5 m/s while for the simulated data it is between 2 and 3 m/s. The observation peak is shifted to the right compared to the simulation peak, this implies that, on average, the observed wind speeds tend to be higher than the simulated wind speeds. The Perkins skill score of 0.69 indicates a relatively moderate agreement between the ERA5 simulations and observed data. For ICON-LAM, the peak of the simulation data is between 4 to 5 m/s, which is closer to the observed wind speeds. The higher Perkins skill score of 0.817 for ICON-LAM compared to ERA5 suggests also a stronger agreement between the ICON-LAM simulations and observed data. This indicates that ICON-LAM's simulations perform better in capturing the observed wind speed distribution.

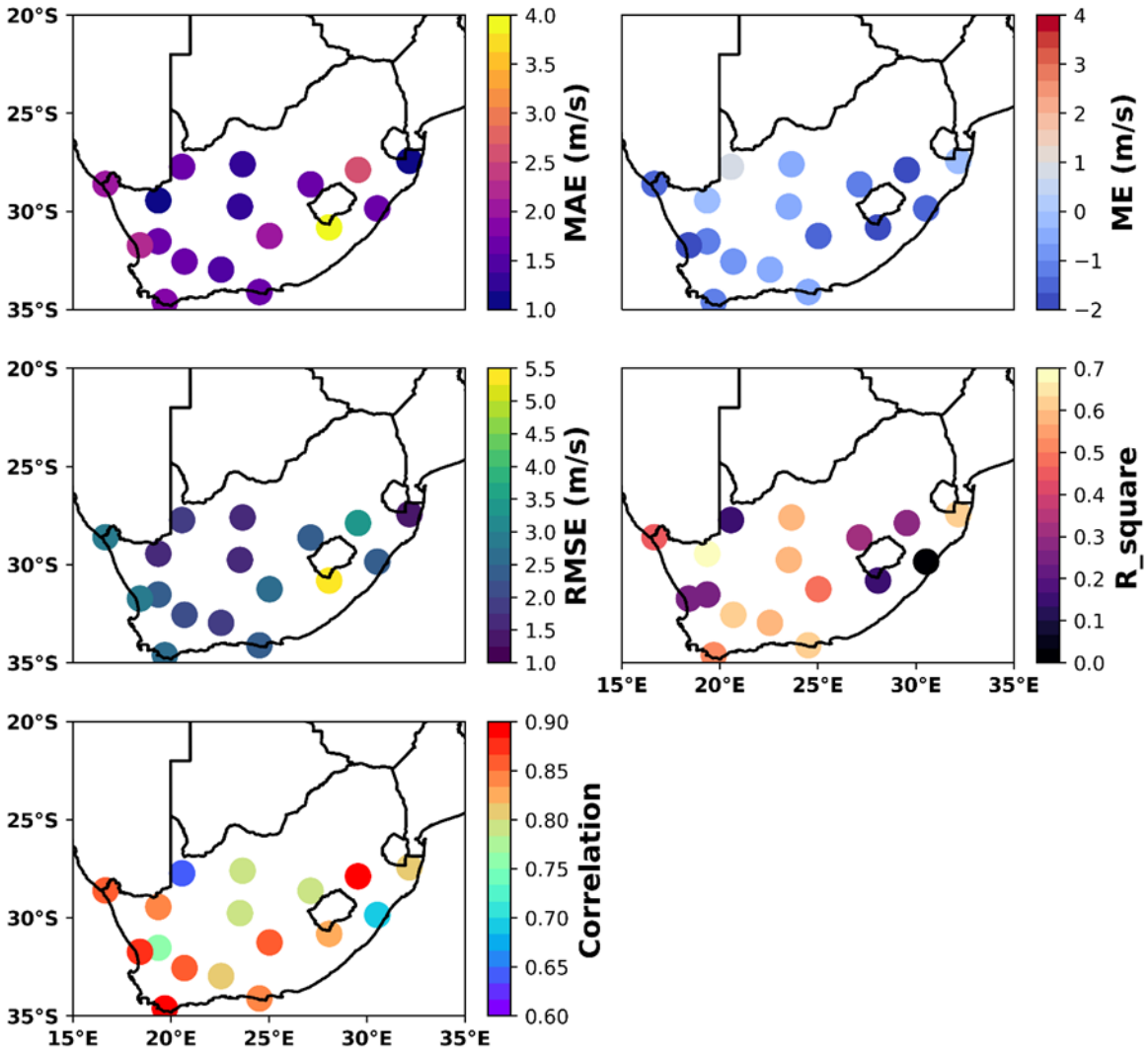


Figure 10: Spatial distribution of statistical metrics for simulated wind speed by ERA5 (compared to measured data) at 10 m.

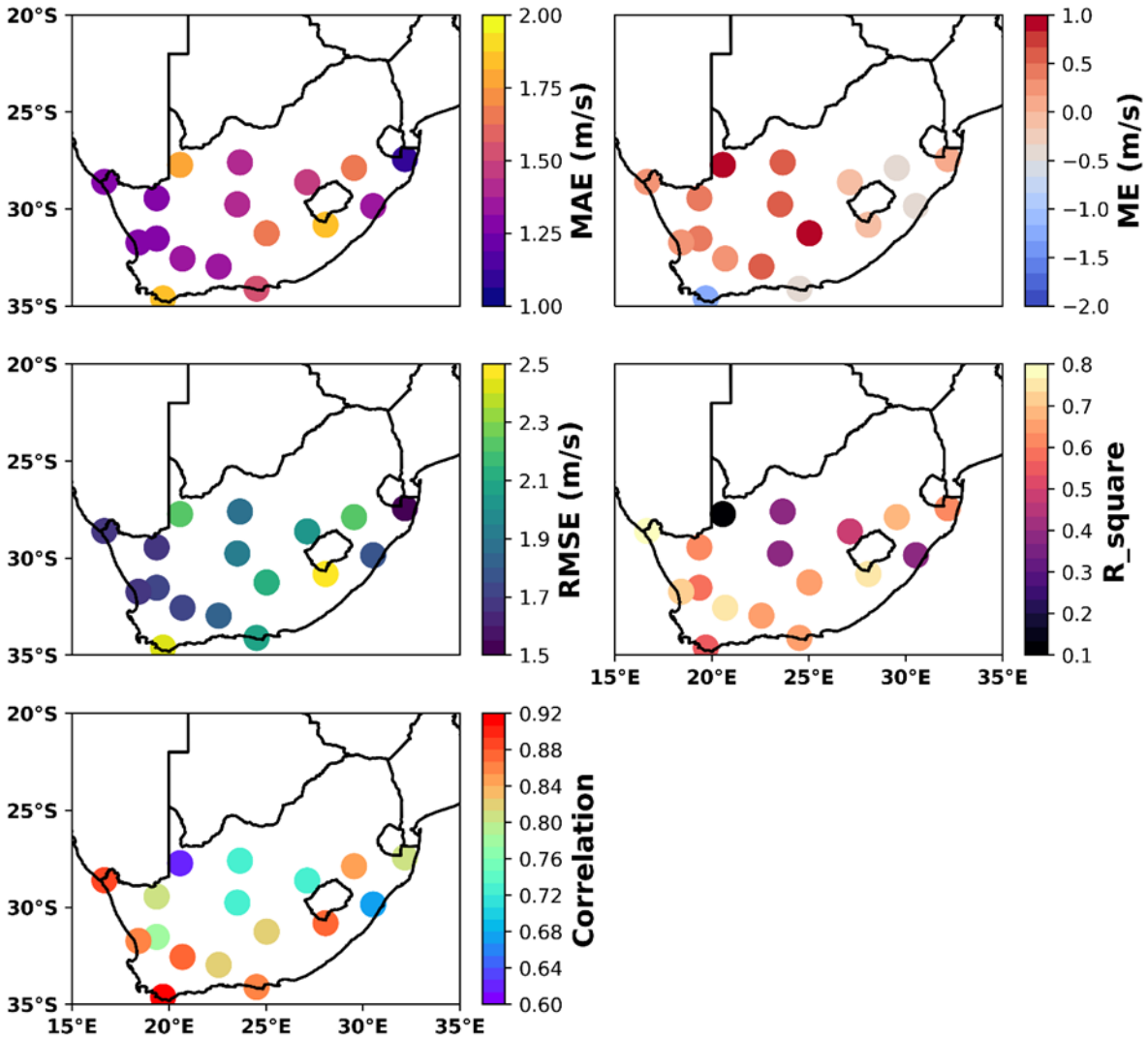


Figure 11: Spatial distribution of statistical metrics for simulated wind speed by ICON-LAM (compared to measured data) at 10 m.

In addition to examining the statistical metrics and probability density functions (PDFs) for ERA5 and ICON-LAM, a spatial dimension further enriches our understanding of the performance of these datasets. Both ERA5 and ICON-LAM show a similar spatial distribution pattern in terms of MAE, with certain locations having relatively high MAE values, indicating larger errors in wind speed simulations. In terms of ME, there is a mix of overestimation and underestimation across different locations for both datasets.

Figure 10 gives insight to the spatial distribution of ERA5 while Figure 11 states the one of ICON-LAM

III.2.1.2 Wind speed at 60 m

Table 3: Statistical metrics for comparing simulated wind speeds (ICON-LAM and ERA5) with observed wind speeds at 60m height, for different wind speed interpolation methods.

	MAE- error (%)			***ME-error (%)***			***RMSE-error (%)***		
Interpolation methods	**ln**	**Pl**	**log**	**ln**	**pl**	**log**	**ln**	**pl**	**log**
ICON-LAM	30.7	27.9	28.0	8.1	7.6	8.3	41.4	37.6	37.7
ERA5	34.3	33.0	33.2	20.4	17.6	18.1	46.2	44.4	44.7

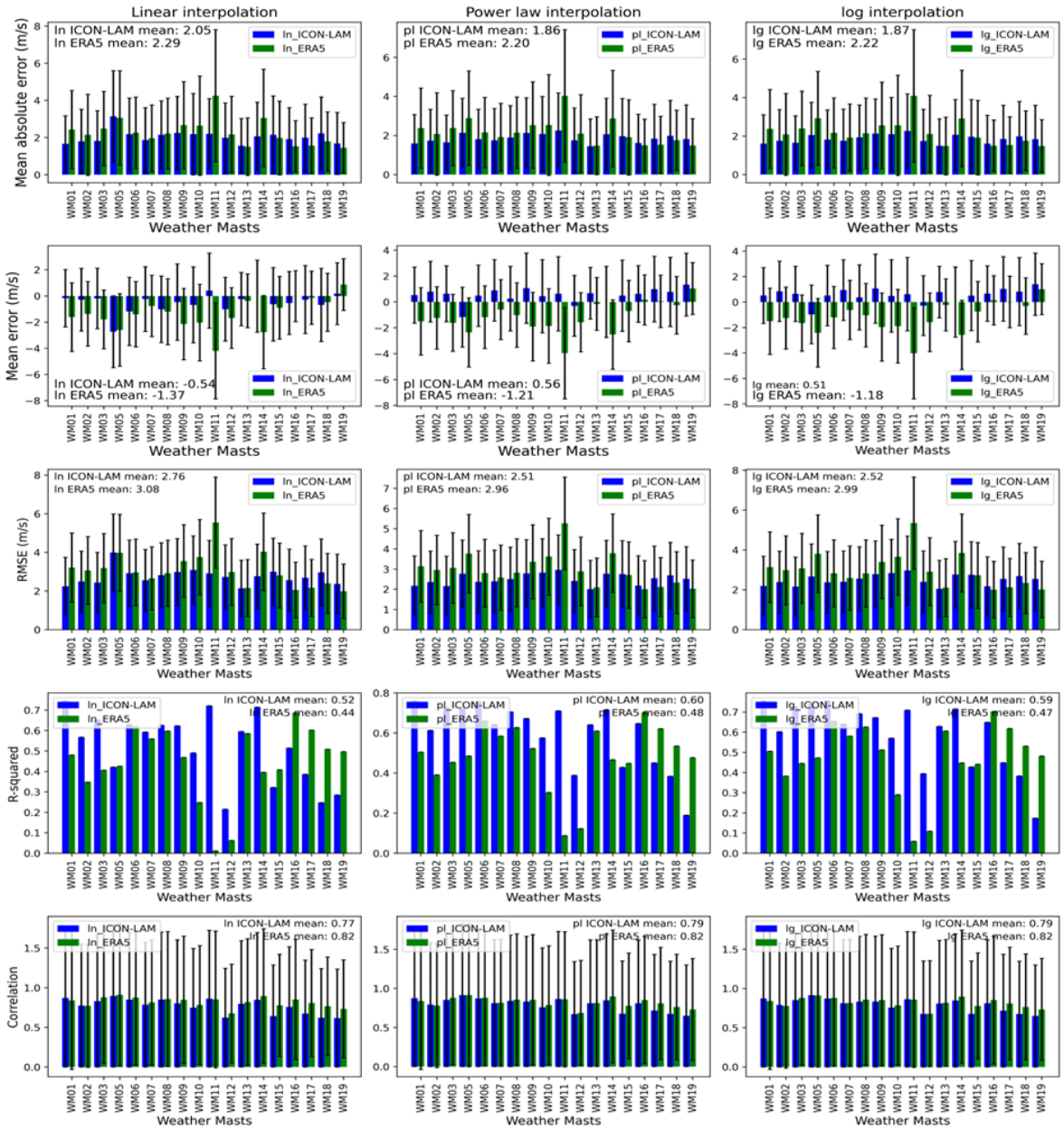


Figure 12: Comparison of statistical metrics of observed and simulated (ERA5, ICON-LAM) wind speeds for different interpolation methods (Colour bars represent metrics values and solid black lines represent standard deviation)

As outlined in the wind extraction section, an interpolation has been employed for scaling the wind speed from the nearest height to 60 meters. The interpolation methods employed are power law (pl), linear interpolation (ln), and logarithmic interpolation (log). Subsequently, an insightful analysis of statistical metrics is conducted as shown in Figure 12 to discern the performance disparities between the two datasets. For ERA5, the mean MAE is around 2.2-2.3 m/s for all interpolation methods (pl, ln, log). In the case of ICON-LAM, the mean MAE is notably lower, with values between 1.86 m/s (pl) and 2.05 m/s (ln). This comparison implies that ICON-LAM consistently achieves lower mean MAE values across all interpolation methods. The ME for ERA5 fluctuates across interpolation methods, ranging from around -1.18 m/s (log) to -1.37 m/s (ln). ICON-LAM, on the other hand, has ME values varying from -0.54 m/s (ln) to 0.56 m/s (pl). ICON-LAM's performance suggests better alignment between simulations and observations, as indicated by the mean ME values closer to zero. For ERA5, the RMSE spans from approximately 2.96 m/s (pl) to 3.08 m/s (ln). ICON-LAM consistently attains lower RMSE values, ranging from 2.51 m/s (pl) to 2.76 m/s (ln). ICON-LAM's simulation accuracy is underscored by its consistently lower mean RMSE values across interpolation methods. The R-squared for ERA5 ranges from around 0.44 (ln) to 0.48 (pl). In contrast, ICON-LAM showcases improved performance, with R-squared values spanning from 0.52 (ln) to 0.6 (pl). ICON-LAM's higher R-squared values indicate that its simulations provide a better fit to the observed wind speed data. ERA5 maintains a stable mean correlation of approximately 0.82 across all interpolation methods (pl, ln, log). ICON-LAM also demonstrates consistent performance, with mean correlation values ranging from 0.77 (ln) to 0.79 (pl). Both datasets exhibit commendable consistency in terms of mean correlation, highlighting their strong linear relationship with observed data. In summary, the statistical metrics' analysis underscores ICON-LAM's superiority in delivering more accurate and closely aligned wind speed simulations, as evidenced by its consistently lower MAE, ME, and RMSE values. Moreover, ICON-LAM's higher mean R-squared values and comparable correlation values further reinforce its enhanced simulation quality and alignment with observed data compared to ERA5. Table 3 highlights the percentage of error between the two datasets. The percentage of error in MAE is lower for ICON-LAM across all interpolation methods compared to ERA5. Among the interpolation methods, the lowest MAE percentage of error is observed for the power law (pl) interpolation method in both ICON-LAM and ERA5. The percentage of error in ME is significantly lower for ICON-LAM compared to ERA5 across all interpolation methods. Similar to the MAE analysis, the power law (pl) interpolation method yields the lowest ME percentage of error for both ICON-LAM and ERA5. The percentage of error in

RMSE is consistently lower for ICON-LAM than for ERA5 across all interpolation methods. Just as in the previous analyses, the power law (pl) interpolation method exhibits the lowest RMSE percentage of error for both ICON-LAM and ERA5. The error analysis clearly demonstrates that ICON-LAM consistently outperforms ERA5 across all interpolation methods in terms of MAE, ME, and RMSE, as evidenced by the lower percentage of error values. Among the interpolation methods, the power law (pl) method consistently yields the smallest percentage of error for both ICON-LAM and ERA5, indicating its effectiveness in minimizing errors in wind speed simulations.

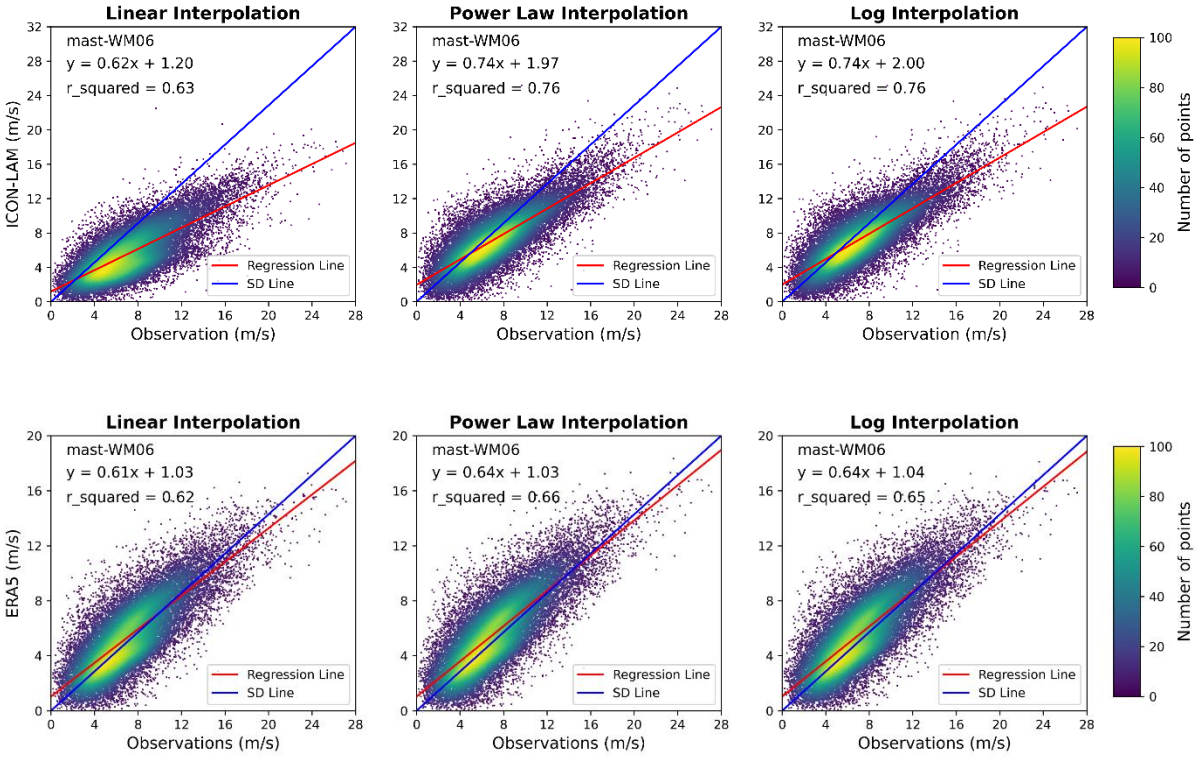


Figure 13: Comparison of observed wind speeds and simulated wind speeds (bottom: ERA5 and top: ICON-LAM) for different interpolation methods including R^2 values.

The scatter plot analysis in Figure 13 on one specific location reveals the comparative performance of wind speed simulations between ERA5 and ICON-LAM using different interpolation methods. For ERA5, the R-squared values are 0.62 (ln), 0.66 (pl), and 0.65 (log), signifying a moderate linear relationship between simulated and observed wind speeds. In contrast, ICON-LAM demonstrates consistently higher R-squared values of 0.63 (ln), 0.76 (pl), and 0.65 (log), highlighting its stronger predictive power. Particularly, the power law (pl) interpolation stands out, showcasing a high R-squared value of 0.76 for ICON-LAM as mentioned above.

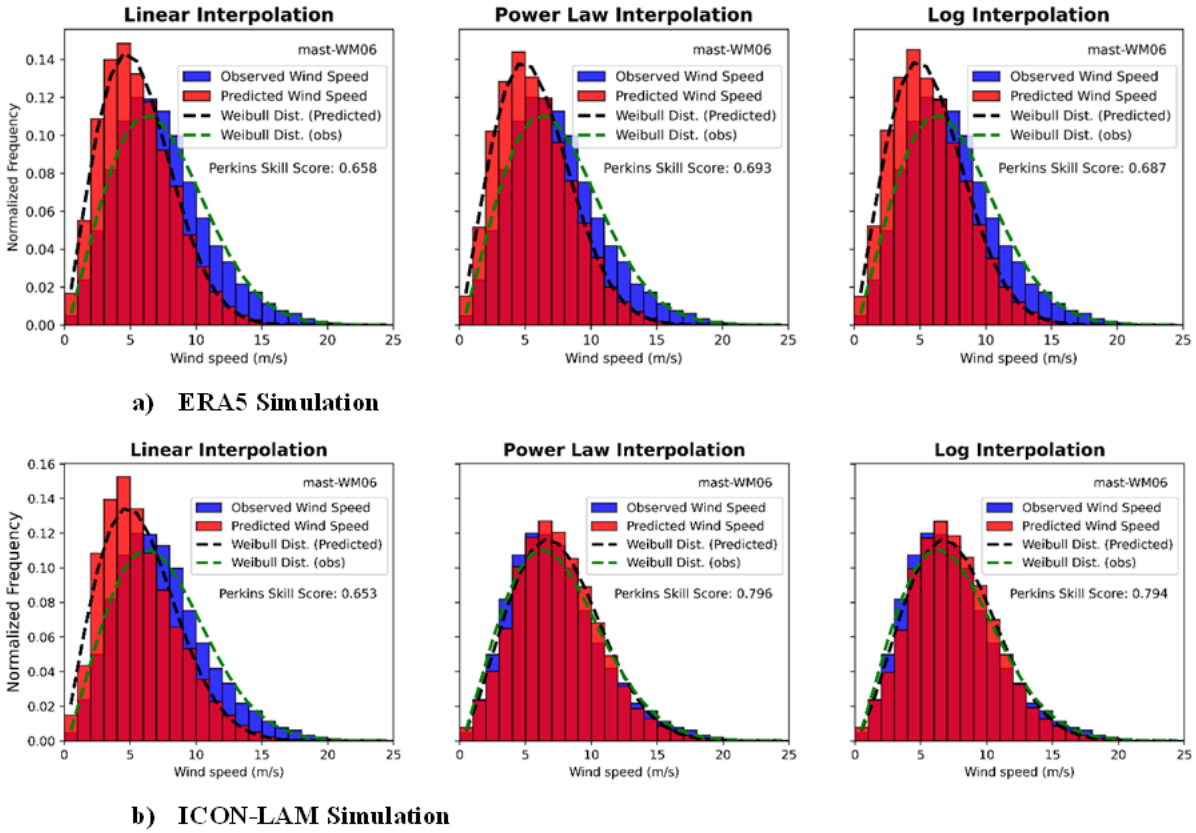


Figure 14: Comparison of probability density function (PDF) of observed and simulation wind speeds based on ERA5(top) and ICON-LAM (bottom) according to different interpolation methods.

In Figure 14 the comparison of probability density function (PDF) plots for wind speed distributions is shown. The PDF plot for ERA5 indicates that the peak of observed wind speeds is for a wind speed range of 5 to 6 m/s while for simulated wind speeds the maximum frequency is for wind speeds ranging from 4 to 5 m/s. A noteworthy observation is that the observed wind speeds are shifted to the right compared to the simulated values. This indicates that ERA5 tends to slightly underestimate wind speeds in this range. The Perkins skill scores for ERA5 simulations exhibit a moderate level of agreement with observations: linear interpolation yields a score of 0.658, power law interpolation achieves 0.693, and logarithmic interpolation scores 0.687. For ICON-LAM, the PDF peak for linear interpolation (ln) is found for wind speeds ranging from 4 to 5 m/s. Interestingly, both power law (pl) and logarithmic (log) interpolations in ICON-LAM result in a PDF peak for slightly higher wind speeds between 6 to 7 m/s. Comparing the Perkins skill scores for ICON-LAM, it is observed that power law (pl) interpolation achieves the highest score of 0.796, followed closely by logarithmic (log) interpolation at 0.794, and linear interpolation (ln) with a score of 0.653.

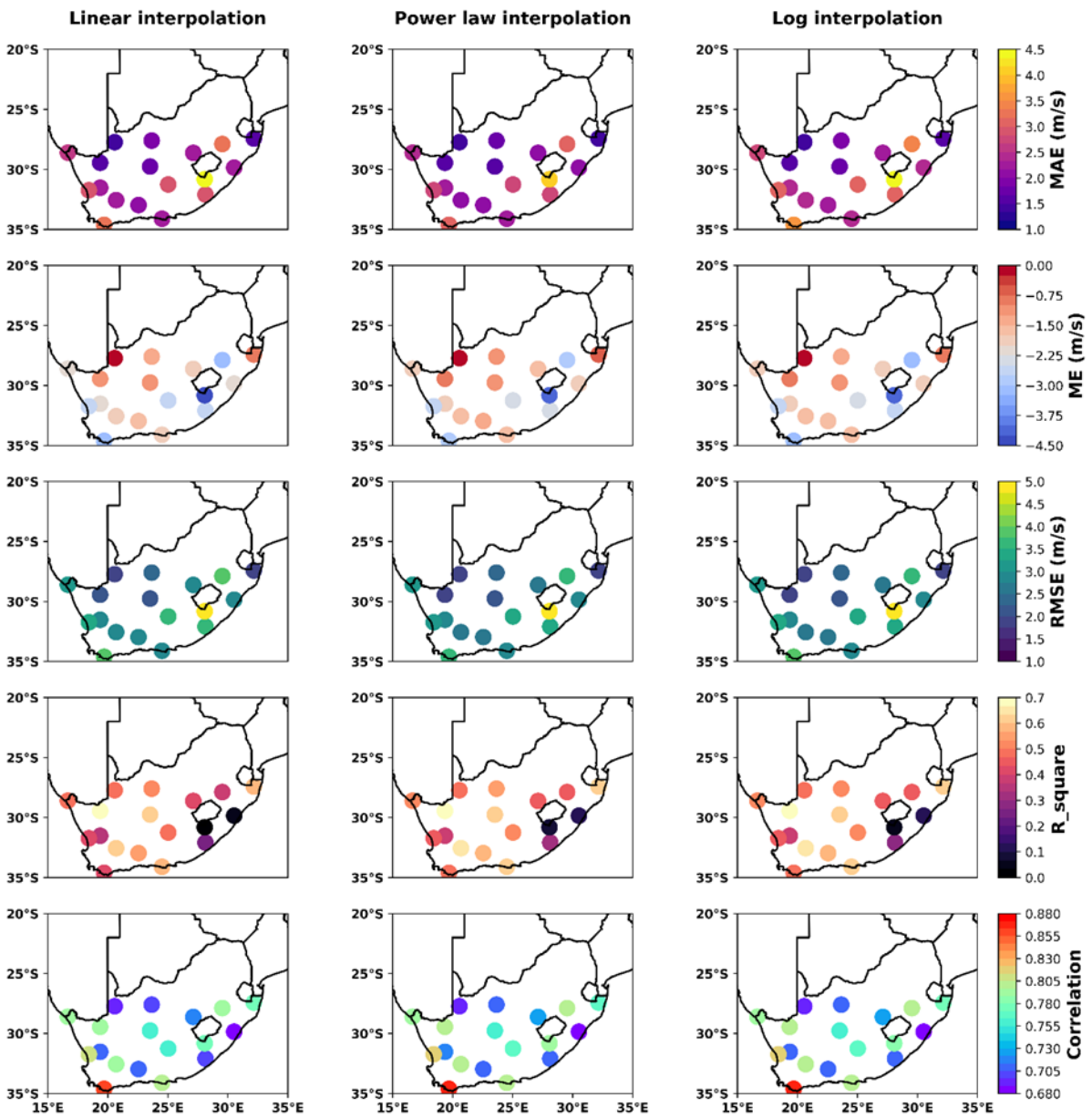


Figure 15: Spatial distribution of statistical metrics for different interpolation methods of ERA5 simulated wind speed (compared to measured data) at 60m.

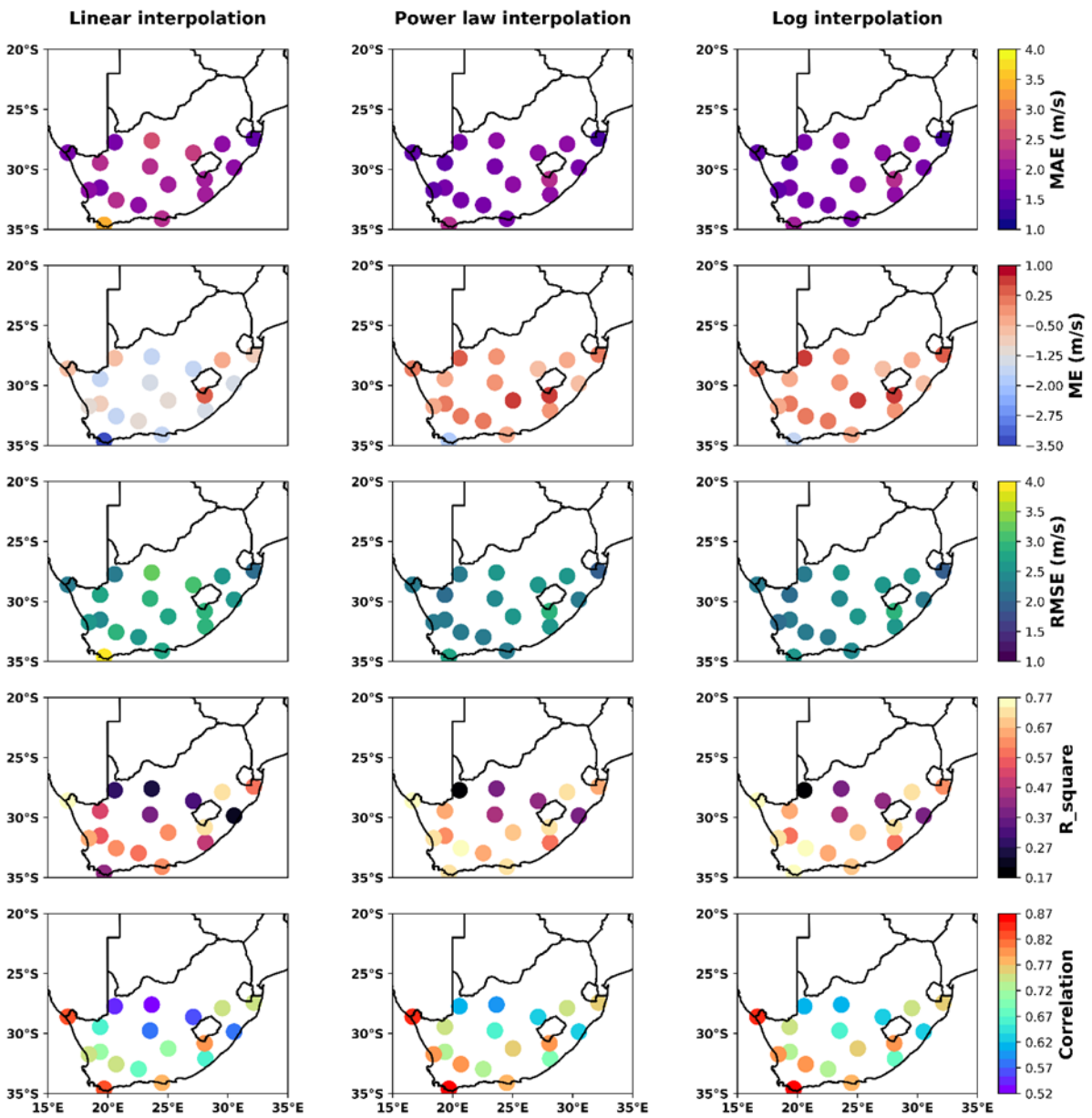


Figure 16: Spatial distribution of statistical metrics for different interpolation methods of ICON-LAM simulated wind speed (compared to measured data) at 60m.

Spatial distribution plots, Figure 15 for ERA5 and Figure 16 for ICON-LAM offer a better insight into the variation and performance of different metrics across geographical regions in relation to various wind speed interpolation methods. This plot provides a dynamic representation of how wind speed simulations diverge or converge from observed values across different locations.

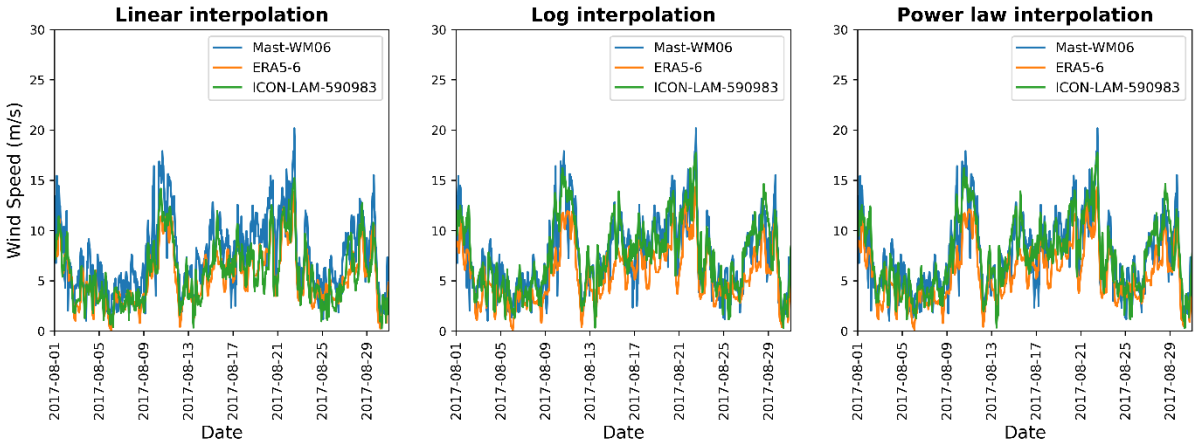


Figure 17: Wind speeds comparison of in-situ-observation (blue line) against simulations (Orange line for ERA5 and green line for ICON-LAM), ERA5-6 and ICON-LAM-590983 are identifiers used to represent the nearest ERA5 and ICON-LAM cells, respectively. The number 6 after ERA5 indicates the nearest ERA5 cell identity, while the number 590983 after ICON-LAM represents the nearest ICON-LAM cell identity.

Following an in-depth analysis of various metrics, a pivotal emerges, facilitating a direct comparison of wind speeds with regard to different interpolation methods. This comparative visualization distinctly illustrates the superiority of ICON-LAM over ERA5 within this specific one-month sample period. By juxtaposing wind speed simulations across interpolation techniques, Figure 17 shows ICON-LAM's enhanced performance. This visual representation effectively highlights the pronounced advantages of ICON-LAM's accuracy and precision in capturing wind speed patterns compared to ERA5 during this particular time frame.

The observed enhancement in the performance of ICON-LAM compared to ERA5 in wind speed modelling may be due to the utilization of a more refined spatial model resolution within the modelling physics and dynamics (Giorgi & Gutowski, 2015; Laprise, 2008; Pham et al., 2021). ICON-LAM's higher level of spatial granularity may allow for better capture of variations in wind speeds, as previous studies have suggested. In contrast, ERA5 has a resolution of about 31 km worldwide (Setchell, 2017), which may not be sufficient to capture fine-scale atmospheric processes. However, this limitation is offset by its exceptional feature of incorporating observed data through data assimilation. This process enhances ERA5's accuracy and comprehensiveness, especially in regions where measurements are available. The integration of observed data allows ERA5 to construct a consistent and coherent record of historical weather patterns and atmospheric conditions, thus contributing to an enriched understanding of atmospheric phenomena. In their exploration of ICON verification in Limited Area Mode, Rieger et al., (2021) discovered a mean error (ME) of approximately 0.3 m/s,

alongside a root mean square error (RMSE) of 1.25 m/s, both attributed to a 10 m wind speed analysis conducted in Poland. Contrasting these findings, our investigation unveiled an average ME of 0.15 m/s and an average RMSE of 1.94 m/s. The slight variance in these values, while not significantly disparate, can be confidently attributed to distinct atmospheric conditions and disparate reference datasets. This disparity further substantiates the efficacy of ICON-LAM.

III.2.2 Wind power

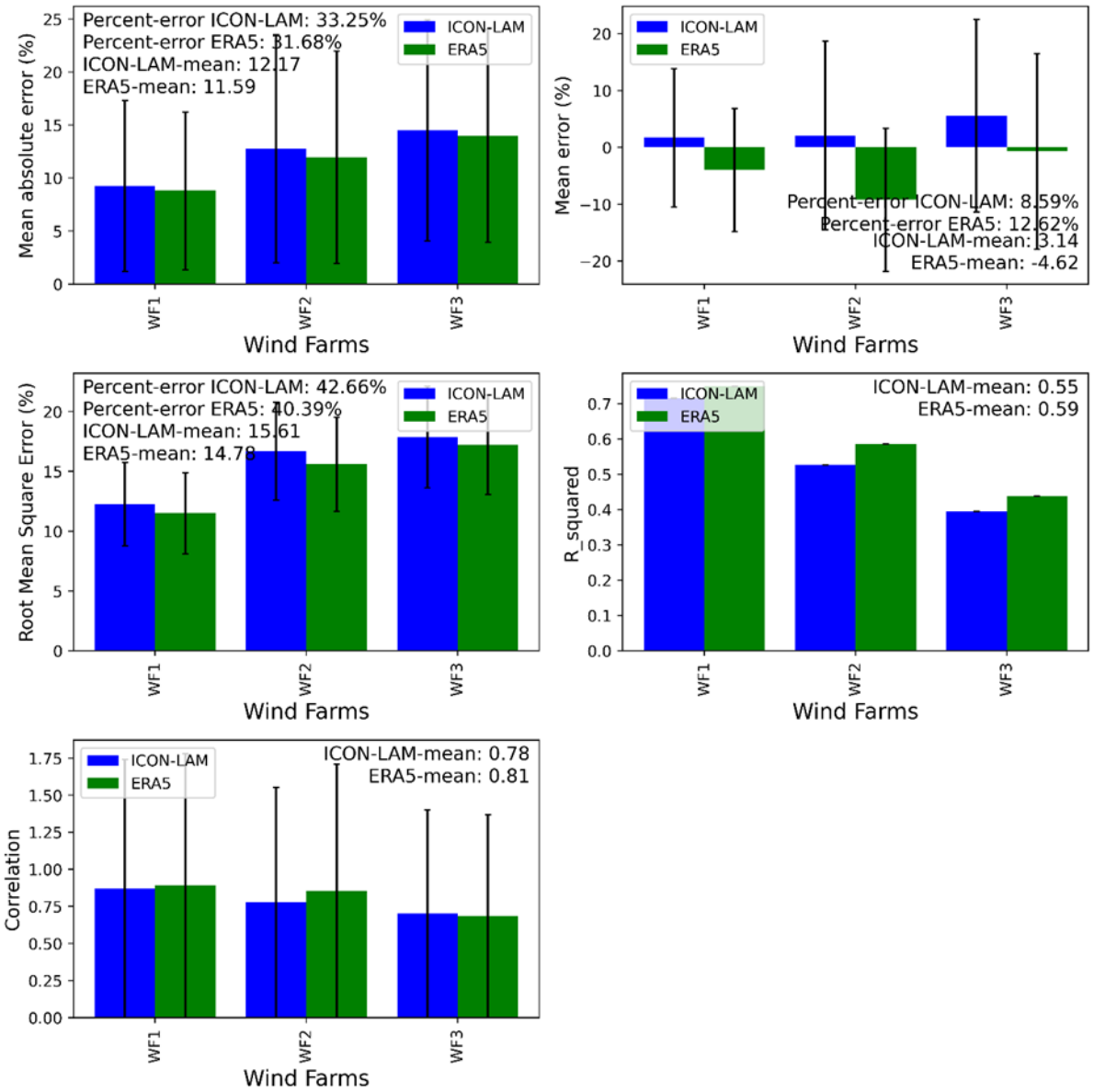


Figure 18: Comparison of statistical metrics of observed and simulated wind power based on ERA5 and ICON-LAM driven RESKit; WF1, WF2 and WF3 stands respectively for Eastern Cape, Northern Cape and Western Cape wind farm (Colour bars represent metrics values and solid black lines represent standard deviation).

The comparison between hourly observed and simulated wind power in terms of capacity factor is conducted at the regional level and shown in the Figure 18. The percent-error for ERA5 in

term of MAE is 31.25%, while for ICON-LAM it is slightly higher at 33.68%. For ME the percent-error for ERA5 is 12.62% with a ME of -4.62 %, signifying an underestimation of wind power by the model. In contrast, ICON-LAM demonstrates a percent-error of 8.59%, indicating better accuracy in capturing wind power levels. Both ERA5 and ICON-LAM exhibit relatively similar performance in terms of RMSE, with a percent-error of 40.39% for ERA5 and 42.66% for ICON-LAM. The Pearson correlation coefficients indicate a moderate linear relationship between observed and simulated wind power. ICON-LAM achieves an average correlation of 0.78, while ERA5 performs slightly better with an average correlation of 0.81. Both models show relatively similar R-squared values, suggesting that they explain a comparable proportion of the variability in the observed wind power data. ICON-LAM has an average R-squared value of 0.55, while ERA5's average R-squared value is 0.59. Overall, the analysis demonstrates that both ERA5 and ICON-LAM, driven by RESKit, provide reasonably accurate simulations of wind power. While ERA5 shows slightly better performance in terms of RMSE, MAE and correlation, ICON-LAM outperforms ERA5 in terms of ME.

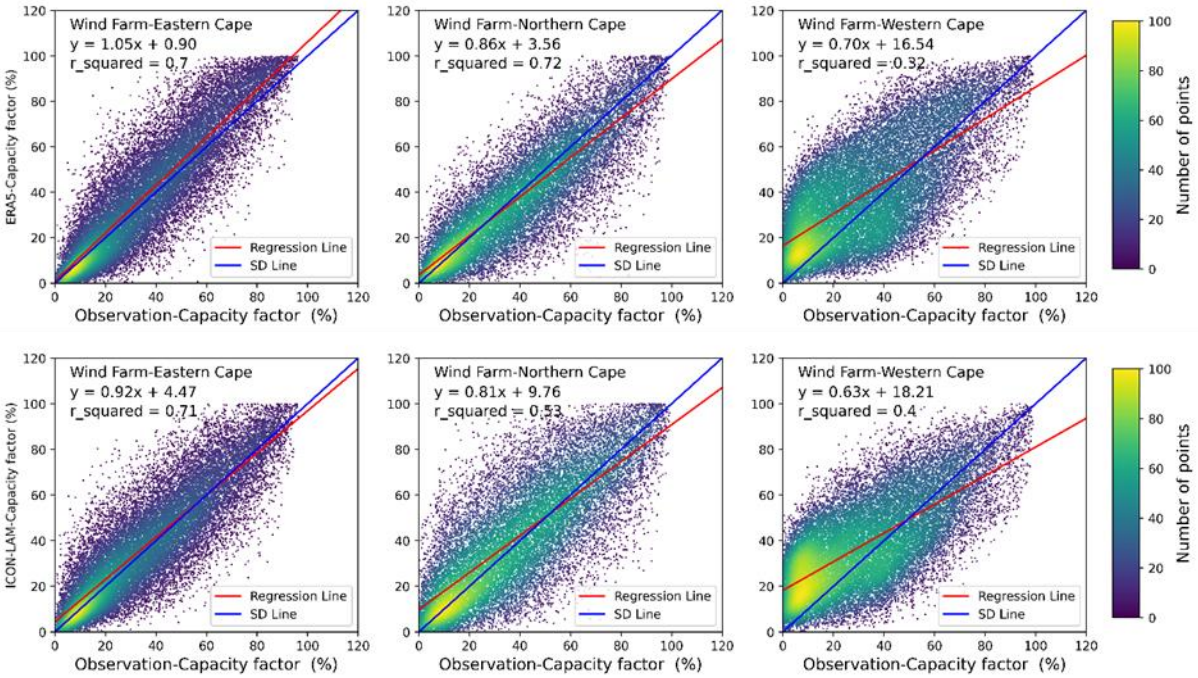
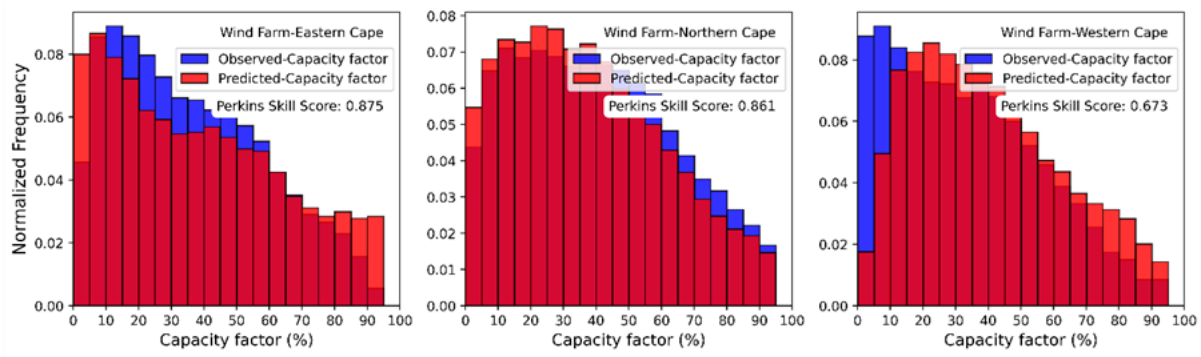


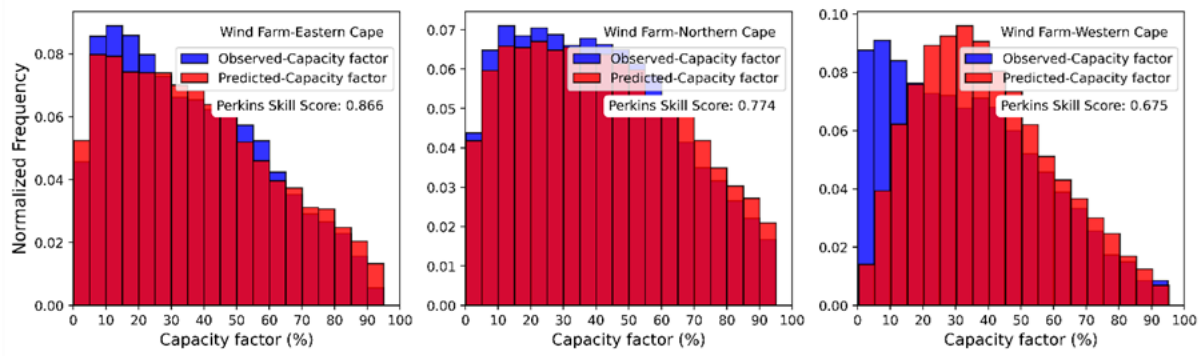
Figure 19: Wind power comparison based on ERA5(top) and ICON-LAM (bottom) driven RESKit simulations with in-situ wind power observations from three provinces.

The scatter plot visualization in Figure 19 presents an insightful comparison. For ERA5, in the Eastern Cape province, the scatter plot showcases a relatively strong linear relationship between the observed power output and the simulated values, with an R-squared value of 0.7. The Northern Cape province exhibits an even higher R-squared value of 0.72, indicating a strong correlation between the observed and simulated power output data. ERA5's simulation

performance is particularly accurate in representing power generation patterns in this area. On the other hand, in the Western Cape province, the scatter plot demonstrates a weaker correlation, with an R-squared value of 0.32. In the Eastern Cape province, the scatter plot for ICON-LAM has a similar pattern as for ERA5, with a slightly higher R-squared value of 0.71. For the Northern Cape province, ICON-LAM has a R-squared value of 0.53, slightly lower than ERA5's performance in this province. In the Western Cape province, ICON-LAM's scatter plot shows an R-squared value of 0.4, suggesting a moderate correlation between the observed and simulated power output data. While this value is higher than ERA5's performance in the Western Cape, it still indicates a relatively moderate agreement. The scatter plot analysis underscores that both ERA5 and ICON-LAM simulations generally align well with observed power output data across the provinces.



a) ERA5 Simulation



b) ICON-LAM Simulation

Figure 20: Comparison of histograms of observed and simulated wind power (by ERA5: top and ICON-LAM: bottom driving RESKit).

The histograms analysis of the power output for corresponding provinces using ICON-LAM and ERA5 simulations is provided by Figure 20. In the Eastern Cape province, the histograms analysis demonstrates a high Perkins Skill Score (PSS) of 0.866. This indicates that ICON-LAM's power output simulation closely matches the observed values, reflecting a strong level

of agreement. In the Northern Cape province, the PSS for ICON-LAM is 0.774, again indicating a notable correspondence between simulated and observed power output. Moving to the Western Cape province, the PSS for ICON-LAM is 0.675, signifying a moderate but still appreciable alignment between the simulated and observed power output data. For the Eastern Cape province, the histograms analysis of ERA5 yields a PSS of 0.875. This signifies a robust agreement between ERA5's simulated power output and the actual observations in this region. In the Northern Cape province, the PSS for ERA5 is even higher at 0.861, showcasing a strong correlation between the simulated and observed power output data, and higher than for ICON-LAM. In the Western Cape province, the PSS for ERA5 is 0.673, indicating a reasonable but slightly lower level of agreement between the simulated and observed power output values. The histograms analysis underscores that both ICON-LAM and ERA5 simulations perform well in capturing power output patterns across the provinces. ERA5 tends to show slightly higher PSS values across all three provinces, suggesting a relatively better match with observed data. In summary we could conclude that ERA5 outperforms ICON-LAM in the power generation.

In terms of wind power generation analysis, ERA5 consistently exhibits superior performance compared to ICON-LAM, as evidenced by key metrics including MAE, RMSE, Pearson correlation, R-squared, and Perkins Skill score. However, definitively establishing ERA5's supremacy over ICON-LAM presents certain challenges, mainly stemming from the nature of the collected wind power generation data, which undergoes extensive aggregation. This aggregated data comprises a network of over 39 wind parks, with 16 located in the Eastern Cape, 10 in the Northern Cape, and 13 in the Western Cape. Each province is represented by a single aggregated power generation time series. This provincial-level aggregation significantly impacts ICON-LAM's performance, as it limits its ability to provide fine-grained insights. Consequently, the substantial aggregation inherent in the observed power data may account for ERA5's more favourable performance results.

III.2.3 Solar radiation

Table 4: Comparison of statistical metrics of observed and simulated GHI

	Mean-MAE(W/m^2)	***Mean-ME(W/m^2)***	***Mean-RMSE(W/m^2)***	***Mean-R_squared***	***Mean-Correlation***
ERA5	278.3	65.9	330.0	0.71	0.91
ICON-LAM	228.4	69.4	283.0	0.68	0.9

Two specific stations have been selected for analysis: SASSCAL WN-361097 and TAHMO WN-TA00486. These station identifiers correspond to distinct model cells within the study area. These two stations have been chosen based on a combination of factors, including data quality, availability, and their ability to offer a representative average depiction of the broader network of stations.

Nighttime data has been omitted from the time series. Specifically, the observation and simulation time series were aligned, after which nighttime data points were extracted based on the nighttime dataframe of the simulation model. Table 4 shows the Global Horizontal Irradiance (GHI) estimation metrics. ERA5 exhibits a higher mean MAE of 278.3 W/m², indicating that, on average, its GHI estimates deviate from observed data by this amount, on the other hand, ICON-LAM displays a lower MAE of 228.4 W/m², suggesting that its GHI estimations have a relatively smaller average deviation from observed data compared to ERA5. Both ERA5 and ICON-LAM display positive mean ME values, indicating a tendency to overestimate GHI on average. ERA5 has a mean ME of 65.9 W/m², while ICON-LAM's mean ME is slightly higher at 69.4 W/m². Similar to mean MAE, ERA5 has a higher mean RMSE of 330.0 W/m², indicating a larger overall error in its GHI estimations compared to observed data and ICON-LAM presents a lower mean RMSE of 283.0 W/m², signifying better accuracy in representing GHI levels. ERA5's mean R-squared value is 0.71, indicating that it captures around 71% of the variability in observed GHI data and ICON-LAM's mean R-squared of 0.68 implies that it explains about 68% of the variability in observed GHI data. Both models exhibit decent R-squared values, indicating a reasonable degree of correlation. ERA5 showcases a higher correlation of 0.91, signifying a strong linear relationship between its GHI estimates and observed data. ICON-LAM's mean correlation is slightly lower at 0.9, but it still demonstrates a robust linear relationship with observed GHI data. The analysis highlights that both ERA5 and ICON-LAM perform reasonably well in estimating GHI. ICON-LAM stands out with

lower MAE and RMSE values, indicating that it might be more accurate in capturing GHI levels. ERA5, while having a higher MAE and RMSE, boasts higher mean R-squared and correlation values, suggesting that it better explains the variability and maintains a strong linear relationship with observed GHI data.

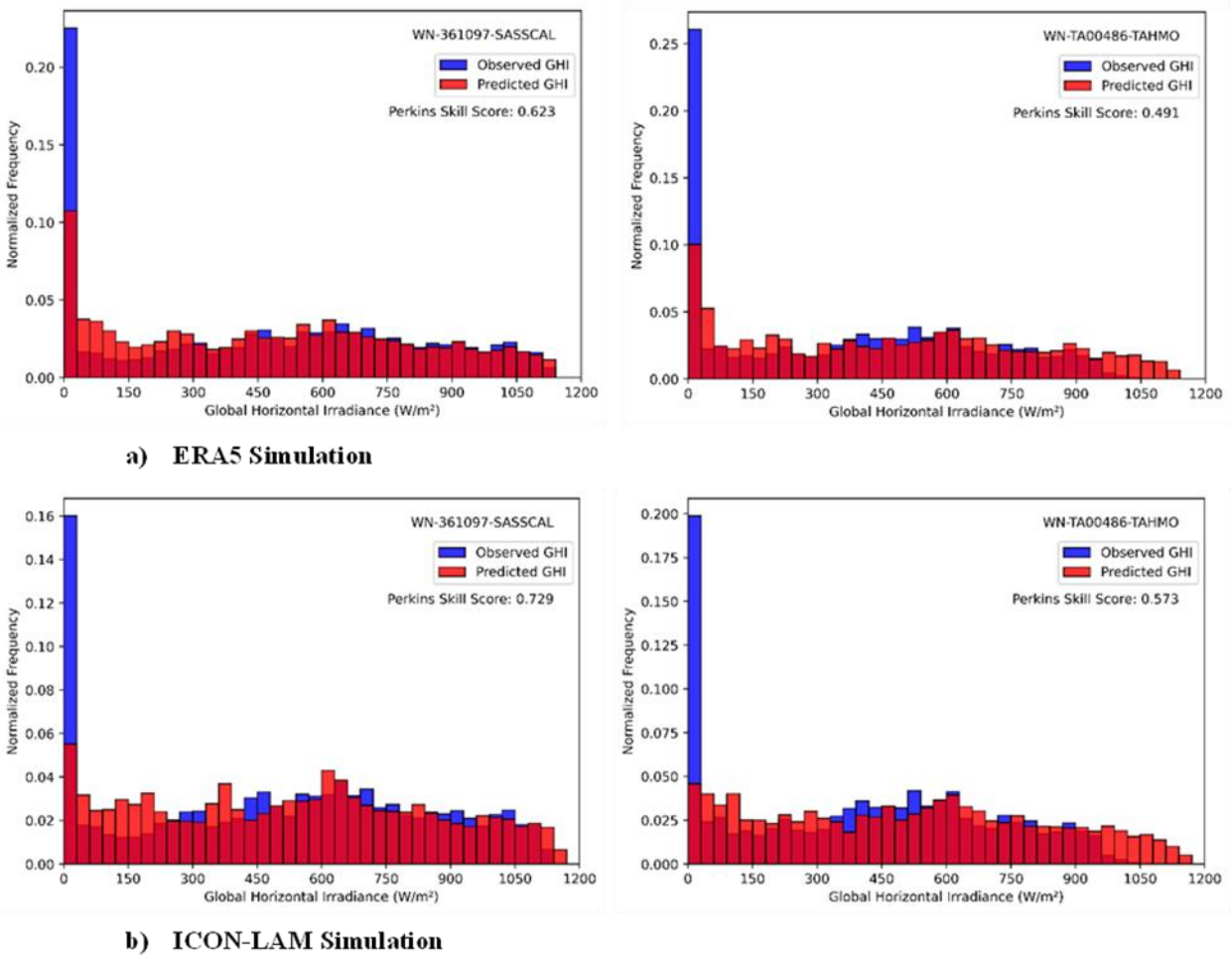


Figure 21: Comparison of histograms of observed and simulated GHI based on ERA5 (top) and ICON-LAM (bottom) driven RESKit.

Figure 21 gives an insight to the GHI distribution. Both models exhibit a frequency peak in GHI estimation between 0 and 30 W/m², indicating a concentration of observations within this range. When considering the Perkins Skill Score (PSS) as a measure of accuracy, it's noteworthy that ERA5's PSS spans from 0.491 to 0.623, reflecting a moderate level of skill in replicating observed GHI patterns. On the other hand, ICON-LAM's PSS range is slightly higher, varying from 0.573 to 0.729. This suggests that ICON-LAM performs consistently better in accurately reproducing GHI values compared to ERA5 across this range of GHI levels. The higher PSS values for ICON-LAM emphasize its stronger proficiency in capturing GHI variations.

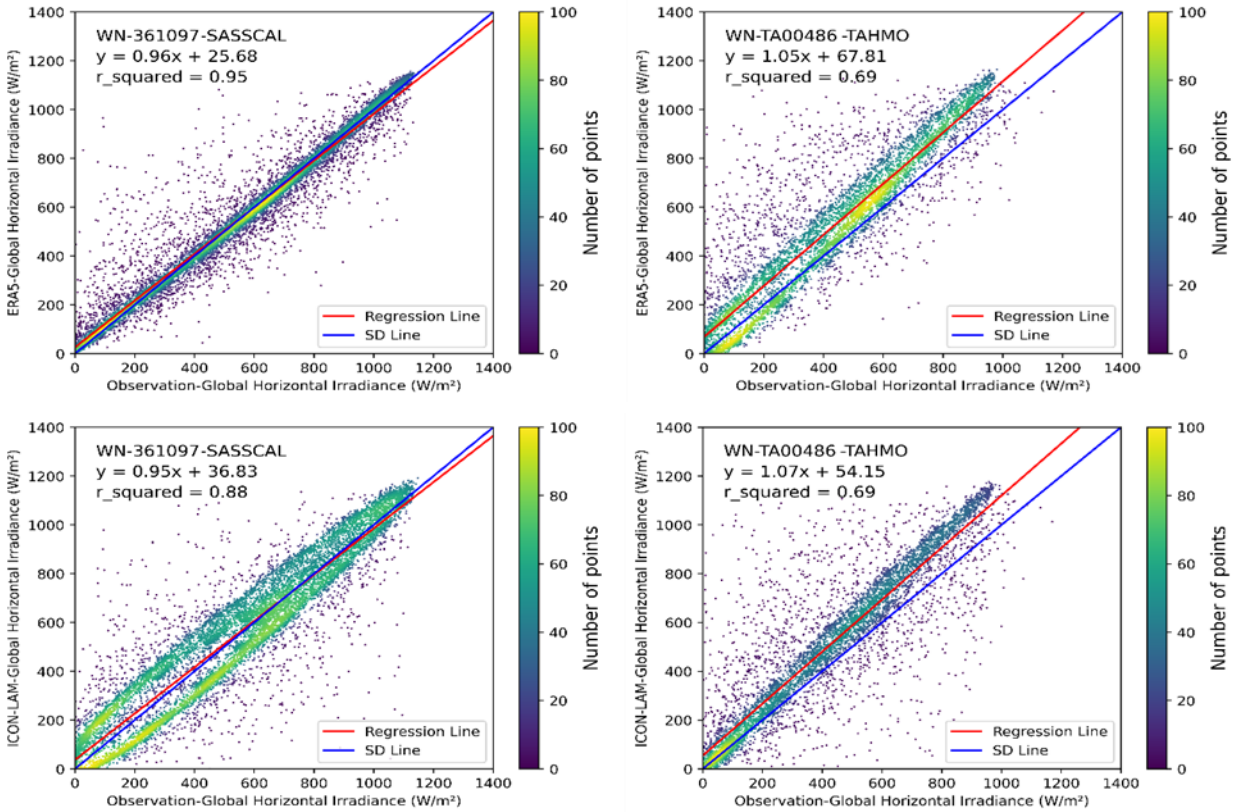


Figure 22: GHI scatter plot comparing simulated values (ERA5 and ICON-LAM) and in-situ GHI observations. The figures also display further statistical metrics.

Figure 22 provides a visual representation of the fluctuations between simulated and observed GHI values. In conjunction with the information from Table 4, a notable trend emerges: at specific stations, ERA5's coefficient of determination (R-squared) values, notably 0.95 for SASSCAL and 0.69 for TAHMO, surpass those of ICON-LAM, which stand at 0.88 for SASSCAL and 0.69 for TAHMO. This suggests that ERA5 performs exceptionally well in establishing linear relationships and explaining the variance in GHI estimations for SASSCAL sites.

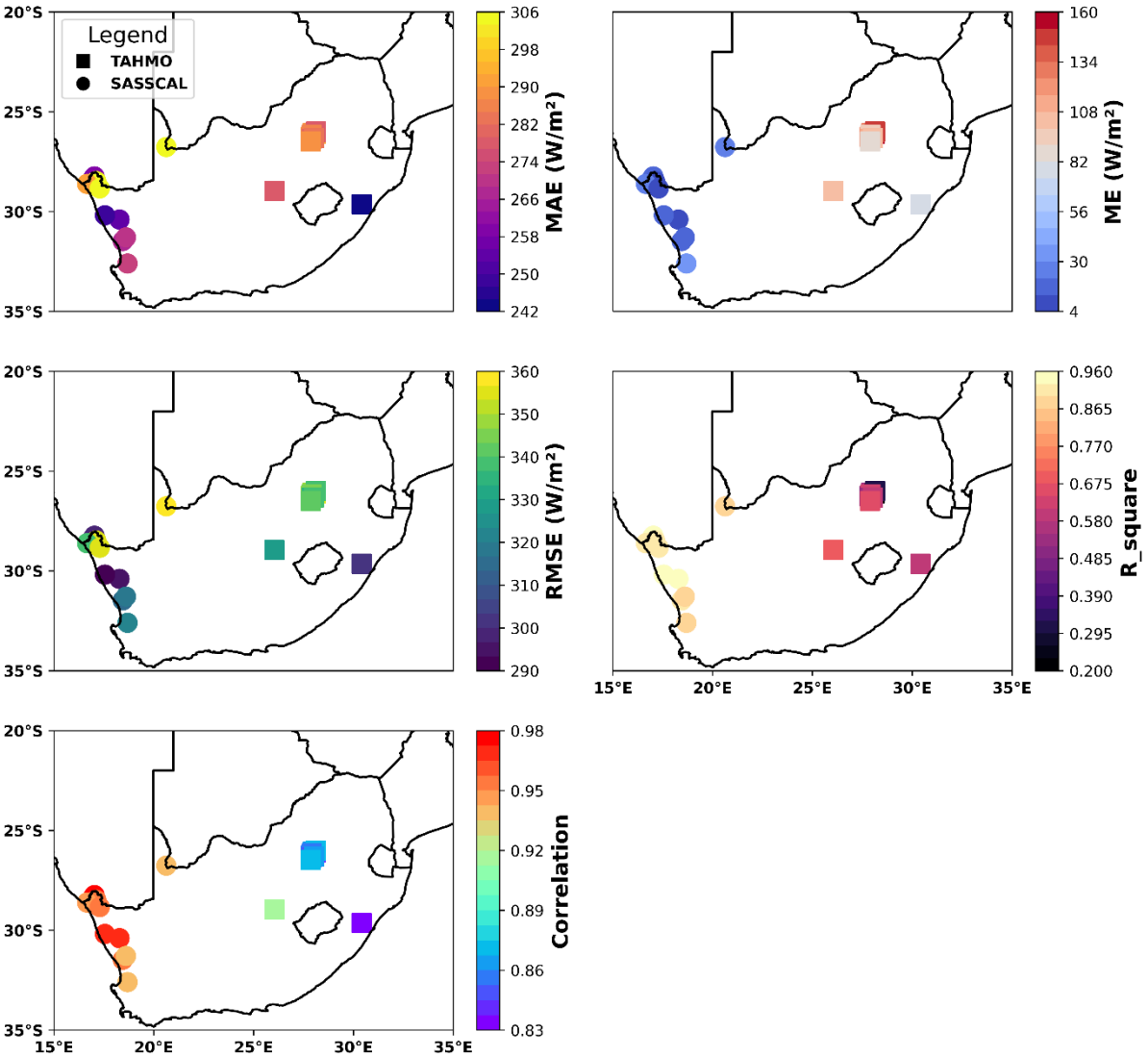


Figure 23: Spatial distribution of statistical metrics for GHI simulated by ERA5, and compared to in situ observations.

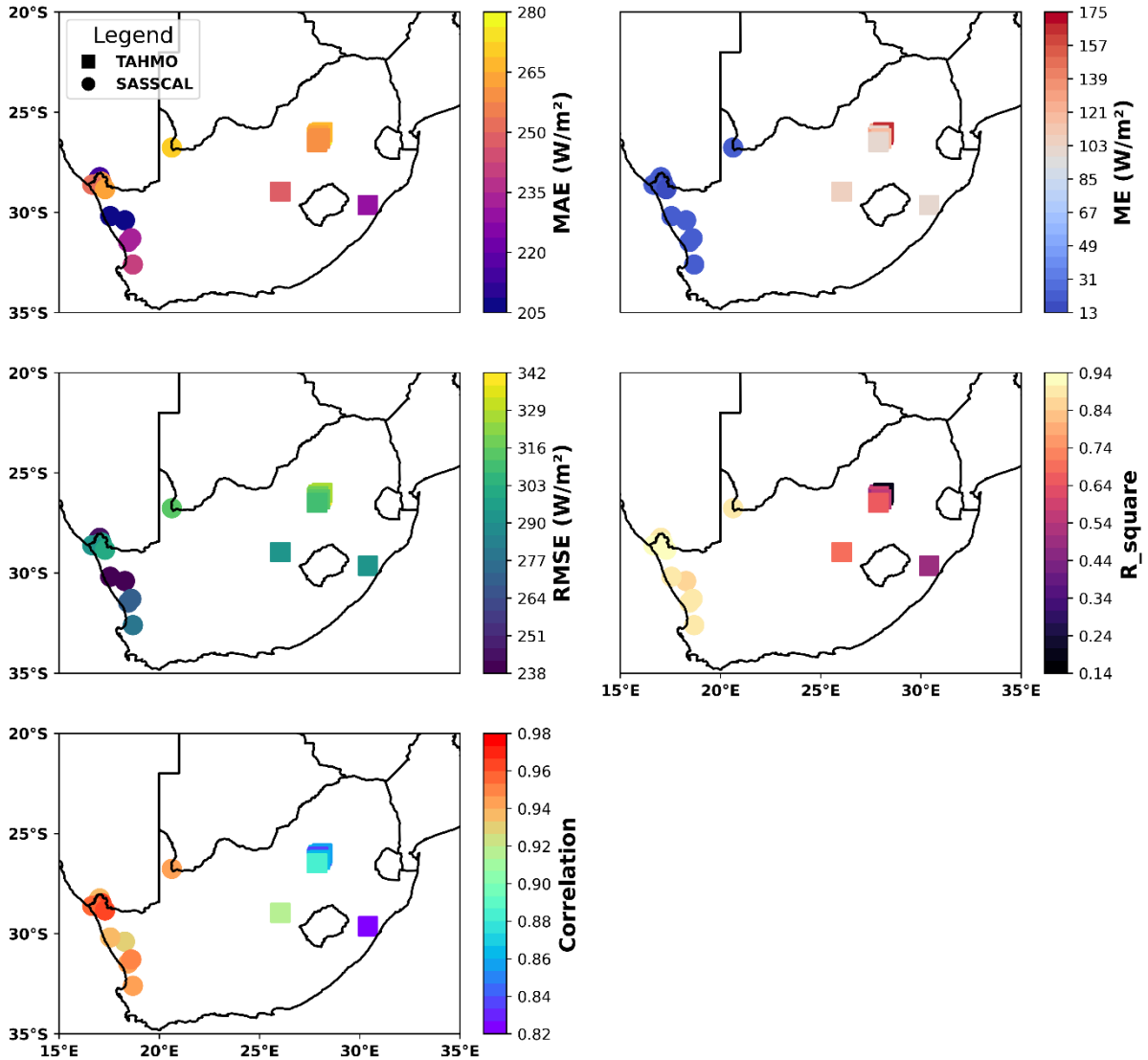


Figure 24: Spatial distribution of statistical metrics for GHI simulated by ICON-LAM, and compared to in situ observations.

Figure 23 and Figure 24 depict the spatial distribution of various metrics, providing valuable insights into their variations across different geographic locations. WASCAL stations along the north-western borders is particularly intriguing; it unveils a gradient of MAE values, ranging from modest to considerable. Interestingly, the ME errors associated with these stations remain relatively subdued. In contrast, the TAHMO stations nestled in the heartland exhibit moderate MAE averages, yet stand out with the most elevated ME values observed. This visual representation not only provides geographical context but also underscores the distinct patterns and trends within the dataset.

The resolution of ERA5 might prove insufficient when applied to regions characterized by substantial fluctuations in solar radiation (Gualtieri, 2015). To highlight this concern,

Sianturi et al., (2020) found in their study of evaluation of ERA5 and MERRA2 that under clear sky conditions or when the clearness index is low, there is a tendency for the measured solar irradiance to be underestimated. Conversely, on cloudy days, both reanalysis estimates often exceed the observed solar irradiance values. This aligns with our study that found that ICON-LAM outperforms ERA5.

III.2.4 Solar power generation

Table 5: Comparison of statistical metrics of observed and simulated solar power (by ERA5 and ICON-LAM driven RESKit).

	MAE (%)		***ME (%)***		***RMSE (%)***		***R-squared***	***Correlation***
	Error	**Mean**	**Error**	**Mean**	**Error**	**Mean**	**Mean**	**Mean**
ERA5	16.91	8.37	7.84	-3.88	22.71	11.24	0.87	0.95
ICON-LAM	17.38	8.6	9.31	-4.61	23.68	11.72	0.86	0.94

The analysis of solar power generation data unveils important insights into the performance of both ERA5 and ICON-LAM simulation models as shown in Table 5. The percent-error for ERA5 in term of MAE is 16.91%, while for ICON-LAM, it is slightly higher 17.38%. This indicates that both models exhibit a similar level of error in predicting solar power generation, with ERA5 showing a slightly lower error. The percent-error of ME for ICON-LAM is 9.31%, coupled with a mean capacity factor of -4.64% implying that the model tends to underestimate solar power generation. Conversely, ERA5 showcases a lower percent-error of 7.84% with a mean capacity factor of -3.88%, suggesting better accuracy in capturing solar power levels. Both ERA5 and ICON-LAM show comparable performance in terms of RMSE, with a percent-error of 23.68% for ICON-LAM and 22.71% for ERA5. Both models perform exceptionally well in terms of Pearson correlation coefficients, indicating a strong linear relationship between observed and simulated solar power generation. ICON-LAM achieves an impressive average correlation of 0.94, while ERA5 demonstrates a slightly higher average correlation of 0.95. Both models exhibit relatively high R-squared values, indicating their ability to explain a substantial portion of the variability in the observed solar power generation data. ICON-LAM boasts an average R-squared value of 0.86, while ERA5's average R-squared value is slightly higher at 0.87.

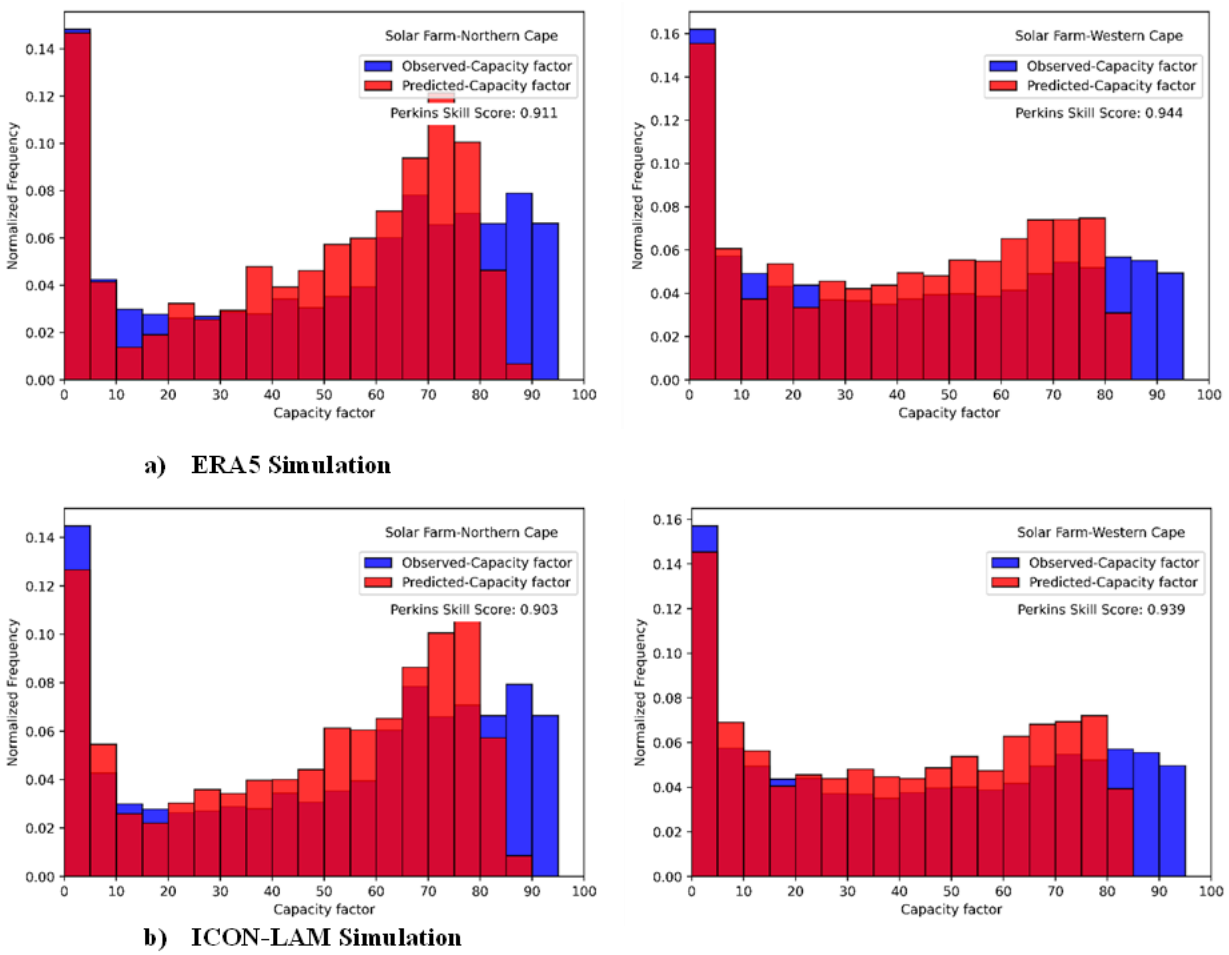


Figure 25: Comparison of histograms of observed and simulated solar power based on ERA5 (top) and ICON-LAM (bottom) driven RESKit.

Figure 25 shows ERA5 and ICON-LAM histograms. In the Northern Cape, the Perkins Skill Scores (PSS) are high, indicating strong agreement between the simulated and observed solar power generation data. ERA5 achieves a PSS of 0.91, while ICON-LAM obtains a slightly lower but still robust PSS of 0.903. In the Western Cape province, both ERA5 and ICON-LAM exhibit even higher PSS values, indicating an excellent agreement between simulated and observed solar power generation. ERA5's PSS is 0.944, and ICON-LAM's PSS is 0. 939.The peak of the histogram for both models is located between 0-5% of the capacity factor. An important trend is observed where the observed solar power generation consistently exceeds the simulation results by 80% of the capacity factor. This suggests that both ERA5 and ICON-LAM slightly underestimate the actual solar power generation in both provinces.

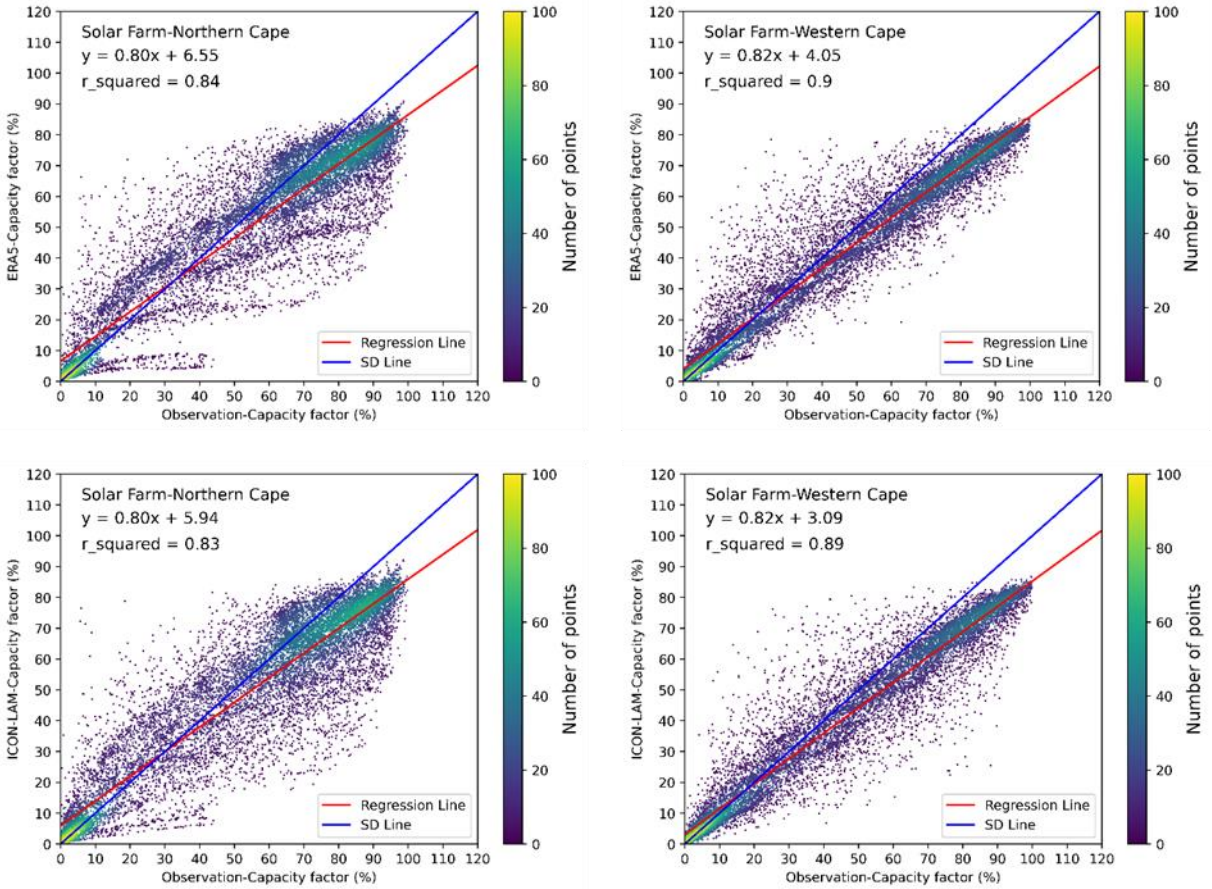


Figure 26: Solar power scatter plot comparing simulated values (ERA5 and ICON-LAM) and in-situ GHI observations. The figures also display further statistical metrics.

The scatter plots in Figure 26 exhibits respectable good R-squared values for the Northern Cape province, with ERA5 achieving an R-squared of 0.84 and ICON-LAM achieving a slightly lower but still robust R-squared of 0.83. In the Western Cape province ERA5's R-squared is 0.9, while ICON-LAM's R-squared is 0.89. These R-squared values indicate a strong linear relationship between the simulated and observed solar power generation data in both regions.

When it comes to solar power generation, ERA5 also consistently outperforms ICON-LAM across various key metrics like MAE, RMSE, Pearson correlation, R-squared, ME and Perkins Skill score. However, definitively establishing ERA5's superiority over ICON-LAM is challenging due to the way we aggregate the solar power generation data.

This data encompasses a network of 44 parks, with 39 located in the Northern Cape and 5 in the Western Cape. For each province, we have a single aggregated time series for power generation. This provincial-level aggregation significantly impacts ICON-LAM's performance, limiting its ability to provide detailed insights. As a result, the substantial aggregation in the observed power data may explain why ERA5 consistently shows better results.

Additionally, it's crucial to acknowledge the distinctive attributes of ERA5 as a reanalysis and ICON-LAM with its enhanced spatial resolution. In instances where ICON-LAM's finer resolution advantage is not applicable, the significance of calibrating the model against measurement data becomes even more pronounced. This underscores the paramount importance of harmonizing model predictions with real-world observations, particularly in scenarios where the benefits of heightened spatial resolution are not manifest.

Conclusion

In conclusion, our analysis reveals interesting trends in the performance of ICON-LAM and ERA5. ICON-LAM demonstrates superior accuracy when it comes to wind speed and solar radiation modelling. However, when we shift our focus to power generation, ERA5 takes the lead in terms of performance. This apparent shift in supremacy between the two models can likely be attributed to the extensive aggregation of observed power data. The high level of aggregation introduces certain limitations, particularly in providing fine-grained insights, which may have contributed to ERA5's superior performance in power generation analysis.

GENERAL CONCLUSION AND PERSPECTIVES

GENERAL CONCLUSION AND PERSPECTIVES

This study delves into the assessment of wind speed and solar radiation as potential sources for further power generation simulation, using data sources across South Africa. Addressing the research inquiries, the following conclusions can be drawn: (1) Regarding wind speed analysis, ICON-LAM exhibits superior performance compared to ERA5 across all locations. Various indicators indicate this superiority: ICON-LAM showcases a notably improved performance in terms of mean error (ME), with an average reduction of 1.24 m/s, a mean absolute error (MAE) lowered by 0.35 m/s, a raised R-squared value of 0.15, and root mean square error (RMSE) reduced by 1.05 m/s. (2) In the context of solar radiation, ICON-LAM emerges as the frontrunner over ERA5. This is substantiated by a Perkins Skill score that is 0.094 points higher, a significantly reduction of 49.88 W/m² in MAE between the two, a decrease of 47 W/m² in RMSE, all in favour of ICON-LAM. (3) In the domain of wind power estimation, ERA5 demonstrates superior performance over ICON-LAM, showcasing a ME that is 0.58% lower, a 0.04 higher R-squared value, a RMSE that is 0.82% lower, a correlation heightened by 0.03, and a Perkins Skill Score (PSS) increased by 0.033. (4) In the realm of solar power estimation, ERA5 maintains its edge over ICON-LAM, exhibiting a lower MAE by 0.26%, a decreased ME by -0.73%, a reduced RMSE by 0.48%, a slightly higher R-squared value by 0.01%, an increment in correlation by 0.01, and a PSS elevated by 0.0315. It's important to emphasize that while ERA5's performance is noteworthy, the level of certainty is somewhat constrained due to the substantial observation data aggregation in power generation, introducing a considerable degree of uncertainty that impacts ICON-LAM's performance assessment.

Continuing research in this field could entail the evaluation of offshore wind energy potential and the analysis of concentrated solar power within the same geographical region. This future research direction aligns with the notion that the performance of high-resolution models like ICON-LAM can be enhanced through data assimilation, leading to the development of a high-resolution reanalysis. By addressing these avenues, we can enhance our understanding of renewable energy potential and contribute to more informed energy-related decisions for South Africa's sustainable future.

BIBLIOGRAPHY REFERENCES

BIBLIOGRAPHY REFERENCES

Águas, H., Mateus, T., Vicente, A., Gaspar, D., Mendes, M. J., Schmidt, W. A., Pereira, L., Fortunato, E., & Martins, R. (2015). Thin Film Silicon Photovoltaic Cells on Paper for Flexible Indoor Applications. *Advanced Functional Materials*, *25*(23), 3592–3598. https://doi.org/10.1002/adfm.201500636

Akinbami, O. M., Oke, S. R., & Bodunrin, M. O. (2021). The state of renewable energy development in South Africa: An overview. *Alexandria Engineering Journal*, *60*(6), 5077–5093. https://doi.org/10.1016/j.aej.2021.03.065

Akor, S. (2021). *Assessment of the Wind Energy Potential over Africa based on ERA5*.

Albani, A., & Ibrahim, M. (2017). Wind Energy Potential and Power Law Indexes Assessment for Selected Near-Coastal Sites in Malaysia. *Energies*, *10*(3), 307. https://doi.org/10.3390/en10030307

Aliyu, A. K., Modu, B., & Tan, C. W. (2018). A review of renewable energy development in Africa: A focus in South Africa, Egypt and Nigeria. *Renewable and Sustainable Energy Reviews*, *81*, 2502–2518. https://doi.org/10.1016/j.rser.2017.06.055

Amillo, A. G., Ntsangwane, L., Huld, T., & Trentmann, J. (2018). Comparison of satellite-retrieved high-resolution solar radiation datasets for South Africa. *Journal of Energy in Southern Africa*, *29*(2). https://doi.org/10.17159/2413-3051/2017/v29i2a3376

Ay, M., & Kisi, O. (2015). Investigation of trend analysis of monthly total precipitation by an innovative method. *Theoretical and Applied Climatology*, *120*(3–4), 617–629. https://doi.org/10.1007/s00704-014-1198-8

Bailey, B. H., McDonald, S. L., Bernadett, D. W., Markus, M. J., & Elsholz, K. V. (1997). *Wind resource assessment handbook: Fundamentals for conducting a successful monitoring program* (NREL/SR--440-22223, ON: DE97000250, 486127; p. NREL/SR--440-22223, ON: DE97000250, 486127). https://doi.org/10.2172/486127

Birol, D. F. (2022). *World Energy Outlook 2022*.

Boccard, N. (2009). Capacity factor of wind power realized values vs. Estimates. *Energy Policy*, *37*(7), 2679–2688. https://doi.org/10.1016/j.enpol.2009.02.046

Copeland, A. W., Black, O. D., & Garrett, A. B. (1941). *The Photovoltaic Effect.*

Daoudi, M., Mou, A. A. S., & Naceur, L. A. (2022). Evaluation of wind energy potential at four provinces in Morocco using two-parameter Weibull distribution function. *International Journal of Power Electronics and Drive Systems (IJPEDS)*, *13*(2), 1209. https://doi.org/10.11591/ijpeds.v13.i2.pp1209-1216

De Myttenaere, A., Golden, B., Le Grand, B., & Rossi, F. (2016). Mean Absolute Percentage Error for regression models. *Neurocomputing*, *192*, 38–48. https://doi.org/10.1016/j.neucom.2015.12.114

Diabaté, L., Blanc, Ph., & Wald, L. (2004). Solar radiation climate in Africa. *Solar Energy*, *76*(6), 733–744. https://doi.org/10.1016/j.solener.2004.01.002

Elsner, P. (2019). Continental-scale assessment of the African offshore wind energy potential: Spatial analysis of an under-appreciated renewable energy resource. *Renewable and Sustainable Energy Reviews*, *104*, 394–407. https://doi.org/10.1016/j.rser.2019.01.034

Fant, C. & UNU-WIDER. (2016). *Wind turbine and photovoltaic generating efficiency in Africa* (125th ed., Vol. 2016). UNU-WIDER. https://doi.org/10.35188/UNU-WIDER/2016/169-7

Fetting, C. (2020). *THE EUROPEAN GREEN DEAL*.

Fujiwara, M., Wright, J. S., Manney, G. L., Gray, L. J., Anstey, J., Birner, T., Davis, S., Gerber, E. P., Harvey, V. L., Hegglin, M. I., Homeyer, C. R., Knox, J. A., Krüger, K., Lambert, A., Long, C. S., Martineau, P., Molod, A., Monge-Sanz, B. M., Santee, M. L., … Zou, C.-Z. (2017). Introduction to the SPARC Reanalysis Intercomparison Project (S-RIP)

and overview of the reanalysis systems. *Atmospheric Chemistry and Physics*, *17*(2), 1417–1452. https://doi.org/10.5194/acp-17-1417-2017

Giorgi, F., & Gutowski, W. J. (2015). Regional Dynamical Downscaling and the CORDEX Initiative. *Annual Review of Environment and Resources*, *40*(1), 467–490. https://doi.org/10.1146/annurev-environ-102014-021217

Goetzberger, A., Luther, J., & Willeke, G. (2002). Solar cells: Past, present, future. *Solar Energy Materials & Solar Cells 74 (2002) 1-11*.

Goodrich, A., Hacke, P., Wang, Q., Sopori, B., Margolis, R., James, T. L., & Woodhouse, M. (2013). A wafer-based monocrystalline silicon photovoltaics road map: Utilizing known technology improvement opportunities for further reductions in manufacturing costs. *Solar Energy Materials and Solar Cells*, *114*, 110–135. https://doi.org/10.1016/j.solmat.2013.01.030

Govindasamy, T. R., & Chetty, N. (2018). Quantifying the global solar radiation received in Pietermaritzburg, KwaZulu-Natal to motivate the consumption of solar technologies. *Open Physics*, *16*(1), 786–794. https://doi.org/10.1515/phys-2018-0098

Green, M. A. (2000). Photovoltaics: Technology overview. *Energy Policy*, *28*(14), 989–998. https://doi.org/10.1016/S0301-4215(00)00086-0

Gruber, K., Regner, P., Wehrle, S., Zeyringer, M., & Schmidt, J. (2022). Towards global validation of wind power simulations: A multi-country assessment of wind power simulation from MERRA-2 and ERA-5 reanalyses bias-corrected with the global wind atlas. *Energy*, *238*, 121520. https://doi.org/10.1016/j.energy.2021.121520

Gualtieri, G. (2015). Surface turbulence intensity as a predictor of extrapolated wind resource to the turbine hub height. *Renewable Energy*, *78*, 68–81. https://doi.org/10.1016/j.renene.2015.01.011

Hersbach, H., Bell, B., Berrisford, P., Hirahara, S., Horányi, A., Muñoz-Sabater, J., Nicolas, J., Peubey, C., Radu, R., Schepers, D., Simmons, A., Soci, C., Abdalla, S., Abellan, X., Balsamo, G., Bechtold, P., Biavati, G., Bidlot, J., Bonavita, M., … Thépaut, J. (2020). The ERA5 global reanalysis. *Quarterly Journal of the Royal Meteorological Society*, *146*(730), 1999–2049. https://doi.org/10.1002/qj.3803

https://www.facebook.com/electrical4u. (2020, October 27). *Working Principle of Wind Turbine | Electrical4U*. Https://Www.Electrical4u.Com/. https://www.electrical4u.com/working-principle-of-wind-turbine/

https://www.facebook.com/linquip. (2021, March 16). *How Do Wind Turbine Generators Work? | Linquip*. https://www.linquip.com/blog/how-do-wind-turbine-generators-work/

Huleihil, M., Mazor, G., Huleihil, M., & Mazor, G. (2012). Wind Turbine Power: The Betz Limit and Beyond. In *Advances in Wind Power*. IntechOpen. https://doi.org/10.5772/52580

International Energy Agency. (2022). *Africa Energy Outlook 2022, IEA*.

Kantar, Y. M., & Usta, I. (2008). Analysis of wind speed distributions: Wind distribution function derived from minimum cross entropy principles as better alternative to Weibull function. *Energy Conversion and Management*, *49*(5), 962–973. https://doi.org/10.1016/j.enconman.2007.10.008

Khare, A., & Rangnekar, S. (2014). Optimal sizing of a grid integrated solar photovoltaic system. *IET Renewable Power Generation*, *8*(1), 67–75. https://doi.org/10.1049/iet-rpg.2012.0382

Laprise, R. (2008). Regional climate modelling. *Journal of Computational Physics*, *227*(7), 3641–3666. https://doi.org/10.1016/j.jcp.2006.10.024

Lenton, T. M., Rockström, J., Gaffney, O., Rahmstorf, S., Richardson, K., Steffen, W., & Schellnhuber, H. J. (2019). Climate tipping points—Too risky to bet against. *Nature*, *575*(7784), 592–595. https://doi.org/10.1038/d41586-019-03595-0

Libanda, B., & Paeth, H. (2023). Modelling wind speed across Zambia: Implications for wind energy. *International Journal of Climatology*, *43*(2), 772–786. https://doi.org/10.1002/joc.7826

Load Factors for Provincial Data – The Renewable Energy Data and Information Service. Retrieved 28 July 2023, from http://redis.energy.gov.za/electricity-production-details/

Mentis, D. (2013). *Wind Energy Assessment in Africa A GIS-based approach*.

Michael, E., Tjahjana, D. D. D. P., & Prabowo, A. R. (2021). Estimating the potential of wind energy resources using Weibull parameters: A case study of the coastline region of Dar es Salaam, Tanzania. *Open Engineering*, *11*(1), 1093–1104. https://doi.org/10.1515/eng-2021-0108

Miller, L. M., & Keith, D. W. (2018). Observation-based solar and wind power capacity factors and power densities. *Environmental Research Letters*, *13*(10), 104008. https://doi.org/10.1088/1748-9326/aae102

Morgan, M. J., Jakovidis, G., & McLeod, I. (1994). An experiment to measure the *I—V* characteristics of a silicon solar cell. *Physics Education*, *29*(4), 252–254. https://doi.org/10.1088/0031-9120/29/4/014

Mukasa, A. D., Mutambatsere, E., Arvanitis, Y., & Triki, T. (2015). Wind energy in sub-Saharan Africa: Financial and political causes for the sector's under-development. *Energy Research & Social Science*, *5*, 90–104. https://doi.org/10.1016/j.erss.2014.12.019

Newell, P., & Bulkeley, H. (2017). Landscape for change? International climate policy and energy transitions: evidence from sub-Saharan Africa. *Climate Policy*, *17*(5), 650–663. https://doi.org/10.1080/14693062.2016.1173003

Njiri, J. G., & Söffker, D. (2016). State-of-the-art in wind turbine control: Trends and challenges. *Renewable and Sustainable Energy Reviews*, *60*, 377–393. https://doi.org/10.1016/j.rser.2016.01.110

Norquist, D. C., & Meeks, W. C. (2010). A comparative verification of forecasts from two operational solar wind models: SOLAR WIND FORECAST MODELS. *Space Weather*, *8*(12). https://doi.org/10.1029/2010SW000598

Nwaigwe, K. N., Mutabilwa, P., & Dintwa, E. (2019). An overview of solar power (PV systems) integration into electricity grids. *Materials Science for Energy Technologies*, *2*(3), 629–633. https://doi.org/10.1016/j.mset.2019.07.002

Ogbomo, O. O., Amalu, E. H., Ekere, N. N., & Olagbegi, P. O. (2017). A review of photovoltaic module technologies for increased performance in tropical climate. *Renewable and Sustainable Energy Reviews*, *75*, 1225–1238. https://doi.org/10.1016/j.rser.2016.11.109

Ohunakin, O. S. (2011). Assessment of wind energy resources for electricity generation using WECS in North-Central region, Nigeria. *Renewable and Sustainable Energy Reviews*, *15*(4), 1968–1976. https://doi.org/10.1016/j.rser.2011.01.001

Perkins, S. E., Pitman, A. J., Holbrook, N. J., & McAneney, J. (2007). Evaluation of the AR4 Climate Models' Simulated Daily Maximum Temperature, Minimum Temperature, and Precipitation over Australia Using Probability Density Functions. *Journal of Climate*, *20*(17), 4356–4376. https://doi.org/10.1175/JCLI4253.1

Pham, T. V., Steger, C., Rockel, B., Keuler, K., Kirchner, I., Mertens, M., Rieger, D., Zängl, G., & Früh, B. (2021). ICON in Climate Limited-area Mode (ICON release version

2.6.1): A new regional climate model. *Geoscientific Model Development*, *14*(2), 985–1005. https://doi.org/10.5194/gmd-14-985-2021

Physics of Wind Turbines | Energy Fundamentals. Retrieved 9 July 2023, from https://home.uni-leipzig.de/energy/energy-fundamentals/15.htm

Pownuk, A. M., & Kreinovich, V. (2017). Why Linear Interpolation? *№3(43) (2017)*. https://doi.org/10.25513/2222-8772.2017.3.43-49

Ramirez Camargo, L., Gruber, K., & Nitsch, F. (2019). Assessing variables of regional reanalysis data sets relevant for modelling small-scale renewable energy systems. *Renewable Energy*, *133*, 1468–1478. https://doi.org/10.1016/j.renene.2018.09.015

Ramon, J., Lledó, L., Torralba, V., Soret, A., & Doblas-Reyes, F. J. (2019). What global reanalysis best represents near-surface winds? *Quarterly Journal of the Royal Meteorological Society*, *145*(724), 3236–3251. https://doi.org/10.1002/qj.3616

RESKit—Renewable Energy Simulation toolkit for Python. (2023). [Python]. FZJ-IEK3. https://github.com/FZJ-IEK3-VSA/RESKit (Original work published 2020)

Rhodes, C. J. (2017). World Meteorological Organisation (WMO) Report: Global Greenhouse Gas Concentrations Highest in 800,000 Years. *Science Progress*, *100*(4), 428–433. https://doi.org/10.3184/003685017X15105820753148

Rieger, D., Milelli, M., Boucouvala, D., Gofa, F., Iriza-Burca, A., Khain, P., Kirsanov, A., Linkowska, J., & Marcucci, F. (2021). *Reports on ICON 006: Verification of ICON in Limited Area Mode at COSMO National Meteorological Services*. https://doi.org/10.5676/DWD_PUB/NWV/ICON_006

Ritchie, H., Roser, M., & Rosado, P. (2020). CO_2 and Greenhouse Gas Emissions. *Our World in Data*. https://ourworldindata.org/co2-emissions

Roehrkasten, S., Thielges, S., & Quitzow, R. (2016). *Sustainable Energy in the G20*.

Ryberg, D. S., Caglayan, D. G., Schmitt, S., Linßen, J., Stolten, D., & Robinius, M. (2019a). The future of European onshore wind energy potential: Detailed distribution and simulation of advanced turbine designs. *Energy*, *182*, 1222–1238. https://doi.org/10.1016/j.energy.2019.06.052

Ryberg, D. S., Caglayan, D. G., Schmitt, S., Linßen, J., Stolten, D., & Robinius, M. (2019b). The future of European onshore wind energy potential: Detailed distribution and simulation of advanced turbine designs. *Energy*, *182*, 1222–1238. https://doi.org/10.1016/j.energy.2019.06.052

Santamarina, J. C. (2006). *The Worldwide Energy Situation*.

Sawadogo, W., Bliefernicht, J., Fersch, B., Salack, S., Guug, S., Diallo, B., Ogunjobi, Kehinde. O., Nakoulma, G., Tanu, M., Meilinger, S., & Kunstmann, H. (2023). Hourly global horizontal irradiance over West Africa: A case study of one-year satellite- and reanalysis-derived estimates vs. in situ measurements. *Renewable Energy*, *216*, 119066. https://doi.org/10.1016/j.renene.2023.119066

Schneider, S. H., & Dickinson, R. E. (1974). Climate modeling. *Reviews of Geophysics*, *12*(3), 447. https://doi.org/10.1029/RG012i003p00447

Schneider, T., & Sobel, A. H. (2022). *The Global Circulation of the Atmosphere*. Princeton University Press.

Setchell, H. (2017, October 4). *ERA5: The new reanalysis of weather and climate data* [Text]. ECMWF. https://www.ecmwf.int/en/about/media-centre/science-blog/2017/era5-new-reanalysis-weather-and-climate-data

Shu, Z. R., & Jesson, M. (2021). Estimation of Weibull parameters for wind energy analysis across the UK. *Journal of Renewable and Sustainable Energy*, *13*(2), 023303. https://doi.org/10.1063/5.0038001

Sianturi, Y., Marjuki, & Sartika, K. (2020). *Evaluation of ERA5 and MERRA2 reanalyses to estimate solar irradiance using ground observations over Indonesia region*. 020002. https://doi.org/10.1063/5.0000854

Silva, A. R., Stosic, T., & Stosic, B. (2021). Wind speed persistence at the Fernando de Noronha archipelago, Brazil. *Theoretical and Applied Climatology*, *144*(1–2), 723–730. https://doi.org/10.1007/s00704-021-03571-7

Solar Cell I-V Characteristic and the Solar Cell I-V Curve. (2021, February 16). Alternative Energy Tutorials. https://www.alternative-energy-tutorials.com/photovoltaics/solar-cell-i-v-characteristic.html

South Africa—Wind farms—Online store and quote request—The Wind Power. Retrieved 29 July 2023, from https://www.thewindpower.net/store_country_en.php?id_zone=58

Sterl, S. (2021). A Grid for all Seasons: Enhancing the Integration of Variable Solar and Wind Power in Electricity Systems Across Africa. *Current Sustainable/Renewable Energy Reports*, *8*(4), 274–281. https://doi.org/10.1007/s40518-021-00192-6

Szewczuk, S., & Prinsloo, F. (2010). *Wind atlas for South Africa (WASA): Project overview and current status*. ResearchSpace. http://hdl.handle.net/10204/5866

Tizgui, I., Bouzahir, H., Guezar, F. E., & Benaid, B. (2017). Wind speed extrapolation and wind power assessment at different heights. *2017 International Conference on Electrical and Information Technologies (ICEIT)*, 1–4. https://doi.org/10.1109/EITech.2017.8255215

Tong, W. (2010a). Fundamentals of wind energy. In *WIT Transactions on State of the Art in Science and Engineering* (1st ed., Vol. 1, pp. 3–48). WIT Press. https://doi.org/10.2495/978-1-84564-205-1/01

Tong, W. (2010b). *Wind Power Generation and Wind Turbine Design*. WIT Press.

Van Der Linde, H. A. (1996). Wind energy in South Africa. *Renewable Energy*, *9*(1–4), 880–883. https://doi.org/10.1016/0960-1481(96)88420-5

Walsh, J. E., Chapman, W. L., Romanovsky, V., Christensen, J. H., & Stendel, M. (2008). Global Climate Model Performance over Alaska and Greenland. *Journal of Climate*, *21*(23), 6156–6174. https://doi.org/10.1175/2008JCLI2163.1

WASAData. Retrieved 17 August 2023, from http://wasadata.csir.co.za/wasa1/WASAData

Wright, J., Landwehr, G., & Chartan, E. (2019). Assessing the value of improved variable renewable energy forecasting accuracy in the South African power system. *Journal of Energy in Southern Africa*, *30*(2), 98–109. https://doi.org/10.17159/2413-3051/2019/v30i2a6293

Yang, D., & Bright, J. M. (2020). Worldwide validation of 8 satellite-derived and reanalysis solar radiation products: A preliminary evaluation and overall metrics for hourly data over 27 years. *Solar Energy*, *210*, 3–19. https://doi.org/10.1016/j.solener.2020.04.016

Zängl, G., Reinert, D., Rípodas, P., & Baldauf, M. (2015). The ICON (ICOsahedral Non-hydrostatic) modelling framework of DWD and MPI-M: Description of the non-hydrostatic dynamical core. *Quarterly Journal of the Royal Meteorological Society*, *141*(687), 563–579. https://doi.org/10.1002/qj.2378

Zheng, C., Xiao, Z., Peng, Y., Li, C., & Du, Z. (2018). Rezoning global offshore wind energy resources. *Renewable Energy*, *129*, 1–11. https://doi.org/10.1016/j.renene.2018.05.090