

The copyright of this thesis rests with the University of Cape Town. No quotation from it or information derived from it is to be published without full acknowledgement of the source. The thesis is to be used for private study or non-commercial research purposes only.

University of Cape Town

Department of Environmental and Geographical Science

Mesoscale Wind Atlas of South Africa

Thesis presented for the degree of Doctor of Philosophy

Kilian Hagemann

July 2008

Contents

Preface	vii
Abstract	vii
Acknowledgements	ix
List of Abbreviations	x
List of Symbols	xiii
List of Figures	xx
List of Tables	xxi
1 Introduction	1
1.1 Objectives and Scope	1
1.2 Wind Power Fundamentals	2
1.2.1 Terminology and Definitions	2
1.2.1.1 Wind Speed and Direction	2
1.2.1.2 Scales of Space and Time	4
1.2.2 Sources of Wind	5
1.2.3 Important Surface Characteristics Affecting the Wind	6
1.2.4 Energy Conversion of a Wind Turbine	7
1.2.5 Essential Statistics for Wind Power	9
1.3 Wind Resource Assessment Methodologies	15
1.3.1 Nature of Assessments	15
1.3.2 Assessments Based on Measurements	16
1.3.3 Assessments Based on Mesoscale Modelling	19
1.3.4 Steady State vs Fully Dynamic	20
1.4 Wind Resource Assessments in South Africa	22
1.4.1 Up to 1995	22
1.4.2 Subsequent Works	24
1.4.3 Total Wind Generation Potential Estimates	26

1.5	The MM5 Model	27
1.6	Methodology Overview	29
2	Station Data and Validation Techniques	31
2.1	The Observations	31
2.2	Validation Methodology	33
2.2.1	Core Comparison Statistics	34
2.2.2	Comparison of Cycles	36
2.2.3	Evaluation of Wind Direction Performance	38
3	Selection of Optimal Simulation Period	41
3.1	Overview	41
3.2	Methodology	42
3.3	Application to South Africa	44
3.4	Uncertainty Analysis	51
3.5	Caveats and Pitfalls	54
4	Selection of Optimal Model Configuration	57
4.1	Overview	57
4.2	Parametrisations in MM5	58
4.2.1	Constant Configurations	59
4.2.2	Variable Configurations	60
4.2.3	International Experience	64
4.3	Design of Sensitivity Simulations	66
4.3.1	Domain, Nesting and Surface Configuration	66
4.3.2	FDDA, Initial and Boundary Conditions	71
4.3.3	Vertical Resolution	72
4.3.4	General Model Configuration	73
4.3.5	Validation Procedure and Pitfalls	74
4.4	Evaluation of LSM and PBL Parametrisations	76
4.4.1	Average Wind Speed and Standard Deviation	76
4.4.2	Daily and Seasonal Cycles	79
4.4.3	Directional Performance	84
4.4.4	Summary	89
4.5	Evaluation of Cumulus Schemes and FDDA	90
4.5.1	Average Wind Speed and Standard Deviation	92
4.5.2	Daily and Seasonal Cycles	94
4.5.3	Directional Performance	99

CONTENTS

4.5.4	Summary	104
4.6	Conclusion	105
5	Wind Atlas of South Africa	107
5.1	Design and Execution of Simulations	107
5.1.1	Domain Configuration	107
5.1.2	Model Preparation and Configuration	109
5.1.3	Execution	110
5.2	Accuracy of Wind Atlas	110
5.2.1	Average Wind Speed and Standard Deviation	111
5.2.2	Daily and Seasonal Cycles	115
5.2.3	Directional Performance	124
5.2.4	Summary and Uncertainty Analysis	126
5.3	Wind Atlas Maps	131
5.3.1	Annual Average Wind Speeds	131
5.3.2	Seasonal Average Wind Speeds	134
5.4	Quantification of Total Wind Resource	135
5.4.1	Data Preparation	136
5.4.2	Total Potential Calculation	139
6	Conclusion	145
6.1	Fulfilment of Objectives	145
6.1.1	Primary Objective	145
6.1.2	Secondary Objective	147
6.2	Discussion, Shortcomings and Future Research	148
6.2.1	Choice of Integration Period	148
6.2.2	Model Validation	149
6.2.3	Model Configuration	151
6.2.4	Model Resolution	154
6.2.5	Wind Atlas Data	154
6.2.6	Wind Potential Estimate	155
6.3	Concluding Remarks	156
A	Combining MM5 Nudging Techniques	159
A.1	Average Wind Speed and Standard Deviation	159
A.2	Daily and Seasonal Cycles	161
A.3	Directional Performance	163
A.4	Summary	167

B	MM5 Model Configuration	169
B.1	Sensitivity Runs for the Eastern Cape	169
B.1.1	TERRAIN	169
B.1.2	REGRID	172
B.1.3	INTERPF	173
B.1.4	MM5	174
B.2	Wind Atlas Runs for South Africa	184
B.2.1	TERRAIN	184
B.2.2	REGRID	186
B.2.3	INTERPF	186
B.2.4	MM5	186
C	MM5 Source Code Modifications	191
C.1	Pleim Xiu LSM Roughness Length Output	191
C.2	Observational Nudging Patch	192
C.3	Output Reduction Patch	198
	Bibliography	215

Preface

Abstract

This thesis explores the utility of the MM5 regional climate model in producing a detailed wind climatology for South Africa in the context of wind power applications.

Given the background of limited available computational resources a sophisticated temporal subsetting algorithm is developed and applied. The goal here is to allow a high resolution 18 by 18km simulation to be carried out over only 365 days of historical records (20 March 1996 until 19 March 1997), yet ensure that it is possible to construct a reasonably accurate climatology from such a short time period because it represents the most typical year available.

Throughout the research the model's sensitivity towards different boundary layer and cumulus parametrisations as well as observational and grid nudging is investigated in depth. New error metrics are specifically designed and applied based on 17 reliable 10m weather stations around the country with the specific purpose of evaluating model skill with respect to the wind patterns that are important for wind turbines.

An optimal model configuration is established by using the Pleim Xiu land surface scheme coupled with the Atmospheric Convective Model for the planetary boundary layer, in addition to the Kain Fritsch II cumulus parametrisation and observational wind nudging. The model is then integrated at 18km for the entire country and the results validated in detail. It is found that average wind speeds are within $0.2 \pm 0.7\text{ms}^{-1}$ (68% confidence interval) and that the daily cycle, which constitutes the largest part of the observed variance, is predicted to within $18\% \pm 20\%$ of amplitude and 1h of timing at coastal sites.

Wind directions are within about 3.5% of frequency of occurrence per 30° wind sector. The chosen parametrisations of MM5 reveal their weaknesses es-

pecially over the highveld where diurnal cycles are often unrealistic. Nocturnal winds are frequently under-predicted, even near the oceans, pointing to systematic errors in the schemes themselves.

In terms of the resulting wind atlas of South Africa a significant inland wind resource is discovered over the three Cape provinces which was previously unknown. Overall the country's wind resource is higher than some previous studies have suggested and is comparable to some of the windiest markets overseas.

In total it is estimated that wind power may realistically provide 81TWh or 35% of the 2007 electricity demand. This figure is based purely on sites with capacity factors higher than 30%, 80m hub heights and 4km maximum distance to suitable roads and transmission lines. This neglects capacity constraints of Eskom's national grid but certainly proves that wind resources in South Africa are plentiful and not a limiting factor to the development of a sustainable wind power industry.

Acknowledgements

My gratitude for this work lies with the following individuals and institutions (in no specific order):

- my supervisor Bruce Hewitson and my unofficial co-supervisor Mark Tadross for providing me with valuable guidance and feedback throughout my entire postgraduate career
- Chris Jack for initially helping out with some hands-on MM5 modelling experience and tips as well as other support
- my research group, the Climate Systems Analysis Group, for generously supporting this work financially in the form of bursaries
- my family and friends whose moral support has been instrumental in making this task much less mentally taxing
- the South African Weather Service for providing me with some much needed weather station data for the validation and fine-tuning of the model
- Shirley Butcher for providing initial tips on how best to conduct the GIS based total wind generation potential calculation
- The Centre for High Performance Computing (CHPC) for allowing me to access some of their supercomputing resources necessary to run the model, store and analyse its output in reasonable time frames
- the ever growing open source and MM5 community for supplying all the great software I used in this research to me free of charge, including the typesetting of this document (LyX)

After four years of working on this dissertation my long standing ambition of contributing to the establishment of a wind power industry in South Africa is finally becoming a reality. I am hoping that this work will convince politicians, policy makers and stakeholders alike that wind power is an electricity generation technology which deserves serious consideration in the country's strategic energy planning, not only due to its merits of avoiding carbon dioxide emissions to mitigate climate change but also in terms of the contribution towards energy security that it offers

List of Abbreviations and Acronyms

The following list is meant to be a useful and condensed list of all the abbreviations used throughout the text as an aid and quick reference for the reader.

5LYR Five Layer (land surface model)

ACM Asymmetric Convective Model (planetary boundary model)

AEP Annual energy production

agl above ground level

AK Anthes-Kuo (cumulus parametrisation)

amsl above mean sea level

AS Arakawa Schubert (cumulus parametrisation)

BL Blackadar (planetary boundary model)

BM Betts Miller (cumulus parametrisation)

BT Burk-Thompson (planetary boundary model)

CAE (daily) cycle amplitude error

cdf cumulative distribution function

CFD computational fluid dynamics

CSIR Centre for Scientific and Industrial Research

DANCED Danish Cooperation for Environment and Development

DFPD (daily) Fourier peak difference

DME Department of Minerals and Energy

ECMWF European Centre for Medium-Range Weather Forecasts

EL East London

FC Fritsch Chappell (cumulus parametrisation)

FDDA four dimensional data assimilation

FFT fast Fourier transform

FR Force Restore (land surface model)

GCM global circulation model

GIS Geographical Information System

GS Gayno-Seaman (planetary boundary model)

(HIR)LAM (High Resolution) Limited Area Model

IEA International Energy Agency

IOA index of agreement

KF Kain Fritsch (cumulus parametrisation)

KF2 Kain Fritsch II (successor to KF)

LIDAR light detection and ranging

LSM land surface model

MAE mean absolute error

MMAE monthly (average wind speed) MAE

MM5 Fifth Generation Pennsylvania State University/NCAR Mesoscale Model

MRF Medium Range Forecast (planetary boundary layer model)

NCAR National Centre for Atmospheric Research

NCEP National Centre for Environmental Prediction

NWP numerical weather prediction

PA Port Alfred

PBL planetary boundary layer

PCA principal component analysis

pdf probability density function

PE Port Elizabeth

PPE (daily) peak phase error

- PX** Pleim Xiu (land surface model)
- RCM** regional climate model
- RMSE** root mean square error
- RRTM** Rapid Radiative Transfer Model
- SARERD** South African Renewable Energy Resource Database
- SAWB** South African Weather Bureau (old description for SAWS)
- SAWS** South African Weather Service
- SODAR** sonic detection and ranging
- SOM** self-organising map
- TKE** turbulent kinetic energy
- USGS** United States Geological Survey
- WAsP** Wind Atlas Analysis and Application Program
- WMAESF** weighted MAE of sector frequencies
- WMAESSA** weighted MAE of sector speed averages
- WMAESSD** weighted MAE of sector speed standard deviations
- WMAEPD** weighted MAE of sector wind power densities
- WMO** World Meteorological Organisation
- WRF** Weather Research and Forecasting Model

List of Symbols

The following table presents all symbols used in this research together with a short description and section, equation and/or page cross references. Upright font indicates that a given quantity is a constant while italicised ones represent variables. This list is meant to aid the reader with respect to all the symbols used throughout the text.

Symbol	Variants	Short Description	Section/ Equation	Page
<i>A</i>	A_θ	Weibull scale parameter (ms^{-1}). A_θ is direction specific	eq 1.17	10
<i>A_r</i>		Total swept area of a wind turbine's rotor (in m^2)	eq 1.13	9
<i>b</i>	$b_V, b_u,$ b_v	ordinary model bias. Similarly to σ the subscript indicates which variable is referred to	eq 2.2	34
<i>C_p</i>		power coefficient of a wind turbine (as a function of V)	eq 1.10	8
<i>C_f</i>		capacity factor of wind turbine	eq 1.14	9
<i>c</i>	$c_V, c_u,$ c_v	standard deviation bias. Subscripts as with b	eq 2.4	35
<i>D</i>		rotor diameter of a wind turbine	eq 1.13	9
<i>D_a</i>		Amplitude or strength of the diurnal cycle (ms^{-1})	eq 1.25	12
<i>E</i>		Energy converted by wind turbine (kWh,MWh, GWh or TWh)	eq 1.12	8
<i>F₀</i>		probability of calms in three parameter Weibull distribution	eq 1.21	11
<i>F_w</i>		two parameter Weibull cumulative distribution function (cdf)	eq 1.17	10
<i>F_{wH}</i>		three parameter Weibull cumulative distribution function	eq 1.21	11
<i>F_P</i>		directional power density distribution	eq 1.30	15

Symbol	Variants	Short Description	Section/ Equation	Page
$F_{\overline{V}}$		directional mean wind speed distribution	eq 1.28	14
F_{θ}		directional wind frequency distribution	eq 1.27	13
F_{σ_V}		directional distribution of wind speed standard deviation	eq 1.29	14
f		frequency (Hz)	sec 1.2.5	12
f_d		frequency of 1 day ($f = 1\text{day}^{-1} = 1.1574 * 10^{-5}\text{Hz}$)	sec 1.2.5	12
f_w		two parameter Weibull probability density function (pdf)	eq 1.18	10
f_{wH}		three parameter Weibull probability density function	eq 1.22	11
$G(f)$		generalised Fourier transformation function	sec 1.2.5	12
H	$H_{s,i}$, $H_{s,k}$, $H_{s,l}$	global, low resolution long term atmospheric dataset (e.g. NCEP/NCAR Reanalysis). The variants are all 365 day subsets thereof, $H_{s,i}$ a generic one, $H_{s,k}$ the one that optimally contains H 's variability across a single classification and $H_{s,l}$ the one that is the most representative across all classifications	sec 3.2	42
i		index of a set or series depending on the context		
j		alternative index to i		
k		As a subscript this is also an alternative index to i		
k	k_{θ}	Weibull shape parameter. k_{θ} is direction specific	eq 1.17	10
k_0		von Kármán constant	eq 1.8	6

Symbol	Variants	Short Description	Section/ Equation	Page
l		wind speed correction factor for climatology	eq 3.3	54
M		number of distinct classifications carried out	sec 3.2	42
m		as a subscript it means it refers to a modelled or model derived quantity	sec 2.2	33
N_s		number of possible subsets $H_{s,i}$	sec 3.2	42
N_c	$N_{c,j}$	number of different classes in a given classification. $N_{c,j}$ is the number of classes in the j^{th} classification	sec 3.2	42
n		length of a dataset	eq 1.1	3
o		as a subscript it means it refers to an observed or observation derived quantity	sec 2.2	33
P_d		power density P_d (Wm^{-2})	eq 1.9	7
p	p_0, p_{top}	atmospheric pressure (Pa). p_0 is the surface pressure and p_{top} the pressure at the top of a climate model's 3D domain	eq 1.31	28
Q	Q_i	Q is the climate state frequency distribution of a given subset of H . Q_i is the Q of the i^{th} subset of H ($H_{s,i}$)	eq 3.1	42
q		mixing ratio, measure of humidity (dimensionless)	eq 1.11	8
R		climate state frequency distribution of a long term reference dataset H	eq 3.1	42
R_a		Perfect gas constant ($8.314JK^{-1}mol^{-1}$)	eq 1.11	8
R_b		specific gas constant of dry air ($287.04Jkg^{-1}K^{-1}$)	eq 1.32	28

Symbol	Variants	Short Description	Section/ Equation	Page
r		Pearson's product-moment correlation coefficient	sec 2.2.1	36
S	S_i	generalised series of wind vectors (V, θ) , i.e. speed and direction pairs. S_i is the i^{th} element of S	eq 1.1	3
s	s_i	sigma level (vertical model coordinate), s_i is the i^{th} sigma level	eq 1.31	28
T		absolute temperature (in $^{\circ}\text{K}$)	eq 1.11	8
T_t	T_s	total time span of a given dataset. T_s is the length of the period a subset covers	sec 3.2	42
t	$t_i, \Delta t$	time. t_i is a particular time, Δt is the time step (Δt^{-1} is the sampling frequency)	eq 1.3	3
t_{max}		time of occurrence of the highest wind speed in V_{24}	sec 1.2.5	12
u	$\bar{u}$	west to east wind component, $\bar{u}$ is its mean	eq 1.6	3
v	$\bar{v}$	south to north wind component, $\bar{v}$ is its mean	eq 1.7	3
V	$V_i, \Delta V$	scalar wind speed in ms^{-1} . V_i is the i^{th} speed, ΔV a change	eq 1.1	3
$\bar{V}$	$\bar{V}_w, \bar{V}_{wH}$	mean wind speed of V series ($\bar{V}_w$ and $\bar{V}_{wH}$ being derived from the two and three parameter Weibull distributions respectively)	eq 1.15	9
V_{24}		average 24 hour wind speed series depicting the diurnal cycle	eq 1.24	11
V_{12}		average 12 month wind speed series depicting the seasonal cycle	sec 1.2.5	12

Symbol	Variants	Short Description	Section/ Equation	Page
$\mathbf{V}_*$		friction velocity (ms^{-1})	eq 1.8	6
x		horizontal space coordinate (west-east)	eq 1.4	3
y		horizontal space coordinate (north-south)	eq 1.4	3
z		vertical space coordinate (height above ground)	eq 1.4	3
z_0		surface roughness length (m)	eq 1.8	6
Γ		the Gamma function	eq 1.19	10
θ		wind direction in degrees ($^\circ$)	eq 1.1	3
$\bar{\theta}$		mean wind direction in degrees ($^\circ$). Only defined for small sampling intervals ($\Delta t \leq 1\text{day}$)	eq 1.26	12
ρ		density of air	eq 1.9	7
σ	$\sigma_V, \sigma_u,$ σ_v, σ_w	standard deviation. Its subscript indicates which variable it is applied to (e.g. σ_V for wind speed, σ_w being derived from the Weibull distribution). Mostly in ms^{-1}	eq 1.16	10
Ω	Ω_i, Ω_k	weighted mean absolute error function for R and Q (or Q_i). Ω_k is that of the optimal $H_{s,k}$	eq 3.1	42

University of Cape Town

List of Figures

1.1	Diab map of wind power potential in South Africa	23
1.2	ESKOM revised wind atlas of 2001	25
2.1	The 17 10m SAWS weather stations considered for this study . .	32
3.1	3x2 classification of long-term NCEP surface winds	46
3.2	Frequency of occurrence map of the 3x2 classification	48
3.3	Standardised Ω scatter plot for all classifications	49
3.4	Aggregate of individual Ω curves for all subsets	50
4.1	Domain configuration of the Eastern Cape sensitivity runs	66
4.2	Standard minus PX surface roughness lengths for both domains .	70
4.3	Height above ground of the lowest half sigma level	73
4.4	Goal plot of c_V vs b_V for all LSM/PBL runs	77
4.5	Goal plot of CAE vs PPE for all LSM/PBL runs	80
4.6	Average diurnal cycles for all LSM/PBL runs	81
4.7	Goal plot of $MMAE$ vs $DFPD$ for all LSM/PBL runs	83
4.8	$WMAESSA$ vs $WMAESSD$ goal plot for all LSM/PBL runs .	85
4.9	$WMAEPD$ vs $WMAESF$ goal plot for all LSM/PBL runs . . .	88
4.10	Frequency windroses for the standard PX/ACM run	89
4.11	Goal plot of c_V vs b_V for all FDDA/cumulus runs	93
4.12	Goal plot of CAE vs PPE for all FDDA/cumulus runs	95
4.13	Average diurnal cycles for all cumulus/FDDA runs	97
4.14	Reference and 4xFDDA V , u and v East London daily cycles . .	98
4.15	Goal plot of $MMAE$ vs $DFPD$ for all FDDA/cumulus runs . .	100
4.16	$WMAESSD$ vs $WMAESSA$ goal plot for FDDA/cumulus runs	101
4.17	$WMAEPD$ vs $WMAESF$ goal plot for FDDA/cumulus runs .	103

5.1	Domain configuration for the South African Wind Atlas	108
5.2	Goal plot of c_V vs b_V for both wind atlas runs	112
5.3	Goal plot of CAE vs PPE for both wind atlas runs	116
5.4	Standard and 4xFDDA V , u and v spatial average daily cycles	117
5.5	Daily cycles for the six windiest stations	119
5.6	Daily cycles for the six intermediate stations	120
5.7	Daily cycles for the five calmest stations	121
5.8	Goal plot of $MMAE$ vs $DFPD$ for both wind atlas runs	123
5.9	$WMAESSD$ vs $WMAESSA$ goal plot for wind atlas runs	125
5.10	$WMAESF$ vs $WMAEPD$ goal plot for wind atlas runs	127
5.11	Average annual wind speeds at 10m above ground in ms^{-1}	131
5.12	Average seasonal wind speeds at 10m above ground in ms^{-1}	133
5.13	Power Curve and C_p values for a Vestas V80 2MW turbine	138
5.14	Road and power transmission network of South Africa	140
A.1	Goal plot of c_V vs b_V for all four combination runs	160
A.2	Goal plot of CAE vs PPE for all combination runs	162
A.3	Goal plot of $MMAE$ vs $DFPD$ for all combination runs	164
A.4	$WMAESSD$ vs $WMAESSA$ goal plot for combination runs	165
A.5	$WMAEPD$ vs $WMAESF$ goal plot for combination runs	166

List of Tables

2.1	Meta-data for the selected weather stations	32
3.1	Results for optimum period within individual classifications	51
3.2	$R - Q_k$ residual matrices for best and worst individual periods .	52
3.3	Uncertainty introduced by optimal time period	53
4.1	Land surface models available in MM5	60
4.2	High resolution PBL parametrisations in MM5	61
4.3	Cumulus parametrisations available in MM5	63
4.4	USGS 24 land use categories and their z_0 values	69
4.5	Configuration of the six LSM/PBL sensitivity runs	76
4.6	Configuration of the eight cumulus and FDDA sensitivity runs .	91
5.1	Bias statistics for the standard wind atlas run	113
5.2	10m uncertainty of the wind atlas results for each error statistic .	130
5.3	Wind power capacity factors for the 14 largest IEA markets . . .	141
5.4	Assumptions for three wind power penetration scenarios	142
A.1	Configuration of the four optimisation runs	161

University of Cape Town

Chapter 1

Introduction

1.1 Objectives and Scope

First and foremost this thesis is driven by two objectives in the context of wind power applications:

Primary: to explore in detail the value of using the regional climate model MM5 to develop a high resolution wind climatology of South Africa

Secondary: to map and quantify the wind resource of South Africa as a result

On the outset regional climate models have several advantages over traditional methods of analysing average wind conditions but are also widely known to suffer from various deficiencies. Thus the primary purpose above presents itself with several challenges that have to be carefully addressed. The sub-objectives of this point are:

- finding out whether and how a valid climatology may be generated by only integrating over a single year in order to minimise the computational requirements of the problem
- establishing the model sensitivity to different parametrisations and other options in terms of being able to reproduce the fundamental characteristics of near surface winds in both space and time
- judging the model's ability to faithfully reproduce the boundary layer air flows relevant for wind power by thoroughly measuring and understanding

the simulation biases and uncertainties using specialised error reporting techniques

The secondary purpose is in the domain of a traditional wind atlas. It is highly dependent on the primary in that it requires its output and needs the same comprehension of the model's limitations in order to be meaningful. It consists of two sub-goals itself:

- presenting and discussing maps of climatological wind speeds around South Africa in comparison to previous works to ascertain what new discoveries, if any, are made
- quantifying the amount of electricity that the country can potentially generate by converting the kinetic energy content of the wind

Much less emphasis is placed on the secondary objective in terms of detailed analysis because it is of much higher importance to build a solid mesoscale modelling foundation on top of which to conduct in-depth wind resource assessments.

The following sections of this introduction seek to inform the reader about the basics of wind, power extraction, relevant statistics, different types of wind resource assessments and previous work specific to South Africa. A brief description of the employed regional climate model is also given. The last section of this chapter provides an overview of how the above objectives are achieved by outlining the broad methodology as well as a chapter summary.

1.2 Wind Power Fundamentals

Before we perform a complete literature review it is necessary to lay down some fundamentals governing wind and the power that is contained therein, from basic definitions and terminology to important physical equations. This is done in the subsections that follow.

1.2.1 Terminology and Definitions

1.2.1.1 Wind Speed and Direction

A particular wind dataset at a particular location is a two dimensional series of individual “readings” which might be either calculated or observed. We write this series S as

$$S = ((V_1, \theta_1), (V_2, \theta_2), \dots, (V_n, \theta_n)) \quad (1.1)$$

where V is the speed (in ms^{-1}), θ is the direction (in degrees clockwise from North the wind is coming from) and n the length of the series. Individual values are addressed as

$$S_i = (V_i, \theta_i) \quad i \in \{1, n\} \quad (1.2)$$

As wind is a function of time we could also write

$$S_i = S(t_i) = (V(t_i), \theta(t_i)) \quad (1.3)$$

where t_i is a given time. In most cases $\Delta t = t_{i+1} - t_i$ is constant and Δt^{-1} is known as the sampling frequency.

In its most generalised form wind is actually also a function of space, both horizontal (x, y) and vertical (z) so that we arrive at the two four dimensional quantities

$$V = V(x, y, z, t) \quad (1.4)$$

$$\theta = \theta(x, y, z, t) \quad (1.5)$$

In the context of computational modelling the wind vector is most often not expressed in terms of a magnitude V and an angle θ but rather in terms of a West to East flow u and a South to North component v . Their relationship is given by

$$u = -V \sin \theta \quad (1.6)$$

$$v = -V \cos \theta \quad (1.7)$$

where the minus sign is necessitated by the fact that we considering the direction the wind comes *from*. For example, a southwesterly wind with $180^\circ < \theta < 270^\circ$ corresponds to positive u and v .

1.2.1.2 Scales of Space and Time

Any meteorological research operates on a certain physical boundary and with respect to this research four *scales of space* need to be defined:

micro scale, small scale, local scale order of metres to hundreds of metres

mesoscale, regional scale order of kilometres to tens of kilometres

synoptic scale, large scale order of hundreds of kilometers

global scale order of thousands of kilometers

The above definition is, however, subject to much debate and varies across scientific and commercial contexts. There are many wind energy consultants and wind simulation software providers who claim that 100x100m is mesoscale while some atmospheric scientists would consider 50km to be synoptic. Despite such discrepancies the above specification is used in this research; the grey areas shall not be of further concern as the spatial validity of the data at hand is what matters most.

We also need to define time scales before diving into the specifics of this research:

ultra high frequency order of microseconds to seconds

high frequency order of minutes

medium frequency order of hours

low frequency order of days

seasonal order of months

inter-annual order of years

decadal order of decades

Variations of the wind at ultra high frequency scales do not influence the electricity production of a modern wind turbine due to its large rotational inertia. Similarly, because the life span of a typical wind farm is 20 years, decadal fluctuations are not relevant either. Hence only high frequency to inter-annual fluctuations are relevant in this context.

1.2.2 Sources of Wind

Wind arises in principle from pressure differentials between two adjacent parcels of air. In order to correct this imbalance the air moves so as to decrease this gradient. On the synoptic scale this translates to a flow from high pressure to low pressure systems. The Coriolis force caused by the rotation of the Earth, however, deflects this wind to the left in the southern and to the right in the northern hemisphere. A balance between the two forces is reached once the flow follows the isobars, a condition known as the *geostrophic wind*, an idealised and theoretical construct which nevertheless determines a large part of the wind's variability. Due to the nature of evolution of such large scale occurrences the geostrophic wind varies with low frequency.

On the other hand there are dynamic mesoscale characteristics with medium frequencies of the wind which are of major importance to wind energy but are often overlooked in traditional treatments of the subject. The most important phenomenon here is the daily cycle of the wind. There are two kinds of 24h cycles that are frequently observed: land/sea breezes and topographically induced (slope) winds [Peixoto and Oort 1992].

A sea breeze develops during the day near the boundary of an ocean or a very large water body and a large land mass. The sun heats up the surface and causes the warm air to rise. The ocean surface's temperature on the other hand stays relatively constant and cool due to its much larger specific heat. Since air density is inversely proportional to its temperature this creates a positive pressure differential between the sea and land air. This sets up a circular motion of air from the sea to the land where it rises and at altitude moves back towards the sea to sink there and start anew. The strength of such flows is directly proportional to the difference in temperature between the two areas. Many efforts have been made in the past to represent such systems accurately in numerical forecasting [Simpson 1994].

A land breeze occurs at night in the opposite direction, i.e. from the land to the sea near surface, caused by relatively warmer sea air. As the magnitude of the nocturnal land/sea temperature differential is typically much smaller and because of the greater static stability that prevails at night the land breeze is less important to local wind characteristics.

Slope winds are a lesser form of this phenomenon occurring in mountainous terrain. During the day the air above the high altitudes heats up much quicker than the air around it at similar altitudes while the valleys below are

often shaded, creating a pressure differential and causing a breeze up the slope (*anabatic* winds). At night, the opposite again holds true and the air sinks downhill (*katabatic* winds). Since South Africa has many mountain ranges and regions of up to 3400m amsl slope winds can become quite significant [Tyson and Preston-Whyte 2000].

While at stronger geostrophic winds neither of these localised daily cycles tend to develop, it has been shown that they nevertheless contribute significantly to the total wind available at any given site. At some coastal sites of South Africa the 24h component can easily account for a large percentage of the total variance of the wind and many of these locations have very clearly defined average daily wind speed fluctuations [Hagemann 2003]. In fact, Alexander Bay for example has a difference between minimum and maximum hourly wind speeds in a day of no less than 9 ms^{-1} [Diab 1995].

1.2.3 Important Surface Characteristics Affecting the Wind

There are some very important land surface characteristics that can have a large impact on how the wind behaves in a local or even regional environment. These phenomena affect the planetary boundary layer only, that is, the first kilometre or so above the surface. This discussion of such influences is kept on a summary level; the reader is referred to more detailed treatments of this subject in the literature upon which this review is largely based [Diab 1995; Peixoto and Oort 1992; Troen and Petersen 1987].

Firstly, the roughness of a given land surface caused by buildings, bushes, trees, etc. causes the wind to be decelerated. This effect increases as one moves closer to the surface. It can be quantified by using a parameter known as the *roughness length* of the surface z_0 , a quantity that may be empirically determined by measuring vertical wind profiles at certain stability conditions. It is used in an approximately logarithmic behaviour of the variation of the wind speed $V(z)$ with height z according to

$$V(z) = \frac{1}{k_0} V_* \ln \frac{z}{z_0} \quad (1.8)$$

where k_0 is the von Kármán constant (≈ 0.4) and V_* is the so-called friction velocity (typically $0.2\text{-}0.4 \text{ ms}^{-1}$) related to surface stress (and thus vertical momentum flux) and the air density [Justus and Mikhail 1976]. This equation is valid for neutral stability regimes only, neglects the zero plane displacement and

hence is the most basic method for vertical wind extrapolation.

Secondly differences in terrain elevation can also have a large impact on the wind speed. As the wind moves over hills and mountains, the flow can be redirected, separated, accelerated, slowed down, form eddies or, in general, a combination of these, depending on the exact configuration of the local topography. For example, on top of an idealised ridge the wind can be twice as strong as at its base. Valleys, on the other hand, are usually calmer and are thus avoided for the erection of wind turbines.

Such terrain effects together with roughness influences represent the most important localised wind perturbations that all modelling efforts must reproduce as accurately as possible, as explained later in section 1.3.

Lastly, obstacles such as large buildings can have significant microscale sheltering effects on the wind. It is found that their presence can influence its speed and direction up to three times the height of the object vertically and up to 40 times this height in the downwind direction [Diab 1995]. Therefore wind measurements in the vicinity of such structures are often very unreliable. Special techniques and algorithms are available to calculate and account for obstacle induced disturbances [Perera 1981]. However, this study is concerned with the mesoscale wind conditions and hence these are of no further interest in this context.

1.2.4 Energy Conversion of a Wind Turbine

The total power content contained in wind varies with the third power of the magnitude of the wind, i.e.

$$P_d = \frac{1}{2} \rho V^3 \quad (1.9)$$

where P_d is the power density (in Wm^{-2}) and ρ the density of air. This is known as the *third power law* and can easily be derived using simple Newtonian mechanics [Justus 1978].

Only about 59% of the theoretically available power density may be converted to rotational or (eventually) electrical energy according to the Betz efficiency limit [Betz 1966]. But over and above that each wind turbine has its own mechanical and aerodynamic losses so that for real world problems equation 1.9 is written as

$$P_d = \frac{1}{2} C_p(V) \rho V^3 \quad (1.10)$$

where $C_p(V)$ is a wind and generator model dependent power coefficient, a simple measure of efficiency, i.e. how much theoretically available power is converted to usable electricity. In practice every wind turbine does not operate below (“cut-in”, V_{in} , typically around 4ms^{-1}) and above (“cut-out”, V_{out} , typically around 25ms^{-1}) certain wind speed thresholds for mechanical reasons where C_p is effectively zero. Between those limits $C_p(V)$ is a smooth curve smaller than the Betz criterion with its maximum between V_{in} and the rated wind speed V_{rated} which corresponds to the peak power output P_{rated} of the machine. The reader is referred to Figure 5.13 on page 138 for a C_p curve of a typical modern wind turbine.

What is important here in the context of this research is that in terms of a wind climatology we are only interested in the occurrence of winds between V_{in} and V_{out} , the rest is irrelevant for electricity generation.

The air density ρ is dependent on atmospheric pressure p , absolute temperature T and mixing ratio q as determined by

$$\rho = \frac{p M_a (1 + q)}{R_a T (1 + 1.608q)} \quad (1.11)$$

where M_a is the molar mass of dry air (0.029kgmol^{-1}) and R_a is the perfect gas constant ($8.314\text{JK}^{-1}\text{mol}^{-1}$) [Guyot 1998]. At standard temperature (15°C) and pressure (1013.25Pa) with a mixing ratio of zero (dry air) the density is 1.2265kgm^{-3} accordingly.

Just like V the quantities q , T and p are four dimensional, i.e. functions of space and time. Therefore P_d and ρ are too but seeing that neither T nor p vary much in time on an absolute scale at a given point ρ also stays fairly constant (varies by no more than a few percent in the near-surface layer). In fact this is the reason wind power calculations usually approximate $\rho(t)$ by its time average $\overline{\rho(t)}$ [Burton 2001].

Now in order to arrive at the total amount E that is converted from kinetic to electrical energy we have to integrate equation 1.10 at a given point (x,y,z) in space using

$$E = \frac{1}{2} A_r \int_{t_1}^{t_2} \rho(t) C_p(V(t)) (V(t))^3 dt \quad (1.12)$$

where t_1 and t_2 are the start and end times of the generation period and A_r is the rotor swept area orthogonal to the wind direction given by

$$A_r = \pi \left(\frac{D}{2} \right)^2 \quad (1.13)$$

D being the rotor diameter (twice the radius). Normally $t_2 - t_1$ is set equal to one year and then E is referred to the “Annual Energy Production” or AEP of a given wind turbine at a specific site. Note, however, that the AEP is usually calculated based on a statistical distribution rather than an integration over time [Burton 2001]. Also, the term AEP is - strictly scientifically speaking - incorrect since energy can never be *produced*, only converted from one form into another, but it is so common in the industry that we shall use it regardless.

Finally equation 1.12 on the previous page implies that the wind turbine on average only operates at a fraction of its maximum power output P_{rated} . To measure this one uses the capacity factor C_f defined by

$$C_f = \frac{E}{P_{rated}(t_2 - t_1)} 100\% \quad (1.14)$$

which is a percentage of the actual energy output compared to what the turbine would have generated if it ran at peak power for the entire period ($t_2 - t_1$). Clearly one has to be careful here in terms of units, best is to use hours and W throughout.

In either event it should be clear that the cubic dependence implies that a small change in wind speed ΔV has a large impact on P_d , the related energy conversion E , capacity factor C_f and ultimately the economic feasibility of wind farms as developers sell the electricity into the grid. It is therefore imperative that areas with high wind resources are identified, quantified and mapped accurately in order to elucidate particular regions of interest. The non-linearity also makes it more difficult to obtain meaningful statistics as we shall explore further in the next section.

1.2.5 Essential Statistics for Wind Power

When talking about wind data at a site the most frequently quoted statistic is the average wind speed $\bar{V}$ defined simply as

$$\bar{V} = \frac{1}{n} \sum_{i=1}^n V_i \quad (1.15)$$

with its standard deviation defined as usual like

$$\sigma_V = \sqrt{\frac{1}{n} \sum_{i=1}^n (V_i - \bar{V})^2} \quad (1.16)$$

In order to get a representative figure for the mean n should be large and span as many years as possible, ideally a decade or more to smooth out variability at all relevant time scales, while Δt^{-1} should be of high frequency to ensure that short term fluctuations are properly accounted for.

Although the arithmetic mean remains an important metric when gauging the potential of a given site, by itself it would only be adequate as a useful criterion if wind speeds were normally distributed. It turns out, however, that the magnitude of the wind is a Weibull variate which is asymmetric in probability about its mean [Justus and Mikhail 1976]. Combined with the cubic dependence described by equation 1.9 the exact nature of the distribution becomes more important than a plain average.

The Weibull cumulative distribution function(cdf) $F_w(V)$ and probability density function(pdf) $f_w(V)$ are defined by

$$F_w(V) = 1 - e^{-\left(\frac{V}{A}\right)^k} \quad (1.17)$$

$$f_w(V) = \frac{k}{A} \left(\frac{V}{A}\right)^{k-1} e^{-\left(\frac{V}{A}\right)^k} \quad (1.18)$$

where again V is the wind speed (in ms^{-1}), A is a scaling parameter (also in ms^{-1}) strongly related to the mean and k is a dimensionless shape parameter [Evans et al. 2000]. This is in fact a very general distribution, which the exponential and Rayleigh distributions are merely special cases of with $k = 1$ and $k = 2$ respectively. In some treatments of wind the Rayleigh distribution is applied though due to its limitations it will not be used in this research.

The mean $\bar{V}_w$ of the Weibull distribution, or first moment of its pdf about the origin, can be calculated using

$$\bar{V}_w = A\Gamma\left(1 + \frac{1}{k}\right) \quad (1.19)$$

Γ being the gamma function. Also of interest though less important in this

discussion is the standard deviation given by

$$\sigma_w = A \sqrt{\Gamma\left(1 + \frac{2}{k}\right) - \Gamma^2\left(1 + \frac{1}{k}\right)} \quad (1.20)$$

It should be noted that equations 1.17 and 1.18 do not allow for calms, i.e. zero wind speeds since $\lim_{V \rightarrow 0} f_w = 0$. As some datasets contain lots of zeros, this assumption breaks down and is often corrected using the three parameter Weibull distribution expressed as

$$F_{wH}(V) = 1 - (1 - F_0)e^{-\left(\frac{V}{A}\right)^k} \quad (1.21)$$

$$f_{wH}(V) = (1 - F_0) \frac{k}{A} \left(\frac{V}{A}\right)^{k-1} e^{-\left(\frac{V}{A}\right)^k} \quad (1.22)$$

where F_0 is now the non-zero probability of calms [Takle and Brown 1978; Tuller and Brett 1983]. The mean then becomes

$$\bar{V}_{wH} = (1 - F_0)A\Gamma\left(1 + \frac{1}{k}\right) \quad (1.23)$$

By fitting the data at hand to one of these equations as appropriate, we can also obtain other statistical properties such as higher order moments (e.g. skewness and kurtosis). Therefore we can describe most of the wind statistics in terms of only two or three parameters.

Apart from such time independent statistics there are also important dynamic characteristics such as the daily cycle which must be taken into account as noted in section 1.2.2 on page 5. To quantify this we first need to calculate a new 24h wind speed series V_{24} as follows

$$V_{24} = \frac{24}{n} \left(\left(\sum_{i=1,25,49,\dots}^n V_i \right), \left(\sum_{i=2,26,50,\dots}^n V_i \right), \dots, \left(\sum_{i=24,48,72,\dots}^n V_i \right) \right) \quad (1.24)$$

assuming that $\Delta t = 1\text{h}$ which, in essence, is nothing other than the 24 element series consisting of averages at every hour of the day. Small adjustments have to be made to 1.24 to account for series that do not start at midnight and end at 23:00.

After we have visually verified that a clear cycle exists we may define the

amplitude D_a as the difference between maximum and minimum of V_{24} 's elements

$$D_a = \max(V_{24}) - \min(V_{24}) \quad (1.25)$$

while the time the peak occurs shall be denoted by t_{max} . In order to obtain even more information one breaks this up further into seasonal 24h oscillations provided the timeseries is long enough.

Both D_a and t_{max} are essential wind statistics for a wind farm operator as it determines how closely the available wind matches supply and demand patterns on the electrical grid. As electricity is often more valuable when sold during load peak (late afternoon to early evening) this has a major influence on the commercial viability of a given project.

Another technique that presents us with useful information on the significance of the diurnal cycle is the calculation of the Fourier transform $G(f)$ of $V(t)$, f being the frequency. At $f_d = 1\text{day}^{-1}$ the value of $G(f_d)$ should be much larger than its surrounding noise for any coastal site. The difference of $G(f_d)$ for two different locations or better, model vs station as seen in subsection 2.2.2, can speak volumes.

We can also take a closer look at seasonal time scales to obtain the 12 month series V_{12} whose values consist of the averages of a given month (1-12). While there is no specific scalar quantity to be derived from V_{12} , a graphical representation of it can deliver vital information on how well the wind regime matches the seasonal demand patterns of high load in winter and lower strain on the grid in summer, with similar implications for the viability of any given wind farm.

So far we have only discussed statistics on the wind *speed*, though any treatment of the subject would be incomplete without an in-depth look at the wind *direction*. Processing of θ is not as straightforward because it is a *circular* variable, i.e. $\theta \in \{0, 360\}^\circ$ but it wraps around 0, so if $\theta_1 = 359^\circ$ and $\theta_2 = 1^\circ$ then their difference $\Delta\theta = 2^\circ$, not 358° , while their average is 0° , not $\frac{1^\circ + 359^\circ}{2} = 180^\circ$.

The study of circular statistics is devoted to such variables and there are plenty of available approaches to deal with them in the literature [Gaile and Burt 1980; Fisher 1993; Mardia 1975]. The standard way define an average direction is in terms of the sums of its x and y components, that is

$$\bar{\theta} = \arctan \left(\frac{\sum_{i=1}^n \sin \theta_i}{\sum_{i=1}^n \cos \theta_i} \right) \quad (1.26)$$

Despite its elegance equation 1.26 on the preceding page only really gives meaningful results for a small sample of wind directions within one day or two, assuming similar conditions. So an entire day of NNW (338°) to NNE (23°) flows would be correctly averaged to $\theta = 0^\circ$.

However, consider a 10 year long series at a site at which the dominant wind direction in summer is SE (135°) and NW (315°) in winter with all other directions hardly ever occurring. Taking the mean according to equation 1.26 on the previous page would yield SW (225°) as the predominant wind direction which certainly has no resemblance with reality. In a case like this it would be wiser to state that a mean direction is not defined.

Even more difficult is a definition of standard deviation for a circular variable. Many authors have attempted to arrive at useful algorithms though opinion remains divided over which one the best and no single approach has to date been agreed upon [Bruce Turner 1986; Mori 1987; Fisher 1987]. It is for this reason that a treatment of standard deviation of wind direction is not carried out in this research.

One way out of this dilemma would be to revert to the use of u and v components of wind directly as both of these are linear and distributed in some way around 0. That, however, implies losing the benefits of the often described and standardised Weibull treatment of wind speed as per equations 1.17 to 1.20. Nevertheless, parts of this thesis will look at u and v individually to complement other statistics.

To get a better idea of where the wind typically comes from the series of wind *directions* is often binned in directional sectors of equal width, customarily 30° resulting in 12 sectors as done in the European Wind Atlas [Troen and Petersen 1987]. The resulting distribution is then plotted on so-called *rose diagrams*, also called *rose plots* or *wind roses*. We shall encounter examples of these later in chapter 4 on page 57 but for now we are concerned with the principal data they contain.

Firstly and most fundamentally we can define the frequency of winds occurring from a certain sector giving rise to the 12 value distribution F_θ as

$$F_\theta = \frac{100}{n} \left(\left| \{ \theta | 0^\circ \leq \theta < 15^\circ \} \cup \{ \theta | 345^\circ \leq \theta \leq 360^\circ \} \right|, \quad (1.27) \right. \\ \left. \left| \{ \theta | 15^\circ \leq \theta < 45^\circ \} \right|, \dots, \left| \{ \theta | 315^\circ \leq \theta < 345^\circ \} \right| \right)$$

where the vertical bars around the sets denote mathematical cardinality of each sector which gives us a percentage if multiplied with $100n^{-1}$. The above angle specification ensures that the sectors are centred around the angle which they're supposed to represent, such that e.g. North is defined as being in the $[345^\circ, 15^\circ)$ range. Unfortunately a 12 sector treatment does *not* map to the usual N, NNE, NE, ENE, E... classification one might intuitively expect since the latter would require 16 sectors, however this research bows to convention as the wind industry takes 12 as standard.

Then there is the distribution $F_{\overline{V}}$ of mean wind speed in each directional bin which can be defined analogously to equations 1.15 and 1.27 as

$$F_{\overline{V}} = \left(\overline{\{V|0^\circ \leq \theta < 15^\circ\} \cup \{V|345^\circ \leq \theta \leq 360^\circ\}}, \right. \quad (1.28) \\ \left. \overline{\{V|15^\circ \leq \theta < 45^\circ\}}, \dots, \overline{\{V|315^\circ \leq \theta < 345^\circ\}} \right)$$

where the over line denotes the average of each set. It must be noted that there are cases where the mean wind speed of a particular sector is high but its frequency of occurrence is low or vice versa, thus making $F_{\overline{V}}$'s importance for wind power applications somewhat questionable, at least on its own. Therefore for any given sector a combination of both frequency and average magnitude must always be considered in order to gauge its significance.

We may also define an effective wind rose of wind speed standard deviation F_{σ_V} and 1.27

$$F_{\sigma_V} = \sigma \left(\{V|0^\circ \leq \theta < 15^\circ\} \cup \{V|345^\circ \leq \theta \leq 360^\circ\}, \quad (1.29) \right. \\ \left. \{V|15^\circ \leq \theta < 45^\circ\}, \dots, \{V|315^\circ \leq \theta < 345^\circ\} \right)$$

where we used the shorthand $\sigma()$ to denote the standard deviation function 1.16 on page 10 as applied to each set. F_{σ_V} provides us with a good indication of which sectors have fairly steady winds and which have a wide range of magnitudes.

We can further transform $F_{\overline{V}}$ into a power density distribution F_P using equation 1.9 by

$$F_P = \frac{1}{2} F_{\rho V^3} \quad (1.30)$$

which helps to place more emphasis on the significant sectors in a wind rose. ρ should as indicated also be averaged over time but according to equation 1.11 on page 8 the absolute values of ρ typically do not vary by much (of the order of 10%) so it is often approximated as a constant.

Finally as yet another combination of these fundamental statistics the Weibull distribution of wind speeds is often computed for each sector. As long as the goodness of fit is high (e.g. by standard chi-square statistics) we can again pull out mean, standard deviation and a host of other results directly from the two parameters A_θ and k_θ , an approach often taken in the wind industry.

1.3 Wind Resource Assessment Methodologies

1.3.1 Nature of Assessments

Wind resource assessments are nothing less than a quantification of an area's or a location's wind potential in terms of the essential wind statistics introduced in section 1.2.5. Site specific studies are quite different from their region wide counterparts, not only in scale but in many other respects too. As such they are discussed separately in the next two sections. The reader is strongly urged to review section 1.2.1.2 for our definition of spatial scales.

As becomes clear from subsection 1.2.3 individual site measurements must be considered microscale. Beyond a few metres vertically and perhaps a few tens of metres horizontally such data are no longer representative in hilly or even mountainous terrain. In most practical applications, however, the exact 3 dimensional position of the anemometer is unlikely to be the only place where a wind turbine will be erected as a typical wind farm consists of several machines.

This is why a lot of research has focused on how best to extrapolate a point measurement to yield good estimates for the wind conditions of a larger area. Some common methods to achieve this are discussed in more detail in subsection 1.3.2. Taking the opposite, top-down approach from a relatively reliable synoptic or global scale dataset to a regional or even micro scale has become increasingly possible with the exponential growth of computational resources in recent decades. This is explained in more detail in subsection 1.3.3. It needs to be noted though that due to the complexity of the wind, its potentially large

temporal and spatial variations as well as the high costs of measurement campaigns none of these techniques can avoid resorting to some form of numerical modelling.

Finally, a *wind atlas* in this context is meant to be a collection of wind statistics for a geographical region covering one or more countries. This is the intuitive interpretation and is the one implied in country or even continent specific studies. Unfortunately the term is sometimes confusingly assigned more location specific connotations as the reader will see below.

1.3.2 Assessments Based on Measurements

The easiest way to obtain reliable wind data for a specific site is to erect some measurement instrumentation at the desired height above ground. For meteorological applications the World Meteorological Organisation (WMO) recommends 10m though for agricultural purposes as low as 2m is often implemented [Oke 2004]. For wind energy applications, in turn, the ideal height is that of the towers of the planned wind turbines (60-120m). Given the complexity of such high masts 50m or 60m is often used though in some countries like Germany or the USA they are often much higher.

In terms of instruments the wind industry still predominantly uses the classic cup anemometers with three or more plastic or metal cups rotating in the horizontal plane. These measure speed only, a separate wind vane is required for wind direction. A similar setup is not uncommon in weather observations, though often more compact devices such as propeller anemometers are employed which do not necessitate a separate vane. In either case, regular calibration of the equipment is important in order to ensure the accuracy thereof. Many newer and more expensive solutions exist and include those based on ultrasonic, SODAR and LIDAR technologies but are mostly only deployed in specialised experiments [Antoniou et al. 2004].

As mentioned before the spatial validity of empirically derived data are somewhat limited and techniques need to be devised in order to make extrapolations, even in the case of a large number of measurements already present. This is of utmost importance for areas without any pre-existing weather stations. Some works solely rely on limited vertical extrapolation based on equation 1.8, “experience and professional insight” to make subjective judgements about the horizontal distribution of the wind. Examples are the classification schemes that have been used extensively on the North American continent [Ilinca et al. 2003;

Elliott et al. 1987].

In order to extrapolate more rigorously though some modelling must be carried out which most of the time implies having to solve the physical equations governing fluid flow, namely the Navier-Stokes equations. Owing to their complexity this cannot be done analytically for almost all real-world problems and thus numerical integration must be carried out. This topic is commonly called computational fluid dynamics or short CFD.

Various approaches to simulating the behaviours of flow over complex terrain exist and have been around for decades [Wood 2000]. Most commonly the Navier-Stokes equations are simplified to such an extent that they can be regarded as linear and thus relatively fast and inexpensive to solve. Perhaps the most well-known software package that achieves this is WAsP (Wind Atlas Analysis and Application Program) developed at the Danish research institution Risø. WAsP is based on the well known principles first outlined by Jackson and Hunt and uses very simple stability corrections [Jackson and Hunt 1975]. For over 20 years, ever since the publication of the European Wind Atlas this software package has been used in countless local and regional studies including the Wind Atlas of South Africa [Troen and Petersen 1987; Diab 1995].

WAsP's fundamental concept is as follows: It takes several micro scale factors such as nearby obstacles, surrounding terrain roughness and topography into account in order to subtract such influences from the data to arrive at a *generalised regional wind climatology* which is meant to be representative on the mesoscale [Troen et al. 1987]. This climatology can then be spatially transferred to another nearby location (on the same scale) where the reverse process is applied: site specific surface characteristics are applied to the "cleaned" data in order to arrive at micro scale statistics. WAsP has 5 sub-models for this purpose, the core of which is the linearised Navier-Stokes solver with the others being completely or partially empirical models.

A collection of such general and site-independent information can then be termed as a wind atlas, most of the time combined with maps of the region with coloured contours of average wind speed and other parameters. However, in WAsP terms a generalised climatology is - somewhat confusingly - also referred to as a wind atlas [Langhans 2003]. The latter meaning is not used any further from now on except when explicitly referred to as a "WAsP wind atlas".

The largest and perhaps most critical limitation of linear models like WAsP is that it is valid only for mostly flat terrain. As soon as surface elevations of about 5% or more are encountered - the case with most hilly and especially

mountainous terrain as found in South Africa - the assumptions break down and large errors can arise due to separation of flow [Llombart et al. 2006]. Unfortunately quite a few consultants and professionals in the industry are unaware of this fact and therefore some studies have yielded unexpected results. An example of such discrepancies is given in subsection 1.4.2 on page 24.

To overcome this problem a more general (non-linear) treatment of the Navier-Stokes equations becomes necessary and coupled with recent advances in raw computational power fully three dimensional algorithms and software packages such as Windsim and Meteodyn have been specifically developed for micro and mesoscale wind resource research [Delaunay 2007; Gravdahl and Harstveit 2000]. They are by necessity much more complex and demanding than WAsP but can offer better estimates of certain wind statistics in real world terrain [Moreno et al. 2003].

The main usage of all these techniques is for site assessments as a crucial part of the standard wind project development cycle. The following basic principle applies: Take some reliable measurement data from as close to the site as possible, input the exact coordinates and hub heights of the planned wind turbines and apply some modelling in order to work out what the wind conditions should be at those positions. Combined with the wind turbine's power curve (power output as a function of wind speed) one would then calculate the expected annual energy production of the envisaged wind park.

The secondary and in this context much more important utilisation is for wind atlas generation. As already stated above WAsP has been used many times for this purpose before, in fact for dozens of national atlases all over the world which include, but are not limited to, the Russian Wind Atlas, the Wind Atlas for Sweden and the Wind Atlas of Algeria [Starkov et al. 2000; Hammouche 1991; Krieg 1992]. The non-linear counterparts, in turn, have only been used in very few wind atlas studies such as the Norwegian Wind Atlas and the Umbria Region Wind Atlas for that part of Italy [Gravdahl 2003; F.Castellani and Franceschini 2005]. The reason is that these techniques are only a few years old and also were never designed with country wide, large scale wind atlas applications but rather site assessments in mind.

What is important to note in summary is that all of these techniques *rely directly on measurements* and cannot be applied meaningfully without them. This limits them to countries or regions in which a dense network of weather stations or other public domain data is readily available. Thus this type is generally not applicable to continents such as Africa where this requirement

cannot be fulfilled.

1.3.3 Assessments Based on Mesoscale Modelling

This method is the alternative to the above and *does not directly rely on measurements* to generate regional wind resource assessments. Besides it offers a number of other advantages which will be outlined later. The principle here is, as mentioned in subsection 1.3.1 on page 15, the other way round: instead of starting from an observation and applying some CFD to obtain regionally representative statistics (*up-scaling*) this approach takes a global scale dataset to initialise and drive a mesoscale climate model at its boundaries in order to arrive at smaller scale, usually regional or country-wide results (*down-scaling*).

The model itself is usually more complex and comprehensive in its treatment of atmospheric variables and interactions. Instead of taking only incompressible fluid flow with its three dimensional components as its only dynamic quantity like the non-linear CFD algorithms mentioned above, climate models fully solve for all other fundamental quantities such as pressure, temperature and different forms of water (vapour, liquid and ice) while handling many more boundary conditions (e.g. surface heat and moisture fluxes in addition to roughness). That of course makes them more computationally intensive and thus these models are typically run at much lower resolution.

The software packages available for this purpose range from public domain mesoscale models such as the Weather Research and Forecasting Model (WRF) and its predecessor the Fifth Generation Penn State/NCAR Mesoscale Model (MM5) to the more restrictive models like the High Resolution Limited Area Model (HIRLAM) which run almost exclusively at certain national weather offices [Skamarock et al. 2005; Grell et al. 1995; Burridge et al. 2005]. They are used in operational forecasting and numerical weather prediction but are also extensively applied in academic contexts and climate research.

These models are forced on the boundaries and initialised by global gridded datasets which tend to have a very significant impact on the resulting wind climate [Pryor et al. 2005]. The most common sources in use today are the National Centers for Environmental Prediction/National Center for Atmospheric Research (NCEP/NCAR) Reanalysis Project and the European Centre for Medium-Range Weather Forecasts (ECMWF) [Kalnay et al. 1996; Kallberg et al. 1997]. NCEP /NCAR data is typically available at $2.5^\circ \times 2.5^\circ$ resolution or about 278km at the equator while ECMWF offers resolutions of up to 50km.

Which one of these is more suitable for wind energy purposes is unclear so some scientists recommend using both in ensemble techniques for short term wind power output prediction [Giebel et al. 2003; Nielsen et al. 2004]. ECMWF is, however, difficult and expensive to acquire while NCEP/NCAR is freely available and has already been archived at the author's institution and therefore forms the natural choice as a basis for this research.

Both assimilate data from a vast network of both surface (land and sea) and upper air observations using advanced objective analysis, quality control and error mitigation techniques. So in essence measurements do find their way into mesoscale modelling applications but only *indirectly*. Besides waiving the requirement for site observations, it becomes obvious that the model has the potential to deliver useful data even for areas where no weather stations are available. Therefore for a country like South Africa with a fairly limited observation network this method becomes highly attractive.

Wind atlases that have been based on mesoscale modelling efforts have become more common in recent years, again due to the vast progress made in raw computational power. Examples include the Brazilian Wind Atlas Project and a wind climate study for Ireland [Feitosa et al. 2001; Frank and Landberg 1997]. The former employed MM5 while the latter uses perhaps the most well known model for this purpose, the Karlsruhe Atmospheric Mesoscale Model (KAMM) [Adrian and Fiedler 1991]. KAMM has, in combination with WAsP processing of complementary observations, been made a standard tool by Risø [Frank et al. 2001].

1.3.4 Steady State vs Fully Dynamic

Most CFD algorithms and mesoscale modelling methodologies to date are based on *steady state* concepts. In this context this means that the simulated variables are only space (3D) but not (or only to a very limited extent) time dependent (4D). This often described “statistical dynamical approach” involves the selection of typical atmospheric states by means of a cluster analysis or classification scheme, complete with respective frequencies of occurrence Frank and Landberg [1997].

For each case a mesoscale model is then driven with constant boundary conditions and run either until the integrations converge to arrive at a constant solution or simply for a short integration period (e.g. 6h). All the data thus collected is then weighted accordingly to yield overall statistics of wind speed

and direction. Another good example of this is given by Mengelkamp for a particular region of Germany [Mengelkamp 1999].

This has the added benefit of significantly cutting down on the required simulation time and thus computer resources. However, the major disadvantage here is that some very important dynamic effects are completely neglected, most notably the daily cycle of the wind. The latter can - as has been pointed out before - have a very significant impact on the wind patterns of any particular site, especially near the coast where land and sea breezes are common [Ayotte et al. 2001; Hagemann 2003; Källstrand et al. 2000].

Furthermore for commercial and power engineering reasons, electricity utilities are most interested how closely the daily evolution of the wind resembles the demand patterns on their grid. If the highest wind speeds typically occur at times of peak usage such as the late afternoon and evenings, wind energy projects can ease supply shortages which are currently experienced in countries like South Africa at the time of this writing.

Therefore it becomes necessary and important to include the treatment of such dynamic features of the flow and this can only be done by performing fully dynamic, time dependent three dimensional experiments. This in turn raises a different problem: In order to capture inter-annual variability and obtain a representative climatology one needs to span a very long period of time, typically of the order of ten years or more. The strain on available computer capacity, however, is enormous if any resolution useful for a wind atlas is desired.

There are two standard ways to deal with this: Either use the world's most powerful supercomputers or restrict the domain in question. In the case of South Africa neither alternatives are feasible so for this research a third, innovative solution is sought: Intelligently reduce the integration period to a single year *without* sacrificing long-term representation. It is important to note that this is where most of the uniqueness in this thesis lies - a detailed description can be found in chapter 3 on page 41.

The chosen method can be thought of as a trade-off between a 10-20 year simulation and the steady state approach described above. While our way introduces some more redundancy than the latter since certain synoptic states and their influence on smaller scales are calculated several times instead of just once, the added value of fully and consistently resolving medium to low frequency features (in a time sense) cannot be disputed.

1.4 Wind Resource Assessments in South Africa

1.4.1 Up to 1995

The majority of the work in this period can be attributed to Diab who started with the first South Africa specific study in the late seventies [Diab 1979]. Since then, she published several general and province specific research [Diab 1985, 1981] culminating in her most widely quoted work in 1995, the Wind Atlas of South Africa prepared for the Department of Minerals and Energy (DME) [Diab 1995].

It is a classic measurement based country wide assessment that contains detailed descriptions of and statistics for no less than 79 individually screened long term weather stations from various different geographical settings around South Africa and various institutional data sources (mostly SAWB). It presents mean wind speed, frequency distributions, fitted Weibull A and k parameters for each 30° sector and overall values as well as the average monthly and daily wind fluctuations in tabular and graphical form.

The stations' anemometer heights above ground vary between 2 and 23m and a conservative vertical extrapolation based on equation 1.8 was used in order to normalise the figures to 10m. 6 of these stations had some WAsP modelling applied to them and the figures are quoted in standard WAsP format with 4 roughness classes.

Diab goes further to classify South Africa into areas of good ($\bar{V} > 4\text{ms}^{-1}$ and $P_d > 200\text{Wm}^{-2}$ at 10m agl), moderate and low wind power potential. The former consists of the entire coastal belt as well as the Drakensberg escarpment while the latter includes most of the highveld. Based on "subjective" considerations Diab produces a national wind resource map as shown in Figure 1.1.

To date this atlas remains the most authoritative, comprehensive and publicly available country wide resource assessment and has been used as the foundation for many related renewable energy and wind power studies [Brebøl et al. 2003; Kotzé 1998; Schäffler 2001; van der Linde 1996].

However, the reliability and integrity of its figures are somewhat questionable. Despite a detailed selection process the atlas over-relies on technologically outdated Dines anemometers as used at the majority of considered stations, most notably the South African Weather Bureau ones [Hagemann 2003]. Due to their high mechanical inertia these instruments were inaccurate and incurred a high frequency of recorded calms which skews the distributions and means

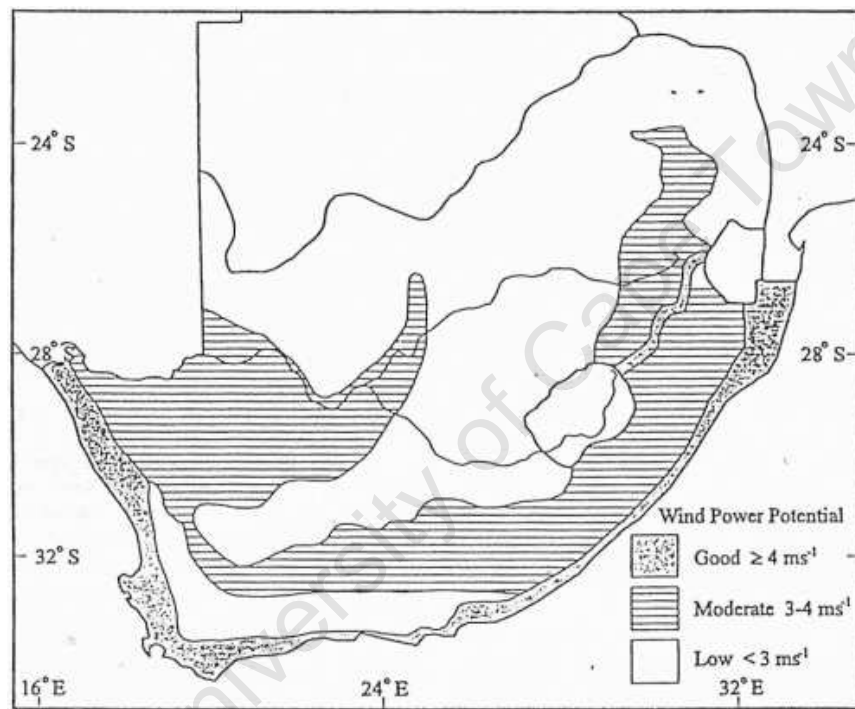


Figure 1.1: Generalised map of wind power potential in South Africa after [Diab 1995]

significantly [Hagemann 2003]. Also 45 of the stations had their anemometers mounted lower than 10m above ground causing them to record winds affected by unnecessarily high surface friction and related turbulence.

These measurements were also written down by human operators at specific instants in time which not only unnecessarily introduced data capture errors into the datasets (e.g. misplaced decimal points, mixed up digits etc.) but also lacked important short term smoothing of high frequencies. Apart from this no analytical and objective methods appear to have flown into Figure 1.1 to the left as the creation of it is completely undocumented. Hence its accuracy leaves room for improvement.

The only other study in this time period worth mentioning is the work of Dutkiewicz et al which focuses mostly on individual wind measurements and not on broad wind resource assessments [Dutkiewicz and Denison 1991]. As such it is not relevant to this thesis and will not be further commented on.

1.4.2 Subsequent Works

After 1995 Diab did not publish any further works on the subject of wind power. The majority of subsequent studies were conducted as part of locally and internationally sponsored projects run by consultants. Unfortunately most of these are not in the public domain such as the report of the UK based consultancy Garrad Hassan of 1999 and the work by the Centre for Scientific and Industrial Research(CSIR) and ESKOM on the “South African Renewable Energy Database” (SARERD) [Fellows and Gow 1999; D’Abreton and Law 1999]. As such they cannot be studied in detail by the author.

Given the unreliable nature of the first wind atlas, ESKOM, CSIR, the DME and DANCED (Danish Co-operation for Environment and Development, now DANIDA) jointly funded a revised wind atlas around 2001 which was supposed to supersede Diab’s work and overcome some of its problems [Ragoonathun et al. 2001]. This “South African Wind Atlas” became part of ESKOM’s South African Bulk Renewable Energy Generation project (SABRE-Gen) but unfortunately never got publicly released even though it is mentioned in various publications [Martens et al. 2001].

Information on this “revised” wind atlas remains scarce to this day though some sources point towards strong concerns about data accuracy and correct application of WAsP modelling methodology [EDRC 2003]. The only available details are published on SABRE-Gen’s website (<http://www.sabregen.co.za>),

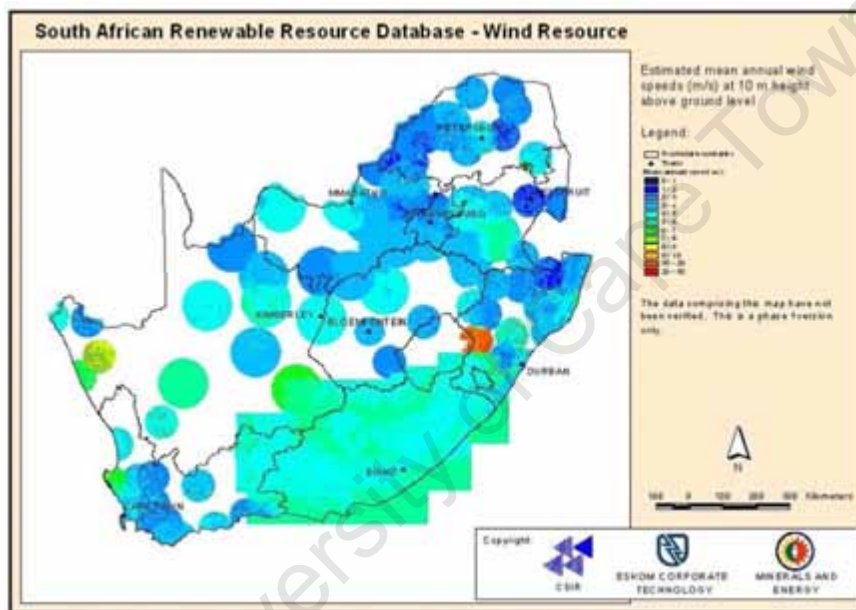


Figure 1.2: ESKOM revised wind atlas of 2001 (taken from <http://www.sabregen.co.za/sarerd%20database/wind.htm>, accessed June 2008)

see Figure 1.2 to the left.

The most comprehensive review of both Diab's and ESKOM's works was carried out in 2003 by the Danish consultants Tripod Wind Energy ApS for the DME [Langhans 2003]. Its report states that ESKOM put forth a "very overall and inaccurate conclusion" that was impossible to logically relate to the hundreds of individual WAsP wind atlas analyses carried out for meteorological wind monitoring sites. It is noted that also in the Diab atlas there is a distinct lack of information on how the detailed analyses have been processed in order to arrive at the overall conclusions given.

It further finds that both studies used stations with "low standard with respect to wind atlas application", that is, with questionable proximity to nearby objects, complex terrain and other flow influencing factors which cannot be accurately modelled by WAsP. For example, the average wind speed for the area around Giant's Castle in the southern Drakensberg has been modelled to be greater than 10ms^{-1} , a clearly unrealistic figure which can be attributed to the steep ground along the Great Escarpment which severely violates the most fundamental assumptions of WAsP (slopes with gradients of less than about 5%).

Some subsequent research suggests that Diab's total South African wind resource estimate is too low [Austin et al. 2003] while some lament the fact that the "lack of adequate wind resource assessments" is a major barrier to wind energy development [EDRC 2003]. Furthermore there is not a single work known to the author to date which has incorporated a mesoscale modelling methodology for South Africa, therefore giving even more justification for this thesis.

1.4.3 Total Wind Generation Potential Estimates

As part of the 1995 wind atlas Diab suggests a total wind energy potential in South Africa of about 7.9 TWh/year or about 2.4% of South Africa's electricity requirements in the year 2000 according to an earlier study [Diab 1985]. This calculation is not based on any wind data but rather on a string of conversions to get from first principles such as total solar insolation, its effect on the kinetic energy of air in the planetary boundary layer, the available landmass of the country and various efficiency factors to a total potential wind power output.

Three subsequent studies have attempted to improve on this initial estimate though only one of these is in the public domain. The Eskom Wind Atlas data

that later flowed into SARERD suggests a possible wind generation of 26 TWh annually but this could not be verified by the author as this figure is contained within a third party report [Loy 2004].

Furthermore a DME study of 2004 approximates the capacity to lie around 60 TWh; this is claimed by Banks and Schäffler but unfortunately incorrectly referenced [Banks and Schäffler 2006, 2007]. The latter two studies do, however, include updated assumptions based on the same wind data which yields a total yearly potential of 106 TWh.

It is interesting to note that with increasing complexity and decreasing age of the wind power analyses the total wind resource is also significantly increasing, from an initial 7.9 TWh to the latest 106 TWh. While this could be attributed to more integrated approaches it is also indicative of a very wide range of uncertainty, hence it is necessary to corroborate or disprove these results which is done as part of this doctorate.

Finally these figures must be viewed in the perspective of South Africa's total electricity generation. In 2007, Eskom sold 218.1 TWh to various customers and countries in Southern Africa according to their annual report [Maroga 2007]. It further states that this represents roughly 95% of all electricity used domestically and hence we can assume a national total of 229.6 TWh. Therefore the four wind resource estimates to date translate into 3.4%, 11.3%, 26.1% and 46.2% of 2007 consumption figures respectively.

1.5 The MM5 Model

The model chosen for this research, MM5, is referred to as the "Fifth-Generation NCAR/Penn State Mesoscale Model" [Grell et al. 1995], the latest evolution of a long term project with its roots in the seventies [Anthes and Warner 1978]. It was designed initially to be a regional ("limited area") climate model (RCM) for short term numerical weather prediction and research. It operates at a mesoscale and is driven at the boundaries by global scale data to predict atmospheric circulation. This is in contrast to a global circulation model (GCM) which has no boundaries as it covers the entire planet.

MM5 supports complex setups with multiple nesting, non-hydrostatic dynamics, four-dimensional data assimilation (Newtonian nudging), a wide number of physics options as well as good portability to a wide range of computer platforms. Even though it is developed at Penn State University it is freely

available and community driven with many of its improvements, patches and add-on functionality contributed by researchers around the world.

MM5 has a horizontal grid which comprises Arakawa-Lamb B-staggering of the velocity(u and v) variables with respect to the scalars (temperature, mixing ratio etc.). The wind fields are defined at the corners of each grid cell (the *dot* points) while the others are given on the centre of each grid point (the *cross* points).

In the vertical MM5 operates on the well known *sigma* coordinate scheme which is terrain following, i.e. avoids the problem of intersection with the ground over mountainous regions. Each vertical level s_k (not to be confused with standard deviation) is given by

$$s_k = \frac{p_k - p_{top}}{p_{s0} - p_{top}} \quad (1.31)$$

where p_k is the (constant) reference state pressure at level k , p_{s0} a reference surface pressure and p_{top} the pressure at the topmost level of the model, usually at 10hPa, i.e. the mid stratosphere at 30km altitude. Accordingly s is 1 at the surface and 0 at the top of the 3d domain. As p_k and p_{s0} are constant and only set up once by the model's pre-processors, it is also possible to convert from a sigma level to an absolute height z above sea level in metres independent of time using

$$z_k = -\frac{R_b T_{lp}}{2g} \left(\ln \frac{p_k}{p_{00}} \right)^2 - \frac{R T_{s0}}{g} \left(\ln \frac{p_k}{p_{00}} \right) \quad (1.32)$$

where R_b is the specific gas constant of dry air ($287.04 \text{ J kg}^{-1} \text{ K}^{-1}$), T_{lp} is a temperature related to the lapse rate (50°K), g the gravitational acceleration constant (9.81 ms^{-2}), p_{00} the reference sea level pressure taken to be 10^5 Pa and T_{s0} the reference temperature at p_{00} based on a typical sounding over the target domain [Dudhia et al. 2004]. Subtracting the terrain height of a given grid cell from z will then yield the height of the given sigma level above ground.

To date MM5 has been largely superseded by the Weather Research and Forecasting Model (WRF) for its greater flexibility and power [Skamarock et al. 2005]. It has, however, only become widely adopted across research disciplines within recent years, *after* the bulk of this research had already been carried out using MM5. Therefore the choice of using MM5 instead of WRF should be obvious and no further discussion relating to WRF shall be undertaken.

MM5 has been used in countless studies worldwide for both operational

forecasting and climate research. It has also found application in the Southern African context on various occasions [New et al. 2003; Tadross et al. 2006]. Important to note is that although MM5 is often mentioned in the context of wind energy [Landberg et al. 2003], the Brazilian Wind Atlas Project remains the only large scale wind resource estimation study based on MM5 apart from a few smaller investigations in Japan and Germany [Feitosa et al. 2001; Teruo et al. 2003; Jiménez et al. 2006]. However, the suitability of MM5 for a South African wind atlas has already been verified in the preparatory study to this thesis by the same author [Hagemann 2003].

It is noted that the latest officially supported version v3.7.2 of MM5 is utilised in this research.

1.6 Methodology Overview

The following is a very brief overview of the components of this thesis which in combination achieve the primary and secondary objectives set out in Section 1.1.

- chapter 2 describes measurement stations and a validation methodology for evaluation of the model's skill
- chapter 3 develops an algorithm to find the most typical year to be used for integration
- chapter 4 tests the performance of various MM5 parametrisations and options to find an optimal configuration for a South African wind resource assessment at the target resolution
- chapter 5 describes the final nationwide simulation at 18x18km grid spacing

First of all as argued in sections 1.3.3 to 1.3.4 and due to the lack of such work in the national context the task at hand is to carry out a fully dynamic mesoscale simulation for the whole of South Africa. Given the computational resources available to the author for this research it is deemed feasible to achieve this at a resolution of 18x18km or 324km² grid cells over a 365 day integration period. The next three chapters are devoted to laying the groundwork for this.

Every model output is almost useless - just like any other physically measured or computed quantity - without giving an estimate of the associated uncertainty. The only reliable way to do this is to compare it with observed values

measured in the atmosphere or at least a gridded dataset that very closely resembles such measurements. Therefore every output needs to be checked against reality but this is not a simple process.

Chapter 2 on the following page goes into the detailed nature of the stations and their data considered as well as the development of a thorough and robust validation methodology specialised for the evaluation of winds. This will in effect help achieve the primary objective by making it possible to objectively measure model strengths and weaknesses in the first place.

As described and justified in subsection 1.3.4, the period of one year is chosen in order to capture a complete annual cycle while not exhausting available computing power required for a multi-year or even decadal simulation at a similar spatial resolution. From the outset this would imply losing a lot of inter-annual variability vital for wind farms. However, chapter 3 on page 41 describes the development and application of an innovative algorithm designed to choose a 365 day time span that is as representative of the climatology as possible, thereby largely mitigating this problem. This in essence fulfils the first sub-objective of our primary goal.

Like many other mesoscale models MM5 has many tunable parameters and parametrisations available to its user. Which combination to choose is a consequence of a mixture of experience, resolution and location of the domain in question as well as some experimentation. As the reader shall find out in chapter 4 on page 57 there is a distinct lack of prior knowledge so many tests need to be carried out in order to find an optimal configuration of the model for the highly specialised application of a South African wind resource assessment. That task aims to achieve the second component of the primary purpose.

Only once all these three sub-projects have been successfully carried out can the actual wind atlas work begin. Chapter 5 on page 107 describes the final setup and results of the full scale integration. The results are thoroughly validated as before which elucidates even more detail about the model's skill and shortcomings.

Chapter 5 also derives a limited amount of useful statistics and collates some main findings in graphical form. The total potential for electricity generation in South Africa is then computed based on various criteria and assumptions. These last two remaining projects represent the fulfilment of the secondary objective as stated in section 1.1.

The conclusion in chapter 6 on page 145 then wraps up the discussion and recommends areas of future research.

Chapter 2

Station Data and Validation Techniques

As outlined in section 1.6 on page 29 this chapter deals with the selection of wind monitoring stations utilised for this study (section 2.1) as well as a detailed validation methodology for evaluation of the model's ability to represent reality based on such observations (section 2.2 on page 33).

2.1 The Observations

As described in subsection 1.3.3 on page 19 a wind resource assessment based on a mesoscale modelling methodology in principle may be carried out without the presence of any particular observations. However, there are two reasons reliable station data must still be included in the analysis: Firstly it is of utmost importance to gauge the accuracy and uncertainty of the simulation output based on the best representation of "reality". Secondly as the reader shall see in chapter 4 on page 57 a technique known as "observational nudging" may improve the quality of the results.

Figure 2.1 to the right shows the 17 South African Weather Service (SAWS) stations that are deemed to be of sufficiently high quality for this purpose. Geographical details and the start and end period of the available wind speed and direction datasets are outlined in Table 2.1 to the right.

With the exception of Bloemfontein all of these have met the fairly stringent selection criteria of both Diab's and ESKOM's wind atlases. However the list is

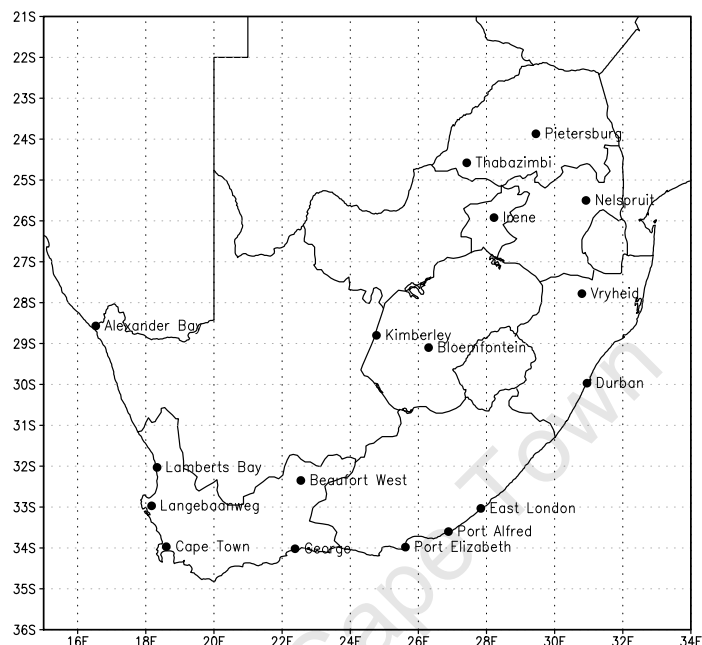


Figure 2.1: The 17 10m SAWS weather stations considered for this study

Station	Latitude	Longitude	Height	Start	End
Alexander Bay	28.57° S	16.53° E	29m	1994/10/19	2005/03/15
Beaufort West	32.35° S	22.55° E	902m	1994/10/19	2005/03/15
Bloemfontein	29.1° S	26.3° E	1359m	1994/10/19	2005/03/15
Cape Town	33.97° S	18.6° E	42m	1994/10/19	2005/06/21
Durban	29.97° S	30.95° E	8m	1994/10/19	2005/03/15
East London	33.033° S	27.833° E	124m	1994/10/19	2005/06/20
George	34.02° S	22.38° E	191m	1994/10/19	2005/06/20
Irene	25.92° S	28.22° E	1524m	1994/10/19	2005/06/21
Kimberley	28.8° S	24.77° E	1197m	1994/10/19	2005/03/14
Lamberts Bay	32.03° S	18.33° E	93m	1994/10/31	2005/03/15
Langebaanweg	32.97° S	18.17° E	31m	1994/10/19	2005/03/16
Nelspruit	25.5° S	30.92° E	883m	1994/10/19	2005/03/16
Pietersburg	23.87° S	29.45° E	1237m	1994/10/19	2005/03/15
Port Alfred	33.6° S	26.883° E	36m	1994/10/19	2005/06/20
Port Elizabeth	33.98° S	25.62° E	59m	1994/10/15	2005/03/15
Thabazimbi	24.58° S	27.42° E	977m	1994/10/19	2005/03/16
Vryheid	27.78° S	30.8° E	1163m	1994/10/19	2005/03/16

Table 2.1: Meta-data for the selected weather station data as supplied by SAWS. Height is given in metres above mean sea level (amsl).

not a mere intersection of the two but rather comprises exclusively airport data (except Irene), all of which use the RM Young 05130 propeller anemometer which is calibrated once a year and mounted on a mast 10m above ground according to SAWS. This ensures not only high reliability of the wind readings but also uniformity in terms of possible measurement biases. Furthermore this data has been tested for consistency before in terms of a fit to the Weibull distribution (equation 1.17 on page 10) and shown to be much more accurate than the pre-1993 measurements [Hagemann 2003].

All datasets use the smallest available sampling interval of 5 minutes which in itself is actually a time average over high and ultra high frequencies. These high frequencies are further smoothed by averaging hourly centred on the hour. That is, the 13:00 value is the arithmetic mean of the 12:30 - 13:25 series. The rationale behind this smoothing is to take out the variability that the model is highly unlikely to be able to resolve at 18km grid size and which would cause unnecessary validation errors. Besides, for grid stability reasons the power output of modern wind turbines is not affected by very short term fluctuations (see e.g. the technology behind the Vestas V80 2MW wind turbine).

It is noted that the directions θ_i are properly averaged using equation 1.26 on page 12 which is acceptable because we are dealing with winds that do not deviate by more than a few dozen degrees over such a small period of time.

It is also important to point out that the hourly wind series that SAWS can supply are ill suited for this application: it consists of values averaged over only a few minutes around the full hour. This means that most of the readings taken in between are discarded and thus these “hourly” time series are not suitable.

The reader may find out in chapter 3 in Table 3.3 on page 53 what the long-term average wind speeds are at these 17 weather stations. These by themselves already indicate a fairly healthy wind resource, larger than what Diab suggested in the 1995 wind atlas [Hagemann 2003]. A more detailed analysis of these stations, however, is not carried out at this stage because this would exceed the scope of this thesis.

2.2 Validation Methodology

Any model output on its own is fairly insignificant without a proper comparison against real world data just like the physical measurement of a given quantity is meaningless without an associated uncertainty. For that reason simulation

results should not be taken at face value unless it is impossible to validate against observations, be it for lack of available stations or otherwise.

In the context of wind energy we would thus like to compare as many of the core statistics outlined in section 1.2.5 as possible, model vs station. Most of the time some sort of difference between the two is sought; the higher this difference the worse the performance of the model and vice versa. Other quantities compare the time evolution or other dynamic effects.

In either event, in this discussion we will use the same terminology as before with the exception of the subscript $_o$ and $_m$ meaning observed and modelled values respectively. Furthermore, as the wind is a vector quantity there are in fact four different components that we can look at according to the principles set out in subsection 1.2.1.1. Accordingly we shall use the subscripts $_V$ for speed, $_\theta$ for direction, $_u$ for west to east and $_v$ for south to north.

It must be stressed that the goal in this section is to seek *quantitative* ways that are *independent* from one another to validate our model.

This means firstly that a set of single values describing characteristics is sought rather than a large selection of graphical comparisons from which conclusions are drawn overly subjectively; this enables us to condense the verification and increase objectivity. As will become clearer in chapter 4 on page 57 there are also too many individual simulation results to contend with for any approach not based on summary statistics to be efficient. Secondly we want non-congruent metrics so that a given simulation is not penalised twice by two statistics which may have different names but essentially report on the performance regarding the same underlying aspect of the wind.

2.2.1 Core Comparison Statistics

Traditionally there are a few standard core metrics we can use for validation purposes [Davis 1973]. Most well known and widely used for wind speed at least are the root mean square error $RMSE_V$ (also known as standard error) and the bias b_V (also referred to as mean error or mean bias error) which are defined as

$$RMSE_V = \sqrt{\frac{1}{n} \sum_{i=1}^n (V_{m_i} - V_{o_i})^2} \quad (2.1)$$

$$b_V = \overline{V_m} - \overline{V_o} \quad (2.2)$$

$RMSE_V$ gives us an idea of how far the two values typically lie apart at any given instant in time. It is very important for forecasting and time series analysis but less important in this context since we are interested in the climatology of the wind, not in its finer dynamics and time evolution with the exception of its cyclical behaviour which we deal with separately in subsection 2.2.2 to the right. As the bias b_V provides the difference in the overall averages it is much more important for our analysis.

As an alternative to the $RMSE_V$ the mean absolute error MAE_V is considered by some as one of the more useful difference measures in the comparison of geophysical datasets [Willmott et al. 1985]. Instead of getting rid of the mixture of positive and negative deviations - which on their own may average to zero - by squaring them, the MAE takes the absolute value thereof:

$$MAE_V = \frac{1}{n} \sum_{i=1}^n |V_{m_i} - V_{o_i}| \quad (2.3)$$

Both root mean square error and mean absolute error can equally well be applied to wind direction by employing circular averaging as described in 1.2.5 on page 9 to arrive at equivalent $RMSE_\theta$ and MAE_θ . However, this would again elucidate merely how well the time evolution matches which is of peripheral concern for a regional wind resource assessment (such directional variability only becomes important at much smaller scales).

While applying the bias equation 2.2 on the preceding page to direction is possible, this would be of little value since it would violate the principles set out earlier which mandate a short averaging period for any meaningful result. However what *can* be done is to compute and interpret b_u and b_v , despite the possibility that the averages of the orthogonal wind components $\bar{u}$ or $\bar{v}$ may be zero, i.e. distributed equally around the origin.

What is also important is the bias in wind speed standard deviation of the two datasets c_V as defined by

$$c_V = \sigma_{V_m} - \sigma_{V_o} \quad (2.4)$$

which essentially tells us if the model is under- or over-predicting the variability in the data, an important aspect of the wind distribution for medium and lower frequencies. Both “ordinary bias” b and “standard deviation bias” c are in the units of the physical quantity measured, in this case ms^{-1} . For similar reasons outlined above it makes little sense to compute c_θ but c_u and c_v could

prove helpful to point out where the model's pitfalls are.

In the context of wind power we regard b and c together as the two most important metrics describing the match of wind climatology's since they fully describe the distribution in a more compact form than a histogram is able to, even if we work with a non-Gaussian variate such as a Weibull one in this case. Higher order statistics such as skewness and kurtosis are merely dependent variables of these two, even the Weibull A and k parameters can ultimately be derived from them (see equations 1.19 to 1.20 on pages 10–11).

There exist a few other core quantities for measuring the agreement between observed and modelled data that are not difference based, most notably statistical correlation, also known as the Pearson's product-moment correlation coefficient. Sometimes its square r^2 is used instead of r . Unfortunately as many have pointed out before it is insensitive to additive or proportional differences between the two series [Willmott 1982; Legates and McCabe Jr 1999]. Therefore it is deemed to be unsuitable for this type of model evaluation. The shortcomings of correlation have been addressed through the use of other measures such as the index of agreement (IOA, [Willmott 1981]), however, even this is of more interest to forecasting in which correct dynamic evolution is more important than absolute predictions and climatology.

2.2.2 Comparison of Cycles

According to section 1.2.2 on page 5 the periodic behaviour of certain wind patterns, especially the daily cycle, is very important in the context of wind resource assessments. Thus we require certain metrics by which we can judge the performance of our model against observations. The literature lacks standard approaches as comparisons are seldom carried out at the level of detail done here, therefore we proceed to define a few of our own.

Firstly let us look at the diurnal cycle of wind. The reader is reminded of equation 1.24 on page 11 which defines a 24h average speed series V_{24} . Assuming there is some kind of cycle clearly present which is likely to be much more irregular than a simple sinusoid (something that must be ascertained visually on the data first), we remember the amplitude of this cycle D_a as given in equation 1.25 on page 12 and based on the observed D_o and the modelled D_m we can define a (relative) cycle amplitude error (CAE) of

$$CAE = \frac{D_o - D_m}{D_o} * 100 \quad (2.5)$$

which is effectively a percentage deviation of their magnitudes. The *CAE* is positive if the model under-predicts the strength of the cycle and negative if it is over-predicted.

To describe not only the size but also the timing of this cycle - at least its peak which is most important for wind power - the *CAE* has to be complemented by the peak phase error (*PPE*). Using the time t_{max} as the time the maximum wind speed occurs ($t_{max} \in [0, 24]$) we can define the *PPE* as

$$PPE = t_{max_o} - t_{max_m} \quad (2.6)$$

whose units are hours. Thus a positive *PPE* indicates that the model's average daily wind pattern peaks too early while negative means it is too late. Both *CAE* and *PPE* approach zero as the model's skill improves.

Given the importance of the diurnal cycle a third comparison statistic is employed here which uses a Fourier technique to distill strong periodicity. As the reader shall see in chapter 4 on page 57, the Fourier transform $G(f)$ actually spikes at $f_d = 1\text{day}^{-1}$. While the absolute value of $G(f_d)$ is not interesting we can use the daily Fourier peak difference *DFPD* given by

$$DFPD = \frac{G_o(f_d) - G_m(f_d)}{G_o(f_d)} * 100 \quad (2.7)$$

as another percentage measure of shortfall in power of the daily peak, where positive and negative values are interpreted identically to the *CAE*.

It must be noted that both $G_o(f)$ and $G_m(f)$ are the results of the Fast Fourier Transform (FFT) based on the detrended wind speed series, that is, the original values with their mean subtracted (slopes are assumed to be zero). This ensures that any bias which is assessed by b above is filtered out correctly. Even before this is done any missing values in the observations have been replaced by their average to avoid running into problems with the continuity assumptions of the FFT [Storkey 2004].

The *DFPD* is highly correlated with the *CAE* although it is derived totally differently, which violates the concept of independence laid out above. For reasons of graphical presentation which will become clear in chapter 4 we will use both regardless.

On a seasonal timescale it is also worth considering how well the monthly averages match up. For this purpose we devise a measure based on the 12 month time series V_{12} called the monthly mean absolute error *MMAE* defined as

$$MMAE = \frac{1}{12} \sum_{k=1}^{12} |V_{12_m,k} - V_{12_o,k}| \quad (2.8)$$

Its value is also given in ms^{-1} and is to be interpreted like any other *MAE*. Similar to the pre-filtering applied to the *DFPD* the V_{12_m} series has to have the model bias removed before computation to be able to judge the model purely on its seasonal skill.

2.2.3 Evaluation of Wind Direction Performance

For validation of the model's wind direction we have seen in subsection 2.2.1 on page 34 that the two core error statistics *b* (ordinary bias) and *c* (standard deviation bias) cannot be meaningfully applied to θ . Therefore we have to come up with different metrics that *are* valuable and we shall do so based on the treatment of the twelve distinct 30° sectors outlined in subsection 1.2.5 on page 9.

For each of the four windrose types that were discussed - frequency F_θ (equation 1.27 on page 13), average speed F_V (equation 1.28), speed standard deviation F_{σ_V} (equation 1.29) and power density F_P (equation 1.30) - we come up with a carefully constructed comparison metric which sums up the model's performance with respect to observations.

First we have the weighted *MAE* of sector frequencies (*WMAESF*)

$$WMAESF = \sum_{j=1}^{12} F_{\theta_{o,j}} |F_{\theta_{m,j}} - F_{\theta_{o,j}}| \quad (2.9)$$

As the subscripts are getting complex by now let us explain this equation in a bit more detail. The summation is performed with index *j* over all twelve 30° sectors. The right part is the standard *MAE* difference of modelled and observed percentages of occurrence. The $F_{\theta_{o,j}}$ achieves a weighting by the measured frequencies to enhance the dominant sectors and de-emphasise the sectors of minor importance. Thus this will us gauge how well the directional wind distribution is simulated; the lower the *WMAESF*, the better the agreement with reality.

Secondly we define a weighted *MAE* of sector speed averages (*WMAESSA*) as

$$WMAESSA = \sum_{j=1}^{12} F_{\theta_{o,j}} \left| F_{\nabla_{m,j}} - F_{\nabla_{o,j}} \right| \quad (2.10)$$

where we now use F_{∇} for the computation of the mean absolute error term but still weight by the observed frequencies in order to achieve the same significance ranking of all sectors as above.

Similarly we propose a comparison metric based on F_{σ_V} called the weighted *MAE* of sector speed standard deviations (*WMAESSD*) given by

$$WMAESSD = \sum_{j=1}^{12} F_{\theta_{o,j}} \left| F_{\sigma_{m,j}} - F_{\sigma_{o,j}} \right| \quad (2.11)$$

where we have made the simplification of $F_{\sigma} = F_{\sigma_V}$ for readability's sake. The *WMAESSD* shall provide us with a good measure of how well the model simulates not the overall but the sector specific variability of the respective wind speeds.

Finally there is the weighted *MAE* of sector wind power densities (*WMAEPD*) based on F_P , arguably the most crucial value for wind energy applications, defined as

$$WMAEPD = \sum_{j=1}^{12} F_{\theta_{o,j}} \left| F_{P_{m,j}} - F_{P_{o,j}} \right| \quad (2.12)$$

based on equation 1.30 on page 15 where air density ρ is approximated by 1.29kgm^{-3} .

Equations 2.9 to 2.12 shall from now on be referred to as the *WMAE* family of comparison metrics.

It is noteworthy that all of them are, unlike the cycle statistics of subsection 2.2.2, based on series that have *not* been detrended or filtered in any way. This means that model bias (both *b* and *c*) are contained in them again, but this time in much more detail, directionally dependent and thus subtly different. On the surface it appears that this violates the principle of independence of the derived statistics as outlined at the beginning of this section. However, if one contemplates this in more detail one realises that the *WMAE* group is *complementary* to the core bias measures and cannot be discarded. The following simplified example illustrates this concept well.

Assume that for a given location model run “A” gets the directional frequency distribution right but significantly over-predicts its easterlies while its westerlies

are too low by the same amount. However, a 50/50% east/west distribution ($\bar{u} = 0$) ensures that “A” still manages to produce a zero bias in both b_V and c_V . Simulation “B” in turn exhibits a constant non-zero bias for all sectors because its treatment of surface roughness is too high. “A” will score perfectly for b_V and c_V but get penalised for high values of $WMAE$ sector speed average and $WMAE$ power density. While model “B” is worse in terms of biases b_V and c_V , its $WMAESSA$ and $WMAEPD$ are much lower. At least in terms of micro-siting of a wind farm which has to be optimised in terms of accelerating winds from dominant directions and minimise mutual turbine shadow effects, the outcome “B” is a lot more useful despite its bias.

Chapter 3

Selection of Optimal Simulation Period

3.1 Overview

As described in subsection 1.3.4 on page 20 of the introduction our time dependent or dynamic modelling approach normally requires integration periods of at least a decade or longer in order to obtain an accurate climatology. However, in regional climate simulations for exploring long-term climate statistics this implies significant computational requirements. Given finite and often limited computer resources, especially in developing nations like South Africa, this often restricts the possible resolution or complexity that the model can be run with.

Instead of using steady state techniques or limiting resolution and/or domain extents we unveil a novel third solution to this common problem. We seek to intelligently reduce the required integration time of a *single* model simulation by selecting an optimal temporal subset to encompass median climate modes representative of the long term state.

While this cannot reflect the complete range of inter-annual variability it does allow for RCM simulations that capture the central tendency of the climate envelope. A steady state approach which keeps the boundary conditions constant is even more likely to miss the inter-annual complexity and is additionally vulnerable to longer term temporal lead and lag processes (e.g. the evolution of cold fronts), besides the reasons already given against it in subsection 1.3.4.

The challenge remains on how to reduce the long term 10-20 year period

to something more manageable, in our case a 365 day complete yearly cycle, *without* sacrificing the long-term representation. This chapter presents both the methodology for this process and the application to circulation patterns over South Africa.

3.2 Methodology

The basic concept is to select from a long term reference period the 365 day sub-period whose distribution of synoptic states best matches that of the reference period. How this is done is now discussed in conceptual terms and the following section will focus on how this is achieved in practice.

A suite of relevant variables is chosen from a global, low resolution atmospheric long term dataset H of duration T_t ($T_t \geq 20\text{years}$) which force any meso- and microscale winds that we are interested in. From H we only consider those fields that are strongly deterministic for the application at hand. This daily time series is then analysed with a classification algorithm to generate a set of typical atmospheric states along with their associated frequency of occurrence.

It is important to note that in this exercise it is irrelevant what each of the individual states mean. The objective is only to define the histogram of synoptic modes of H against which the equivalent sub-period histograms will be compared.

Using the same classification method we now compute the histogram of all possible contiguous subsets $H_{s,i}$ with time span $T_s = 365$ days wholly contained within the reference dataset ($H_{s,i} \subset H$) and see how well their respective histograms compare to the reference. The subscript s here indicates the fact that it is a 365 day subset while $i \in [1, N_s]$ where N_s is the total number of possible contiguous subsets H_s .

For this purpose, let the frequency distribution of the long term reference be denoted by R and that of the particular subset by Q_i . We then define a weighted mean absolute error function Ω as the performance metric for the comparison as

$$\Omega_i = \left| \frac{R - Q_i}{R} \right| \quad (3.1)$$

The residuals $R - Q_i$ are weighted by R to account for the fact that a 5% difference in frequency is more severe in a common state than in an infrequent

one. Also note that we are using element by element arithmetic here, not matrix operations, as both R and Q_i are actually multi-valued quantities (we left out the extra indexes so that the discussion below can be simplified).

Ω_i is then minimised to find the best fitting subset $H_{s,k}$ such that $\Omega_k \leq \Omega_i$ for all $i \in [1, N_s]$. In other words

$$\Omega(H_{s,k}) = \min(\Omega_i) \quad (3.2)$$

By further looking at the residual $R - Q_k$ it is possible to determine just how different the frequency distribution of $H_{s,k}$ is to that of the reference set H and thus how well the algorithm performs.

Any classification undertaken will remain to some degree subjective to the investigator's algorithm of choice and especially in the number of distinct states that we happen to employ which determine the degree of generalisation. $H_{s,k}$ clearly becomes a function of N_c , the number of classes. In the example domain used in this thesis, it was found that a low (<5) N_c tends to generalise too much while large N_c (>30) place too much emphasis on specific cases.

In order to properly account for this it is advisable to explore and average over the "classification space" by repeating the above method for a set of M classifications, each with its own number of states $N_{c,j}$ with $j \in [1, M]$. We can then aggregate all the individual $N_s \times M$ Ω_{ij} results into one final list by using a simple arithmetic mean and the period $H_{s,l}$ that is ranked first on a list of ascending Ω_{ij} becomes the overall winner.

Within each iteration j , the range of Ω_{ij} is likely to differ. With small N_c , the range is expected to be greater than with large N_c , owing to the increased level of specialisation and corresponding smaller absolute differences. In order to reliably aggregate the results from the different classifications, it becomes necessary to standardise all Ω_{ij} to the range 0-1. This is easily achieved by means of a linear scaling function which utilises the maximum and minimum of Ω_{ij} for each j .

Each subset $H_{s,i}$ thus receives M error values within a normalised range which are simply averaged to yield a final ranking that elucidates the overall best period $H_{s,l}$. Note, however, that it is unreasonable to expect any of the $H_{s,k,j}$ (optimal subset $H_{s,k}$ for a given number of states $N_{c,j}$) to coincide with $H_{s,l}$ because of the inherent differences of the various classifications, although they should be close in most cases.

3.3 Application to South Africa

To demonstrate and test the methodology proposed above we now apply it to the South African domain. With the focus on wind resource assessments, only surface ($s = .995$) and mid-troposphere (700mb, 600mb and 500mb) u and v-wind fields are taken every 12 hours from the NCEP/NCAR Reanalysis dataset as H [Kalnay et al. 1996]. Only these eight wind fields were selected because together they are deemed to fully drive the planetary boundary layer flows relevant for wind power.

The chosen domain is bounded by 37.5-20°S and 12.5-37.5°E covering South Africa with a generous amount of padding. The input data is rescaled to the range [0-1] using a linear transformation in order to standardise the different variables for the classification. The reference period is chosen to be 1979-2004 as a compromise between the long term and data quality, as pre-satellite era re-analysis before 1979 contains errors over this domain [Tennant 2004]. This gives us $T_t = 26$ years which is deemed sufficient to yield a representative climatology, aside from issues of multi-decadal variability.

The classifier employed is the Self Organising Map (SOM) algorithm, a neural networks technique well suited for the task at hand [Kohonen 2001, 1989]. As Reusch et al. have shown in an assessment of SOMs vs the more well known principal component analysis (PCA) to identify patterns in data, a SOM can minimise the role of subjective choices in application [Reusch et al. 2005]. SOMs have wide application across multiple disciplines, and are gaining acceptance in the climate community as a powerful means of evaluating atmospheric data. Numerous authors to date have illustrated their potential applications to synoptic climatology from the general to specialisations as diverse as clouds, extreme events, daily precipitation, sea ice extent and intra-seasonal variability from regions and latitudes across the globe such as the USA, Puerto Rico, the Balkans, Antarctica and Southern Africa [Hewitson and Crane 2002; Ambrose et al. 2000; Cavazos 2000; Cavazos and Hewitson 2002; Ni et al. 2002; Hudson 1998; Malmgren and Winter 1999; Tennant 2003].

A SOM identifies archetypical modes within the multi-dimensional data space such that the nodes represent a 2-dimensional expression of the continuum of the data. Once properly trained, a SOM maps each instantaneous input onto a specific node in the 2 dimensional array of nodes for symbolising each reference state.

Each node has an associated “codebook vector” which embodies the state

vector for a position within the continuum of the data space. Adjacent nodes have closely related vectors, while those further apart in the matrix have more dissimilar ones. After training, the SOM results in the most disjoint modes from the data space occupying nodes that are separated by the largest distance [Hewitson and Crane 2002]. In this application the components of these vectors are simply the dimensional wind fields u and v at the specified pressure levels for the domain under consideration. The final outcome is analogous to a clustering approach, robust in the presence of noisy data and relatively invariant to the choices made in implementation.

SOM-PAK is a free software package for SOM applications used in this research [Kohonen et al. 1996]. It runs efficiently with a low computational resource overhead even for the large datasets in question.

For ease of visualisation the rectangular topography of nodes is chosen, with array sizes of 3x2, 4x3, 5x3, 6x3 and 6x4 maps ($N_c \in \{6, 12, 15, 18, 24\}$, these are the $N_{c,j}$) used to explore the sensitivity to the degree of generalisation. This set of number of nodes is deemed a reasonable sampling of the classification space in terms of the N_c discussion of the previous section; in other words such a set includes both high degrees of generalisation ($N_c = 6$) with relatively high frequencies per state and cases with more emphasis on specialisation ($N_c = 24$) with low individual occurrence. Apart from this reasoning, it has been found that the SOM results scale robustly as the number of nodes change [Crane and Hewitson 2003].

Two training cycles are used, the one starting with a large kernel to allow the SOM to uncover the coarse scale structure of the data space, the second with a smaller kernel to allow for refinement of the node vectors to the local variance structures. The reference vectors are randomly initialised and trained using ten times the number of available inputs (18994) in stage I and three times this number in stage II training.

With each input vector the SOM determines the most closely aligned node reference vector within the array. This vector is adapted to be more similar to the input vector by a magnitude determined by a “learning factor” which incrementally decreases through the training process. In the SOM-PAK software, the training rates were set to 0.05 and 0.02 respectively.

Nodes which are in proximity to the updated node in the array, within a “neighbourhood kernel”, are similarly updated but to a lesser degree. The size of the kernel decreases as training progresses. In this way adjacent nodes come to represent related regions of the input space.

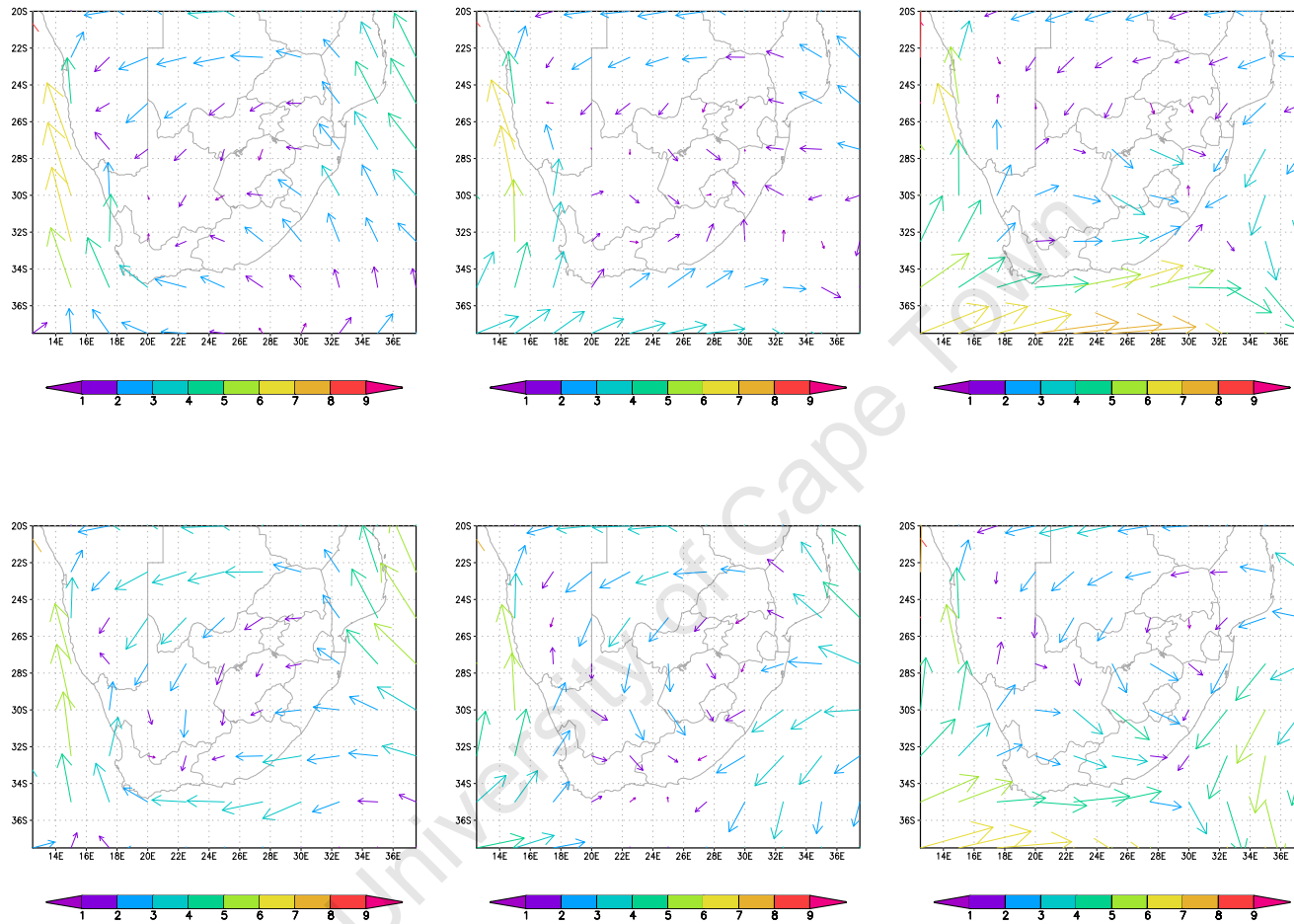


Figure 3.1: Most general 3x2 Classification of long-term (1979-2004) NCEP surface (sigma .995) winds in ms^{-1}

Proper convergence has been verified using ten different random initialisations which all yielded virtually the same map. Stability and integrity of all five sets of classes has been further checked using Sammon's mapping [Sammon 1969].

On completion of the training the input vectors are once again applied to the SOM with no vector updates to obtain the final mapping of input vectors onto nodes, i.e. what input pattern most closely matches each specific cluster. From this information the frequency of occurrence of input vectors onto each node can be determined and a two dimensional histogram is generated which reflects the probability density function (pdf) of the eight dimensional data space (the u and v winds for four pressure levels). The actual node vectors could be informatively mapped for further examination of the atmospheric modes represented by each of them but this is not done here [Hewitson and Crane 2002].

The representation of the codebook vectors (surface wind vector component only) for the 3x2 map are given in Figure 3.1 on the preceding page as an example of the trained SOM identification of modes within the data space. The associated frequency distribution of the nodes is shown in Figure 3.2 to the right. It can be seen that the histogram is quite flat for this high degree of generalisation (3x2), with occurrence ranging from a minimum of 14.8% to a maximum of 17.6%.

While these illustrations only show the most general of all classifications with number of classes $N_{c,1} = 6$, they serve as a good visualisation of the process as Figure 3.2 to the right is essentially the matrix $R_1(j = 1)$, i.e. the pdf of long term reference for the first classification.

Subsequent to the training, the histograms for all subset periods $H_{s,i}$ are determined in an analogous fashion. In this application we have to restrict the subsets $H_{s,i}$ to the period from 1993-2004 because of limited availability of good quality observed wind data which we will use in validation (see section 2.1 on page 31). This gives us a total number of possible contiguous subsets of $N_s = 4018$ (remember $i \in \{1, N_s\}$). The full training period from 1979-2004, nonetheless, better develops the SOM to accommodate the full range of synoptic states.

In comparing the sub-period histograms to the full reference histogram, the rescaled $\Omega_{i,j}$ values as a function of 365 day period are very noisy for all subsets, so only a representative set is presented in Figure 3.3 on page 49 which plots $\Omega_{i,j}$ for each of the SOMs. Upon studying this graphical result carefully it becomes clear that while the different classifications as a function of the SOM array size

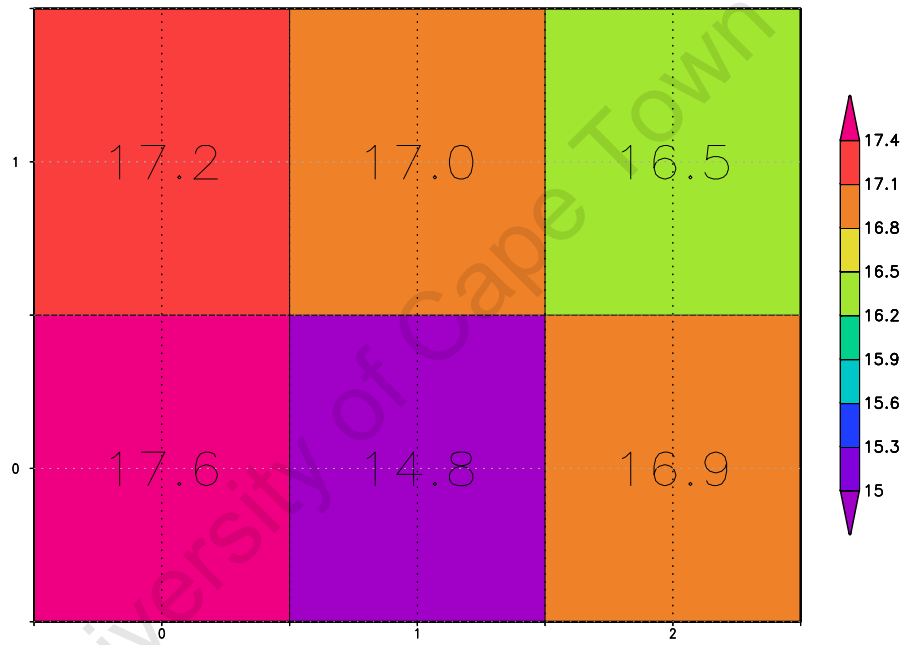


Figure 3.2: Frequency of occurrence map (two dimensional histogram) of the most general 3x2 states of Figure 3.1 on page 46 in %. The labels on the axes indicate the node numbers

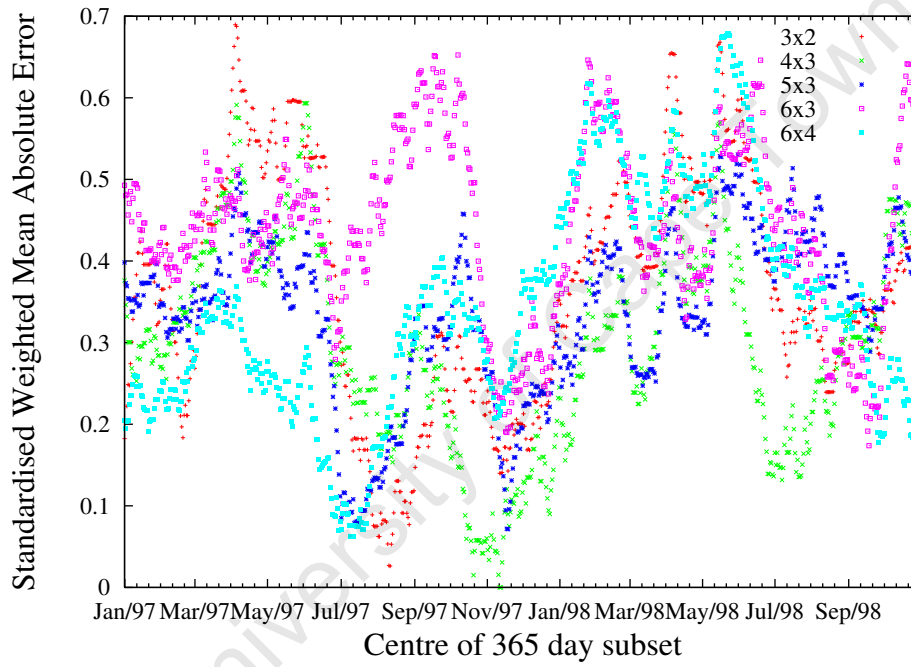
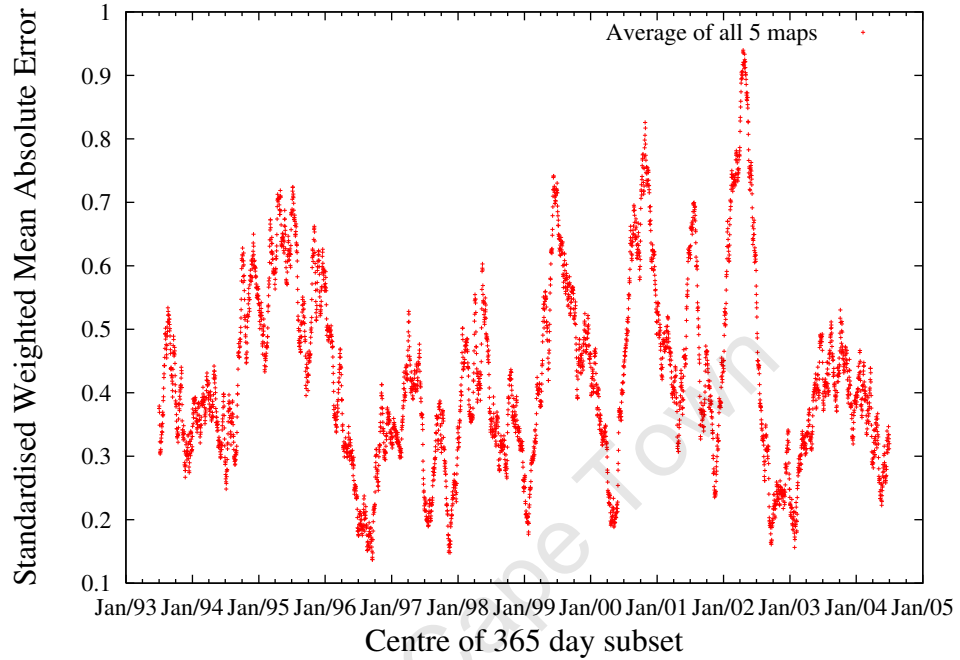


Figure 3.3: Standardised weighted mean absolute error (Ω) scatter plot for select subsets and all 5 classification schemes ($N_c \in \{6, 12, 15, 18, 24\}$)

Figure 3.4: Aggregate of individual Ω curves for all subsets

(effectively N_c) yield different Ω curves in terms of absolute value (as would be expected), they exhibit similar relative behaviour across all sub-periods because a high degree of correlation exists.

The reader may further observe at a casual glance that the 365 day period centred around June 1998 (i.e. January until December 1998) was very atypical of the wind climatology while June 1997 until May 1998 seem to be much closer aligned to the long term average with fairly low values of Ω .

These results give confidence for the last step, the averaging of all the individual curves Ω_{ij} over the different classifications as exemplified in Figure 3.3. The aggregated Ω_i is presented in Figure 3.4. From here it is a trivial exercise to select the optimum subset $H_{s,l}$ as being from 20 March 1996 to 19 March 1997 with a global minimum $\Omega = 0.137$. Note that there are also other near optimum periods that could equally well be selected - in the list of all $H_{s,i}$ ranked according to ascending Ω the first 59 subsets all have $\Omega < 0.18$.

Ideally we would like to see a value much closer to 0 here but we must bear in mind that the differences between the various classifications in Figure 3.3 to the left cause the average to be slightly higher. To quantify the significance of

Classification	Absolute Rank	Relative Rank	Ω_j
3x2	12	0.30%	0.032
4x3	998	24.84%	0.291
5x3	148	3.68%	0.154
6x3	109	2.71%	0.117
6x4	178	4.43%	0.090

Table 3.1: Results of normalised Ω and ranking (out of $N = 4018$) for the overall optimum period 20/03/1996 to 19/03/1997 within the individual classifications

this result we can say that these first 59 periods are, on average, amongst the top 18% of all possible subsets $N_s(4018)$ in terms of how well they represent the long-term climatology as indicated by each individual classification.

It is also interesting to note that this procedure is equally effective to identify the least representative period. In this case the most anomalous 365 day period is from 19 October 2001 to 18 October 2002. With a global maximum $\Omega = 0.94$ it is much less ambiguous and clearly appears above the rest.

For comparison, Table 3.1 shows how well $H_{s,l}$ fares within the individual classifications. With the exception of the 4x3 case, $H_{s,l}$ consistently ranks among the top 180 or 4.5%. The 4x3 SOM only places it amongst the top 25%, although this appears to be an outlier for which an explanation is lacking.

Finally, Table 3.2 to the right allows the reader to get an idea of just how close $H_{s,l}$ actually matches the climatology of the reference period H . It presents the raw residual matrices $R - Q_k$ for both the overall best period as well as the worst individual period for comparison of all five classifications. In the best case the values are all relatively small which means that $H_{s,l}$ never introduces more than a few days extra or less of any given particular synoptic state compared to the long term average. The worst case in turn causes inconsistencies of up to three weeks which in turn cause a much higher 365 day anomaly.

3.4 Uncertainty Analysis

As the above discussions show we cannot claim that 20 March 1996 to 19 March 1997 is a perfect match of the climatology but it is reasonably close. How close is a question of uncertainty which we can actually answer quite easily by looking at the microscale response as given by observations instead of the large scale forcing.

For all available weather stations we compute the averages and standard

Classification	Best overall period $H_{s,l}$	Worst individual period
3x2	0.6 1.1 -1.4	17.6 8.6 -13.4
	3.8 1.3 -5.4	16.4 -4.2 -24.9
4x3	5.0 -4.6 2.8 5.6	-1.5 10.4 6.4 4.4
	2.8 -3.6 -0.2 4.4	10.3 -3.6 -5.2 1.9
	5.2 0.2 -8.1 1.8	12.7 -5.4 -23.1 -7.3
5x3	5.7 -4.8 0.2 -1.5 -7.2	6.7 -4.3 10.2 -3.5 -12.2
	3.8 -0.5 -2.5 8.3 -0.9	-5.2 2.0 10.0 10.3 -3.4
	-3.7 0.5 0.0 1.9 0.7	-7.7 -9.5 4.0 9.9 -7.3
6x3	0.8 -2.2 1.8 -8.1 2.4 3.3	0.3 -1.7 10.3 -10.6 -15.1 -0.7
	3.5 3.8 4.7 -0.1 -0.4 -0.9	-1.0 7.3 6.7 -17.1 -0.4 0.1
	0.0 0.7 -1.9 1.7 0.4 -9.5	11.6 4.2 4.6 -4.3 2.4 3.5
6x4	1.9 0.6 1.8 -2.4 -3.0 -2.3	-6.1 -5.4 7.8 7.6 4.5 1.7
	-1.3 3.6 -1.4 1.3 3.8 -2.9	-6.8 7.6 7.6 5.3 -4.3 4.1
	-1.3 2.9 1.7 5.1 -0.8 0.0	4.7 5.4 -0.3 -3.4 2.2 -4.0
	-4.3 -3.1 -0.5 -1.2 -4.1 5.8	2.8 2.9 -10.5 -14.7 -9.6 0.8

Table 3.2: $R - Q_k$ residual matrices for best overall and worst individual periods. The values are standardised to days out of 365 such that a negative value means that particular node is under-represented by that many days and vice versa (i.e. $\sum_{i=1}^{N_s} (R_i - Q_{k,i}) = 0$)

Station	All-time Mean	Period Mean	Δ Mean	All-time Stdev	Period Stdev	Δ Stdev
Alexander Bay	4.45	4.63	-0.17	3.36	3.41	-0.05
Beaufort West	4.16	4.33	-0.17	2.58	2.64	-0.06
Bloemfontein	2.64	2.72	-0.08	2.11	2.18	-0.07
Cape Town	5.24	5.86	-0.63	3.21	3.62	-0.41
Durban	4.22	4.27	-0.04	2.56	2.61	-0.05
East London	4.57	4.74	-0.17	2.45	2.59	-0.15
George	3.25	3.42	-0.17	2.18	2.32	-0.14
Irene	3.20	3.35	-0.14	1.74	1.75	-0.01
Kimberley	3.78	3.95	-0.16	2.28	2.31	-0.03
Lamberts Bay	3.72	4.00	-0.27	1.93	2.06	-0.13
Langebaanweg	3.83	3.90	-0.07	2.31	2.38	-0.07
Nelspruit	1.94	2.07	-0.13	1.11	1.09	0.03
Pietersburg	3.37	3.32	0.06	2.01	2.07	-0.07
Port Alfred	3.80	4.33	-0.53	2.23	2.35	-0.12
Port Elizabeth	5.34	5.71	-0.37	3.36	3.54	-0.17
Thabazimbi	1.69	1.82	-0.13	1.19	1.25	-0.07
Vryheid	2.25	2.57	-0.31	1.52	1.40	0.12
Average	3.61	3.82	-0.20	2.24	2.33	-0.09
Standard Deviation	1.02	1.10	0.17	0.64	0.71	0.11

Table 3.3: Uncertainty introduced by optimal time period in ms^{-1} . All-time (for all available data) and optimum period wind speed average and standard deviation (stdev) for every station as well as a spatial aggregate are given

deviations on the one hand for all available data and on the other for the reduced time span. The detailed results given in Table 3.3 to the left provide us with two very useful conclusions: Firstly the errors introduced are relatively small, of the order of 0.2ms^{-1} for $\overline{V}$ and 0.1ms^{-1} for σ_V . This tells us that the algorithm works very well indeed. Only on two occasions, namely for Cape Town and Port Alfred, does the discrepancy exceed 0.5ms^{-1} ; the former is probably due to an over-representation of the synoptic state responsible for the south-easters that dominate the wind rose there.

Secondly the errors appear to be clearly systematic - with the exception of Pietersburg the observations recorded during the optimal period were windier than the long term conditions. This means that when evaluating our model output we can safely adjust the winds downwards to correct for this bias.

How such an alteration can be carried out is a bit trickier as we cannot simply subtract 0.2ms^{-1} from the simulated time series - this would not only result in negative wind speeds but also do nothing to reduce the variance which is also too high. Instead before presenting wind atlas statistics a scale factor l should be applied to all wind speeds such that

$$l\overline{V} = \overline{V} - 0.2\text{ms}^{-1} \quad (3.3)$$

l could easily be found in each case by using a standard binary search algorithm as found in the literature between $l = 0.5$ and $l = 1.0$ which terminates at a given specified accuracy, say, less than 0.0001ms^{-1} difference between the right and left hand sides of the above equation [Press et al. 2002]. In effect this will compress the histogram which is exactly what the numbers in Table 3.3 to the left suggest. In this way we ensure that the resulting wind climate really is as representative of the climatology as it could possibly be.

3.5 Caveats and Pitfalls

While the approach described in section 3.2 on page 42 and applied in section 3.3 on page 44 has shown strong potential and has yielded an optimum 365 day subset from 20 March 1996 to 19 March 1997, we need to carefully think again about the strengths and limitations of the algorithm.

Firstly we can, with 90% confidence or more, state that the above period is more representative of the long term average wind regime than any other contiguous 365 day window within 1993-2004. Thus the originally stated goal

has been achieved. Secondly as section 3.4 on page 51 has shown any residual bias can be corrected in a straight forward manner.

Furthermore, while we have achieved a good depiction of inter-annual patterns, the seasonal cycle cannot safely be assumed to be as well accommodated. To illustrate this point, April 1996 might have recorded totally anomalous winds but equal and opposite trends in May 1996 could have made up for the inconsistency such that the entire 1 year span still appears average to our algorithm. This adds extra uncertainty and needs to be borne in mind later when we extract seasonal figures from MM5.

As long as the averaging occurs over the entire year the daily cycle is less likely to be affected by this issue as we have taken 12 hourly values from across all possible synoptic conditions, some of which are of course more conducive to land and sea breezes than others. Therefore the typical diurnal wind oscillation should have been well preserved by our approach.

University of Cape Town

Chapter 4

Selection of Optimal Model Configuration

As outlined in section 1.6 on page 29 this chapter is designed to test the performance of various MM5 parametrisations and configurations to find an optimal configuration for a South African wind resource assessment at the target resolution of 18x18km. This grid length has been chosen from the outset of this study according to the availability of computational resources available to the author in terms of both disk space and CPU power, something which is demonstrated in more detail later in chapter 5 (see section 5.1 on page 107).

4.1 Overview

MM5 like many other comparable regional climate models has a large number of configurable options. Some of them are merely related to the nature and extents of the domain in question - such as number of grid cells in the x, y and z directions, horizontal and vertical resolutions, geographical location etc. Apart from having to adhere to a few basic rules there is little difficulty involved in adjusting these as the nature of the study together with the available computational resources will largely determine what they should be.

Then there are MM5's parametrisations which decide on how certain important physical processes in the atmosphere that occur at *sub-grid scales* are being approximated. There are many of them available though most of them are mutually exclusive and selecting which should be used for a given project is a

non-trivial task as their model specific implementations are often very complex. Which one performs best is, in general, dependent on grid resolution, latitude, season, location of domain (inland or coastal) and also the variables of interest. It often happens that a given parametrisation performs very well at simulating for example the nocturnal planetary boundary layer (PBL) but fails to properly account for the magnitude of daily temperature fluctuations.

Apart from domain settings and physics parametrisations there is also the question of constraining the freedom of the model to develop its own dynamics. For historical simulations (hindcasts) such as those that are performed in this thesis it is often desirable to keep the model reasonably close to available large scale GCM derived gridded datasets such as NCEP/NCAR (*grid nudging* or *four dimensional data assimilation*(FDDA)), especially when very extensive domains are under investigation. It can also prove beneficial to maintain close correlation to direct point observations on the ground and aloft (*observational nudging*). Such techniques may deliver better results as unrestrained runs.

In our case of course we are primarily interested in the wind patterns, but in picking an optimal model configuration we have to rely on partly international best practice as given in the literature and partly on experimentation for the specific purpose of a South African wind atlas. Important to remember is that no model is perfect and can only be expected to reproduce the first order response. Therefore we are looking for the best subset of available options for our task, not to rewrite the model's core in order improve performance.

4.2 Parametrisations in MM5

The MM5 manual gives a useful overview of the various schemes and their specific tunables available to us which are reproduced here only to a limited extent [Dudhia et al. 2004]. Each of them has a name in the corresponding FORTRAN code and is specified in special text files for input to the model. The variables and their values are given in brackets where appropriate.

Some parametrisations are judged to be highly influential on flow characteristics and therefore warrant very careful consideration in terms of both past research and experimentation. These will be discussed in detail in subsection 4.2.2 on page 60. Others have a much smaller impact on the wind and are thus kept constant, chosen according to what is efficient or what has been found to yield good results in the literature. We will turn to these next.

4.2.1 Constant Configurations

Regarding the representation of atmospheric radiation the Rapid Radiative Transfer Model(RRTM) is chosen (FRAD=4) [Mlawer et al. 1997]. RRTM itself employs MM5's cloud-radiation formulation for shortwave cloud interactions. For infrared effects and beyond it uses several spectral bands combined with a correlated-k method. It is designed to accurately simulate processes due to water vapour, ozone, CO₂ as well as accounting for cloud optical depth. For efficiency it is implemented with lookup tables.

An in-depth study over several states of the USA in 2004 revealed that RRTM is superior with respect to the other radiation parametrisations available in MM5 in terms of both accuracy and numerical performance [Tonnesen et al. 2004]. Other research over Houston/Galveston(Texas/USA) confirms that RRTM is indeed helpful in improving air quality modelling results [Tremback 2003].

To calculate explicit moisture the simple ice parametrisation is utilised (IMPHYS=4) [Dudhia 1989]. While not as comprehensive as some of its alternatives it is very efficient as it does not add memory overhead and requires minimal CPU resources thanks to a table based implementation (MPHYSTBL=1). While moisture and related cloud evolution does exhibit a certain influence on the flow the numerical efficiency here is deemed to be of greater importance and hence this is kept constant throughout this research. As an aside, the author did actually test a different scheme (mixed phase/Reisner 1, IMPHYS=5) but it yielded hardly any differences in the resultant winds, apart from increasing simulation time by 16% (detailed results not shown in this thesis).

In terms of boundary conditions a fairly standard approach is taken: Laterally a relaxation inflow/outflow condition is applied for the grid with the coarsest mesh in order to force the model with the global NCEP data (IBOUDY=3). This sets the outermost rows and columns equal to the input field while the next four points are relaxed towards the boundary values linearly decreasing to zero as one moves further inside the domain. Thus only the fifth row or column into the mother domain can be freely determined by the model's dynamics.

Nests use the time dependent approach whereby the outermost two points are determined by the parent's fields (IBOUDY=2). For lower boundary conditions the sea surface temperature(SST) fields are allowed to vary with time (ISSTVAR=1) while at the top of the grid a normal upper radiative boundary calculation is activated which reduces spurious noise from otherwise reflective phenomena (IFUPR=1).

Name [Ref]	Abbr.	Short Description	ISOIL
Force/Restore [Grell et al. 1995]	FR	Also known as Blackadar scheme with a single slab whose temperature is based on energy budget; constant substrate temperature	0
Five Layer [Dudhia 1996]	5LYR	Temperature prediction at approx. 1,2,4,8&16 cm otherwise similar to FR	1
NOAH [Chen and Dudhia 2001]	-	Four soil layers (10,30,60,100cm) with advanced vegetation, albedo and snow/ice interactions. Chooses dominant land use type for grid cell surface parameters	2
Pleim/Xiu [Xiu and Pleim 2000]	PX	two layer model (1&100cm) with different kinds of moisture fluxes. Aggregates all land use types and characteristics present in given grid cell	3

Table 4.1: Land surface models available in MM5. Each has an abbreviation, a corresponding configuration value (ISOIL) in MM5 as well as a reference

Furthermore where applicable the default 1-point feedback with smoother-desmoothing is used for propagation of nest fields to its parent domain (IFEED=3). This passes up the coincident point of the high resolution grid, and then removes two grid length noise by use of a filter.

Finally the way grids are nested can also produce different results. There is the one way nesting approach where the mother domain is integrated without any feedback from anything of higher resolution within it. The alternative is the two way method in which all domains are fully coupled. Most of the time the latter is preferred though because its treatment is physically more sound as confirmed by the literature and hence it is chosen here [Tremback 2003].

4.2.2 Variable Configurations

There are many other MM5 options over which opinion is divided and whose ideal choice with respect to a large scale wind resource assessment is far from obvious. Most important is the selection of land surface model (LSM) for vital contributors such as surface roughness and temperature as well as planetary boundary layer (PBL) representation for characterisation of vertical wind profile, advection effects governing the diurnal cycle etc.

Name [Ref]	Abbr.	Short description	IBLTYP	Possible ISOIL
Blackadar [1979]	BL	Oldest and most tested scheme with four stability regimes and free convective mixed layer; non-local closure	2	0,1
Burk Thompson [1989]	BT	turbulent kinetic energy (TKE) based on Mellor-Yamada formulae. Integrated force/restore ground interaction	3	0
ETA [Janjić 1994]	-	TKE based (similar to BT) but requiring prognostic LSM. Uses similarity theory and implicit diffusion scheme	4	1,2
MRF [Hong and Pan 1996]	-	based on Troen-Mahrt representation of counter-gradient term and K profile as in operational NCEP MRF model	5	1,2
Gayno Seaman [Shafran et al. 2000]	GS	Mellor-Yamada TKE based formulation with better accuracy for saturated conditions	6	0,1
Pleim Chang/Asymmetric Convective Model [1992]	ACM	derivative of BL with enhanced non-local vertical mixing	7	3

Table 4.2: Planetary boundary layer parametrisations in MM5 (IBLTYP) suitable for high vertical resolution. Compatible LSM's are shown in the fifth column

MM5 offers four LSM alternatives as outlined in Table 4.1 on page 60. NOAH and Pleim-Xiu(PX) are the most advanced and comprehensive ones although they require additional computational overhead. These are very closely coupled to the six PBL parametrisations and many of them are limited to certain combinations as given in Table 4.2 to the left.

One of the main differences between the PX LSM and the others is that the latter only consider the *dominant* (most frequent) land use category that occurs within a given grid cell. As the author has pointed out in an earlier study this may lead to inconsistencies near the coast where it often occurs that an incorrectly assigned “open water” category and associated surface characteristics such as roughness length z_0 may lead to large errors in the simulated wind field [Hagemann 2003]. PX instead takes an aggregate of all land use types present in a given grid cell by applying appropriate averaging techniques to all surface parameters (logarithmic for z_0) which is a physically more sound approach [Xiu and Pleim 2000].

Apart from different LSM and PBL simplifications the model’s representation of cumulus processes can have a significant impact on the model’s accuracy too. This is not only true for rainfall prediction skill (irrelevant in this context) but also for other near surface phenomena governed by up- and down-draft dynamics such as the all important daily cycle of wind. MM5 offers a total of seven such schemes applicable to resolutions of 18km and lower, detailed in Table 4.3 on the next page.

In addition to the above MM5 allows for very powerful four dimensional data assimilation (FDDA). FDDA is meant to keep the model’s dynamics closer to reality by “nudging” it towards a given observed dataset, be it gridded or at point locations both on the surface and aloft.

In practice this uses Newtonian relaxation which effectively minimises the error with respect to given wind, temperature and relative humidity fields with individually specified time and space constants to allow for a limited radius of influence. While this does add non-physical terms to the model’s equations and thus violates certain conservation laws it can reduce model errors due to shortcomings in the initialisation, treatment at the boundaries, physics etc. (see below for references). However, if MM5 is constrained too much by nudging that is too strong it limits its ability to develop the full variability in the processes it is supposed to reproduce, hence care is called for.

Specifically, the extent to which FDDA is applied to a simulation is controlled by a nudging factor G that is specified individually for winds (G_V),

Name [Ref]	Abbr	Short Description	ICUPA	Res. (km)
Anthes Kuo [Anthes et al. 1987]	AK	moisture convergence based, mostly convective rainfall	2	>30
Grell [Grell et al. 1995]	-	single cloud method based on rate of destabilisation allowing for both up and down-draft fluxes	3	10-30
Arakawa Schubert [Grell et al. 1995]	AS	similar to Grell but multi cloud	4	>30
Fritsch Chappell [Fritsch and Chappell 1980]	FC	relaxation based with single cloud, up and down-draft as well as subsidence though fixed entrainment rate	5	20-30
Kain Fritsch [Emanuel and Raymond 1993]	KF	similar to FC with several enhancements such as variable entrainment through cloud mixing scheme	6	
Betts Miller [Emanuel and Raymond 1993]	BM	uses relaxation towards reference thermodynamic profile; no down-draft allowance	7	>30
Kain Fritsch 2 [Kain 2004]	KF2	improved version of KF with shallow convection	8	

Table 4.3: Cumulus parametrisations available in MM5 (ICUPA) suitable for the target resolution. Most have an officially recommended horizontal grid size in km given in the second last column

temperature (G_T) and mixing ratio (G_q) and is typically of the order of 10^{-3} or 10^{-4}s^{-1} [Stauffer and Seaman 1990; Stauffer et al. 1991]. The relaxation towards the observed or analysis data happens asymptotically according to an inverse exponential function, accordingly G^{-1} provides a characteristic time scale of assimilation, under simplified conditions an e-folding time of adjustment [Xu et al. 2002]. Thus variables varying with high frequency (such as surface winds) should be nudged using a relatively large nudging factor G and vice versa. What this means in practice will be shown later in section 4.5 on page 90 when the different values of G used in this research are presented.

4.2.3 International Experience

As already mentioned in section 4.1 on page 57 an optimal LSM, PBL, cumulus, FDDA and nesting configuration is highly dependent on the exact nature and stated goals of the research at hand - there is no perfect combination for all applications. Hence it is crucial to consult international experience and best practice in the field first before embarking on a project.

The vast majority of available sensitivity studies of MM5 fall into the field of air quality modelling. Tonnesen et al. have carried out extensive testing of MM5 over the continental US at 36 and 12km using different LSM(5 layer, NOAH and PX), PBL (Blackadar, Eta, MRF and ACM), cumulus (KF2, BM and Grell) and FDDA (grid only, grid+obs, including&excluding surface and PBL nudging) options [Tonnesen et al. 2004].

Their comprehensive report concludes that PX/ACM is superior to other LSM/PBL duos, especially when augmented with Betts-Miller for cumulus processes. Best FDDA setup proved to be 3D grid nudging for wind (surface too), temperature and relative humidity in addition to observational wind nudging. They further state that the US Environmental Protection Agency (EPA) recommends the use of PX/ACM over others. Also they mention that they had to use a modified version of MRF since the original one in MM5 has deficiencies which cause irregularities in the diurnal cycle, something that we shall revisit at a later stage.

Zhang et al. discovered in their treatment of nocturnal low level jets that Blackadar yielded better reproductions of vertical wind profiles than the Gayno-Seaman parametrisation [Zhang et al. 2001]. They note that limiting grid FDDA to above the planetary boundary layer also improved MM5's accuracy and thus recommend to always restrict FDDA in that way, consistent with best practice.

In a different commercial study it is found that the MRF PBL scheme outperformed ETA, GS and BT on grids up to 4km resolution (PX was not investigated) [Tremback 2003]. Interestingly enough it mentions the same shortcomings of MRF as Tonnesen et al. and considerable effort is spent at arriving at a modification of MRF that mitigates this problem. Further it concludes that two way nesting should always be used as it is numerically and meteorologically more consistent. In terms of FDDA it advises that observational nudging is more effective than grid nudging, at least for surface fields owing to the limited resolution and associated issues in the analysis. According to its authors a TKE based scheme such as Gayno-Seaman should in theory be more robust for more general cases (e.g. development of sea breezes), however they state MM5's implementation of the Mellor-Yamada parametrisations leaves much to be desired.

Zhang and Zheng performed 3 day simulations of weak gradient flows over the central US utilising all MM5 PBL treatments except for ACM [Zhang and Zheng 2004]. They find that Blackadar manages to reproduce the oscillations in surface wind patterns best due to its formulation of direct mass and momentum exchanges and its numerous stability regimes. Gayno-Seaman and MRF on the other hand performed very poorly in terms of both timing and amplitude of the daily peak, despite the former's more physically sound TKE algorithm. They further discover that a faithful model representation of the diurnal cycle of temperature does not at all imply that the cycle of surface winds are realistic to the same degree.

As the reader can see there are numerous published works which deal with the subject of MM5 parametrisation sensitivity. However, they do not present a cohesive picture of what should be used in our case as there is not a single openly available study that has used MM5 for wind resource assessments over latitudes and climate as unique as South Africa, not to mention experience at the target resolution of 18km. Furthermore air quality investigations often focus on low wind episodes which are mostly irrelevant in the context of wind power.

Therefore we conclude that it is insufficient to only look at theoretical aspects or overseas experience. We are left with little choice but to carry out our own customised test simulations of MM5 with different options to optimise the model for the specific purpose of a nationwide wind atlas.

Next we shall deal with the basic design of the experiments while the following sections will present the reader with different LSM, PBL, cumulus and FDDA settings in search for the most suitable configuration.

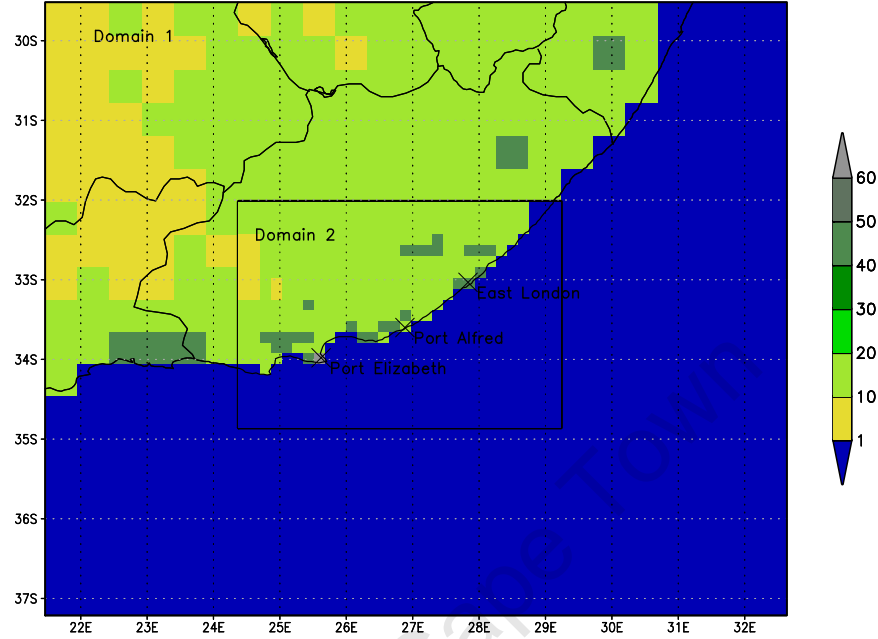


Figure 4.1: Domain configuration of the Eastern Cape sensitivity runs. The colour scheme gives the surface roughness length z_0 in cm of the dominant land use type in each grid cell as employed by all land surface schemes in MM5 except for Pleim/Xiu. Gray/green patches are mostly forest related. The parent domain (#1) and its nest (#2) are individually labelled, the former with the coarse 54km and the latter with the finer 18km resolution as indicated by the relative sizes of the coloured grid boxes.

4.3 Design of Sensitivity Simulations

4.3.1 Domain, Nesting and Surface Configuration

Any model sensitivity study often needs to be limited in extent, resolution and time period due to the fact that dozens of them with slightly perturbed configurations have to be carried out but the available computational resources are very limited.

As we want to find one of the best combinations for all seasons at the same resolution as the final wind atlas we have to limit the size of the domain but keep everything else identical or at least very similar to the envisaged South Africa wide simulation. We choose the Eastern Cape province as the centre of our efforts (33.45°S and 27.05°W) with exact grid cell and nest layout as illustrated

in Figure 4.1 on the preceding page.

There are several reasons for this selection. Firstly the density of stations here is relatively high (see Figure 2.1 on page 32) as three of the SAWS weather stations listed in section 2.1 on page 31 coincide with this region: Port Elizabeth(PE) in the Southeast, Port Alfred(PA) in the centre and East London(EL) in the Northwest. Secondly the region incorporates key physical features; strong topographical forcing, subject to both summer convective rainfall and winter frontal systems as well as a strong influence by regional oceanic features including the Agulhas current.

Thirdly the Eastern Cape is a very important province in terms of South Africa’s rural development priorities and developing wind farms would bring much needed employment to the area. Lastly one might argue that we are effectively optimise the model for only coastal conditions but neglecting inland conditions, however, near the sea is exactly where most of the wind resource has been found to be located in the past [Diab 1995]. Hence a decent model performance here would be of great value.

Two domains are employed on a Mercator map projection, the outer one or “parent” (labelled Domain 1 in Figure 4.1 on the previous page) with 54km and the inner one or “nest” (labelled Domain 2 in Figure 4.1) with 18km horizontal resolution. The choice of 54km as the coarse grid length is due to MM5’s requirement of a 1:3 horizontal staggering ratio. The parent domain is forced on the boundary with NCEP/NCAR Reanalysis data as described in section 4.3.2 on page 71 while the nest obtains its boundary conditions entirely from its parent in the standard two way nesting fashion, i.e. both are integrated computationally at the same time.

The number of cross points are 23x19 and 30x21 for the outer and inner mesh respectively. The lower left corner of the nest is chosen to be at point (7,7) allowing for at least two free cells outside of the nest’s boundary. Similarly there are at least six cells between any station and the outer edge of domain 2 ensuring that we can fully capitalise on the higher resolution. Since both domain extents and number of grid points are quite small the simulations are likely to be tightly driven by the boundary conditions.

Terrain height data is derived from the global United States Geological Survey (USGS) GTOPO30 30 arc second (~ 0.9 km) digital elevation model (available online at <http://edc.usgs.gov/products/elevation/gtopo30/gtopo30.html>) [Gesch et al. 1999]. The mother domain uses MM5’s regridded five arc minute (~ 9 km) dataset while the nest employs the full 30” information to calculate

the elevation mesh with a two pass smoother/desmoother scheme. The full configuration of the TERRAIN sub-component of MM5 for these sensitivity runs is given in Appendix B.1.1 on page 169.

Land use information is also taken from the USGS global 30 arc second dataset using the standard 24 category classification [Anderson et al. 1976]. Each of the 24 defined classes have a specified summer and winter surface roughness length z_0 amongst four other parameters - albedo, moisture availability, emissivity and thermal inertia [Loveland et al. 2000]. As the dominant grouping within a grid cell is taken with all LSM's except for PX this yields very little spatial variability and a somewhat grossly simplified representation of reality as the reader may see in Figure 4.1.

Each category is supposed to have only one summer and winter value while the LSM would then vary over them using a simple sinusoidal representation depending on season. PX does not vary z_0 seasonally but as mentioned above applies logarithmic averaging of all sub-grid scale z_0 to arrive at the final lower boundary condition [Xiu and Pleim 2000]. The lack of such seasonal variability is deemed acceptable since the majority of South Africa's vegetation is perennial due to relatively warm winters, hence land surface friction is expected to be relatively constant throughout the year.

However, the treatment of z_0 is not quite consistent within MM5. The problem is that the default LSM values for the Force/Restore, 5 Layer and NOAH LSM's as determined by the user modifiable "LANDUSE.TBL" file in the model's Run directory are different to those hardcoded in PX's surface routines (vegeland.F) as outlined in Table 4.4 on the following page. Even worse the officially documented variables in the MM5 manual are in places different to both of these though this could be attributable to printing errors.

These changes can be significant, at least for the urban land category. As can be seen in Figure 4.1 on page 66 our Port Elizabeth measurement coincides with such an environment according to MM5. With a much lower z_0 value of 50cm instead of 80cm the PX/ACM combination is likely to simulate stronger winds at this location compared to its other LSM/PBL counterparts.

PX does not in its standard configuration output its aggregated land use parameters for later analysis but the author has patched its FORTRAN source code to allow for this, see section C.1 on page 191 in the appendix. The results are subtracted from MM5's standard z_0 depiction to highlight the differences between the two as given in Figure 4.2 on page 70.

As expected the spatial variability is much higher in the PX surface scheme

#	Description	MM5 manual		LANDUSE.TBL		PX
		Summer	Winter	Summer	Winter	
1	Urban	50	50	80	80	50
2	Dryland Cropland and Pasture	15	5	15	5	10
3	Irrigated Cropland and Pasture	15	5	15	5	10
4	Mixed Dry/Irr Crop and Pasture	15	5	15	5	10
5	Grassland / Cropland Mosaic	14	5	14	5	10
6	Woodland / Cropland Mosaic	20	20	20	20	40
7	Grassland	12	10	12	10	7
8	Shrubland	10	10	10	10	20
9	Mixed Shrubland/Grassland	11	10	11	10	20
10	Savanna	15	15	15	15	20
11	Broadleaf Deciduous Forest	50	50	50	50	50
12	Deciduous Coniferous Forest	50	50	50	50	50
13	Evergreen Broadleaf Forest	50	50	50	50	40
14	Evergreen Coniferous Forest	50	50	50	50	50
15	Mixed forest	50	50	50	50	50
16	Water body	0.01	0.01	0.01	0.01	0.001
17	Herbaceous Wetland	20	20	20	20	15
18	Forested Wetlands	40	40	40	40	45
19	Barren or Sparsely Vegetated	10	10	10	10	5
20	Herbaceous Tundra	10	10	10	10	10
21	Shrub and Brush Tundra	30	30	30	30	10
22	Mixed Tundra	15	15	15	15	5
23	Bare Ground Tundra	0.10	5	10	5	5
24	Perennial Snowfields or Glaciers	0.01	0.01	5	5	5
25	Missing data	n/a	n/a	n/a	n/a	n/a

Table 4.4: USGS 24 category land use classification and associated z_0 values in cm. Values from the MM5 manual (page 68, [Dudhia et al. 2004]) are given for comparison but the important ones are those from the default LANDUSE.TBL (user modifiable table) and the ones in the default PX source code (physics/pbl_sfc/pxpbl/vegeland.F). **Bold** entries highlight those where the latter two are different

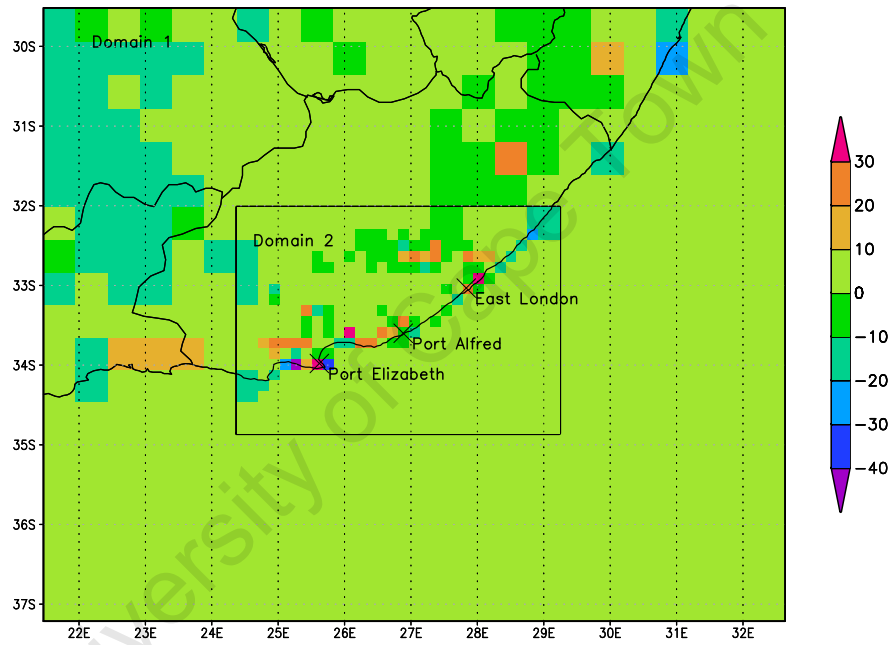


Figure 4.2: MM5 standard minus PX surface roughness lengths z_0 for both domains in cm. The green shaded areas indicate insignificant differences ($|\Delta z_0| < 0.1$) between the two while the other colours show much higher disagreements

and z_0 over the Port Elizabeth observation is around 30cm lower. However, the grid cell just west of PE is treated as much rougher and there are many other large changes around. Port Alfred's coincident square has a similar z_0 though some surrounding (North and West) areas are smoother in the PX LSM. The East London station's surface is modelled with much less friction too though the grid cell to the North East exhibits the same urban anomaly as in PE.

Seeing that the available measurements for validation are all at 10m above ground the influence of z_0 on the wind regime is likely to be quite strong. As such the above discussion indicates that we are likely to obtain higher wind speeds using the PX/ACM parametrisation than with any of the others.

4.3.2 FDDA, Initial and Boundary Conditions

To keep consistency with the recommendations of section 3.3 on page 44 the simulation period of 15 March 1996 00:00 to 20 March 1997 00:00 is chosen allowing for an additional five day spin up period to ensure the model can sufficiently develop its own dynamics before its output is critically evaluated against observed wind patterns.

All sensitivity simulations share the same initialisation which is based on 2.5° resolution NCEP/NCAR Reanalysis data at 15/03/1996 00:00 UTC. The outer boundaries thereof are chosen such that they provide a reasonably sized buffer around the 54km parent domain (10°S-45°S and 5°W-45°W). No objective analysis is performed as it is deemed unnecessary for hindcasts of this type, FDDA is used in some model configurations in order to bring the model closer to observations. The boundary conditions for MM5 are similarly derived from the Reanalysis for the entire 370 day temporal span at six hourly intervals.

The available NCEP data in netCDF format is pre-processed using custom scripts into the intermediate format mandated by REGRID, one file per timestep [Dudhia et al. 2004]. The variables u , v , air temperature, relative humidity, geopotential height, sea level pressure and sea surface temperature (SST) are extracted first as the minimum requirement for REGRID. For the LSM input 0-10cm, 10-200cm and 300cm soil temperature as well as 0-10cm and 10-200cm soil moisture were available. From here REGRID's main processor interpolates these values onto the model grid defined previously. The exact REGRID settings can be viewed in subsection B.1.2 on page 172 in the appendix.

MM5's input data for gridded FDDA is prepared identically just that the pre-processed data comprises all 6h timesteps falling within the entire simu-

lation period, not just the first one. For observational nudging the necessary MM5OBS_DOMAIN2 file (only for nest) is a lot trickier to create since no ready made tools exist for this purpose. Some custom code was written to create this input containing surface wind only with $\Delta t = 1\text{h}$ based on the available stations in the region. Due to the fact that obs nudging creates to some degree a dependence of the model output on the observed inputs which would unnecessarily skew our validation results Port Alfred is excluded so that only Port Elizabeth and East London are utilised for this purpose.

Unfortunately the standard MM5 implementation has problems with long term simulations that span year boundaries and involve station nudging. That's due to MM5 allowing the input format to specify a Julian day and hour only (dddhh - e.g. 16623.5 is Julian day 166 and hour 2330 UTC) which means that if the day wraps around to 001 on a year boundary the model's FDDA code switches off due to a "malformed" date (they are required to be in strictly increasing order). The author had to work around this limitation with a custom made patch to MM5's FORTRAN code so that different years can be accommodated. For details the reader is referred to Appendix C.2 on page 192.

4.3.3 Vertical Resolution

Once the horizontal input adjustments and FDDA preparations are completed the last preparatory step of vertical redistribution of pressure to sigma level data is performed with MM5's INTERPF sub-model. Forty vertical levels are chosen with a very high resolution in the PBL having six of them below 200m above ground. This is done in order to match the levels that we need not only for our validation (10m) but also for general wind resource assessment purposes. Forty is also considered a fair trade-off between high vertical resolution and the amount of computational resources required.

The nature of sigma levels implies that their exact height above ground is dependent on terrain elevation and thus decreases slightly as one moves inland. Also, according to MM5's design u and v are defined on half sigma levels. Taking all this into consideration the first five full s levels are chosen to be 1.000, 0.9975, 0.995, 0.991 and 0.986.

These correspond to the lowest four half s levels of roughly 10m (9.9m - 10.3m), 30m (29.6m - 31.0m), 56m (55.2m - 57.9m) and 93m (91m-95m) above ground over the Eastern Cape parent domain with the lower values over the Drakensberg escarpment and the greater heights over the Indian Ocean as can

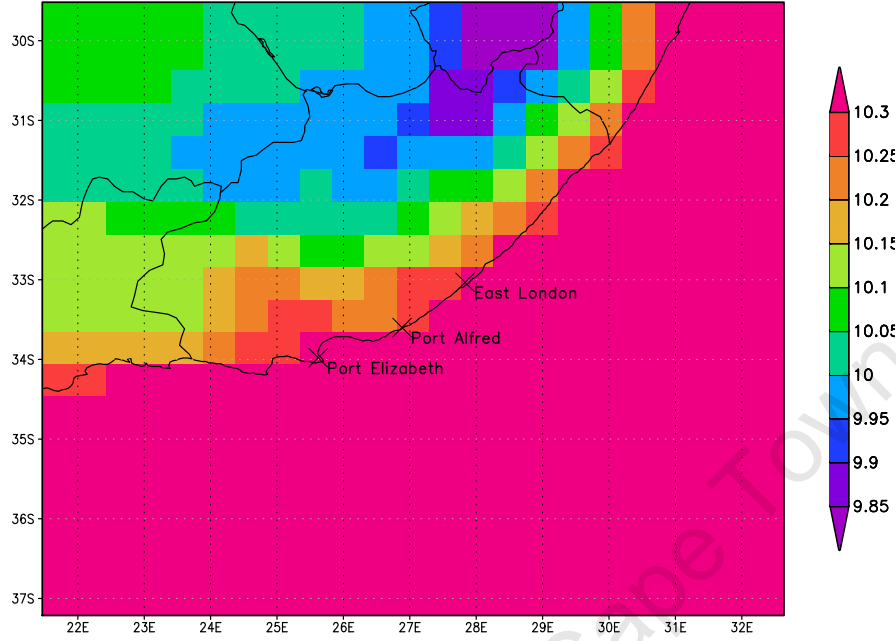


Figure 4.3: Height above ground of the lowest half sigma level in metres ($s = 0.99875$). Only the course grid parent domain is shown

be seen in Figure 4.3 for $z = 1$.

The exact INTERPF configuration with all other s levels, constant parameters and more which are mostly kept at their defaults may be viewed in Appendix B.1.3 on page 173.

4.3.4 General Model Configuration

In order to save on disk space and increase efficiency of the model post-processing MM5's output has been considerably constrained by applying the patch described in Appendix C.3 on page 198. It reduced the file size of the results by a factor of three to about 3GB (from 9GB) which significantly aided in carrying out the large number of simulations.

In terms of compile time options both domains consistently use the simple ice explicit moisture scheme with look up tables (IMPHYS = 4, MPHYSTBL = 1), RRTM radiation (FRAD = 4) and the shallow convection option turned on (ISHALLO = 1). Everything else is varied as described in the next sections.

The integration is run from 15 March 1996 until 20 March 1997, i.e. for

the full optimal period found in section 3.3 on page 44 with an initial 5 day spin up period. Hence the forecast time is 532800 (370 days) with a coarse grid timestep of 90 seconds. The latter defies the rule of thumb of 3 times the lowest horizontal resolution (54km) as higher values kept resulting in CFL violations. Model output is written every hour with restart saves done every 6 hours of model time. LSM initialisation for the PX runs is done via the values from REGRID as opposed to the LANDUSE.TBL figures and the soil temperature and moisture levels are adjusted according to the Reanalysis input.

FDDA, if activated, is always performed over the entire integration. Grid-ded nudging is applied based on the 6 hourly NCEP data only on the coarse domain for wind, temperature and mixing ratio, excluding the PBL. Observational Newtonian relaxation is done only on the high resolution domain at Port Elizabeth and East London for wind only using the hourly input with a time window half-period of 30min. Horizontal and vertical radii of influence are chosen to be 20km and 0.001 s units respectively to prevent excessive spreading of these corrections.

All other tunables are kept that their defaults which are physically consistent. A generic `configure.user` (compile time) and `mm5.deck` (run time) configuration file highlighting exactly how certain settings were kept constant and which were varied may be found in Appendix B.1.4 on page 174.

Each of these simulations took between 46h and 55h to run to completion on a 64bit AMD Opteron dual processor Linux based (Suse) machine with 4GB of RAM and only one CPU active. Actual RAM usage of the `mm5` process was much lower at about 20MB.

4.3.5 Validation Procedure and Pitfalls

From the MM5 output of the high resolution domain we derive the hourly time series from the grid cell that coincides with a particular observation as illustrated well in Figure 4.1 on page 66. The resulting modelled and observed time series S_m and S_o are then analysed side by side.

That approach in itself, however, does have its shortcomings: we are comparing a 18x18km *area average* with *point readings* which are fundamentally different. Only if the measurement happens to be fairly typical of the wind patterns that occur inside the box, i.e. at scales smaller than the model can resolve, can we accurately claim for example that the model over-predicts the wind speed.

In reality, however, each observation is likely to be governed uniquely by highly localised surface roughness and topographical distortions as described in subsection 1.2.3 on page 6, in essence making it more probable that it claims a wind climate that is removed from the real mesoscale average. In theory, should in such a case S_m be equal to S_o we do *not* have a perfect model but rather one that is actually highly erroneous.

Unfortunately it is impossible to find out what the real world grid average is as it would require a measuring campaign involving literally millions of closely interspersed anemometers throughout a given area and then taking the arithmetic mean of all of them. Therefore we are left with slightly inconsistent comparisons; the only way to mitigate this problem is to average the results over as many observations as we can, relying on the fact that their variability with respect to the typical wind climate in their 18x18km cell is smoothed out in the process.

In any event this issue is not disastrous in itself. As long as we are aware of it we can still meaningfully conduct such verifications. For example, a bias between modelled and observed average wind speeds is not necessarily indicative of a shortcoming of MM5.

In order to see how well S_m actually matches S_o we apply all the concepts that were introduced in Section 2.2 on page 33 for validation according to the special requirements of a regional scale wind resource assessment. It is recommended that the reader revises these at this point in order to simplify understanding of the detailed evaluation carried out below. We will show wind speed bias b_V , standard deviation bias of speed c_V , cycle amplitude error CAE , peak phase error PPE , daily Fourier peak difference $DFPD$, 12 monthly(seasonal) mean absolute error $MMAE$, weighted MAE of sector frequencies $WMAESF$, weighted MAE of sector speed averages $WMAESSA$, weighted MAE of sector speed standard deviations $WMAESSD$ and finally the weighted MAE of sector wind power densities $WMAEPD$ for each and every run in graphical format.

The logically related metrics are grouped together and visualised on the x and y axis respectively. For all of them a zero indicates perfect agreement, so for each run we visualise that combination in a *goal plot* in which the axes are dynamically scaled but the exact location of the origin is clearly marked with a black arrow. In that way the reader can immediately differentiate between a low skill and good performance run by looking at the distance of the point to the arrow tip, yet at the same time get a good idea of MM5's total variability

Run Description	Short run name	ISOIL	IBLTYP
NOAH LSM/Eta	NOAH/Eta	2	4
NOAH LSM/MRF	NOAH/MRF	2	5
Five Layer/Blackadar	5LYR/BL	1	2
Five Layer/Gayno-Seaman	5LYR/GS	1	6
Force Restore/Burk-Thompson	FR/BT	0	3
Pleim Xiu/ACM	PX/ACM	3	7

Table 4.5: Configuration of the six LSM/PBL sensitivity runs

across all its parametrisations.

The only disadvantage of this approach is that the scales always change and thus what may look like a large difference may in reality account for little more than noise because the absolute values of the statistics do not vary by much. Hence this is taken into consideration where necessary in the discussion that follows.

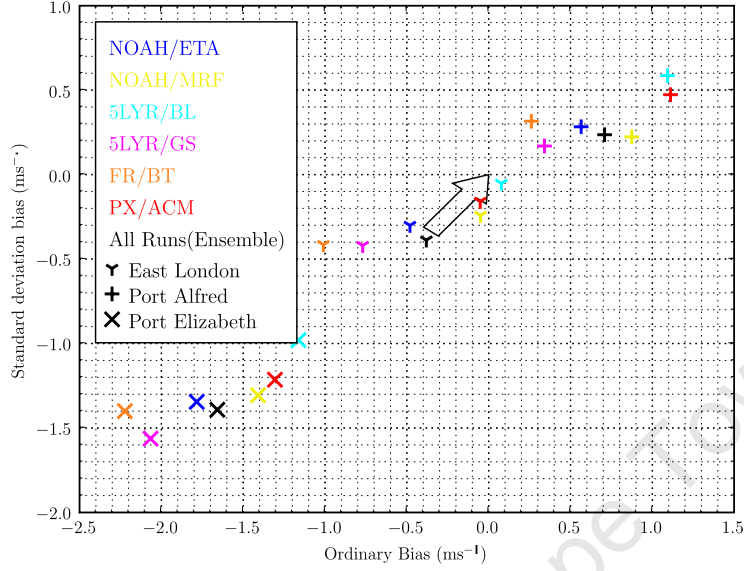
It should be noted that some discussions of model skill take a “reference” forecast and then compare subsequent perturbations to that, however that approach would not work well in this context since we do not yet know how such a base case performs in absolute terms.

4.4 Evaluation of LSM and PBL Parametrizations

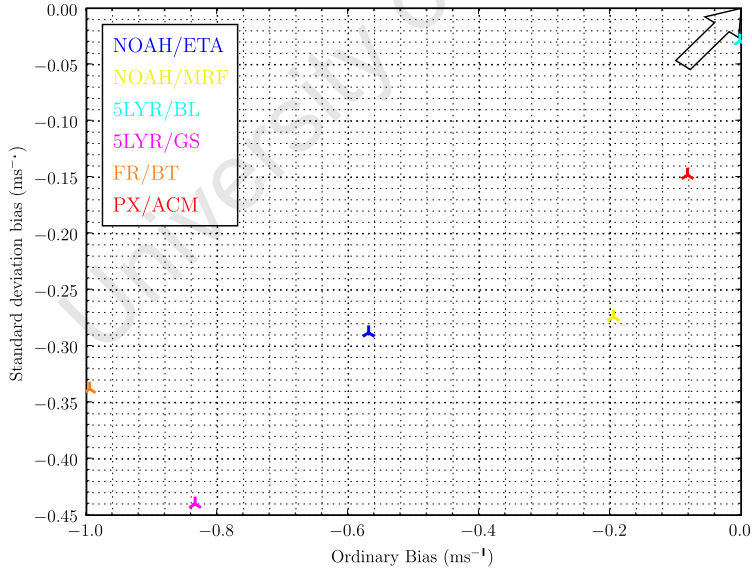
The first set of our actual sensitivity tests investigates the skill of all major LSM and PBL combinations. We test each PBL parametrisation in turn with the most complex and physically sound LSM it supports (see Table 4.2 on page 61) which results in the 6 permutations outlined in Table 4.5. All of them use the Grell cumulus scheme with shallow convection turned on. The rest of the parameters may be extracted from Appendix B.1.4.

4.4.1 Average Wind Speed and Standard Deviation

First we compare the b_V and c_V skill for all stations and all runs in Figure 4.4(a) on the next page (again note that the arrow points at the origin). As the reader can see different symbols have been chosen for the individual observations while different colours have been chosen to depict the various LSM/PBL combinations. An average over all six runs (“ensemble”) has been included in order to provide



(a) Individual stations. Inverted tripods are for EL, pluses for PA and crosses for PE. Black markers represent the ensemble (averaged over all runs)



(b) Spatial average over all three observations (represented by normal tripods)

Figure 4.4: Goal plot of standard deviation bias c_V vs ordinary bias b_V of wind speed for all six LSM/PBL runs, each depicted by a certain colour

a good indication of the inherent, parametrisation agnostic bias of MM5 for a given measurement.

A lot of information may be gleaned from this representation. Firstly the results cluster quite significantly per location irrespective of the LSM/PBL scheme as indicated by the ensemble points. This is what we would expect from comparing point observations with grid cell averages as pointed out in subsection 4.3.5 on page 74.

Port Alfred's $\bar{V}$ and σ are consistently over-predicted by small amounts ($\sim 0.5\text{ms}^{-1}$), for East London it is the opposite (except for Blackadar) while there seems to be a very strong negative model bias for Port Elizabeth. In the latter case this tells us that either the model is performing exceptionally poorly here or, which the author considers more likely, that localised effects at the PE airport skew our results. A plausible explanation would be that MM5 uses a very high roughness length for this grid cell which is unrealistic at the airport since the anemometer is mostly surrounded by short and regularly cut grass in the vicinity of the airport runway (i.e. a low roughness surrounding).

Also interesting to note is a significant general gradient from high $\bar{V}$ and σ of 5LYR/BL to low $\bar{V}$ and σ of 5LYR/GS with FR/BT having by far the lowest speed but slightly higher variance than 5LYR/GS. The three ranks in between are also fairly regularly populated with NOAH/Eta being just left of the ensemble and PX/ACM and NOAH/MRF to the right of it. This is reassuring because it simplifies making a reliable judgement about what kind of coastal winds one can expect from the six LSM/PBL combinations.

Even though we hypothesised in subsection 4.3.1 on page 66 that PX/ACM would yield the highest wind speeds due to its typically lower roughness specification in the LSM (see Figure 4.2 on page 70) this did not quite turn out to be the case. While PX/ACM certainly has stronger winds than NOAH/Eta, 5LYR/GS and FR/BT, the differences are insignificant compared to NOAH/MRF and even inverted in the case of 5LYR/BL. This clearly shows that there are dynamics other than those directly affected by z_0 which play a much more important role in the determination of the flow, perhaps the local and non-local vertical mixing and associated momentum transfer that is handled very differently by the six PBL schemes. This does, however, cast a shadow on the performance of Blackadar.

In order to mitigate some of the grid vs point problems elaborated on in subsection 4.3.5 on page 74 we now turn to the comparison of spatial averages. In essence this means that the three time series for PE, EL and PA are combined

into one with equal weighting and are juxtaposed to the similarly derived MM5 aggregate.

The results are illustrated in Figure 4.4(b) on page 77 where each simulation is now denoted by a coloured square. They confirm the trends we just outlined with perhaps a surprising winner: 5LYR/BL exhibits the best performance with PX/ACM being a close second and followed by NOAH/MRF, though the scales are now much smaller ($\sim 0.3\text{ms}^{-1}$). Only NOAH/Eta as well as to an even larger extent 5LYR/GS and FR/BT have significant problems in simulating the 10m winds well and are thus unlikely candidates for a wind atlas.

4.4.2 Daily and Seasonal Cycles

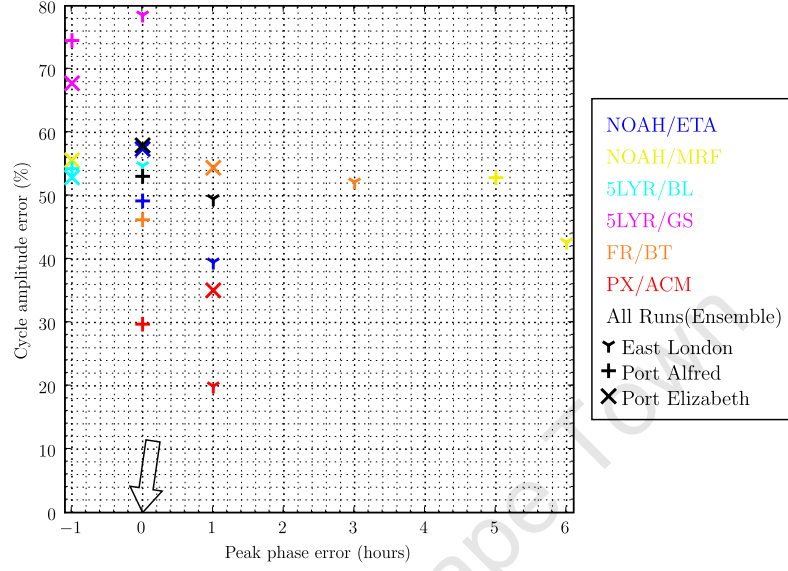
Turning now to the skill evaluation of MM5's diurnal wind cycle, we plot wind speed cycle amplitude error *CAE* vs peak phase error *PPE* in the same style as Figure 4.4 on page 77 (with individual, ensemble and spatial averages) in Figure 4.5 to the right.

Looking at the individual per station results in Figure 4.5(a) to the right one can see that PX/ACM is clearly far ahead of the others as it comes no less than 20% closer to getting the amplitude right than its closest competitor NOAH/Eta, though both have a timing within one hour of occurrence of the observed peak. 5LYR/BL and FR/BT provide mediocre performance with FR/BT's wind maximum for East London being three hours too early. 5LYR/GS has got significant problems getting to simulate any sort of cycle as its *CAE* lies above 65% for all three locations.

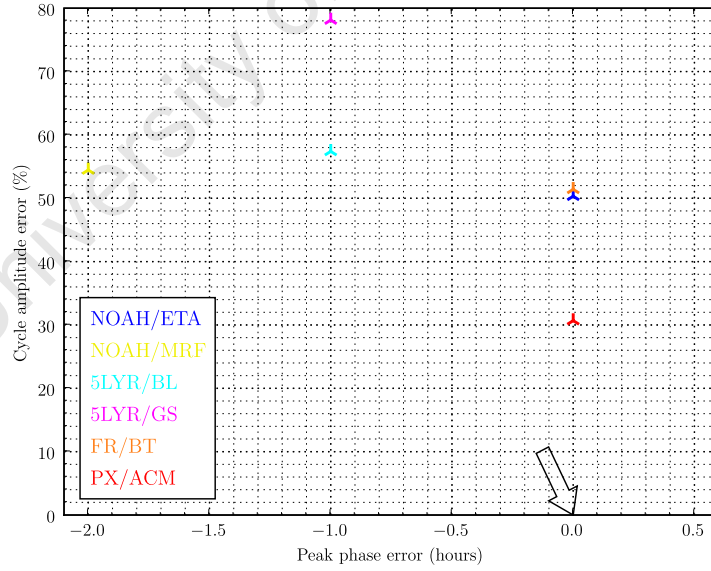
What is interesting to point out is the poor score of the otherwise often praised NOAH/MRF combination. While its *CAE* values of around 50% are in the region of the others its *PPE* is exceptionally large with five and six hours too early for Port Alfred and East London respectively. We shall suggest a logical explanation for this below.

As we pointed out in subsection 2.2.2 on page 36 when we introduced the *PPA* and *CAE* metrics these two are fairly meaningless if a clear cycle is not present in the data in the first place. Therefore as a cross check the 24h wind periodicity needs to be visually verified which is carried out for all runs and all stations in Figure 4.6 on page 81.

A strong cycle is present in nature at all three airports with a D_a of about 3.4, 3.2 and 2.5ms^{-1} for PE, PA and EL respectively. However, all schemes fail to reproduce the full amplitude and especially the calmer night episodes. Only



(a) Individual stations



(b) Spatial average over all stations

Figure 4.5: Goal plot of daily cycle amplitude error CAE vs daily peak phase error PPE of wind speed for all LSM/PBL runs

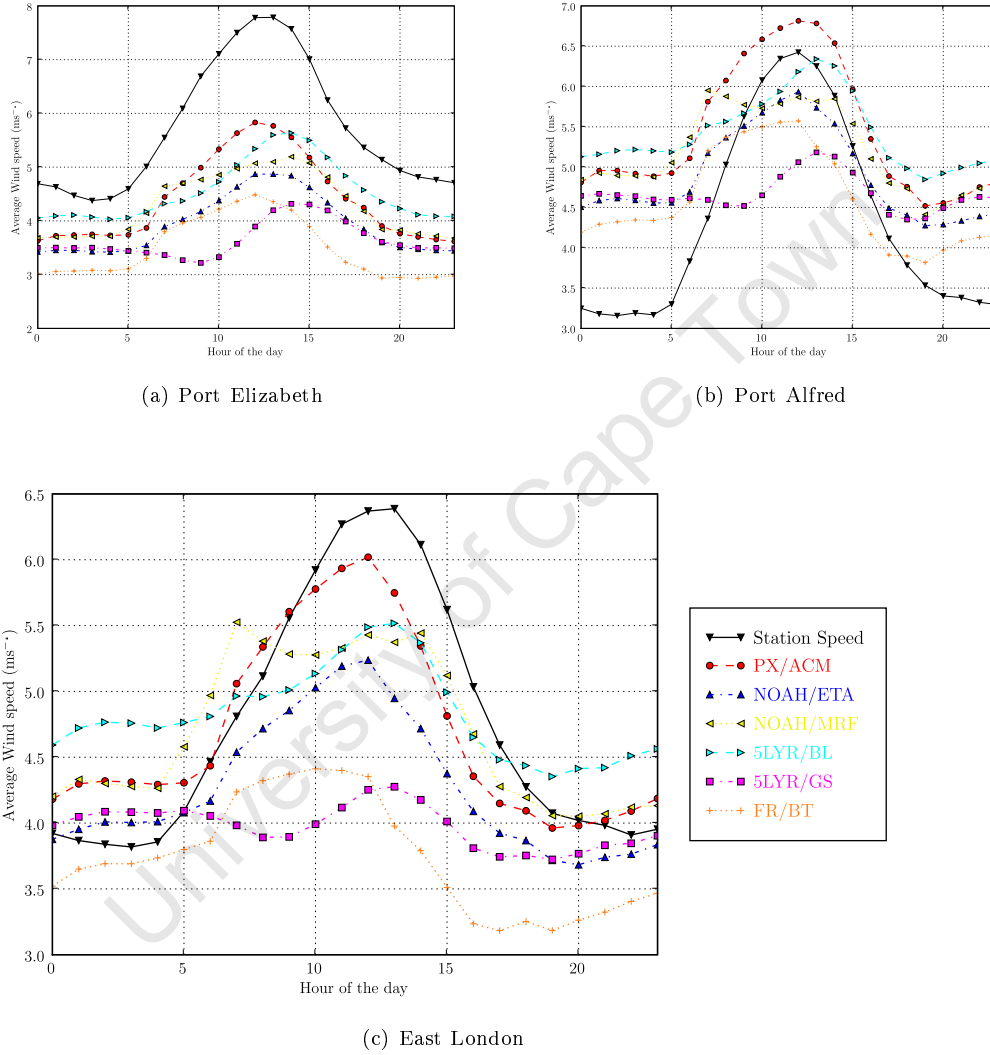


Figure 4.6: Average diurnal cycles for all LSM/PBL runs. Each parametrisation is represented by the same colour as used in Figure 4.4 on page 77 and a different symbol alongside the observed (black) pattern in terms of speed V .

PX/ACM and NOAH/Eta are satisfactory here given the still relatively coarse 18km resolution. Compared to almost non existing 24h variability at 50km resolution of MM5 this is a vast improvement [Hagemann 2003]. Furthermore, this view confirms the findings above: both 5LYR/BL and FR/BT generate a cycle that is too weak or misplaced while 5LYR/GS has almost no harmonics whatsoever.

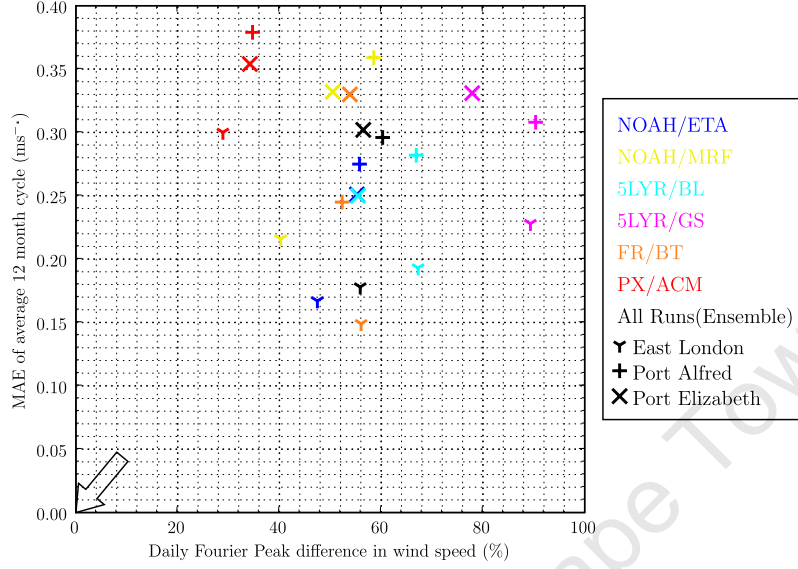
In Figure 4.6 to the left also lies more proof for NOAH/MRF's suboptimal performance. From early morning (highest speed for East London is at 7am) to mid afternoon its wind profile is almost flat (cut off) and no smooth curve ever occurs. This might well be the same artefact of MM5's MRF implementation bugs leading to an abnormal PBL evolution that was noted in other studies and referred to earlier in subsection 4.2.3 on page 64. While the fixes mentioned there could certainly help this is not deemed necessary since PX/ACM has already shown to be quite capable of simulating the 24h component.

Finally the individual cycles provide us with more clues as to MM5's inherent, parametrisation agnostic bias at the three stations. At Port Elizabeth the measured V_{24} is *always* larger, no matter what the time of the day. This confirms the earlier suspicion that very localised roughness length effects at that airport must be responsible for the under-prediction. At Port Alfred, however, it seems to be mostly the model's inability to simulate the much lower wind speeds at night that causes the positive overall bias. It is due to a much lower daily peak that FR/BT and 5LYR/GS are closest to the observed mean at PA - essentially best performance for all the wrong reasons.

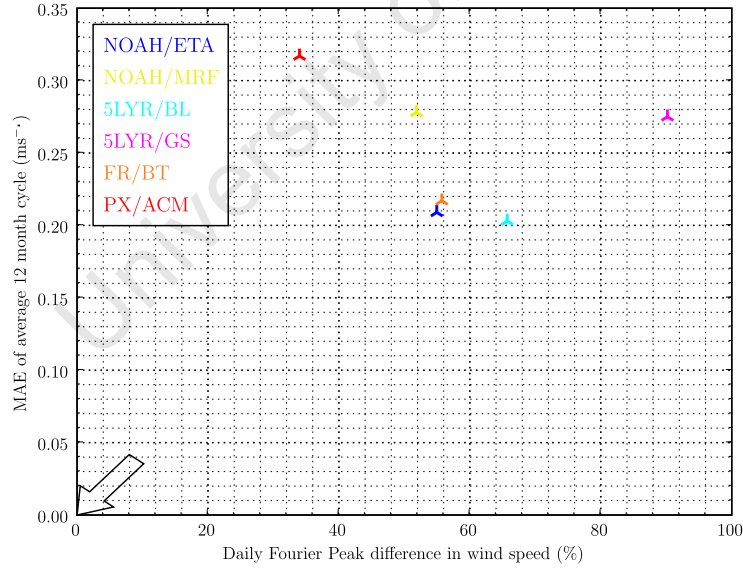
It also becomes clear now why Blackadar consistently gives us the highest wind speeds - its night time V is simply too high driving up its $\bar{V}$ significantly. It suggests that its parametrisation of turbulent exchange in terms of the Richardson number is too simplistic for stable night time regimes. It could also generate a land breeze that is too strong. This is another reason 5LYR/BL should not be recommended for use in the final wind atlas.

We now turn to the remaining two cyclical metrics, namely the daily Fourier peak difference $DFPD$ and the 12 monthly (seasonal) mean absolute error $MMAE$ which are outlined in Figure 4.7 on the following page for all runs.

Before evaluating these results in detail it must be reiterated that the $DFPD$ is just a different way of looking at the 24h cycle skill and thus will likely not add more information than what has already been discovered above (see subsection 2.2.2 on page 36 for a detailed explanation). Thus the y-values ($MMAE$) of the graph are much more interesting. However, for a two dimensional repre-



(a) Individual stations



(b) Spatial average over all stations

Figure 4.7: Goal plot of 12 monthly(seasonal) mean absolute error $MMAE$ vs daily Fourier peak difference $DFPD$ of wind speed for all LSM/PBL runs

sensation the *MMAE* needs to be paired with a second metric and that is the reason the *DFPD* is kept. The reader is also reminded of the fact that both of these statistics are based on time series whose bias has already been subtracted so that it does not get included again in the analysis.

In terms of *DFPD* (x-values) the previous results regarding the daily cycle are indeed very much reaffirmed: PX/PBL comes out on top with a total *DFPD* of less than 35%. The NOAH and FR/BT algorithms share the second place with values of between 50% and 60% while again 5LYR/GS is the worst candidate with a 90% shortfall (see Figure 4.7(b)).

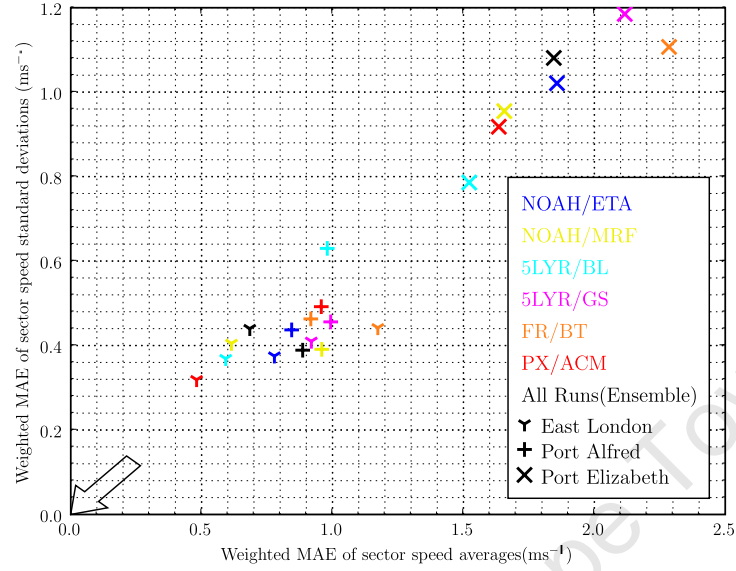
Turning to the *MMAE* and looking at the ensemble results in Figure 4.7(a) to the left (the black points) an interesting pattern emerges: East London's seasonal variability is matched far better than that of Port Alfred and Port Elizabeth who lie fairly close together. However, this peculiarity can be considered statistically insignificant bearing in mind that the full range of the y-axis is only 0.45ms^{-1} and the differences between EL's *MMAE* and the next closest point are between 0.05 and 0.12ms^{-1} .

On first sight it might also appear striking that PX/PBL exhibits the worst performance in terms of *MMAE*, something which is confirmed by the spatial average results given in Figure 4.7(b) to the left. Again though the distance to the three winners NOAH/Eta, FR/BT and 5LYR/BL - who all lie very close together - is a mere 0.12ms^{-1} at most.

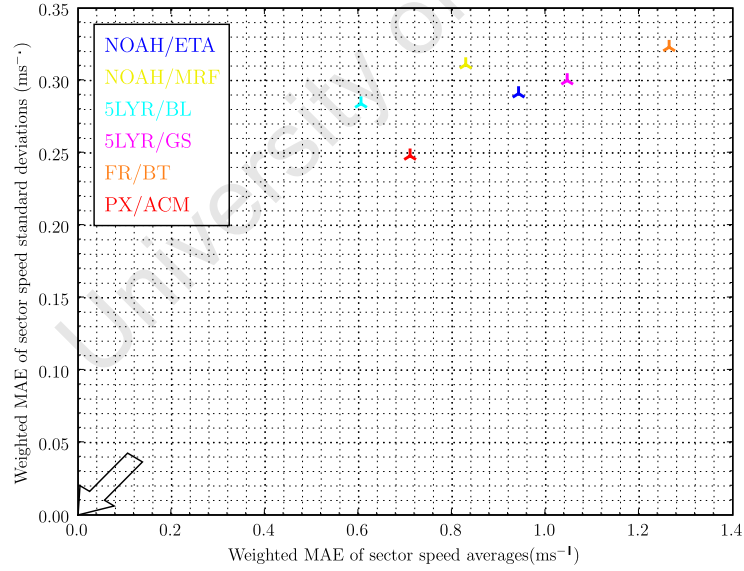
These considerations combined lead us to the conclusion that the differences in seasonal prediction skill of the LSM/PBL schemes in MM5 is insubstantial; any of them would be appropriate for application to the generation of a wind atlas. This could well be because seasonality is dominated by other factors such as boundary forcing, radiative transfers or cloud cover treatments.

4.4.3 Directional Performance

On the surface it would appear that wind direction is entirely determined by the synoptic scale dynamics inherent in the model's boundary conditions as well as topographic forcing to the extent at which MM5 is able to resolve them. Both of these main drivers are kept constant during these sensitivity runs. Hence in theory varying LSM and PBL parametrisations should have very little if any effect on the sectoral distribution of the flow. However, as we shall see below, they do in fact exhibit significantly different behaviours in *both* components of the wind vector, not only V .



(a) Individual stations



(b) Spatial average over all stations

Figure 4.8: Goal plot of weighted MAE of sector wind speed averages ($WMAESSA$) and standard deviations ($WMAESSD$) for all LSM/PBL runs

Therefore in we now turn to the application of the directional skill metrics defined by the *WMAE* family of indicators in subsection 2.2.3 on page 38. It is reiterated that none of these are based on any filtering, detrending or bias removal of the original time series since we want to complement our existing knowledge with some sector specific evaluation.

Analogous to the analysis of wind speed bias b_V vs standard deviation bias c_V we first deal with the weighted *MAE* of speed averages *WMAESSA* and standard deviations *WMAESSD*, both of which are essentially an aggregate of the twelve individual sector differences, weighted by their sector's significance. In Figure 4.8 to the left we illustrate the results for our six LSM/PBL sensitivity runs.

Focusing on the observation specific details of Figure 4.8(a) we see as expected the large Port Elizabeth bias as noted earlier in the discussion of subsection 4.4.1. Comparing Figure 4.8(a) with Figure 4.4(a) on page 77 we discover that *WMAESSA* and b_V have almost identical values for PE (except for the opposite sign introduced by the *MAE*), at least for the FR/BT, 5LYR/GS and NOAH/Eta schemes.

Merely PX/ACM, NOAH/MRF and 5LYR/BL have a *WMAESSA* that is 0.2-0.4ms⁻¹ larger than the absolute values of their b_V , but the distribution of the parametrisations relative to each other is unchanged with 5LYR/BL coming closest to reality and PX/ACM and NOAH/MRF a close second. The same applies for PE's various *WMAESSD* compared to its c_V figures. Hence we have nothing to add to our previous discussion of MM5's different LSM/PBL skills for PE.

For Port Alfred, however, the picture looks very different to our previous observations. While b_V ranged from 0.25 to about 1.1ms⁻¹ all points are now closely clustered around *WMAESSA* $\approx$ 0.9ms⁻¹ with only Blackadar's *WMAESSD* being roughly 0.2 ms⁻¹ larger than all the others. This means that while there were large differences in the b_V skill before the six schemes all perform on par in terms of direction, at least for PA. Indeed this is exactly the kind of complementary information that we alluded to back in subsection 2.2.3 on page 38.

Lastly for East London the picture looks very similar to before with the same relative rankings of the various parametrisations. The only difference is that now PX/ACM is leading instead of 5LYR/BL, albeit by a small margin.

Moving on to the spatially averaged results of 4.8(b) we see that overall PX/ACM and 5LYR/BL are closest to the origin, FR/BT performing the worst

and all others being in between. It must be noted that the *WMAESSD* differences between them are not statistically significant (less than 0.1ms^{-1}) but the *WMAESSA* definitely are (up to 0.1ms^{-1}).

The last pair of metrics that we consider in this section of PBL/LSM evaluations are the sector frequency *WMAESF* and power density *WMAEPD*. The results are presented on the goal plots in Figure 4.9 to the right.

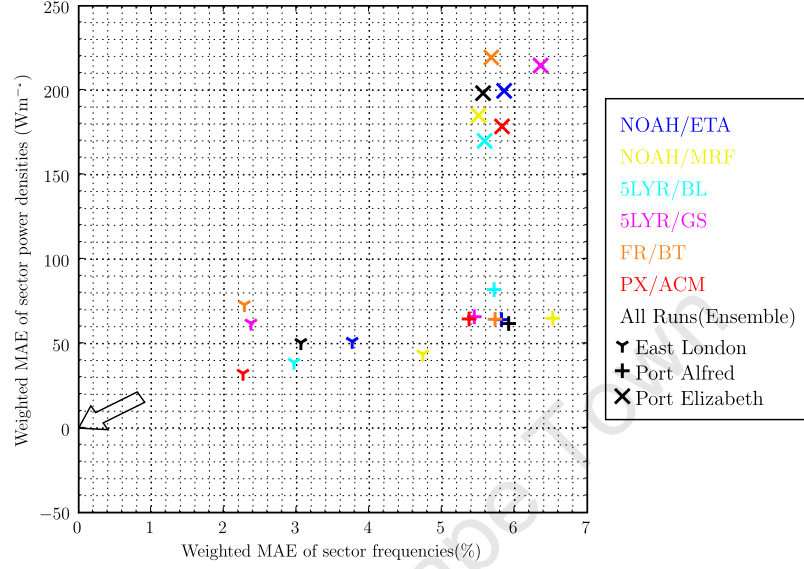
Unfortunately the power density skill as given on the y-axis does not add any more substance to our discussion so far. For Port Elizabeth the model's strong negative wind speed bias manifests itself again with by far the highest *WMAEPD* of around 200Wm^{-2} , a very sizable number in wind industry circles (Figure 4.9(a)). Within that systematic error constraint 5LYR/BL again comes out on top with PX/ACM a close second and FR/BT and 5LYR/GS last. At the East London airport the differences are much lower at about 50Wm^{-2} , a much more acceptable performance, with PX/ACM lying in front this time but the same relative distribution otherwise.

At Port Alfred we witness a similar normalisation as pointed out above: All schemes have a *WMAEPD* of 65Wm^{-2} while 5LYR/BL is the only outlier with 81Wm^{-2} . This correlates with the earlier *WMAESSD* standard deviation mismatch, which makes sense because raising the wind speed to the third power will severely penalise a larger difference in σ_V . The spatial averages of 4.9(b) do not add any more information in terms of power density skill either.

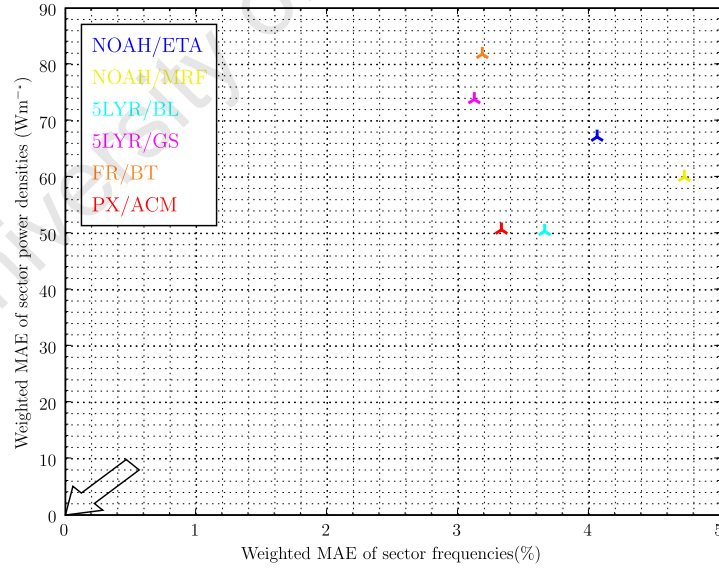
Interesting to mention is the appearance of model's generally good performance in terms of frequencies. The *WMAESF* is below 7% for all stations and runs and less than 5% for the spatial mean. However, the reader must be reminded of the fact that even though we are talking percentages the *WMAESF* is an absolute error metric. It is about the occurrence out of 100% of each of the 12 sectors so the absolute values should be around 8% to begin with assuming a uniform directional distribution. Hence the *WMAESF* is comparable with the magnitude of the actual sector frequencies and thus MM5's skill of θ is not as good as it should be.

To give the reader an idea of what this looks like in practice, Figure 4.10 on page 89 depicts the frequency windroses for the standard PX/ACM run. Compatible with low values of *WMAESF* of around 3% one can clearly see that the match for East London is by far the best, while both PE and PA are not that well simulated with a *WMAESF* closer to 5.5%.

Within the spectrum of the various parametrisations PX/ACM once again does very well, though surprisingly the two up to now poor candidates FR/BT



(a) Individual stations



(b) Spatial average over all stations

Figure 4.9: Goal plot of weighted MAE of sector power densities ($WMAEPD$) and frequencies ($WMAESF$) for all LSM/PBL runs

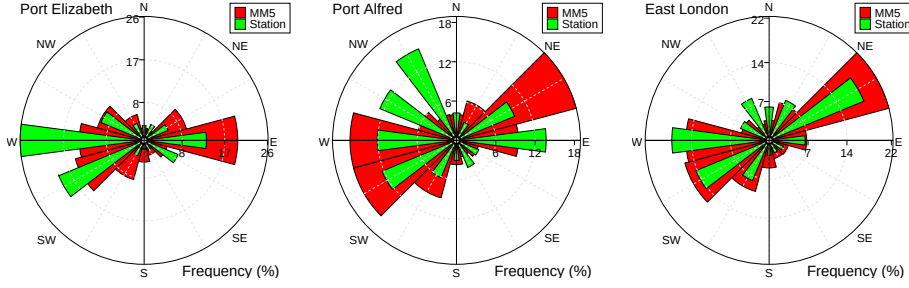


Figure 4.10: Frequency windroses of model vs station for the standard PX/ACM run at all three locations

and 5LYR/GS are slightly ahead in the aggregate results with NOAH/MRF being the clear loser this time (see Figure 4.9(b) on the preceding page). With a total spread of 1.6% and in the light of the above explanation of the nature of *WMAESF* this is deemed statistically significant.

4.4.4 Summary

From the various error metrics and considerations presented in the previous subsections we may conclude that out of all six land surface model (LSM) and planetary boundary layer (PBL) parametrisations that MM5 offers, the Pleim Xiu/ACM one provides the best performance in terms of simulating surface winds. This is especially true for the prediction of the daily wind cycle which has been demonstrated to represent a very large and important component in the variability of the flow.

The differences between the schemes may be attributed to a number of possible reasons:

- model implementation deficiencies as mentioned in the literature for MRF as well as the TKE based schemes BT, GS and ETA [Tonnesen et al. 2004; Tremback 2003]
- inadequate treatment of night time conditions in the case of Blackadar (Richardson number based parametrisation too simplistic)
- over- and underestimated values of surface momentum fluxes and differing treatment of vertical momentum transfer

Even in areas where PX/ACM is not superior (e.g. seasonal skill) it has been shown that the schemes ahead of it are either “better” for all the wrong reasons or exhibit only a small or insubstantial improvement in the results. Its physically more sound treatment of surface roughness length further contributes to its suitability. Therefore the author recommends and uses PX/ACM as the preferred LSM/PBL combination for the rest of this study.

4.5 Evaluation of Cumulus Schemes and FDDA

Having settled on PX/ACM as the LSM/PBL parametrisation of choice, we now turn to investigating other MM5 options which may have a positive impact on our modelling results, namely the use of different cumulus schemes and four dimensional data assimilation (FDDA) applications. In total nine runs were performed and an overview of their configurations is given in Table 4.6 on the next page.

FDDA schemes used are grid and observational nudging both at one, two and four times the default MM5 strengths (variables G_V , G_T , G_Q and G_{IV}) to investigate how the model reacts.

Mesh based FDDA is exclusively applied to the coarse resolution domain for wind, temperature and mixing ratio ($u/v/T/q$, I4D=1,0) and only above the planetary boundary layer (using MM5’s INONBL=1,1 switch) in order to let MM5 develop its own dynamics in both the PBL and the entire 18km nest. The values of $G = 2.5 \cdot 10^{-4}$, $5.0 \cdot 10^{-4}$ and $1.0 \cdot 10^{-3} \text{s}^{-1}$ correspond to characteristic adjustment periods of 67, 33 and 17 minutes respectively which demonstrates that at $G = 1.0 \cdot 10^{-3} \text{s}^{-1}$ the relaxation clearly occurs the quickest (“strongest” nudging). However, it must be borne in mind that the temporal resolution of the input NCEP/NCAR reanalysis data is only 6 hours (see subsection 4.3.2 on page 71).

Observational nudging similarly uses three nudging factors, each increasing by a factor of two, and activated only for the nest (I4DI=0,1) and winds (u and v). A relatively small horizontal and vertical radius of influence of 20km and 0.001 sigma units respectively is applied in order to recognise the microscale nature of the physical measurements. In general the nudging factors are higher here than for grid FDDA because our input station data are not only hourly but also due to the fact that surface winds operate at much higher frequencies than the variables affected above the PBL.

Run Description	Short run name	ICUPA	Grid FDDA G (s^{-1})	Obs Nudging G (s^{-1})
Pleim Xiu/ACM	PX/ACM	3,3	none	none
Single strength grid FDDA	1xFDDA	3,3	$2.5 \cdot 10^{-4}$ ($u/v/T/q$)	none
Double strength grid FDDA	2xFDDA	3,3	$5.0 \cdot 10^{-4}$ ($u/v/T/q$)	none
Quad strength grid FDDA	4xFDDA	3,3	$1.0 \cdot 10^{-3}$ ($u/v/T/q$)	none
Anthes Kuo/Betts Miller	AK/BM	2,7	none	none
Anthes Kuo/Kain Fritsch II	AK/KF2	2,8	none	none
Single strength observational nudging	1xOBS	3,3	none	$4.0 \cdot 10^{-4}$ (u/v)
Double strength observational nudging	2xOBS	3,3	none	$8.0 \cdot 10^{-4}$ (u/v)
Quad strength observational nudging	4xOBS	3,3	none	$1.6 \cdot 10^{-3}$ (u/v)

Table 4.6: Configuration of the eight cumulus and FDDA sensitivity runs. Cumulus schemes are specified as “mother domain/nest”, grid FDDA is only carried out on the coarse mesh (wind, temperature and mixing ratio) and observational nudging is only performed on the nest (wind only). The respective FDDA strengths are given in the last two columns - G_V , G_T and G_Q in the second last and G_{IV} (for observed winds) in the last column. The first row (Pleim Xiu/ACM) is the reference simulation and identical to the one in section 4.4 on page 76

Accordingly, the e-folding times of the observational nudging factors for wind G_{IV} chosen in Table 4.6 are about 42, 21 and 10 minutes respectively (from smallest to largest G_{IV}). Given a half-period of 30 minutes during which the observation is “active” (see setting and meaning of TWINDO in appendix B.1.4 on page 174) this suggests that the higher values should definitely yield improvements.

Both FDDA variants have their input data prepared as described in subsection 4.3.2. For the detailed settings of all parameters the reader is referred to Appendix B.1.4 on page 174.

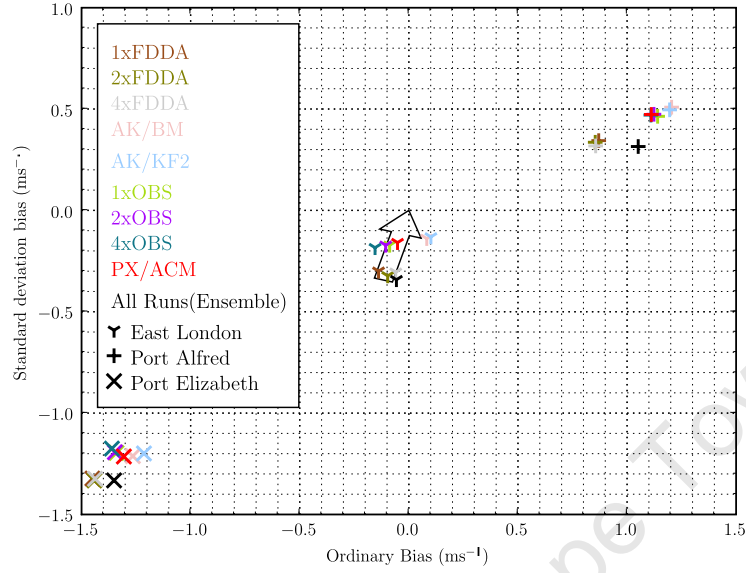
In terms of the two varied cumulus parametrisations we choose the Anthes Kuo(AK) one for the coarse 54km domain as it is officially recommended for spatial discretisations of 50km or lower [Dudhia et al. 2004]. For the nest we evaluate Betts Miller (BM) and Kain Fritsch II (KF2) as they are both often recommended in the literature (see subsection 4.2.3 on page 64). While convection schemes can be strongly influenced by the LSM/PBL configuration in terms of, for example, moisture advection, it was chosen to evaluate these in isolation for simplicity reasons.

For all these eight perturbations the reference is the previous PX/ACM run whose performance is desired to be enhanced further. In the subsections that follow we employ the same validation methods as done in section 4.4 on page 76 for the LSM/PBL combinations. It is assumed that by now the reader is familiar with the various error metrics that have been designed for this study, hence the discussion is kept a bit shorter. However, especially seeing that Port Alfred did *not* have any observational nudging applied in order to see if MM5 introduces any unwanted artefacts between areas that *did* have Newtonian station relaxation enabled, it is still necessary to look at both location specific and area averaged results separately.

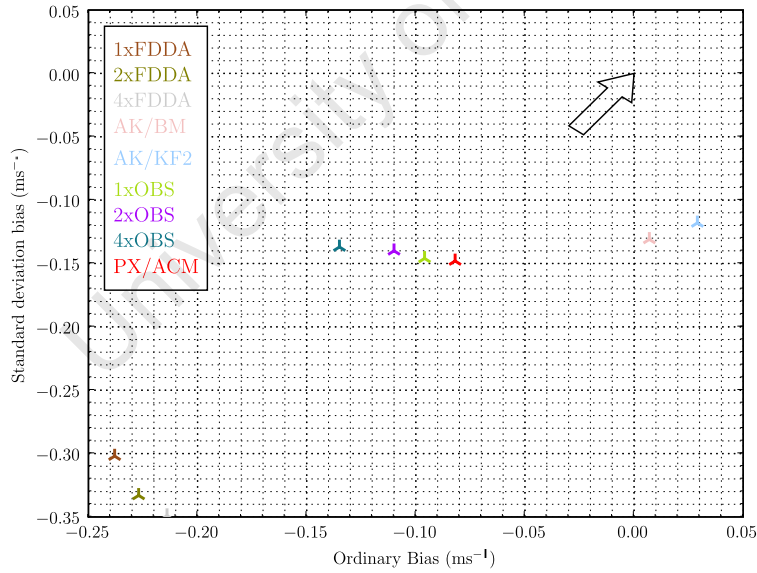
4.5.1 Average Wind Speed and Standard Deviation

Again we first look at the model’s skill at predicting the normal and standard deviation wind speed biases b_V and c_V in Figure 4.11 on the next page. The station specific results in Figure 4.11(a) reveal that the configuration independent systematic error is obviously the same for the three airports as before but also that there is actually fairly little variability over the eight runs.

Whereas with the LSM/PBL runs b_V could easily differ by 1.0ms^{-1} the range is 0.4ms^{-1} at most for Port Alfred and even lower for the other two. For c_V this



(a) Individual stations



(b) Spatial average

Figure 4.11: Goal plot of standard deviation bias c_V vs ordinary bias b_V of wind speed for all eight FDDA/cumulus runs. The red symbols mark the PX/ACM reference whose performance is to be improved upon

is even worse as the spread never exceeds 0.2ms^{-1} . In that light the distinction becomes blurred and any conclusions drawn from this picture are likely to be marginally significant at best.

Nevertheless the effects of FDDA and different cumulus parametrisations are consistent across all three anemometers. Grid FDDA above the PBL unfortunately decreases the skill of the model by causing both lower $\bar{V}$ and σ , effectively compressing the flow's statistical distribution. While this can be explained in terms of the smaller variability owing to the much lower resolution ($\sim 250\text{km}$) of the NCEP Reanalysis, this is a rather disappointing result especially seeing that we attempted to work around this limitation by excluding both the PBL and the nest. Furthermore increasing the applied FDDA strength has little impact on the outcome.

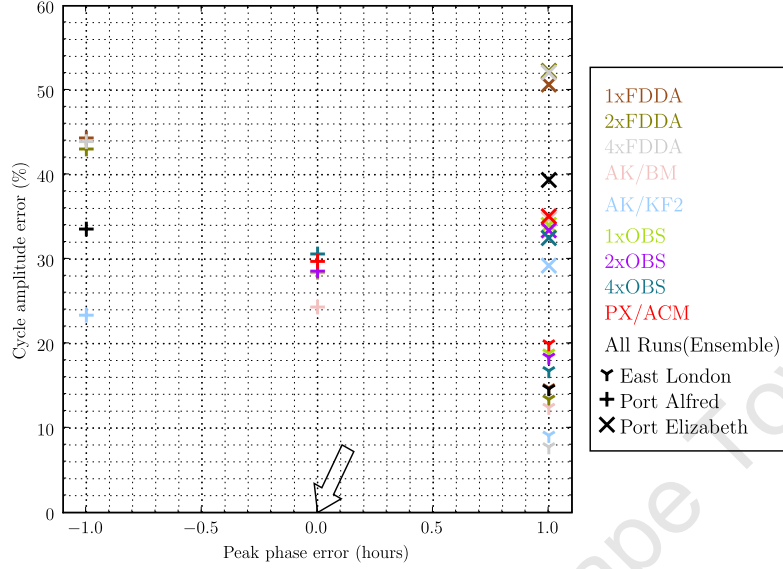
The performance of observational nudging is not much better. Despite the expectation of moving the points closer to the origin their c_V is unchanged with respect to the reference while their b_V actually decreases for both Port Elizabeth and East London. This effect increases with higher strength of the applied nudging. Overall model performance is also thereby slightly decreased. At least there are no negative consequences that can be seen at Port Alfred which has been excluded from the nudging.

Both AK/BM and AK/KF2 are found to increase b_V while keeping c_V constant. Their impact is comparable and while this could be due to a whole number of factors in which they differ with respect to the Grell cumulus scheme, we defer this analysis to the next subsection.

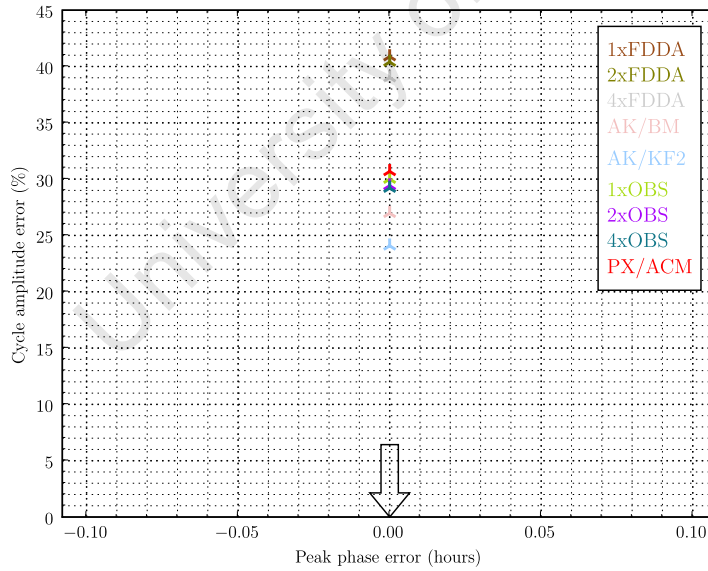
The spatially averaged results illustrated in Figure 4.11(b) to the left confirm the above discussion. Both grid and station FDDA lead us further away from reality though the influence of the former is much larger. Merely the two cumulus perturbations bring MM5 closer to the origin, though none of them seem better at this stage. The scales here are much smaller so as mentioned before the statistical significance of these results is low.

4.5.2 Daily and Seasonal Cycles

Next we consider whether FDDA, BM or KF2 are able to improve the model's ability to reproduce the land and sea breezes. To gauge the match of amplitude and phase of these phenomena we plot the CAE vs PPE in Figure 4.12 on the next page. In all cases the PPE is 1h or 0h so the timing of the peak is very good regardless of configuration, so nothing further can be said about the t_{max}



(a) Individual stations



(b) Spatial average over all stations

Figure 4.12: Goal plot of daily cycle amplitude error CAE vs daily peak phase error PPE of wind speed for all FDDA/cumulus runs. The red symbols mark the PX/ACM reference whose performance is to be improved upon

performance.

The results are much more striking in the *CAE* deltas between best and worst as they are in the region of 20%, at least for Port Alfred and Port Elizabeth (Figure 4.12(a)). In both cases grid FDDA loses us about 15% of *CAE* and effectively flattens the curve - another setback for this technique. Increasing the strength of the nudging has little consequences. This further explains our previous assertion that mesh FDDA introduces a lower standard deviation as the daily cycle gets flattened, despite measures to limit FDDA's hindrance in the model's freedom to develop PBL dynamics.

Surprisingly perhaps at the East London airport this pattern is *reversed*: the stronger the grid nudging, the more faithfully MM5 simulates the 24h periodicity such that at quad strength the model comes within 8% of reality, a 12% improvement over the reference. For a more in-depth consideration we must also look at the actual V_{24} of station vs model for all locations given in Figure 4.13. While the complexity of the plot might conceal this slightly, 4xFDDA manages to increase the peak, decrease the night time speeds considerably and thus achieve highest accuracy.

In order to explain this anomalous performance at East London we have to consider even more detail and break V_{24} up into individual components u_{24} and v_{24} . Note, however, that $V_{24} \neq \sqrt{u_{24}^2 + v_{24}^2}$ because of the averaging of the individual values at each hour of the day. The results visualised in Figure 4.14 on page 98 again show the "better" V_{24} plus a clear cycle in both u and v . While the two runs are virtually identical in the average diurnal evolution, offset and lag in u , grid nudging introduces a much higher mid-day southerly wind component v than exists in reality. That of course increases the wind speeds and ultimately decreases the *CAE* - another case of being more "correct" for all the wrong reasons which means that in fact grid FDDA also deteriorates our 24h oscillation at East London.

The observational nudging runs again have much less of an impact on the *CAE* with a maximum improvement of 4% for 4xOBS and even less for 2xOBS and 1xOBS. For Port Alfred no noteworthy changes can be detected so while the inter-observation space shows no improvement at least the model is consistent by not introducing regressions either. Thus one might argue that it does not hurt to turn on this option for a final wind atlas simulation.

Both AK/BM and AK/KF2 show a marked enhancement of the daily cycle, clearly visible in both Figures 4.12(a) and 4.13, though AK/KF2 is the clear winner of the two. It provides a *CAE* that's 6-12% better than the base run.

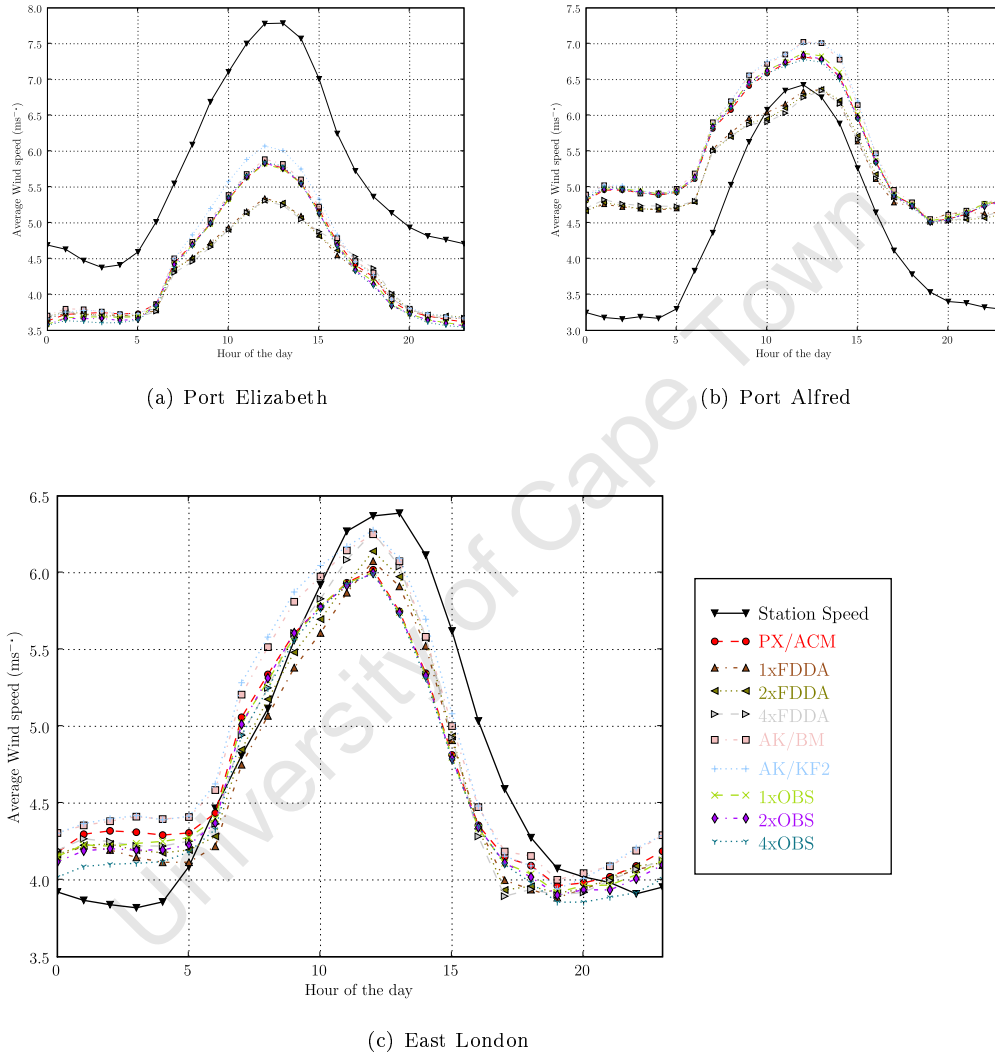
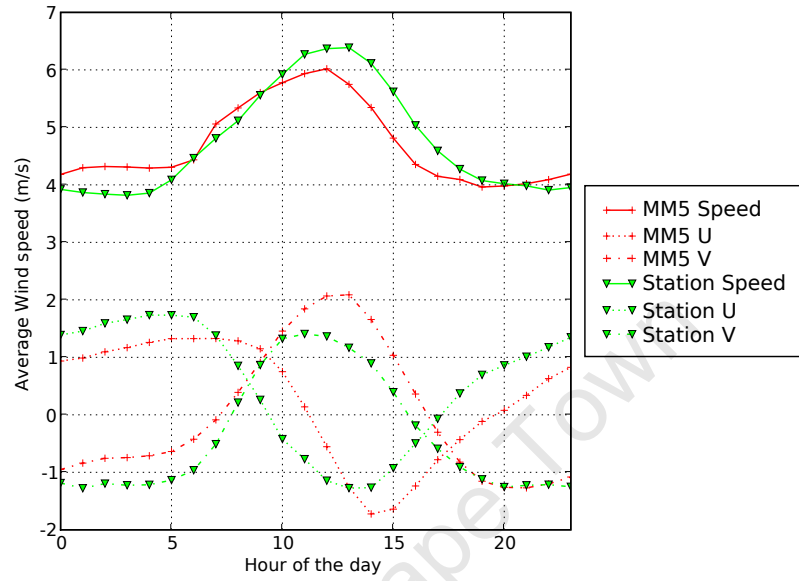
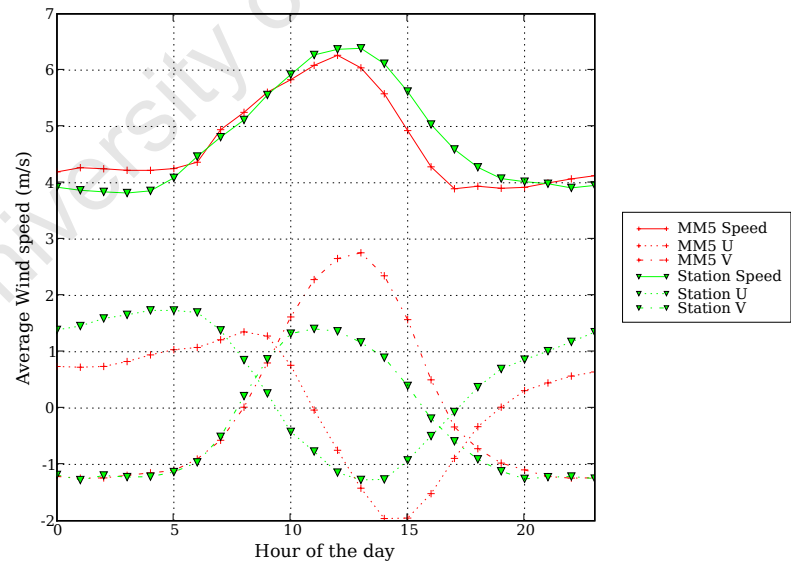


Figure 4.13: Average diurnal cycles for all eight cumulus/FDDA runs. Each parametrisation is represented by the same colour as used before and a different symbol alongside the observed (black) pattern in terms of speed V .



(a) Reference run (PX/ACM)



(b) 4xFDDA grid nudging

Figure 4.14: Reference (PX/ACM) and 4xFDDA V_{24} , u_{24} and v_{24} daily cycles at East London Airport

This may be attributed to the fact that its convection algorithms are much more physically sound at 18km resolution than the standard Grell treatment, an important characteristic which clearly governs the evolution of a land/sea breeze that seems to be well articulated at the coastal airports.

When we average the three time series to obtain a better spatial result we can confirm the statements made above (Figure 4.12(b)). Grid nudging on the whole does more damage than good with an overall loss of around 10% in the amplitude. Per station relaxation does help by a few percentage points while the biggest improvement of about 7% may be achieved by the use of the AK/KF2 cumulus parametrisations.

Turning now to the remaining pair of cycle metrics the *DFPD* and *MMAE* are visualised in Figure 4.15 to the right. Once again the familiar clustering of stations is unfortunately gone making analysis a bit more difficult, but at least this implies that we are dealing with fairly small station specific biases.

As before the *DFPD* does not add any value and merely confirms the diurnal cycle achievements noted above. On the y-axis the *MMAE* is spread out by a mere 0.2ms^{-1} but nevertheless shows an interesting pattern (Figure 4.15(a)). All the xFDDA runs show little or no seasonal amelioration. This is most likely due to the small number of grid cells in the nest which do not permit model drift which the mesh based FDDA is supposed to correct.

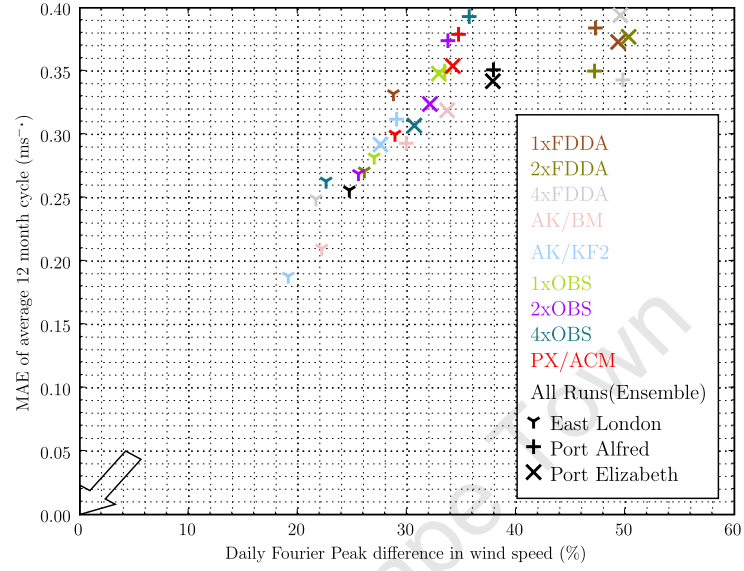
Observational nudging does bring both EL and PE closer to the origin but only by a very small amount. MM5's *MMAE* skill at Port Alfred in turn remains fairly constant. This is also consistent with what we discovered before.

Across all airports AK/KF2 repeatedly exhibits lower values of *MMAE* by up to 0.1ms^{-1} . On the spatial average in Figure 4.15(b) to the right this comes out clearly too. All in all we can safely state that analogously to the LSM/PBL investigations the seasonal results are statistically insignificant as the magnitudes of the variations are too small.

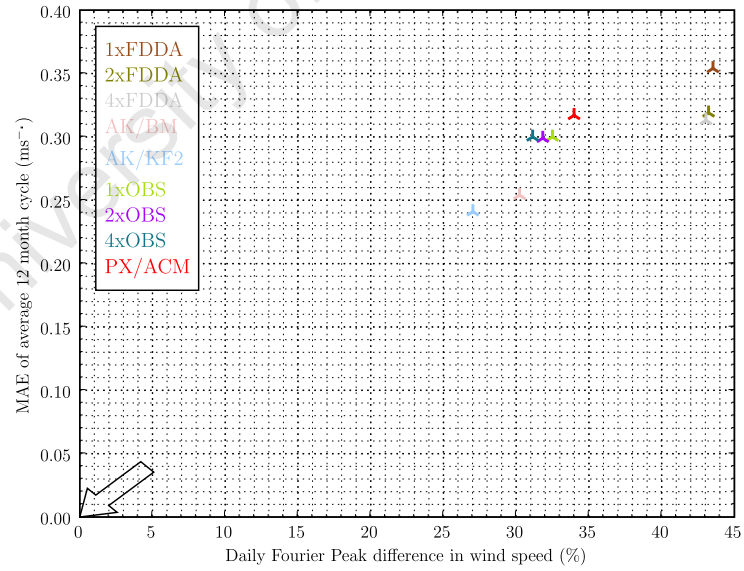
4.5.3 Directional Performance

In this remaining set of evaluation metrics we consider the directional skill of the FDDA/cumulus runs. First the goal plot of *WMAESSD* vs *WMAESSA* which is analogous to c_V vs b_V is presented in Figure 4.16 on page 101.

Unfortunately we detect very little spread across the eight configurations, even less than encountered before in Figure 4.11 on page 93. The *WMAESSA* varies by no more than 0.2ms^{-1} (PE) while the *WMAESSD* moves within the

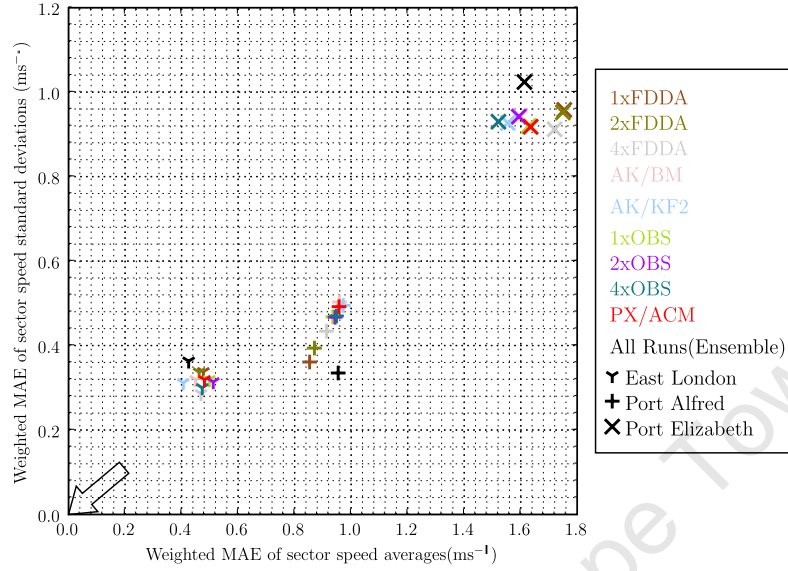


(a) Individual stations

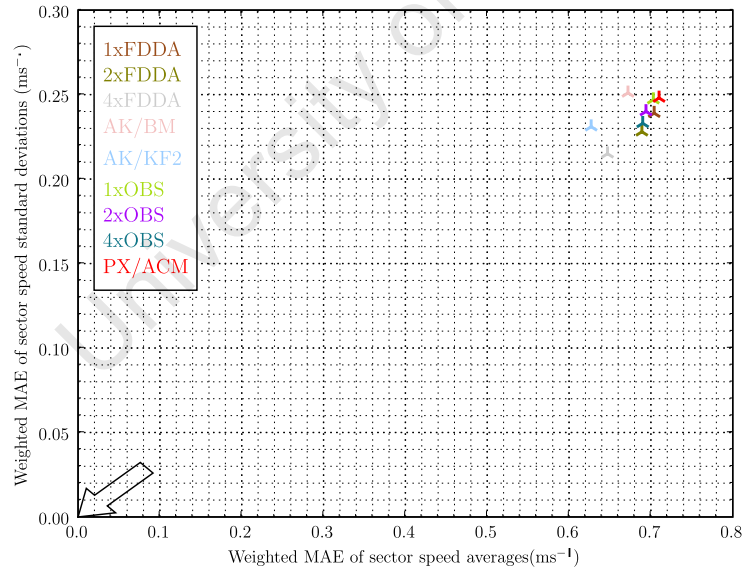


(b) Spatial average over all stations

Figure 4.15: Goal plot of 12 monthly(seasonal) mean absolute error $MMAE$ vs daily Fourier peak difference $DFPD$ of wind speed for all FDDA/cumulus runs



(a) Individual stations



(b) Spatial average over all stations

Figure 4.16: Goal plot of weighted MAE of sector wind speed averages ($WMAESSA$) and standard deviations ($WMAESSD$) for all eight FDDA/cumulus runs

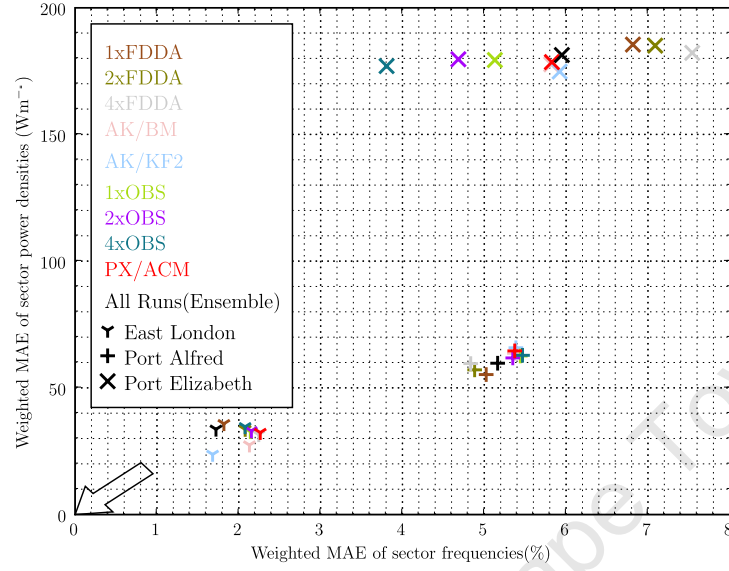
same range (PA) or less (Figure 4.16(a)). For East London both metrics cluster very tightly such that no model option dominates over another. The effect of grid FDDA is equally inconclusive for Port Alfred and Port Elizabeth as in the former case it gives a slight improvement and in the latter a small deterioration, at least in terms of the *WMAESSA*.

The influence of observational nudging or different cumulus parametrisations seems to be small too. The spatially averaged points in Figure 4.16(b) add little information as even the closest pair to the origin, 4xFDDA (followed by AK/KF2) yields a *WMAESSA* of only 0.08ms^{-1} lower than the reference. Hence we may with confidence state that in terms of *WMAESSA* and *WMAESSD* none of the eight FDDA/cumulus runs is any better or worse than PX/ACM in its default setup.

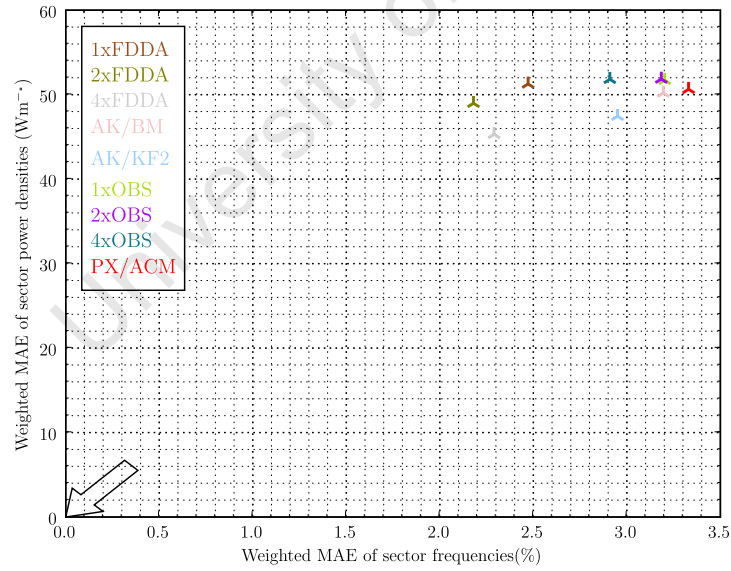
More insightful are the *WMAE* of sector frequencies (*WMAESF*) and power densities (*WMAEPD*) given in Figure 4.17 on the next page. While the *WMAEPD* hardly changes at all, the *WMAESF* clearly does, at least for Port Elizabeth (4.17(a)). While the cumulus variations there have no measurable effect, xOBS clearly pulls the model output closer to reality, the more so the higher the nudging strength to a maximum of 2% off the *WMAESF*, as we would expect of making the simulation a function of the measurements. xFDDA, however, works in the exact opposite way and decreases MM5's skill at PE by up to 1.6% on the *WMAESF*, very much in line with previous analysis.

At the other two airports further northeast along the Eastern Cape coastline the variability of the *WMAESF* is much smaller, though FDDA increases the accuracy of the predictions slightly. xOBS again has a negligible impact on Port Alfred which was excluded from nudging, another aspect underlining the reliability of that method. AK/KF2 only shines at East London by bringing MM5 closer to the origin in both dimensions.

Surprisingly enough *all* the permutations bring an improvement in skill when we consider the spatial mean illustrated in Figure 4.17(b), especially mesh nudging above the PBL which brings the largest difference (1.1% of *WMAESF*). As before AK/KF2 is better than AK/BM, but both beat the reference Grell/Grell combination. Observational nudging has a minor overall effect but certainly is on the positive side.



(a) Individual stations



(b) Spatial average over all stations

Figure 4.17: Goal plot of weighted MAE of sector power densities ($WMAEPD$) and frequencies ($WMAESF$) for all eight FDDA/cumulus runs

4.5.4 Summary

Based on the reference PX/ACM configuration we have evaluated different strength gridded and observational nudging options as well as two more complex cumulus convection schemes. Throughout the analysis it became clear that the margins that separate them all are smaller than those that we came across in our LSM/PBL discussion earlier in section 4.4, but there are certain trends that may be extracted with confidence.

Gridded FDDA above the PBL is found to mostly deteriorate MM5's accuracy. Only select directional validation metrics are minimised but the biggest problem is the flattening and disturbance of the daily cycle. Peak amplitudes and related overall variability are decreased substantially leading to suboptimal performance. This may in all cases be attributed to a much weaker spatial and temporal variability of the NCEP/NCAR Reanalysis data owing to its low resolution. It could also mean that we are seeing evidence of wind specific Reanalysis errors over the Eastern Cape region which have not been discovered before.

However, increasing the FDDA factors from their defaults yielded ameliorations across the board with 4xFDDA being the best. Given the importance of mesh FDDA to prevent model drift over long integration periods and much larger domains than those utilised here it is likely that it could be used in the final wind atlas after all, more about this later.

Observational nudging helps to a certain extent in bringing the model closer to reality, the more so the higher the strength, though overall less so than one might expect. At the very least it does not seem to introduce any errors into areas without surface adjustments as the example of Port Alfred has demonstrated. This is in line with what one would expect of actively decreasing model bias at specific grid points only. Therefore this method can only be recommended for the South Africa wide simulation with four times the default coefficients.

Finally the use of the Anthes-Kuo/Kain Fritsch II parametrisation is highly advised as it enhances the model's skill in terms of all the ten verification statistics considered in this section. Given the importance of the daily cycle in terms of its large contribution to total variance of the wind a particularly good skill here also positively influences other error metrics such as standard deviation bias c_v . The diurnal oscillation in turn is highly dependent on localised land/sea convective processes which the more advanced schemes like KFII are designed to handle much better than the more simplified Grell treatment that the reference

simulation used. That is the reason AK/KFII provides us with a better model performance.

4.6 Conclusion

We have shown that the Pleim Xiu(PX) land surface model in combination with its planetary boundary layer(PBL) scheme, the Atmospheric Convective Model(ACM) is the most skillful LSM/PBL parametrisation in MM5 when it comes to near-surface flows important for wind power. Further it has become clear that Anthes-Kuo for the mother nest in combination with the Kain Fritsch II cumulus scheme provides us with even further enhancements in the model's wind performance. Finally observational nudging yields additional improvements while grid FDDA deteriorates the all important daily cycle but gets closer to reality for directional statistics.

What we have not done yet is to test if the AK/KF2 and various nudging techniques also work well in conjunction and not just in isolation. As it turns out the validation of combination runs does not throw up any surprises which is why this discussion is left for Appendix A on page 159. The results are that their benefits and weaknesses are effectively linearly additive.

Leaves only the question of whether or not to employ relaxation towards NCEP/NCAR data. The above analysis clearly points out that for the Eastern Cape on such a small domain this is not recommended due to the suboptimal daily cycle, however, when moving on to a much larger South Africa wide model integration model drift may become a significant problem. Without knowing in advance which issue will dominate the best way to proceed is to carry out the large simulation twice, once with 4xFDDA and once without and compare the results, based on PX/ACM and AK/KF2 with 4xOBS as the base configuration.

Finally it must be reiterated that we optimised MM5 for the Eastern Cape region and could only consider observations close to the sea for detailed validation. This means that for inland prediction skill or even performance over other coastal regions of the country the above chosen combination might not be ideal, but given the economic importance of the Eastern Cape as well as the predecessor wind atlas findings of highest average wind speeds near the shore this is deemed an acceptable compromise.

University of Cape Town

Chapter 5

Wind Atlas of South Africa

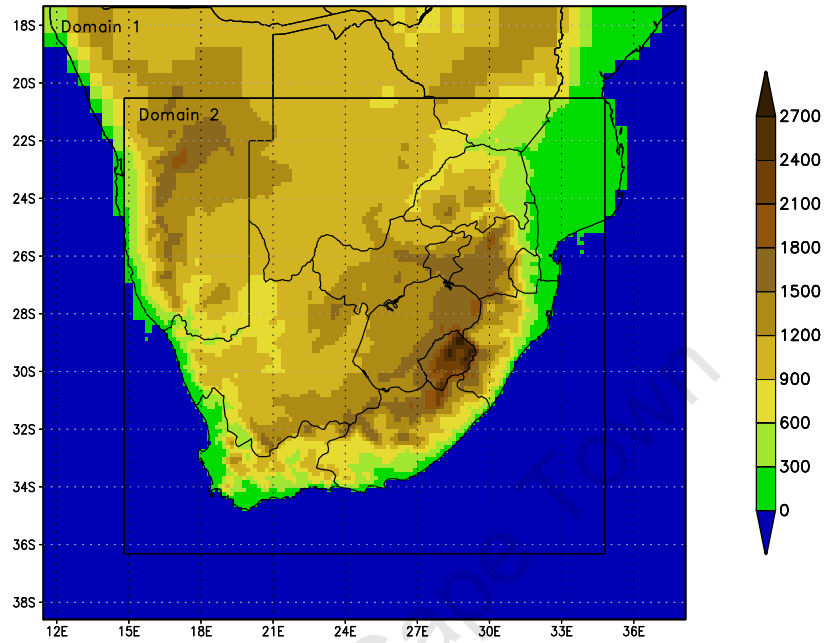
As discussed in section 1.6 on page 29 this chapter describes the final nationwide wind simulations at 18x18km in terms of:

- how they were derived (section 5.1)
- an analysis of how accurate the results are (section 5.2 on page 110)
- discussion of annual and seasonal average wind speed maps (section 5.3 on page 131)
- total wind resource quantification (section 5.4 on page 135)

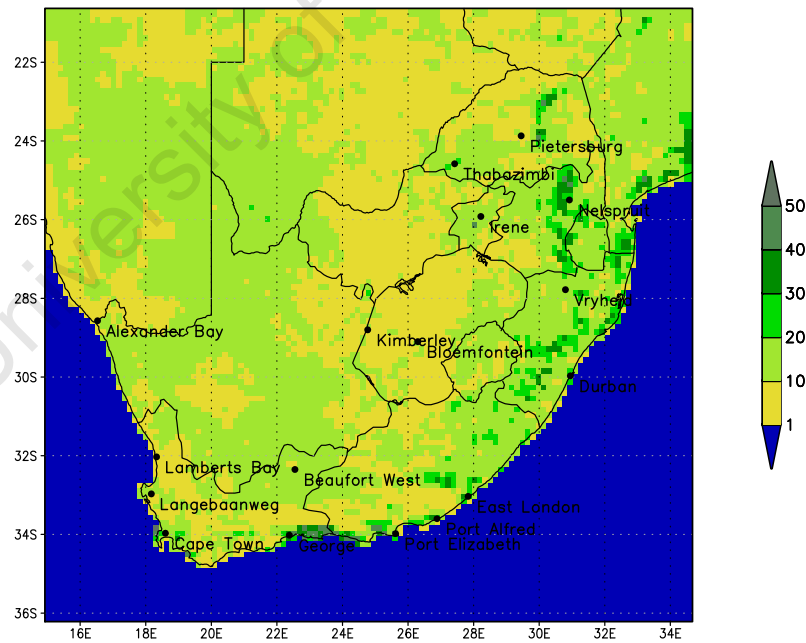
5.1 Design and Execution of Simulations

5.1.1 Domain Configuration

The grid arrangement to cover the whole of South Africa has been designed analogously to the Eastern Cape runs of the previous chapter (see section 4.3.1 on page 66). The 54km parent domain reaches as far north as Namibia's Caprivi strip and on the west, south and east includes a large portion of the Atlantic and Indian Oceans (see Figure 5.1 to the right). Even the 18km nest has a padding of at least 10 mesh cells on either side. This ensures a large enough border zone where the model is relaxed towards the boundary conditions and output is not really useful. As before only the parent is forced on its boundaries with NCEP/NCAR Reanalysis while the nest's boundary conditions are fully determined by its parent.



(a) Height above sea level in m for both domains



(b) Roughness length in cm for domain 2

Figure 5.1: Domain configuration for the South African Wind Atlas. Both topography and roughness length are given, the individual colour pixels represent grid cells

Terrain heights and land use characteristics are again taken from the USGS 30 arc second datasets. As one can see in Figure 5.1(a) only the largest topographical features such as the Highveld and the Drakensberg are represented well while smaller mountain ranges such as the Swartberg or Outeniqua are smoothed out even at the high resolution. Figure 5.1(b) reveals that the Pleim/Xiu roughness length approximation results in low $z_0 \leq 20\text{cm}$ with only the South Coast, parts of Kwazulu-Natal and Mpumalanga being frequented by larger values.

For the exact TERRAIN configuration the reader is referred to Appendix B.2.1 on page 184.

5.1.2 Model Preparation and Configuration

The input data is prepared identically to that for the Eastern Cape in section 4.3.2 on page 71 with NCEP/NCAR Reanalysis data as both initial and boundary conditions for the entire integration over the same representative time period with five days spin up time (15 March 1996 until 20 March 1997) found in section 4.3.4 on page 73.

Like REGRID the settings of INTERPF for the wind atlas simulations are the same as those for our sensitivity runs. Hence the same 40 vertical sigma levels are used which for the high resolution yields the following heights above ground interesting for wind power applications:

1. 10m (9.7-10.25m) - useful for comparison to weather stations
2. 30m (29-31m) - for validation against 30m measurements
3. 56m (54.5-57.5m) - close to a standard 60m hub height
4. 92m (89-94.5m) - between 80m and 100m hub heights
5. 137m (132-141m) - upper bound for hub height of modern wind turbines

Any other desired heights may of course be easily be interpolated from these five using a cubic spline or similar technique.

MM5 itself is configured in the way recommended by section 4.6 on page 105 of the previous chapter, once with and once without grid FDDA being turned on, to be referred to as “4xFDDA” and “Standard” respectively. The only patch applied to the code base was the observational nudging one described in Appendix C.2.

In terms of observational nudging the same findings are applied (4xOBS) though this time all available stations are utilised. This makes a final model validation against the *same* observations somewhat questionable and less meaningful. The justification for this is twofold: firstly, in the interest of maximising accuracy for wind power applications one of the ultimate goals is to minimise overall model error, hence the more stations are included, the better the results. Secondly, when it comes to performing an uncertainty analysis of the wind atlas we will respect the fact that such a comparison will likely yield a lower bound for errors (see subsection 5.2.4 on page 126).

The detailed settings for all model components of these SA wide runs may be obtained in Appendix B.2.4 on page 186.

5.1.3 Execution

The model was run on an IBM p690, a shared memory computer with 32 1.9GHz Power4+ processors and 78GB of RAM at the Centre for High Performance Computing (CHPC) in Mowbray, Cape Town. The two simulations were configured to use all CPU's by utilising MM5's openmp parallelism feature, however, in practice only about 24 CPU's were fully saturated.

Each run only required about 421MB of RAM and 3.5 days of real time to complete with an output for domain 2 of no less than 182GB. Post-processing this output proved to be a bigger challenge and more time consuming than running the model itself, though the latter was complicated by the fact that the process consistently crashed half way through and had to be restarted from a previously saved state once before it could complete successfully.

This meant having to carefully merge the two separate outputs without creating overlaps or even gaps. The physical continuity of all wind and other fields at the seam were thoroughly checked visually at several levels such that the integrity of the model output which the following sections are based on was not affected in any way.

5.2 Accuracy of Wind Atlas

We now proceed to carry out a validation procedure analogous to the Eastern Cape one of the last chapter (see subsection 4.3.5 on page 74) using the comparison statistics developed in section 2.2 on page 33, just using all of the 17 SAWS weather stations described in chapter 2. We will often distinguish between the

stations within 50km of the ocean, namely (from West to East) Alexander Bay, Lamberts Bay, Langebaanweg, Cape Town, George, Port Elizabeth, Port Alfred, East London and Durban as well as the inland stations Beaufort West, Kimberley, Bloemfontein, Thabazimbi, Irene, Pietersburg, Vryheid and Nelspruit.

Before proceeding to the in-depth verification it must be stressed that all of the above observations were utilised for observational nudging with a small radius of influence. This in effect makes the model output a function of the measurements, at least to a certain extent. Yet we use the same for establishing MM5's accuracy which some would argue is an invalid approach because it prejudices the model validation by reducing biases upfront.

However, firstly we have no other observations to compare the model to and secondly we recall that we used all stations for the nudging precisely to maximise accuracy in the first place. All that implies is that the real world discrepancies for grid cells outside of the relaxation zones (the cell coincident with the anemometer and the immediately surrounding ones) are likely to be somewhat higher than those that we shall discuss in the next few sections, something that will be kept in mind and revisited in subsection 5.2.4 on page 126.

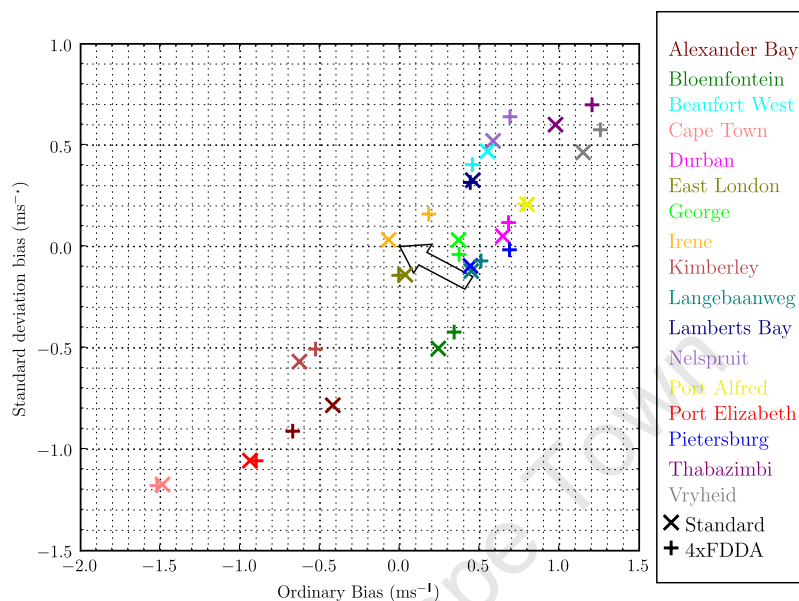
5.2.1 Average Wind Speed and Standard Deviation

The normal and standard deviation wind speed biases b_V and c_V are presented in Figure 5.2 to the right. There is a lot of information embedded within this so we will unpack it one by one.

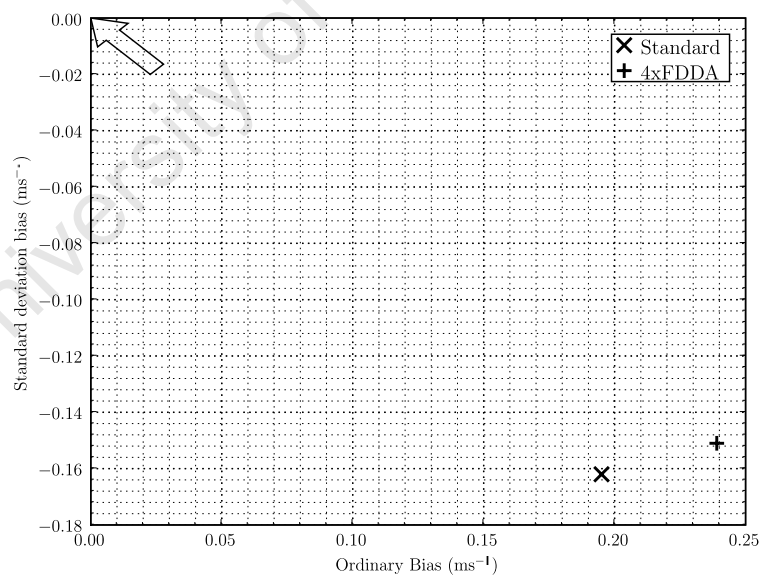
Firstly the station specific results in Figure 5.2(a) suggest that all the points lie close to a straight line from low to high bias with a slope of just less than one and intercept just less than zero. This is consistent with the nature of the underlying Weibull distribution which typically has a larger σ_V for higher $\bar{V}$.

Secondly the effect of turning on grid nudging seems to move the points along the same straight line, increasing b_V and/or c_V by $0.05\text{-}0.25\text{ms}^{-1}$ for nine stations and decreasing them only at two locations (Beaufort West and Alexander Bay), the rest showing no appreciable shift. This is in stark contrast to what we observed before in section 4.5 on page 90 where it was concluded that this movement was the other way round. Therefore we were correct in our suspicion that mesh FDDA can produce a very different response for a large domain than for a very limited Eastern Cape one and justifies the decision to carry out two simulations.

Which one of the two is more accurate is a question that the spatial average



(a) Individual stations



(b) Spatial average

Figure 5.2: Goal plot of standard deviation bias c_V vs ordinary bias b_V of wind speed for both wind atlas simulations. Each individual station is represented by a different colour while each run is depicted by a distinct symbol

Rank	Station	$\overline{V_o}(\text{ms}^{-1})$	$b_V(\text{ms}^{-1})$	$c_V(\text{ms}^{-1})$	$z_0(\text{cm})$
1	Cape Town	5.87	-1.49	-1.17	9.68
2	Port Elizabeth	5.71	-0.94	-1.06	41.16
3	East London	4.75	0.03	-0.14	22.10
4	Alexander Bay	4.63	-0.42	-0.78	5.98
5	Port Alfred	4.33	0.80	0.21	20.31
6	Beaufort West	4.33	0.55	0.47	15.14
7	Durban	4.27	0.65	0.05	17.28
8	Lamberts Bay	4.00	0.46	0.33	9.96
9	Kimberley	3.95	-0.63	-0.57	8.37
10	Langebaanweg	3.90	0.45	-0.12	9.22
11	George	3.42	0.37	0.03	39.32
12	Irene	3.35	-0.07	0.03	10.02
13	Pietersburg	3.32	0.44	-0.10	8.20
14	Bloemfontein	2.72	0.24	-0.50	10.42
15	Vryheid	2.57	1.15	0.46	13.27
16	Nelspruit	2.07	0.58	0.52	16.73
17	Thabazimbi	1.82	0.98	0.60	16.12
	Average	3.82	0.18	-0.10	16.08
	Stdev	1.13	0.70	0.55	10.16

Table 5.1: Bias statistics for the standard wind atlas run at all 17 SAWS weather stations. The entries are ordered from highest to lowest mean wind speed. For reference the roughness length MM5 assigns to the corresponding grid cell is given

of Figure 5.2(b) may answer seeing that we now have a reasonably large number of observations to smooth out area average vs point measurement discrepancies. While the difference is a comparatively small 0.05ms^{-1} of b_V the standard run does have the higher skill.

Turning now to the absolute performance of MM5, both $\overline{V}$ and σ_V metrics lie in the positive bias range except for three coastal (Alexander Bay, Cape Town and Port Elizabeth) and one inland (Kimberley) station. Most of b_V seems to be concentrated between -0.1 and 1.0ms^{-1} with twelve stations appearing here. This means that most of the time the model's area average is higher than what's observed at point locations which could be indicative of a systematic error. However, at this stage it is impossible to judge whether the model's flow is just too strong in general at 10m above ground or if the observations are just too obstructed by nearby buildings and thus often record impeded wind.

By glancing at the error statistics individually in Table 5.1 (for the standard run only) we may get a few other hints as to the nature of these discrepancies.

Cape Town has not only the highest observed mean wind speed but also the largest negative bias of -1.5ms^{-1} . This is despite the fact that the model already assumes a very low surface roughness of 9.7cm here despite the mostly urban surroundings. Perhaps there are wind channelling effects between Table Mountain in the west and the Hottentots Holland mountains to the east which accelerate the wind significantly, though the model is unable to resolve such effects at 18km. This was already hypothesised in the precursor study to this wind atlas [Hagemann 2003].

The next strongest flow is recorded at Port Elizabeth at 5.7ms^{-1} and just like Cape Town the model is a lot weaker, around 0.9ms^{-1} . Here z_0 might be at fault though - a high 41cm might not coincide with the conditions around the anemometer as there are only built-up areas NW to NE and the sea is only 3km to the east. Nevertheless, the application of a much larger domain has resulted in an improvement of about 0.4ms^{-1} relative to our previous Eastern Cape results (see Appendix A on page 159).

Encouraging is the almost perfect match at East London at third place with $\overline{V_o} = 4.8\text{ms}^{-1}$. Not only b_V but also c_V are very close to zero, indicating an optimal speed histogram, in line with our preceding discussions. Fourth in line is Alexander Bay with 4.6ms^{-1} and once again a negative bias of -0.4ms^{-1} regardless of the already low z_0 of 6cm. Perhaps sub-grid scale topographical forcing is responsible for this error.

Fifth is our last Eastern Cape station Port Alfred at a respectable $\overline{V_o} = 4.3\text{ms}^{-1}$ and $b_V = 0.8\text{ms}^{-1}$ which is still quite large but compared to the identically configured run for the provincial domain this is 0.4ms^{-1} closer to reality, most likely because of the addition of observational nudging (Port Alfred was excluded before). At this point we can state that the decision to carry out sensitivity studies over the Eastern Cape have definitely paid off since its three stations are in the top five of the country in terms of importance to wind power.

Beaufort West as the only inland station with $\overline{V_o} > 4.0\text{ms}^{-1}$ exhibits very similar features to Port Alfred with an identical mean wind speed and a somewhat lower bias. Durban and Lamberts Bay rank seventh and eighth respectively with b_V of $0.45 - 0.65\text{ms}^{-1}$.

Finally we must emphasise the meaning of the second last row of Table 5.1 to the left. The reader may recall that the spatially averaged values of Figure 5.2(b) have been derived from an aggregate time series, not an aggregate metric. That last row carries out the latter, that is, averaging all the individual b_V for comparison. This yields an independent confirmation of the fact that the model

is, in general, over-predicting winds by 0.2ms^{-1} .

5.2.2 Daily and Seasonal Cycles

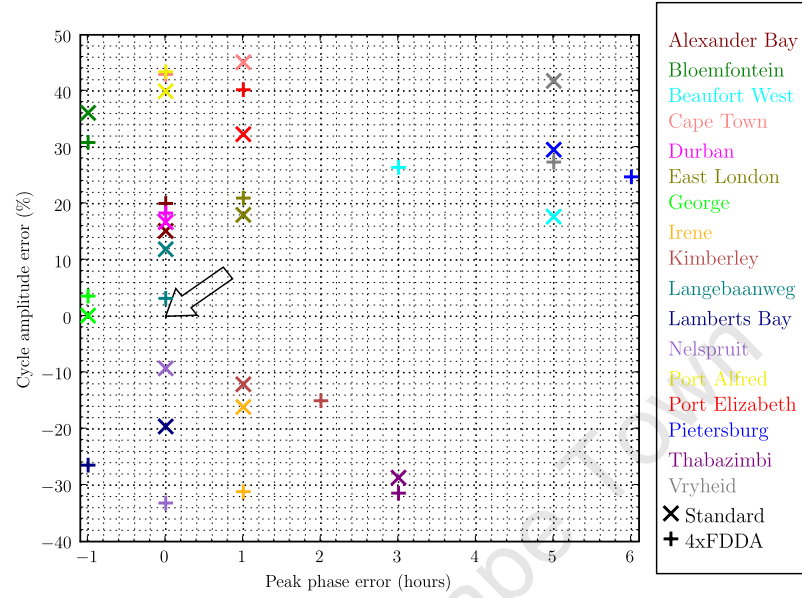
In terms of the diurnal cycle in the wind we now look at the *CAE* vs *PPE* performance in Figure 5.3 to the right. The station specific results of Figure 5.3(a) are by far the most useful and reveal plenty of information about the skill of MM5.

First it is very reassuring that the timing of the peak is well represented to within an hour for all but four measurements, Beaufort West, Pietersburg, Vryheid and Thabazimbi. It thus seems that the coastal land and sea breezes can indeed be modelled fairly well at 18km resolution but the valley and slope winds inland cannot, which is consistent bearing in mind that topographically forced thermal flows due to nearby mountains and hills are mostly microscale and therefore depend on features that cannot be resolved sufficiently.

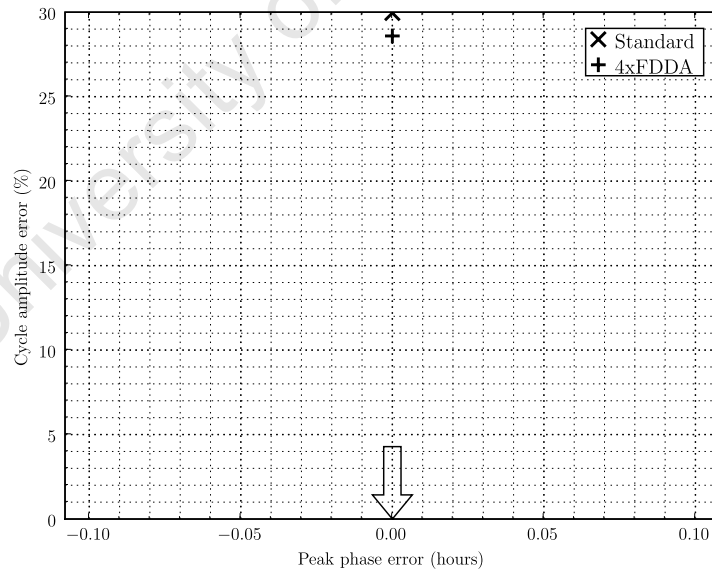
Regarding whether or not grid nudging still has a damaging effect on the development of an appreciable 24h harmonic Figure 5.3(a) reveals that with the exception of the above four towns in which the model is very far off to begin with, only Bloemfontein and Langebaanweg benefit from FDDA. All other 13 sites indicate that the standard run is superior, though the differences are not nearly as large as they were before at our Eastern Cape sensitivity experiments, especially for Port Elizabeth and Port Alfred where large regressions were seen before.

Even though the spatial averages of Figure 5.3(b) to the right seem to suggest otherwise, the difference is not only marginal (1.5%) but we remember that in our Eastern Cape runs grid FDDA minimised the *CAE* for all the wrong reasons at East London. This is in fact the case again for the spatial average as the reader may verify in Figure 5.4 on page 117. Mesh nudging pulls both the *u* and *v* cycles away from reality but in the process manages to achieve a slightly better *V* cycle, an artefact of the errors inherent in the reanalysis. Hence we conclude that the output of the “Standard” run without relaxation towards NCEP/NCAR reanalysis data is definitely preferable.

On the same note it is interesting to see that on average MM5 again fails to properly reproduce the low observed nocturnal winds, the main cause for not being able to get the *CAE* just right which is probably the most important driver of bias. This is likely an issue with PX/ACM’s treatment of radiative forcing, perhaps manifested through too strong night time vertical momentum

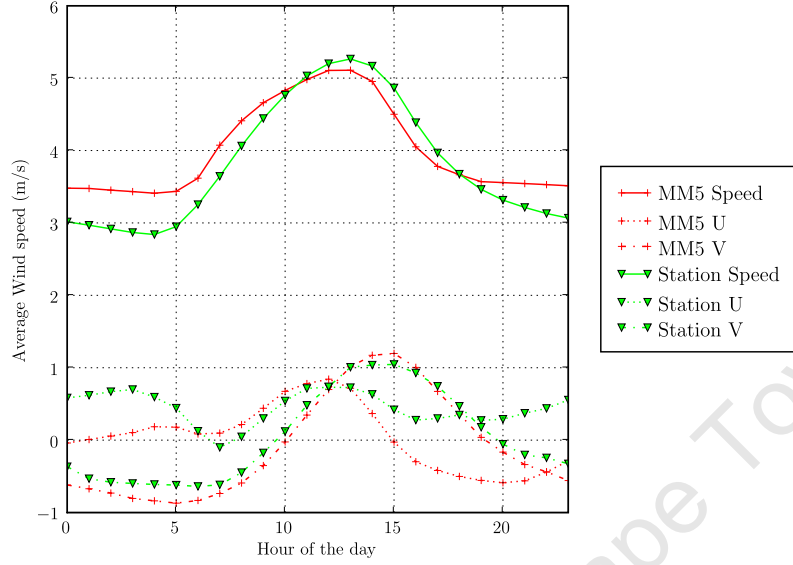


(a) Individual stations

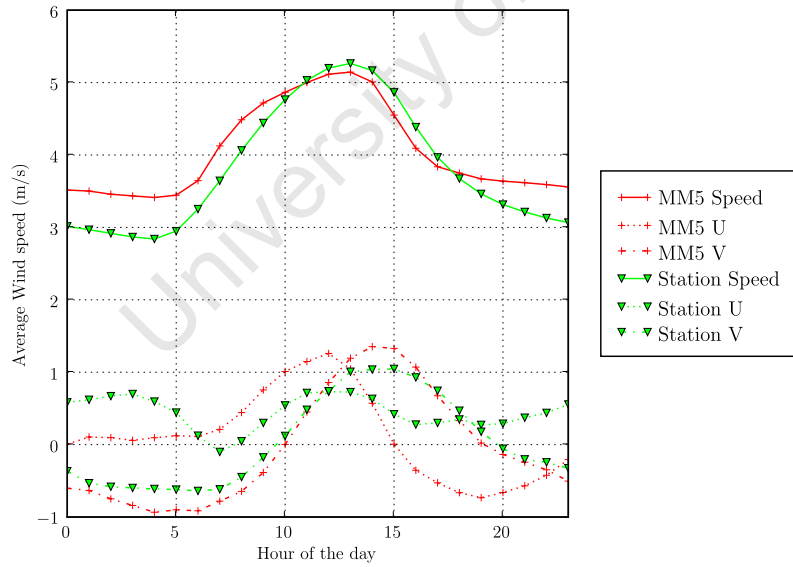


(b) Spatial average over all stations

Figure 5.3: Goal plot of daily cycle amplitude error CAE vs daily peak phase error PPE of wind speed for both wind atlas runs. Again each colour represents a particular observation, different symbols stand for the two separate simulations



(a) Standard Run



(b) Standard with 4xFDDA grid nudging

Figure 5.4: Standard and 4xFDDA V_{24} , u_{24} and v_{24} daily cycles spatially averaged over all weather station locations

exchanges caused by an incorrect stability profile. Also, on average the model is very skilled at simulating the 24h north-south fluctuations but fails on the evolution of the u winds, again mostly at night. We shall revisit this point in subsection 6.2.3 on page 151.

At this stage it is useful to look at the individual periodicities of all 17 locations given in Figures 5.5 to 5.7 on pages 119–121, again in order of decreasing importance to wind power considerations.

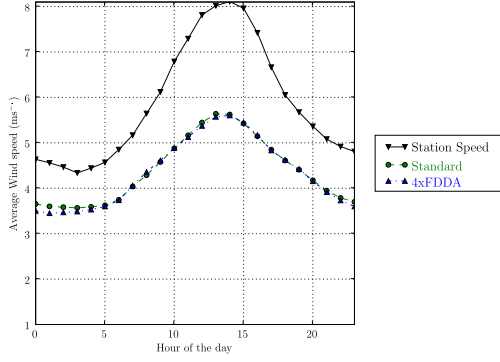
One can clearly see now in Figure 5.5(a) why the match at Cape Town is so poor. Not only is the amplitude under-predicted by around 45% but the modelled winds *always* stay well below what's observed, very similar to what is happening at Port Elizabeth just a lot more accentuated. In both cases as before we can only speculate that this systematic bias is due to a local surface roughness discrepancy combined with microscale orographic forcing that cannot be resolved at 18km.

Regarding our three Eastern Cape anemometers our standard wind atlas simulation unfortunately does not perform nearly as well as the AK/KF2 with 4xOBS combination on the much smaller domain used before (compare to Appendix A on page 159). While t_{max} is mostly identical everywhere, at Port Elizabeth we lose about 5% in CAE , from an original 27% to 32%. Whereas we previously got to within 8% of reality at East London we now have to be content with 18%, but at least the biases are still very small. At Port Alfred this regression is the worst with a fall of no less than 18% from 22% to 40%. Surprising then that b_V and c_V improved so much here (about 0.4ms^{-1}).

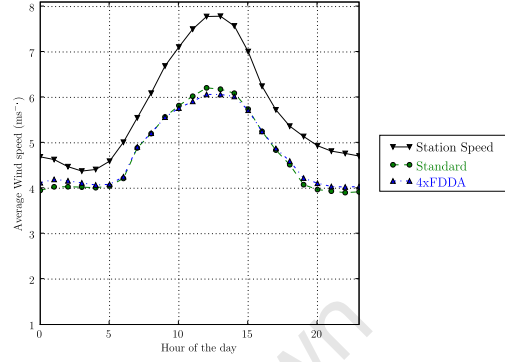
Also perhaps unexpected is the fact that at 18km, which is still a coarse resolution for a sea breeze, we manage to over-predict the diurnal cycle at Lamberts Bay by approximately 20%. But then again it is likely that there are local obstructions around the 10m mast that slow down the wind and flatten the peak, thereby distorting reality.

Finally we must comment on the harmonics at inland stations. Firstly it is counter-intuitive that *all* of them exhibit a well defined 24h oscillation, windy or not, albeit with a lower amplitude. This is despite the fact that quite a few of them, e.g. Bloemfontein, Irene, Kimberley and Pietersburg neither have any large water bodies nearby nor significant surrounding mountains and hills from which thermally induced winds could develop. Such features could well be governed by the periodic growth and decay of the convective planetary boundary layer and the related vertical momentum mixing as is typical of the highveld.

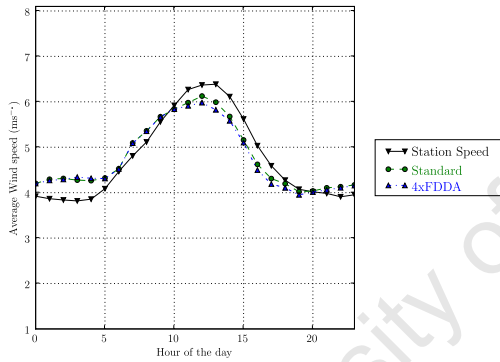
Either way, it tells us that the 24h component is *the* most important dynamic



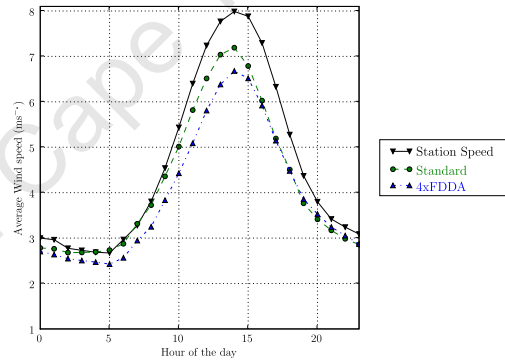
(a) Cape Town



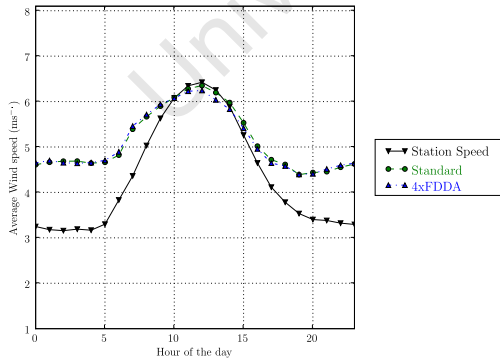
(b) Port Elizabeth



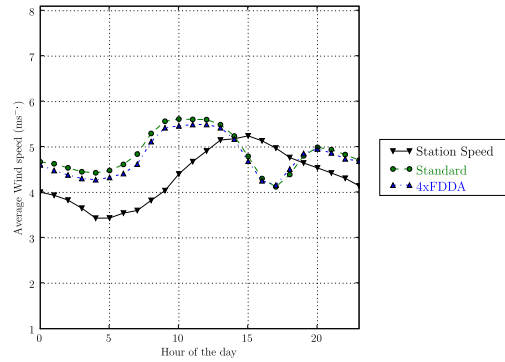
(c) East London



(d) Alexander Bay



(e) Port Alfred



(f) Beaufort West

Figure 5.5: Modelled and observed daily cycles for the six windiest weather stations (five coastal and one inland)

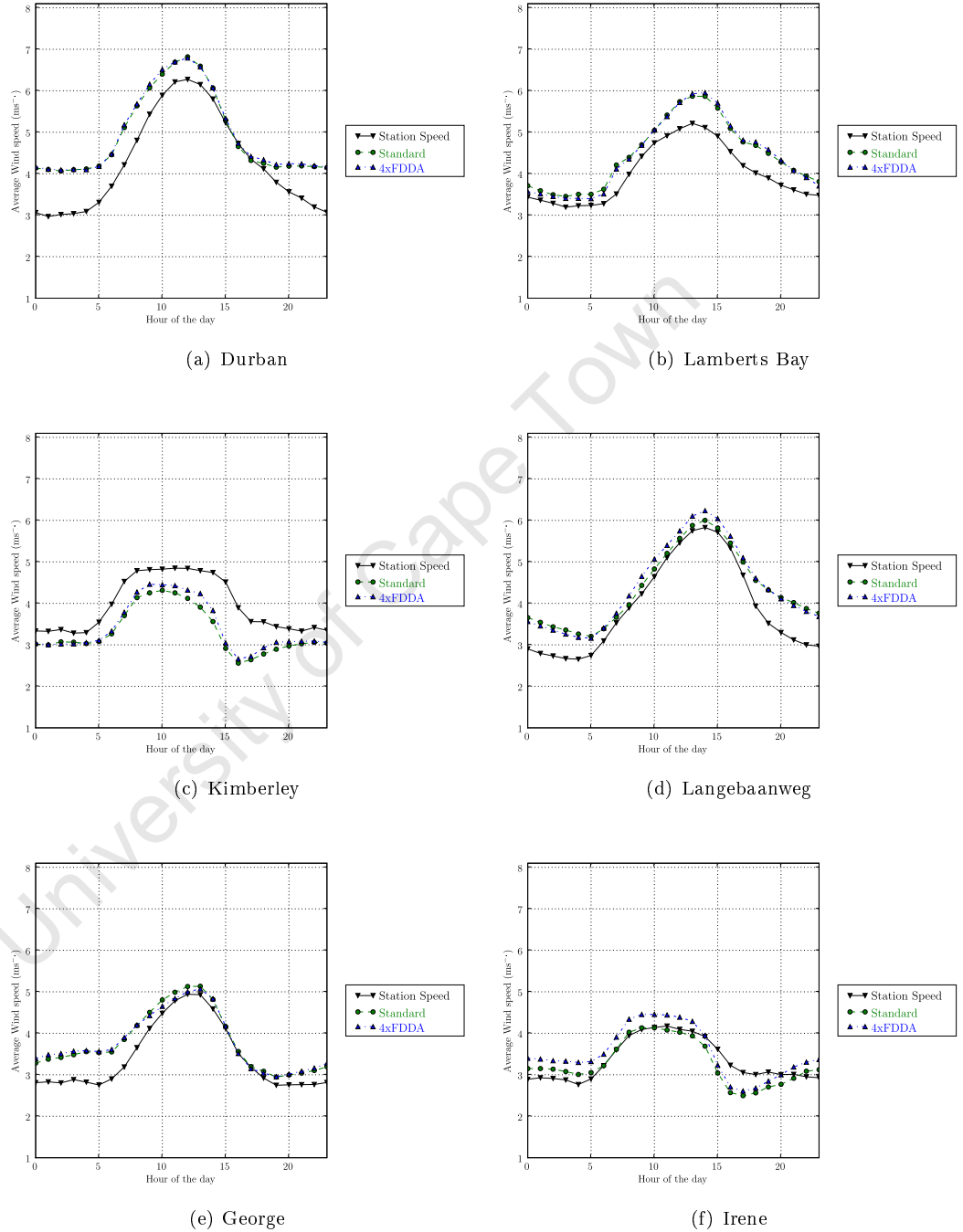


Figure 5.6: Modelled and observed daily cycles for the six intermediate weather stations (rank 7-12, four coastal and two inland)

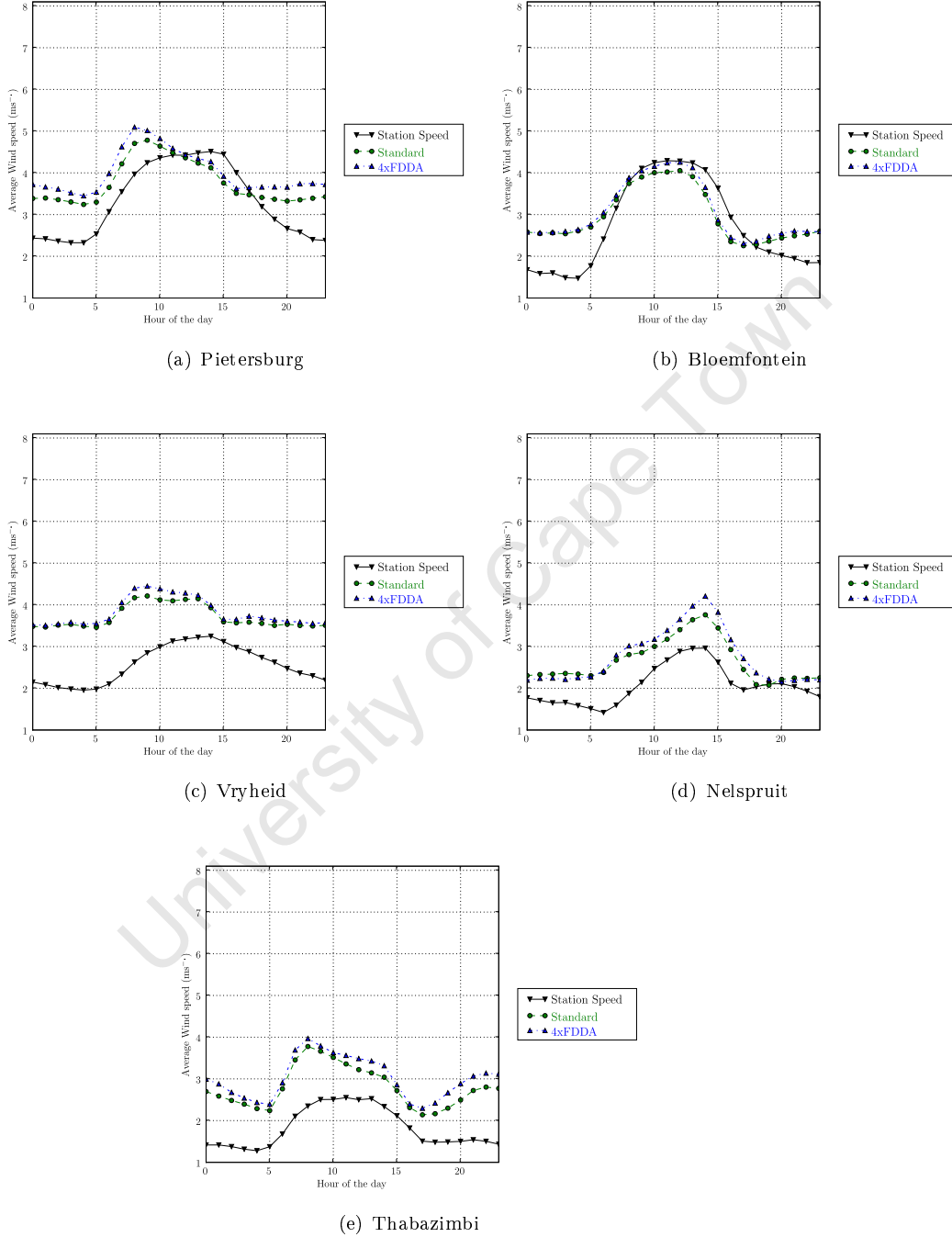


Figure 5.7: Modelled and observed daily cycles for the five calmest weather stations (all inland)

part of the flow's variability and should never be neglected in any wind power study, least of which mesoscale modelling efforts to get as close to reality as possible.

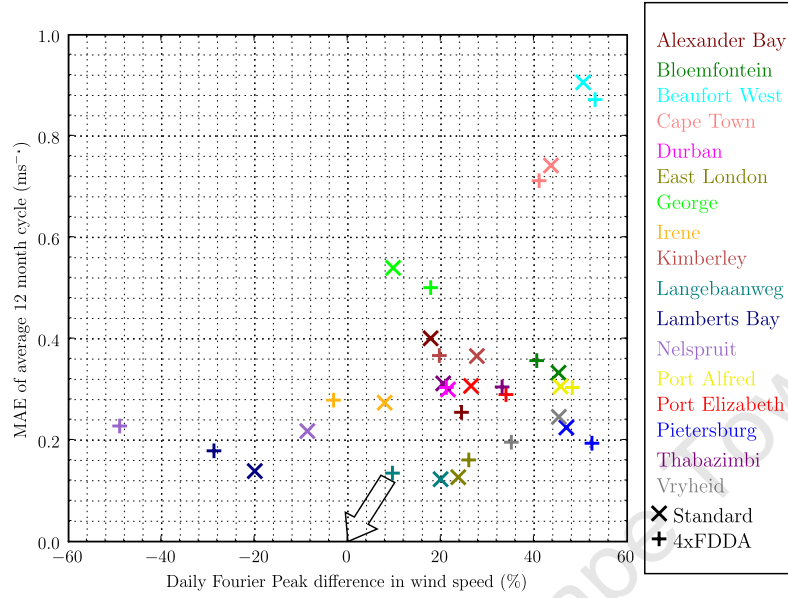
Having said that this aspect of MM5's time-dependent performance at the non-coastal sites is fairly poor. Very often a strong minimum is simulated in the late afternoon between 4pm and 6pm - one that does not actually exist in nature. Also at Beaufort West, Pietersburg, Vryheid and Thabazimbi the model peaks far too early. This can most probably *not* be attributed to lack of topographical detail, it looks more like a flaw within Pleim/Xiu surface fluxes and land surface interaction at those times of the day, coupled with the inadequacies inherent in the employed ACM planetary boundary layer parametrisation as well as the chosen cumulus scheme.

The climatology on the highveld is very much driven by convective processes which we did not optimise our model for. We remind ourselves why this trade-off was accepted in the first place, namely due to previous works suggesting that the winds are not strong there on average [Diab 1995]. Now we discover that this is not quite true, as Beaufort West, the 6th windiest station, demonstrates.

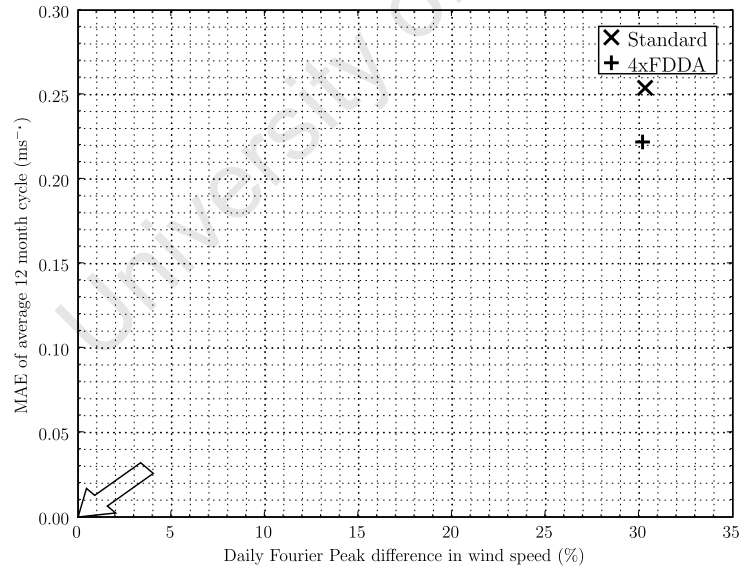
That MM5 has significant difficulties in reproducing the rather complex diurnal cycle over the interior has been noted before in the case of precipitation [Tadross et al. 2006]. However, a detailed discussion and investigation of these effects and why our MM5 configuration is unable to simulate them well is considered to be beyond the scope of this project. All we can state at this stage is that there are issues here that somewhat diminish the value of using MM5 in the context of a high resolution wind climatology and which should certainly be addressed in future research. We shall revisit this point in more detail in the concluding discussions of subsection 6.2.3.

Moving on to the second set of validation metrics we now look at the *DFPD* and *MMAE* pairs which are visualised in Figure 5.8 on the next page. As before we are only interested in the vertical distribution of the points as the *DFPD* does not state anything beyond what the *CAE* already tells us.

Addressing the FDDA question first it turns out by looking at the individual metrics of Figure 5.8(a) that seven stations benefit from grid nudging being turned on (Cape Town, Beaufort West, Port Elizabeth, Alexander Bay, George, Pietersburg and Vryheid). Only three (East London, Lamberts Bay and Bloemfontein) have their seasonal variations better represented in the standard run while the remaining seven show no noteworthy change. Hence mesh FDDA improves the *MMAE* skill of MM5, though only by a very small margin of around



(a) Individual stations



(b) Spatial average over all stations

Figure 5.8: Goal plot of 12 monthly(seasonal) mean absolute error $MMAE$ vs daily Fourier peak difference $DFPD$ of wind speed for both wind atlas runs

$0.025\text{--}0.05\text{ms}^{-1}$ which is underlined by the spatial average statistics of Figure 5.8(b).

For all but George, Beaufort West and Cape Town the model manages to simulate the month to month fluctuations fairly well with an *MMAE* of less than 0.4ms^{-1} . This is an encouraging result suggesting that the seasonal wind atlas output is indeed useful.

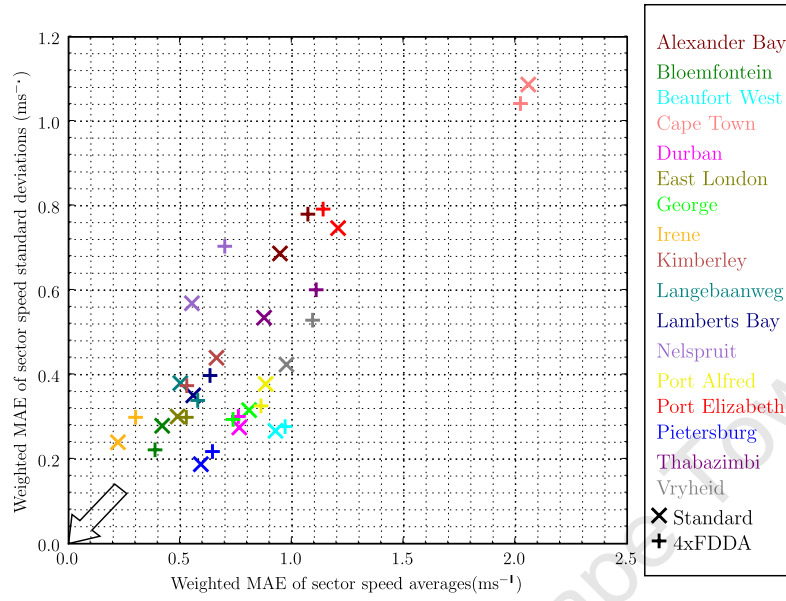
Compared to the Eastern Cape sensitivity runs East London the *MMAE* accuracy has kept mostly constant. East London shows a very minor improvement from 0.17ms^{-1} to 0.13ms^{-1} , Port Elizabeth a slight decrease from 0.24ms^{-1} to 0.31ms^{-1} while Port Alfred remained virtually unchanged from 0.33ms^{-1} to 0.31ms^{-1} .

5.2.3 Directional Performance

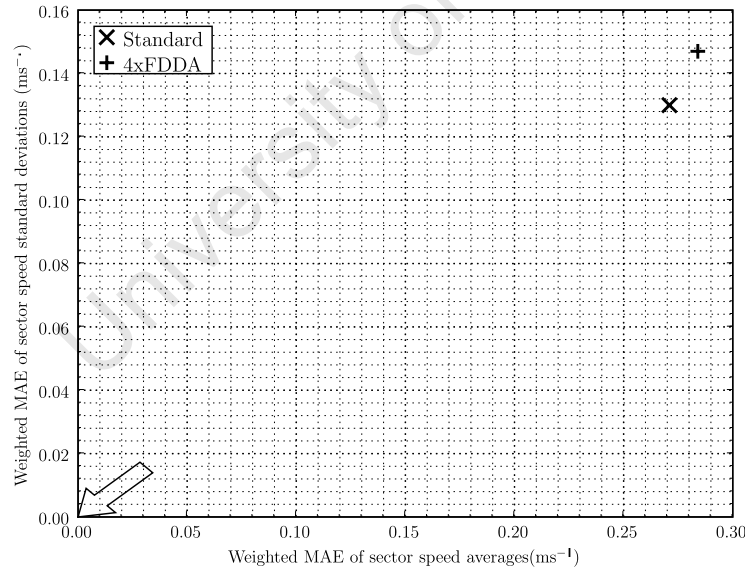
This remaining evaluation deals with the directional accuracy of the two wind atlas runs. The goal plot of *WMAESSD* vs *WMAESSA* (analogous to c_V vs b_V) is presented in Figure 5.9 on the following page. The distribution of the individual observation metrics of Figure 5.9(a) are again very similar to those of Figure 5.2(b) on page 112, just that all values are of course positive now. Cape Town exhibits the largest errors by far, followed by Port Elizabeth, Alexander Bay, Vryheid and Beaufort West, whose *WMAESSA* is greater than 0.9ms^{-1} . All the others are between a moderate 0.2 (Irene) and 0.9ms^{-1} .

For the three Eastern Cape 10m masts this compares favourably with the previous province specific sensitivity runs (see Appendix A on page 159). For Port Elizabeth the *WMAESSA* reduced by 0.2ms^{-1} from 1.44ms^{-1} to 1.2ms^{-1} with similar achievements in the *WMASSD* which changed from 0.93ms^{-1} to 0.75ms^{-1} . Port Alfred's gains are of the order of 0.1ms^{-1} for both error types while East London exhibited virtually no changes as the match is already very good there. Once again, the larger domain simulation has actually enhanced the validity of the model, most likely because it has more freedom to correct for errors in the Reanalysis with increased resolution of surface characteristics such as roughness and topography, not only in the vicinity of the stations but much further upstream too.

The standard model configuration proves to be superior over grid FDDA here as well. For most anemometers (Alexander Bay, Thabazimbi, Vryheid, Nelspruit, Port Alfred, Beaufort West, Durban, Lamberts Bay, Irene and Pietersburg) the latter decreases both skill statistics while only in four cases (Cape



(a) Individual stations



(b) Spatial average over all stations

Figure 5.9: Goal plot of weighted MAE of sector wind speed averages ($WMAESSA$) and standard deviations ($WMAESSD$) for both wind atlas runs

Town, Kimberley, George and Bloemfontein) it improves our results. For the remaining three locations (Port Elizabeth, Langebaanweg and East London) the differences are either mixed or negligible. The spatially averaged scores of Figure 5.9(b) reinforce the claim that the application of mesh nudging does not enhance the wind sector specific performance of the model but rather deteriorates it, albeit by small amounts.

Figure 5.9(b) further shows us that in general MM5 does a better job at predicting directional variability ($\overline{WMAESSD} = 0.13\text{ms}^{-1}$) by a factor of two than directional speed averages ($\overline{WMAESSA} = 0.27\text{ms}^{-1}$). Hence when evaluating any given site in South Africa we can rely on the accuracy of the former more so than with the latter.

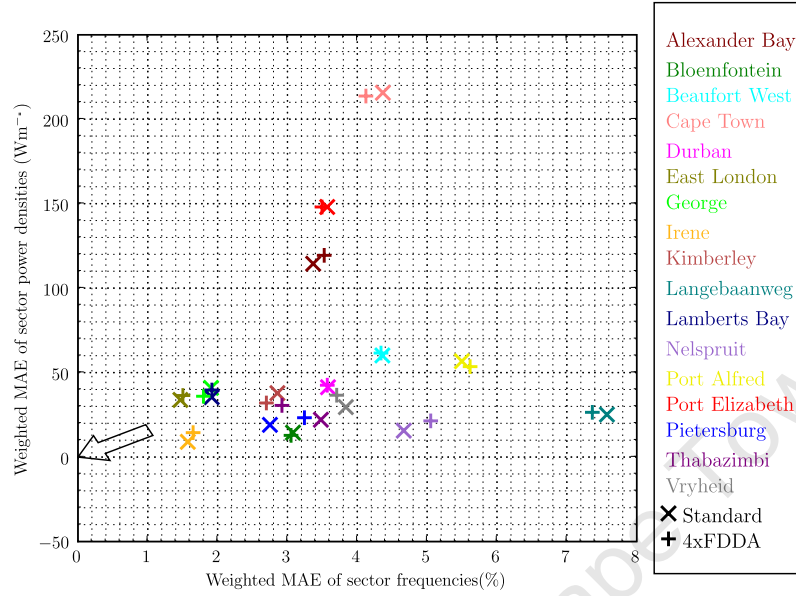
Finally we take a look at the frequency ($WMAESF$) and power density ($WMAEPD$) errors in Figure 5.10 on the next page. The station specific results of Figure 5.10(a) reveal that most of $WMAEPD$ is clustered between 10 and 60Wm^{-2} with only our three outliers Alexander Bay, Port Elizabeth and Cape Town which we have already discussed lying above 100Wm^{-2} . Furthermore all $WMAESF$ lie below 4.7% except for Port Alfred (5.5%) and Langebaanweg (7.6%) which surprisingly enough have not occurred as outliers before. This suggests localised wind vane distortion caused by nearby obstacles such as buildings and trees, however, we cannot say for sure without investigating the detailed surroundings of these sites.

Our Eastern Cape sensitivity run performance marginally improved with respect to frequency; 0.3%, -0.1% and 0.04% for Port Elizabeth, Port Alfred and East London respectively (compare with A on page 159). Regarding the power density the errors are reduced but this is to be expected from lower biases.

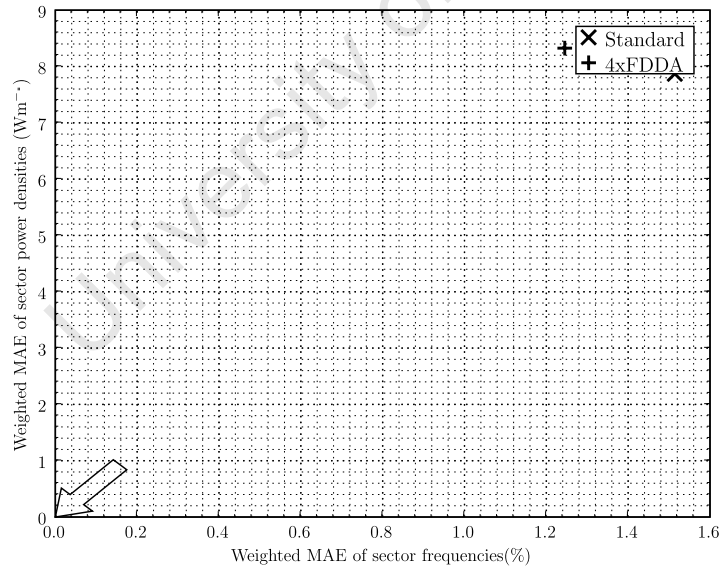
Grid nudging as before has little or no apparent benefits here either. While Cape Town, Kimberley, George, Langebaanweg do show amelioration of both metrics with mesh FDDA turned on, Nelspruit, Pietersburg, Alexander Bay, Irene and Lamberts Bay reveal deterioration with negligible or mixed results at the remaining sites. Hence the global change of about 0.30% that the spatially averaged picture of Figure 5.10(b) suggests is likely to be insignificant if meaningful at all.

5.2.4 Summary and Uncertainty Analysis

To sum up the discussion of the previous subsections we may state the following for our 18km results:



(a) Individual stations



(b) Spatial average over all stations

Figure 5.10: Goal plot of weighted MAE of sector frequencies ($WMAESF$) and power densities ($WMAEPD$) for both wind atlas runs

- Despite identical core integration parameters the mere presence of a much larger domain has significantly affected the model's skill at the three Eastern Cape stations that were investigated in detail in chapter 4. All bias and *MAE* related metrics (*MMAE* and the four *WMAE*) have improved appreciably while the errors in the strength of the daily cycle as indicated by the *CAE* have increased by a sizeable margin. This is indicative of a strong sensitivity of the model to its grid layout and could be explained by the influence of certain topographical teleconnections that were not present before (e.g. the entire Drakensberg)
- As before grid nudging has mostly detrimental effects on the validity of the simulation although the differences are now much smaller (positive or negative) than they were before in the sensitivity experiments of chapter 4. The reason may be found in the errors inherent in the NCEP/NCAR Reanalysis over the South African domain. This confirms our initial statement that mesh FDDA should *not* be applied in order to get as close as possible to reality. Hence the utility of this wind atlas lies in the run that was relaxed towards observations only (referred to as “standard” above)
- The daily cycle turns out to be the most important component of variability in the wind speed. For coastal sites the model is very good at predicting this oscillation while for inland measurements MM5 has difficulties in matching the evolution of the observed 24h periodicity. The problems in the interior are not well understood and should be investigated in detail in future research
- On average the model is seen to have a positive night time bias which may be attributed to an incorrect treatment of nocturnal radiative forcing and/or inadequate parametrisation of nocturnal turbulence in stable stratifications

What remains to be done is to estimate the fundamental uncertainties that are inherent in our generated 18km dataset. Looking back at the bias statistics of Table 5.1 on page 113 we may check if b_V and c_V do not follow a Gaussian distribution by applying the non-parametric Anderson-Darling test which is particularly useful for this purpose with such a small sample of only 17 values [Stephens 1974]. It provides us with p-values of 0.0995 and 0.2492 for b_V and c_V respectively which imply that despite the presence of a long left tail and

associated outliers (the large negative biases for Cape Town and Port Elizabeth) the departure from normality is not statistically significant at the 5% level.

While this does not prove that b_V and c_V are Gaussian variates it does make this assumption somewhat safer, meaning we may utilise some of the properties of this distribution to quantify our uncertainty. Therefore the standard deviation figure of both b_V and c_V of Table 5.1 on page 113 can be taken as an estimate of the error that arises when comparing our wind atlas's 18km area average with a 10m weather station. This equates to 0.70ms^{-1} for b_V and 0.55ms^{-1} for c_V with a standard 68% confidence interval. Do note, however, that the expected error in the grid cell values themselves are probably lower than that but we have no way of telling because there is no gridded wind dataset that reflects "reality", one of the reasons we have embarked on this project in the first place.

Even though the *PPE* is strictly speaking a continuous statistic in our case it is discrete owing to the limited temporal resolution of one hour for the underlying time series. Hence finding a correct pdf to describe it may be a difficult undertaking, however, we have already stated with confidence that for coastal sites MM5 predicts the timing of the daily wind speed peak to within one hour and that for inland locations the model's daily cycle should not be relied upon due to a poor match in general.

The *CAE* would again be expected to follow a normal distribution. An Anderson-Darling test using the 17 value sample yields a p-value of 0.341 so again we cannot reject the null hypothesis of a Gaussian. For coastal areas then where this makes sense the the average *CAE* of 17.8% may be framed with an error of 20% on either side within 68% certainty.

Our remaining five skill scores (*MMAE* and the four *WMAE*) are all based on a mean absolute error measure which by definition cannot assume normality since it is restricted to the interval $[0, \infty)$. Instead these probably follow a gamma or Weibull distribution though a curve fitting exercise would be severely limited by the fact that our sample is so small. Hence we shall use their averages and standard deviations as very crude approximations of the expected error and its variability.

All of these considerations are collated and distilled in Table 5.2 to the right together with all values not yet explicitly mentioned as well as additional explanations. Important to note is that on average the model's wind speed is in line with reality near the coast but at the expense of a large standard error of 0.8ms^{-1} while the magnitude of the inland flow is over-predicted by around 0.4ms^{-1} with a somewhat smaller error of 0.6ms^{-1} , in both cases within a 68%

Metric	Coastal		Inland		Overall	
	mean	σ	mean	σ	mean	σ
b_V (ms^{-1})	-0.01	0.78	0.41	0.57	0.18	0.70
c_V (ms^{-1})	-0.30	0.56	0.12	0.47	-0.10	0.55
CAE (%)	17.75	20.00	not reliable		n/a	
PPE (hours)	within 1 hour		not reliable		n/a	
$MMAE$ (ms^{-1})	0.33	0.21	0.36	0.23	0.35	0.21
$WMAESSA$ (ms^{-1})	0.91	0.49	0.65	0.26	0.79	0.41
$WMAESSD$ (ms^{-1})	0.50	0.28	0.37	0.14	0.44	0.23
$WMAESF$ (%)	3.69	1.94	3.32	0.99	3.52	1.53
$WMAEPD$ (Wm^{-2})	79.07	66.04	25.98	16.52	54.09	55.19

Table 5.2: Uncertainty of the Wind Atlas results for each error statistic at 10m above ground based on aggregates from 17 model-station comparisons. Means may be regarded as *systematic* expected errors, standard deviations in **bold** represent a standard error for a 68% confidence interval, a crude estimate for its variability otherwise. The reader is referred to section 2.2 on page 33 for the definition of these metrics; note that the $DFPD$ has been omitted for lack of additionality

confidence interval.

It must be emphasised that the above only applies to 10m above ground because these were the only available measurements at this stage. For a wind farm development that height is of course of very little value but in the absence of other publicly available readings within the PBL for the period under consideration we cannot objectively determine the model's error in the vertical.

However, we can make some rudimentary and subjective assessments of MM5's error above 10m. As we move higher up in the PBL the microscale shelter effects due to nearby obstacles decreases which should in practice improve the match between modelled and observed values. This is confirmed by the findings of detailed MM5 evaluations over the continental United States [Tonnesen et al. 2004; Jia and Morris 2003].

This advantage is likely offset by the slightly larger errors that would be encountered outside of relaxation zones since we utilised observational nudging at *all* stations that we considered for the validation. Hence when extracting data at 60m or 80m hub height we assume that the accuracy of the simulated wind resource anywhere within the mesh is about the same as that at 10m.

Finally we must remind ourselves of the long term correction of -0.2ms^{-1} that was stipulated in section 3.4 on page 51 owing to the way that our optimal 365 simulation time span is slightly windier than the observed climatology at

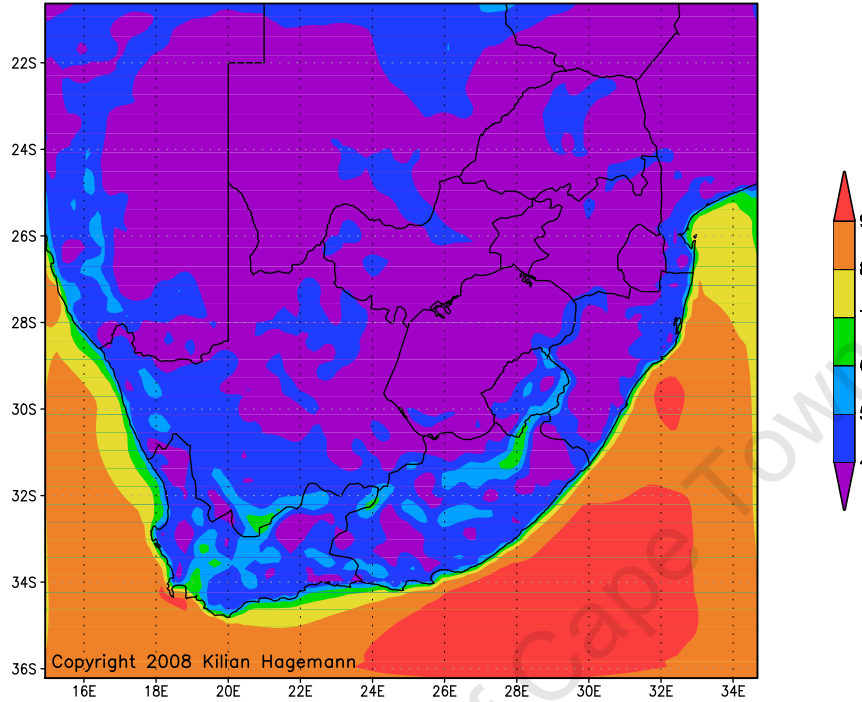


Figure 5.11: Average annual wind speeds at 10m above ground in ms^{-1}

10m. Note that this applies only for annual statistics, for seasonal ones it is more likely a lower bound because of the reasons given in section 3.5. While this residual must certainly be ignored for the purpose of any model vs measurement comparison over identical time periods it should be applied when evaluating the economic wind farm feasibility of a given site.

5.3 Wind Atlas Maps

5.3.1 Annual Average Wind Speeds

In this section we finally present the results which justify the title of this thesis. A map of average annual wind speeds at 10m for South Africa is given in Figure 5.11 which has been generated from the lowest sigma level winds (about 10m above ground) with the -0.2ms^{-1} residual long-term correction of section 3.4 on page 51 added. It should be assessed with respect to the previous two wind atlases, the 1995 “Wind Atlas of South Africa” by Diab (Figure 1.1 on page 23)

and the 2001 “South African Wind Atlas” by Eskom (Figure 1.2 on page 25).

The differences are substantial. Diab found only the areas within about 50km of the coast as well as the Drakensberg escarpment between Lesotho and Swaziland to be of “good” potential with wind speeds of greater than 4ms^{-1} . Accordingly almost all inland areas appear as “moderate” ($3\text{--}4\text{ms}^{-1}$) or mostly “low” ($<4\text{ms}^{-1}$) on her map. This is in stark contrast to Figure 5.11 where the lowest colour contour has been chosen to be Diab’s highest (4ms^{-1}) to emphasise the fact that according to the *same* classification almost half the country may actually be termed to be of “good” wind resource.

On top of that there are sizeable inland regions of the Northern, Western and Eastern Cape which could be labelled “excellent” as they exceed 5ms^{-1} at 10m. Such areas even seem larger than those of comparable resources close to the sea and hence the latter acquire a much lower relative significance in our map.

Comparison with the Eskom atlas suggests that our results provide for much more detail over the Eastern Cape with an average wind speed range of $3\text{--}7\text{ms}^{-1}$ instead of $4\text{--}6\text{ms}^{-1}$. Certain high potential areas in Mpumalanga and the Northern Cape as well as many other points which could be regarded as “outliers” (e.g. Giant’s Castle) are not reproduced in MM5 because it does not directly rely on individual measurements and the large errors associated with microscale WAsP modelling over complex terrain.

Needless to say that our mesoscale atlas is free of any data voids and provides detailed, seamless and continuous coverage of the entire country. This is arguably one of the biggest advantages that the utilisation of a mesoscale model such as MM5 for generation of a wind climatology embodies.

It must be noted that none of the previous works have considered or calculated overall uncertainties in their respective results. Section on page 110 in comparison accommodates a fairly detailed evaluation of the model strengths and weaknesses when compared to relatively reliable point measurements at 10m. It yields an in-depth understanding of the errors in the simulated wind speeds and directions which helps to clarify the utility of this wind atlas. To this extent Figure 5.11 is accurate to within about 0.7ms^{-1} with a 68% confidence interval.

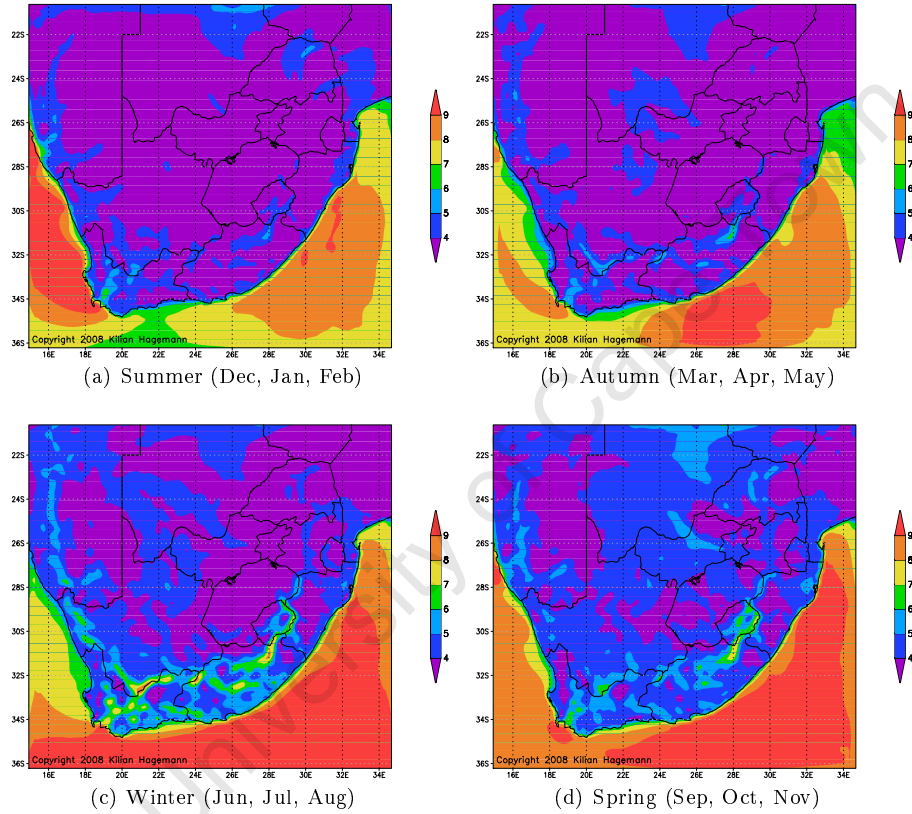


Figure 5.12: Average seasonal wind speed maps of South Africa at 10m above ground in ms^{-1} . Many detailed features may likely be attributed to topographical influences (*not* observational nudging, compare with Figure 2.1 on page 32)

5.3.2 Seasonal Average Wind Speeds

The average wind speeds according to season presented in Figure 5.12 to the left provide us with even more information that was contained in neither of the two predecessor wind atlases. It looks as if most of South Africa's strong winds occur during wintertime which is consistent with the fact that the South Atlantic low pressure and frontal systems characterised by high winds move much closer northwards into the country during June, July and August. The same "hotspots" in the Eastern, Northern and Western Cape apply as discussed above.

While on the whole the results appear to be consistent, some of the details could prove to be unrealistic. There are three locations near the boundaries of the respective Cape provinces that exhibit $8\text{ms}^{-1} < \bar{V} < 9\text{ms}^{-1}$ at 10m which seems almost too high. Some caution is thus advised not to accept these values on blind faith, especially considering that we have shown that MM5 has trouble simulating the diurnal cycle inland with our chosen parametrisation. Of course only future in-situ measurements could establish whether our model is really correct here or not.

The second windiest season is Spring, likely because both September and October still see cold fronts, cut-off lows and other high pressure gradient features pass through the southern parts of the country. Summer is the calmest with only the southwestern parts of the Western Cape as well as the West Coast experiencing significant flows, particularly the Cape Town area and Peninsula. This is to be expected with the South Atlantic High moving south into a position west of the country which results in very strong and sustained south-easterlies over that region during December, January and March. Autumn is somewhere in between with high wind resources only on the coastal edges of the Western Cape as well as the Eastern Cape interior.

In terms of the viability of wind power in the country this is a very significant result as these patterns *exactly* match electricity demand patterns which peak in winter and are at their lowest in summer [Maroga 2007]. It makes wind generated electricity that much more valuable in the context of strategic energy supply planning and capacity provisioning.

In terms of the accuracy the reader is reminded of the average *MMAE* figures of subsection 5.2.4 on page 126 which are in the region of 0.4ms^{-1} . Combining this additively with the ordinary bias uncertainty b_V of 0.7ms^{-1} using standard error propagation theory yields $\sqrt{0.4^2 + 0.7^2}$ or 0.8ms^{-1} . That

is the estimated error for the values given in Figure 5.12 on page 133.

Considering the fact that our optimal 365 day time subset was never chosen with time of the year accuracy in mind (see section 3.5 on page 54) the seasonal statistics are probably even less reliable than that, up to $\pm 1.0 \text{ms}^{-1}$ or more (pure estimate). This could well be the reason for the rather high wind speeds discussed above, perhaps 1996 was a particularly windy winter. However, the central seasonal tendencies are still likely to be valid, especially seeing that we averaged over three months in each of the four cases.

All in all we realise that while our mesoscale modelling methodology has delivered interesting results, their utility is not as high as one might expect from a wind atlas. That in itself does not diminish the value that an RCM such as MM5 may provide in terms of a wind climatology, especially considering how easily seasonal maps could be created, but rather emphasises a flaw in the way we restricted our 365 day integration period.

5.4 Quantification of Total Wind Resource

Now that we have generated a national wind resource dataset for South Africa and ascertained its accuracy the remaining question to address is how much wind power potential there is in the country. How much of the country's total electricity demand may be satisfied by energy converted by wind turbines in a renewable and carbon dioxide free manner? In the light of climate change mitigation and associated reduction of greenhouse gas emissions, South Africa's electricity crisis and security of supply concerns stemming from ever increasing fossil fuel costs this consideration has a significant strategic importance for both private and public stakeholders.

As discussed in more detail in our introduction (subsection 1.4.3 on page 26), the literature provides a very wide estimate of total exploitable wind capacity varying from 7.9 TWh to 106 TWh per annum - 3.4% to 46.2% of 2007 consumption. In this section we shall use the mesoscale wind data described above as well as some infrastructure data in order to arrive at our own evaluation of South Africa's wind potential.

We use the following broad approach. In order for a wind farm at a given site to be economically viable it has to fulfil the following criteria:

- its wind regime must be favourable
- it must be close to roads for transport, installation and maintenance of

the turbines

- it must be close to power lines to transmit the generated electricity to demand centres
- it must be located *onshore* and within South African boundaries (estimation of the offshore potential and that of other Southern African countries is left for future work)

The last one is straight forward to judge - as long as the location is within our borders and inside of coastlines it qualifies. The other three are much more difficult to pin down, especially the first one (more on this later). But in terms of infrastructure we assume that at a minimum access to secondary or main roads is needed to safely transport the large tower sections and blades required for the construction of the turbines. In addition we require that wind farms be connected only to high voltage transmission lines greater or equal to 66kV. These assumptions are very straight forward and in line with similar studies [Banks and Schäffler 2007].

Broadly speaking we will employ the services of a Geographical Information System (GIS) to evaluate the suitability of each cell across South Africa at a new grid of 1km resolution. For those regions that qualify the total annual energy production (AEP) is calculated and the individual values summed up to yield a total estimate.

How exactly this is performed and what assumptions are made in the process is dealt with in subsection 5.4.2 on page 139. First, however, comes the preparation of various input data so that we can easily use it for the stated purpose.

5.4.1 Data Preparation

We begin by reformatting the wind data that was prepared in section 5.1 on page 107 and evaluated in section 5.2 on page 110. As the foundation we have four dimensional quantities - hourly timeseries at approximately 10m, 30m 56m, 92m and 137m above ground of wind speed and direction (plus many other variables like T , p , q etc.) spanning 365 days every 18km for the entire country. In order to restructure this into something more useful for a two dimensional GIS to handle we carry out the following tasks:

1. post-processing the raw model output into four dimensional values of u , v and ρ

2. vertical interpolation of these quantities to standard 60m, 80m and 100m hub heights
3. calculation of wind speed V based on u and v
4. transformation of winds and air density into a grid of estimated annual energy production (AEP or E) and capacity factor C_f
5. horizontal interpolation of E and C_f to a 1km by 1km grid

Step 1. is performed by firstly converting MM5's various perturbation, surface and model lid pressure fields to an absolute $p(x, y, s, t)$ (s being the sigma level) in Pa based on equation 1.31 on page 28. That p is then combined with temperature $T(x, y, s, t)$ and mixing ratio $q(x, y, s, t)$ to arrive at the total 4d air density $\rho(x, y, s, t)$ according to equation 1.11 on page 8.

The resampling onto standard hub heights as step 2. is then effected by converting each sigma level to heights above ground for each grid cell using equation 1.32 on page 28, subtracting the model terrain height and then using a cubic spline defined over levels 2-5 to arrive at the standard 60, 80 and 100m above ground for ρ , u and v .

Step 3. is straightforward as

$$V = \sqrt{u^2 + v^2} \quad (5.1)$$

which follows from equations 1.6 to 1.7 on page 3.

The fourth task is a bit more complicated. We first of all have to choose a wind turbine whose power curve and associated C_p values we may use. For this purpose we take a Vestas V80 2000/80 2MW (onshore) model which is deemed a typical generator these days considering that the international trend is to upscale - the average turbine size in the UK in 2006 was 1.953MW [BTM Consult 2007]. Its power curve and C_p characteristics are given in Figure 5.13¹.

We use these values to calculate our energy yield per turbine per pixel according to equation 1.12 on page 8 (see subsection 1.2.4). The integration itself is approximated with a trapezoidal rule summation over a timestep of one hour. Capacity factors C_f are analogously evaluated using equation 1.14. As a third variable we compute the error in E that is caused by taking $\bar{p}$ as a constant outside of the integration, though it is found that the discrepancy is of the order of 1% or less.

¹The data was provided by the Danish Wind Industry Association, <http://www.windpower.org/en/tour/wres/pow/index.htm>

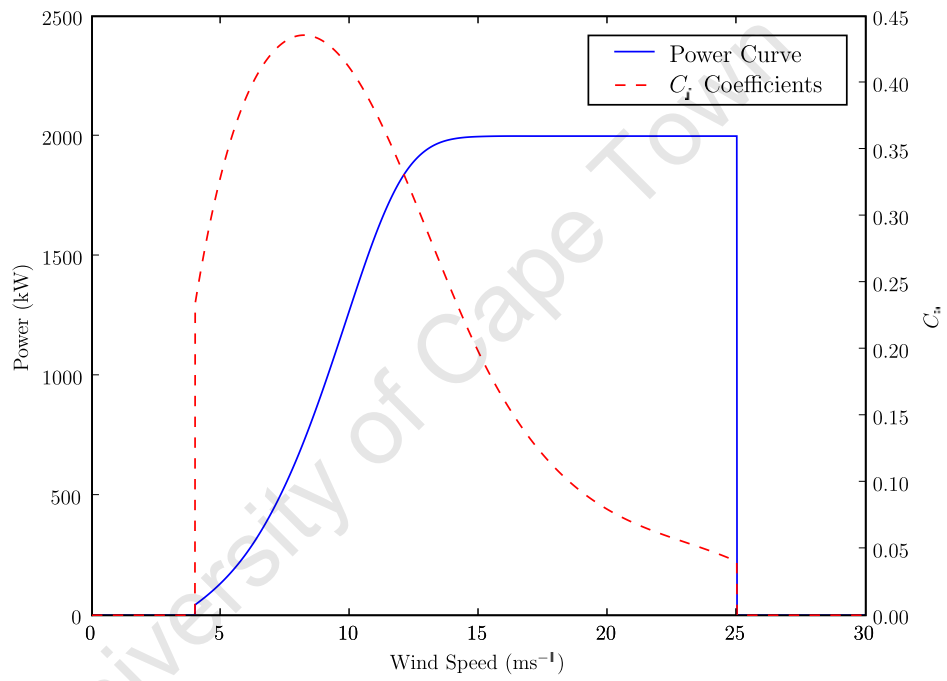


Figure 5.13: Power Curve and C_p values for a Vestas V80 2000/80 2MW turbine. The P values apply for $\rho = 1.225 \text{ kgm}^{-3}$. Note that the cut-in and cut-out wind speeds are 4.0 ms^{-1} and 25.0 ms^{-1} respectively

The final horizontal interpolation is carried out onto a lat/lon grid that measures 1593 by 1427 cells using increments of 0.01060° longitude and 0.0090° latitude and starting at the southeast corner $35^\circ\text{S } 16.1^\circ\text{E}$. At 32°S where most of the wind resource is found this corresponds to exactly 1km by 1km but even at 22°S , the northernmost part of the country, the east-west spacing is 1.09km, a very reasonable approximation.

In terms of infrastructure datasets we use vector layers from Surveys and Mapping in standard GIS (Geographical Information System) format which have been provided based on the 1:50000 maps². Together with some 1:500000 scale vector files this provides an adequate foundation in terms of roads, transmission lines, borders and coastlines.

The tiled inputs are first concatenated and filtered according to the above requirements, yielding a high detail nationwide result as shown in Figure 5.14 to the right. The processed vector layers are then converted onto the same raster as E and C_f above and finally transformed into distance maps in degrees.

5.4.2 Total Potential Calculation

To determine South Africa's total wind power potential along the lines of what was discussed above we have to rely on a set of assumptions: proximity to roads and transmission lines in km, number of 2MW turbines that can be expected to fit on the area covered by each 1km^2 grid cell, which hub height can be utilised and most importantly some sort of wind threshold above which electricity generation is deemed economically feasible.

For the latter criterion we base our evaluation on capacity factor C_f because it is an easily understood quantity and may be readily compared with *existing* (not potential) wind farms overseas. The International Energy Agency(IEA) reported in 2004 installed wind power capacities and total wind generation for all its member states [Malyshev et al. 2004]. These figures along with the ratio of actual to maximum theoretical generation (installed capacity * 365 days * 24 hours), i.e. per definition of C_f , are outlined in Table 5.3 on page 141.

It has been noted before that these statistics are likely a lower bound of reality because of the way that Table 5.3 disregards the fact that some new capacity might have only come online a month or two before year end, yet they are assumed to have contributed for the electricity generation for the entire

²Chief Directorate: Surveys and Mapping, Rhodes Avenue, Mowbray, 7700 Cape Town. Obtained July 2008

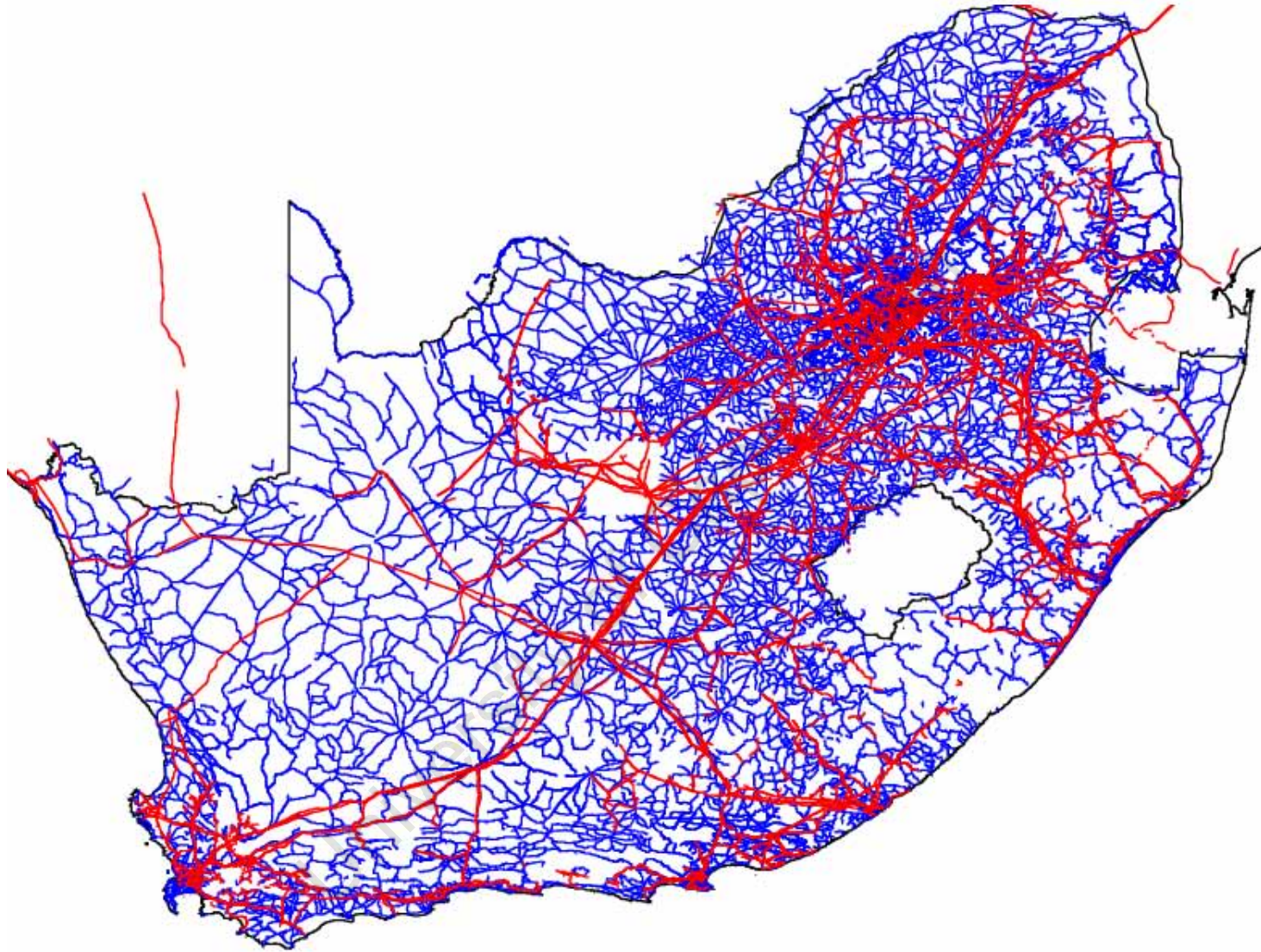


Figure 5.14: Secondary roads (blue) and power transmission network (red) of South Africa

Country	Installed Capacity (MW)	Annual Generation (GWh)	Capacity Factor
Germany	8700	10700	14.04%
US	3864	6800	20.09%
Spain	3200	7000	24.97%
Denmark	2556	4300	19.20%
Italy	664	1179	20.27%
Netherlands	490	820	19.10%
UK	427	970	25.93%
Sweden	290	481	18.93%
Portugal*	227	691	34.75%
Japan	175	252	16.44%
Ireland*	166	465	31.98%
Canada	131	375	32.68%
France	81	125	17.62%
Australia	76	210	31.54%
Average			23.40%

Table 5.3: Installed wind power capacity, annual generation and capacity factors for the 14 IEA countries with the highest wind energy penetration [Malyshev et al. 2004]. Countries with marked with * represent national 2003 figures, the others are for the year 2001

year [Pitt et al. 2005]. However it is still important to point out that the top six IEA countries all have capacity factors of less than 25% and that the global average for 2001 lies at around 23.4%. Also, a source independent of the IEA reported figures of 21% for Denmark, 16% for Germany and 18% for Sweden in the second quarter of 2006 which is in broad agreement with the above [Canter 2006].

For South Africa the data is very sparse owing to the lack of operational wind farms but for Eskom's Klipheuwel farm a C_f of 10% for winter and up to 35% for summer, i.e. an annual average of around 22.5% has been published [Maroga 2007]. Various other projects, including the Darling wind farm, have been predicted to lie at around 30% [Langhans 2003].

Based on these considerations we create three wind power penetration scenarios - an optimistic, realistic and pessimistic one loosely based on economic factors and probability of materialising. Each of them uses different assumptions which are summarised in Table 5.4 to the right; the minimum acceptable capacity factors are 35%, 30% and 25% respectively which are still fairly conservative compared to the International norm described above. The different

Scenario	Maximum Distance to Nearest Road	Maximum Distance to Nearest Power Line	Employed Hub Height	Minimum Capacity Factor
optimistic	$0.05^\circ (\sim 5\text{km})$	$0.05^\circ (\sim 5\text{km})$	100m	25%
realistic	$0.04^\circ (\sim 4\text{km})$	$0.04^\circ (\sim 4\text{km})$	80m	30%
pessimistic	$0.03^\circ (\sim 3\text{km})$	$0.03^\circ (\sim 3\text{km})$	60m	35%

Table 5.4: Assumptions for three wind power penetration scenarios of South Africa.

hub heights are important too as the higher off the ground the rotor is mounted the more energy it may capture according to equation 1.8 on page 6. However, higher towers are more costly to manufacture, transport and erect and thus might not make sense in all cases.

For each scenario we calculate the total number of grid cells that satisfy the respective criteria. Then we work out how many turbines could fit into a single pixel. Given a perfect site one could possibly accommodate up to 6 turbines of rotor diameter $D = 80\text{m}$ per km^2 using industry accepted minimum spacings to avoid excessive wake shadow effects [Burton 2001]. But so far we have not yet accounted for important land use constraints such as environmental sensitivity, ruggedness of terrain, proximity to urban or residential areas etc. Hence we conservatively assume a turbine density of unity per km^2 .

Summing up the resulting energy yields E per grid cell provides us with the following total potential wind power generation for South Africa in terms of TWh and percentages of total 2007 electricity sales of 229.6 GWh [Maroga 2007]:

- optimistic - 157.18 TWh (68.46%)
- **realistic - 80.54 TWh (35.08%)**
- pessimistic - 20.06 TWh (8.74%)

As the reader can see the results vary drastically. The pessimistic scenario is surprisingly low at just 20.06 TWh per annum which this is due to the fact that there are many more sites in the 30-35% capacity factor range than locations with a C_f of greater than 35%. In the course of the development of a wind industry in South Africa these sites are likely to get saturated first as economic factors would naturally dictate.

Just by going to the slightly lower but still internationally competitive C_f bracket of greater than 30%, allowing for an extra kilometre to nearest infrastructure and moving up from 60m to 80m hub height our total exploitable wind resource increases fourfold. Such a substantial difference can also be attributed to the significantly higher winds achieved at 80m. This scenario is judged by the author to be the most realistic and likely one for South Africa in the long term and in fact lies halfway between the most recent estimates and thus corroborates them [Banks and Schäffler 2006].

Going one step further to a minimum C_f of 25%, a maximum transmission line and roads distance of 5km and moving up to a hub height of 100m in turn only doubles the available wind generation potential. It is unlikely that these locations will be utilised for wind farms in the near future unless there are significant regulatory incentives in place that provide for a premium at lower wind speed sites such as is the case in Germany [BTM Consult 2007].

On a final note it must be stated that no country in the world has yet managed to integrate more than 20% of wind capacity into their national transmission grids owing to the variable nature of the atmosphere in general and the resulting supply management problems. Thus even the realistic scenario is likely to be constrained by the capabilities of Eskom's infrastructure until significant advancements have been made in terms of electricity storage and dispatch. However, what we can say with great certainty based on the figures above is that until about 2050 the total available wind resource of South Africa will *not* be a limiting factor to the development of renewable electricity generation.

University of Cape Town

Chapter 6

Conclusion

6.1 Fulfilment of Objectives

The very first section on page 1 of this thesis defined the primary and secondary objectives of this thesis as well as respective sub-goals. Let us first review our main findings with respect to what extent these *were* fulfilled and what the main findings are.

Section 6.2 on page 148 then deals with the ways in which they were *not* satisfied, discusses the results (or lack thereof) in more detail and suggests how the encountered issues could be addressed in future.

6.1.1 Primary Objective

We primarily set out to explore MM5's merits for developing a high resolution wind climatology.

To satisfy the first sub-objective we have found that it *is* possible to generate a valid climatology using an integration period of one year only by using the temporal sub-sampling technique developed in chapter 3. Residual biases are fairly small (-0.2ms^{-1}) and computational resource requirements are thereby significantly reduced. There are some caveats though which somewhat diminish the value of this method in certain aspects. These will be elaborated upon in subsection 6.2.1.

Regarding the remaining two sub-tasks of this main goal we discovered at a resolution of 18km that:

- the main sensitivity of the model with respect to faithfully reproducing

near surface wind fields lies in the parametrisation of the planetary boundary layer and the cumulus convection processes. The most important driver in the variability of the wind turns out to be the diurnal cycle which the Pleim Xiu/ACM LSM/PBL combination with the Kain Fritsch II cumulus scheme are most skilled at for South African latitudes, at least for coastal regions

- observational and grid nudging has a much smaller effect on the accuracy of the model output though the inclusion of the former and exclusion of the latter tends to yield better results on average. NCEP/NCAR Reanalysis was used for the grid nudging and its inherent errors over the South African domain are suspected to be responsible for rather poor FDDA results
- with the above options enabled MM5 can achieve a fairly close agreement with observations to within $18\% \pm 20\%$ (68% confidence interval) in terms of the amplitude and 1h with respect to the timing of the daily cycle, at least at coastal sites. At 18km this is a significant result as other studies have found a useful cycle to appear only at much higher resolutions of 5km [Källstrand et al. 2000]
- while a diurnal oscillation is still reproduced for inland locations the model has difficulties to match the natural evolution thereof and often produces unrealistic features. This is likely to be caused by the inability of the parametrisations to capture the features inherent in the dominant convective processes there
- the chosen parametrisations on average exhibit a positive night time bias which cause many of the other bias and cycle mismatches. This could be attributed to a problematic treatment of radiative forcing with the associated errors in nocturnal stability regimes
- overall (coastal and inland) average wind speeds are predicted within $0.2 \pm 0.7 \text{ms}^{-1}$ (68% confidence interval)
- overall wind directions are within about 3.5% of frequency of occurrence per 30° wind sector

Coupled with the fact that an RCM always delivers spatially continuous output we may conclude that with certain limits (defined above and below) MM5 is a valuable tool to develop a detailed wind climatology of South Africa. For

coastal regions it is able to simulate the fundamental processes governing the near-surface wind patterns to within the accuracy that one might expect for the 18km resolution.

However, there are numerous other shortcomings that we have so far only touched on throughout the various chapters, especially regarding the model's inabilities over inland regions. There are also a number of factors which have limited the extent to which we were able to achieve our primary objective. These need to be put in perspective, treated in more detail and discussed together with the above findings in a critical light. This is done in subsections 6.2.2 to 6.2.4 on pages 149–154.

6.1.2 Secondary Objective

The secondary purpose of mapping and quantifying South Africa's wind resources has been achieved by providing the following:

- a national wind dataset at 18x18km resolution and arbitrary heights above ground (pre-interpolated to 60m, 80m and 100m). Its accuracy, strengths and weaknesses have been outlined in detail in section 5.2 but will be revisited in due course
- maps of annual and seasonal wind speed climatology which have not only revealed previously unknown areas of high wind resources inland but also the fact that the wind resource is higher than expected. We further discovered that the strongest winds occur in winter when electricity demand is highest. This satisfies the first sub-objective
- a realistic quantification of wind power's total electricity generation potential to be 81TWh or 35% of 2007's total consumption based purely on sites with capacity factors higher than 30%, 80m hub heights and 4km maximum distance to suitable roads and transmission lines. Despite neglecting current transmission grid constraints this fulfils the second sub-goal and while this knowledge is not entirely new, it reaffirms recent independent research findings which put the total figure at 60 and 106TWh respectively [Banks and Schäffler 2006]

In the light of the secondary objective we have thus made significant scientific contributions in terms of both where the good wind resources are in South Africa and how large the overall potential is. In terms of our primary goal this

further underlines the value that MM5 can provide in the context of a wind atlas because of the way that the nature of the regional climate model and the utilised resolution have made such analyses possible in the first place.

Of course the above points are attached to concerns in their own right. The caveats of the wind atlas are noted in section 6.2.5 on page 154 while the problems underlying the total resource estimate are dealt with in subsection 6.2.6 on page 155, together with more detailed discussions of the findings.

6.2 Discussion, Shortcomings and Future Research

This thesis does suffer from various shortcomings which somewhat limit the degree to which we were able to serve the purposes that we originally committed ourselves to. Hence we need to state to what extent the originally stated objectives were *not* fulfilled and what may be done in future to address this. This section also serves to discuss the above findings in some more detail than has been done up to now.

6.2.1 Choice of Integration Period

The time sub-sampling algorithm to yield an “optimal” 365 day period that is still representative of the climatology does not come without trade-offs. “Optimal” does not imply perfect but rather as good as we have been able to achieve, meaning there is some residual loss in variability that we have to accept and which may add some uncertainty to the wind maps that were produced. Also as we mentioned in section 3.5 on page 54 it clearly prejudices the validity of seasonal statistics since only the match in variance of the entire year was optimised, not the climatology of each individual month. This is the reason no detailed analysis of the seasonal cycle (like the diurnal one) was ever carried out.

There are various approaches one might choose to take in future to mitigate this problem:

- reduce the model resolution
- carry out the integration for further, non-overlapping 365 day subsets that similarly come close to the long term reference (highly recommended)
- analogously find an optimal 2 year sub-period which is likely to get much closer in terms of being able to represent the climatology

- take into account four years of new data which has become available since this calculation was carried out in order to make the results more robust
- alter the algorithm to pick out four “ideal” seasons or even twelve optimal months within the reference period and run the model for each of them in turn
- with the right budget and storage capacities simply integrate the model over ten or even twenty years to avoid this pitfall altogether

The last possibility clearly stands on its own and could be considered more of a “brute force” approach. The others present more of a scientific challenge and may be combined in creative ways to come up with a better solution. Even though it would double computational requirements the second alternative is attractive because of the way that a continuous timeseries is preserved.

Breaking up the integration period into months or seasons though is equally promising and probably the better way forward as it would allow for a much more detailed analysis of wind patterns at these time scales without the disproportionate uncertainties caused by the lack of captured variability. Annual statistics would still be accurate due to the fact that all of its constituencies are close to the climatology.

Finally we must point out that the frequency of synoptic events which the method is based on is not static. Climate change indications are that they will alter in future which adds even more uncertainty to what we have discussed so far.

However, this aspect was left out on purpose because one has to evaluate what the climate is right now (i.e. the *status quo*) before one may even begin to investigate how it is likely to be perturbed in decades to come. Only once current wind resources are fully understood may one embark on research that aims to make certain projections into the future.

6.2.2 Model Validation

The wind atlas accuracy assessment of section 5.2 made it clear that the overall error statistics could have benefited from a sample much larger than 17 weather stations. While the standards for anemometer consideration should not be lowered, the sourcing of data outside of SAWS and subsequent utilisation in model evaluation might prove useful in enhancing the robustness of the various performance metrics. In addition a detailed investigation into the surroundings in

terms of buildings and other obstacles, topography and roughness influences in the near vicinity of each station could prove beneficial to determine which biases can be explained by microscale effects and which are attributable to deficiencies in the model itself.

More significantly though we need to reiterate that the model was validated at *10m above ground only*. This was mostly due to the lack of publicly available measurements at greater heights but considering SODAR/LIDAR, rawinsonde or similar sounding data would provide us with an understanding of the model's ability to reproduce vertical profiles, at least in the first 150m or so which is of importance to wind power.

The parametrisation of the planetary boundary layer complete with local and non-local mixing schemes as well as the formulation of convective processes almost exclusively govern the model's skill in this respect. With different vertical momentum fluxes higher up the horizontal wind components may also change drastically.

It might well be that our chosen Atmospheric Convective Model (ACM) that is coupled with the Pleim Xiu land surface implementation has much greater shortcomings higher above ground where the wind turbine rotors are situated. While the literature provides some reassurance by suggesting the opposite [Tonnesen et al. 2004; Jia and Morris 2003], this has never been verified or explored before in the case of South African climate. This is especially important for the inland regions which as we have seen present a challenge to MM5's wind simulation abilities.

However, we were unable to investigate any of this which restricts the way we were able to address our primary objective and in particular its third sub-goal. Therefore future studies should emphasise conducting detailed comparisons with observations between about 40m and 200m above ground and perhaps even higher. This should aid greatly in order to dissect the employed regional climate model's behaviour with respect to stability and winds aloft. Furthermore including such soundings in observational nudging could yield valuable performance increases.

We have noted in subsection 5.2.4 on page 126 that with increased height above ground the flow disturbances caused by nearby obstacles and discontinuous roughness changes decrease. Hence many of the measurement inherent biases which are due to phenomena the model may never resolve become much smaller. As such comparisons above 10m would add much more value to a detailed exploration of the model's strengths and weaknesses, both in terms of its

overall skill and its different parametrisations (cumulus and PBL).

In terms of evaluating the significance in skill improvement or deterioration as measured by our various validation statistics we have only employed subjective judgements which represents another shortcoming of this research, albeit small. In order to investigate whether a given change in model parametrisation for example yields a statistically significant improvement or to generate uncertainty bounds we could in future rely on bootstrapping techniques which have been successfully used in other modelling studies [Schoof and Pryor 2006].

It should also be pointed out that in retrospect, directional averaging as performed by equation 1.26 on page 12 can easily be improved upon by calculating the resultant vector of $(\bar{u}, \bar{v})$ which is arguably better suited for wind power applications because it effectively weights each direction by its wind speed. Following on from this is the concept of wind *persistence*, as defined by the magnitude of the average vector divided by the mean wind speed, i.e. $\frac{\sqrt{\bar{u}^2 + \bar{v}^2}}{\bar{V}}$. Also useful could be a performance metric around the *mean vector error*, defined in terms of a bias in $(\bar{u}, \bar{v})$. Future treatments of this subject should certainly consider these aspects.

Finally we have seen in places that considering the daily wind cycle on speeds alone (V_{24}) can sometimes be misleading as a larger mismatch in the underlying u and v components may lead to lower discrepancies on V_{24} for the wrong reasons. Hence it would be beneficial to define and apply more robust diurnal cycle error metrics based on two wind components and not speed alone.

6.2.3 Model Configuration

It was noted in subsection 5.2.2 on page 115 that the Pleim Xiu/ACM LSM/PBL parametrisation coupled with the Kain-Fritsch II scheme provided a fairly good match of diurnal cycles at coastal locations. As we witnessed this was key to getting overall wind speed averages and standard deviations correct as it accounts for most of the variance in the flow. To within station specific biases caused by sub-grid scale effects we may state that the model is quite capable to simulate reality.

However, the evaluation of most inland locations have alerted us to the fact that the model has significant difficulties here to reproduce average daily fluctuations and thus other vital wind statistics. In the discussion of subsection 5.2.2 we realised that this was likely not caused by lack of topographical detail but rather by fundamental shortcomings in the employed PBL and/or cumulus

sub-models over this very different climatic zone.

The highveld is characterised by important convective processes which drive other patterns such as rainfall and clearly have an impact on winds too. The utilised cumulus parametrisation likely has a very strong influence on model skill and the same may be said about PBL implementations. We did not investigate their respective roles in any way as the sensitivity runs of chapter 4 only focused on coastal stations for various reasons given in subsection 4.3.1 on page 66.

This represents a significant shortcoming of this research, especially in the light of our wind atlas implying that a lot of the country's wind resources lie in the interior. Future work should clearly focus on this area, ideally around those sites that this doctorate has identified as having strong winds on average.

But before one would dive into the detailed model behaviour over such areas one would first have to satisfy two important prerequisites:

1. Having reliable measurements of wind vectors available at *several* heights above ground
2. Understanding what natural processes drive the diurnal wind variations in the interior in the first place

The first one has been discussed in some detail in subsection 6.2.2 on page 149 but the second one we have not yet touched on at all. Back in section 5.2.2 we noted how a well defined daily wind speed cycle was observed at 10m even though neither land/sea breezes nor slope winds could be responsible for these (though the periodic growth and decay of the convective PBL might be).

So there must be a third type of circulation that arises from convective processes but what exactly this constitutes is unknown at this stage. Considering both V and θ (or u and v for that matter) as noted in the previous subsection could aid in such an investigation.

In terms of grid nudging we have further discovered that it actually deteriorates model skill over the South African domain. We have attributed this to the errors present in the NCEP/NCAR Reanalysis values themselves and not to problems with nudging per se seeing that the vast majority of studies in the US recommend grid FDDA. The reanalysis issues in turn are likely to be caused by the low density of surface and upper air observations in Southern Africa and the large data voids over our surrounding oceans which drive a lot of our climate.

This raises a different problem as we used the very same data to drive MM5 on the boundaries, propagating much of these biases right down to the 18km

nest which we based our entire analysis on. Therefore it might be worth looking into using other global climate datasets such as ECMWF or perhaps the MRF forecast data to ascertain which one of them provides the best results, both with and without grid nudging. Most importantly, perhaps, the effects of using the updated NCEP Reanalysis II should be investigated, especially seeing that it claims to have fixed many issues in the subtropics of the Southern Hemisphere where South Africa is located [Kanamitsu et al. 2002].

Then there is the question of diurnal temperature cycles driving wind patterns. Theoretically all daily winds (land/sea breezes, slope winds etc.) are strongly influenced by temperature differentials even though in practice MM5 has been seen to decouple the two to a certain extent [Zhang and Zheng 2004]. In any event surface characteristics other than topography and roughness length such as albedo and thermal inertia may become quite important to the development of the flow.

Again, this is an aspect that has not been considered in this research at all. We recommend that future studies should investigate this in some depth, not only with respect to the validity of the values assigned to the various land use categories but also in terms of how they are handled in the various land surface parametrisations of the regional climate model.

We have investigated only two cumulus configurations other than the standard Grell scheme but have seen significant improvements with the two more complex ones (Betts Miller and Kain Fritsch II). Hence follow up research could look into more such combinations, especially those in which the convection is treated the same in the parent domain and the nest for the sake of physical consistency.

Lastly we noted on a number of occasions that on average MM5 exhibits a positive night time bias. This was discovered during the evaluation of both individual station and spatially averaged diurnal cycles in subsection 5.2.2 on page 115. We have not offered much in terms of explanations in this regard but at least we acknowledged that this is indeed problematic. As other authors have pointed out in the past, however, this seems to be a recurring theme in modern boundary layer meteorology: no current day PBL scheme is capable to accurately reproduce stably stratified turbulence that often occurs at night [Zhang et al. 2001].

In this research we have merely chosen between bad and worse and therefore this warrants a future in-depth look at the model's parametrisation of the nocturnal PBL which is probably tightly coupled with the developments of 24h

temperature fluctuations alluded to above. Advancements on this aspect could aid greatly in improving the skill of the RCM.

6.2.4 Model Resolution

It must be reiterated that many of South Africa's climatologically important mountain chains are not at all resolved at 18km, especially those in the Western Cape where a large proportion of the wind resource is located. This would, for example, significantly increase average wind speed biases on either side of a mountain should the predominant geostrophic wind direction be perpendicular to it. An increase of resolution to 6km or even higher in future research would help alleviate this problem.

Simulations with much lower grid sizes are recommended in any event in order to reduce the problems associated with area average vs point measurements explained in subsection 4.3.5 on page 74. More sophisticated and up-to-date models like WRF, MM5's successor, might be able to significantly improve accuracy while at the same time revealing even more detail about the spatial distribution of the country's wind resource. A detailed comparison of its PBL and cumulus schemes with respect to what we have done would greatly aid in understanding the practical benefits and scientific advancements, if any, the transition to WRF has brought about.

Such a task is actually simplified now that our research already provides 18km input on the boundaries which saves the time, CPU and storage intensive step of successively nesting inside global and very coarse reanalysis data such as NCEP/NCAR. In practice this would probably entail carrying out model runs for each of the Cape provinces and perhaps Kwazulu-Natal in turn and thereby completely skipping the regions which are not interesting according to the findings of this work.

6.2.5 Wind Atlas Data

The model output, even when used to nest even higher resolutions, will likely never be accurate enough to achieve the low error margins required by a reliable annual energy production projection of a commercial wind farm. That is due to the errors which will always be present in the boundary conditions, no matter how sophisticated the model itself may be.

This is an important fundamental limitation of the mesoscale modelling approach and limits the value that may be derived from a methodology such as the

one applied in this thesis. What MM5 *can* provide, however, is a very useful *first guess* or *rough estimate* of the available wind resources through various statistics for any given region of South Africa. This is where the value of the results really lies, notwithstanding of course the various uncertainties and problems that were elaborated above. The fact that we did to a certain extent investigate what the strengths and weaknesses of the three dimensional timeseries are provides us with a much more solid basis than wind energy consultancies might be able to achieve¹.

Hence we point out that the dataset underlying the wind atlas should provide a more than adequate foundation for wind farm pre-feasibility studies of various locations around the country. We can then take informed decisions on where the highest average wind speeds coupled with good access to infrastructure are likely to be situated.

This in turn determines where to place future site-specific measurements to reduce uncertainties to an economically acceptable level and thus ultimately where to erect wind turbines destined to power the future of this country. The higher the resolution of the model input, the easier it should be to locate such points. Especially the directional information contained in our results are invaluable in the search for suitably oriented hills and the detailed mounting layout of anemometers.

6.2.6 Wind Potential Estimate

We duly note the large differences in the optimistic, pessimistic and realistic scenarios that were studied in section 5.4 on page 135. While plausible explanations for the significant spread have been provided there already, an even more detailed total wind generation potential estimate should be derived in future which could minimise uncertainties more than we have been able to achieve.

Such an estimate should pay closer attention to other land use constraints such as population density, location of nature reserves and other ecologically sensitive sites as well as exclusion of very rugged and mountainous terrain which clearly cannot accommodate any substantial numbers of wind turbines. These are some important factors that we did not incorporate in our somewhat simplified calculation in section 5.4 but rather estimated by assuming a wind turbine density of unity for each 1km^2 grid cell.

¹See for example the lack of detailed validation and explanation of commercial mesoscale wind datasets such as <http://firstlook.3tiergroup.com>

It should also make use of robust economic considerations such as cost per kWh of electricity generated instead of capacity factors which are highly turbine model and rotor specific. Incorporating further pricing projections of fossil energy carriers that South Africa has traditionally been relying upon as well as full cost based accounting of various alternatives would add further value to such an undertaking.

Furthermore, what was not pointed out in section 5.4 is that the air density calculation relied on modelled pressure, mixing ratio and temperature values that were not validated against observations at all. However, as we have seen the variability of ρ is relatively low and therefore the error that we incurred in the process is likely to be fairly low.

We also acknowledge that neither offshore wind capacity nor turbines in other countries of the Southern African Development Community have been taken into account in this study. Namibia for example is currently investing in various wind farms to reduce its dependence on imported power from Eskom so an in-depth look at the wind resources here might prove to be quite useful.

European markets are increasingly pushing for wind farms out at sea in response to space constraints and land availability. Even though this will likely remain a negligible concern for South Africa in the foreseeable future, our offshore potential is set to promise much higher energy yields than the 35% of demand estimated above given the vast coastline that this country enjoys.

6.3 Concluding Remarks

Even though the following thoughts are only of peripheral interest to the primary objective of this work the author would like to take this opportunity to relate its significance to the larger framework of the future economic and social development of South Africa.

Judging from the available wind resource compared to developed countries overseas with a much larger and more mature wind power industry it seems as though South Africa is poised for significant growth in this sector. Considering our country's current electricity shortages which are projected to last several years into the future coupled with the very steep increases in fossil fuels such as oil and coal in recent times the future of wind energy looks increasingly bright, even in a country that still relies on coal for 88% of its electricity and supposedly has plenty of it left to mine. This does not even mention the obligations of

climate change mitigation that we must commit ourselves to in the next decades to come that will push us even stronger in the direction of renewables.

While other “green” electricity generation technologies such as solar and wave are on the outset very promising wind is currently still the most cost-effective way to generate electricity in bulk from the sun. But taking a long-term view to 2050 all renewables will play hand in hand to compensate each other’s variability in such a way that coupled with various storage technologies a grid integration of up to 100% will be achievable by then.

The author will help realise this vision and this PhD is but a minor stepping stone in the roadmap to achieve this ambitious goal.

University of Cape Town

University of Cape Town

Appendix A

Combining MM5 Nudging Techniques

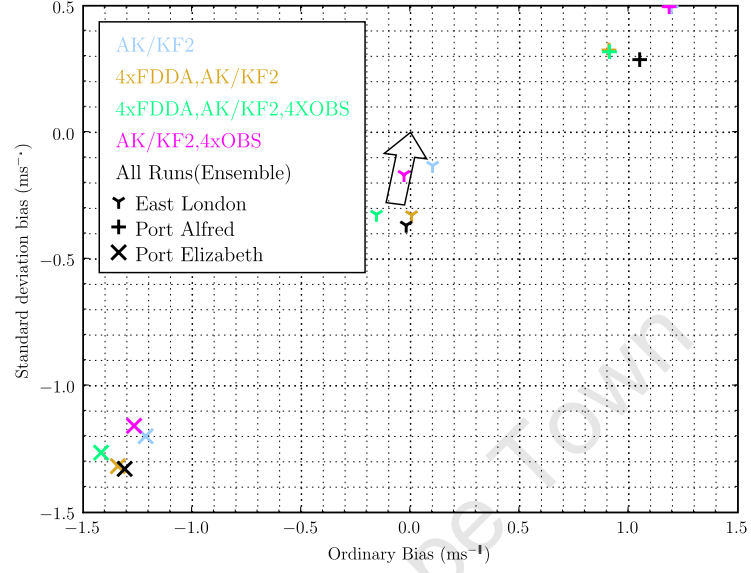
So far we have settled on PX/ACM as the preferred LSM/PBL scheme and have concluded that AK/KF2 as well as observational nudging are desirable, with grid FDDA being of doubtful use. However, the latter three have so far been used only in isolation and it is not necessarily given that they work well together too as they are highly non-linear in nature.

Therefore in this section we deal with the four combinations of Kain Fritsch II, quad strength grid and station FDDA as outlined in Table A.1 on page 161. The reference simulation is now taken to be the standalone PX/ACM with AK/KF2 concluded in subsection 4.5.4 on page 104 to be the one with the clearest and most consistent benefits.

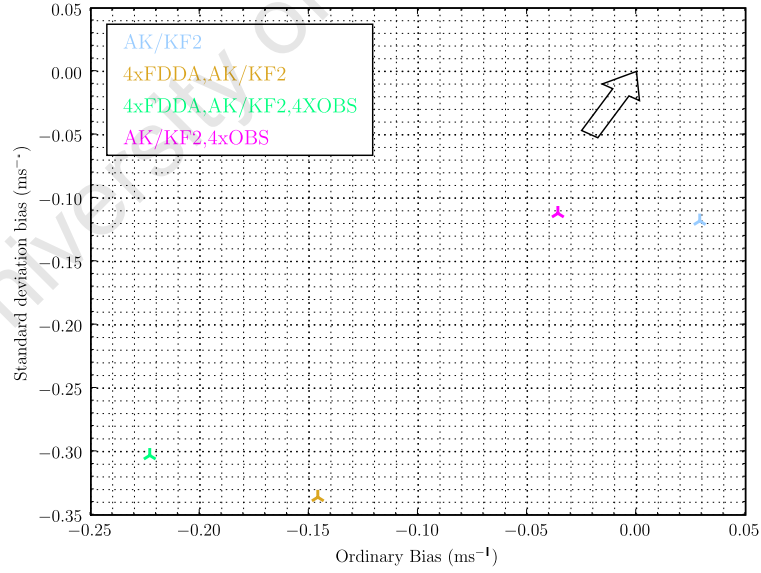
This time round the analysis is fairly condensed as individual diurnal cycles are excluded as the reader by now should have a fair idea of what the model can and cannot achieve in this respect.

A.1 Average Wind Speed and Standard Deviation

Wind speed mean b_V and standard deviation bias c_V are given in Figure A.1 to the right. The addition of grid FDDA as before consistently compresses the histogram by decreasing both $\overline{V}$ and σ across all stations (Figure A.1(a)) while



(a) Individual stations



(b) Spatial average

Figure A.1: Goal plot of standard deviation bias c_V vs ordinary bias b_V of wind speed for all four combination runs. The blue symbol marks the PX/ACM with AK/KF2 reference whose performance is to be improved upon

Run Description	Short run name	Grid FDDA	Obs Nudging
Anthes Kuo/Kain Fritsch II	AK/KF2	none	none
Quad strength grid FDDA with AK/KF2	4xFDDA, AK/KF2	0.0010 x $u/v/T/q$	none
Quad strength observational nudging with AK/KF2	AK/KF2, 4xOBS	none	0.0016 x V
Quad strength grid and station FDDA with AK/KF2	4xFDDA, AK/KF2, 4XOBS	0.0010 x $u/v/T/q$	0.0016 x V

Table A.1: Configuration of the four optimisation runs. All of them use Pleim Xiu/ACM and the AK/KF2 cumulus scheme and differ only in their use of FDDA. The respective FDDA coefficients are given in the last two columns - GV, GT and GQ in second last and GIV in last column. The first entry is the reference simulation and identical to the one in section 4.5 on page 90

observational nudging slightly increases σ but decreases the magnitude of the flow by $0.1-0.15\text{ms}^{-1}$ for PE and EL respectively.

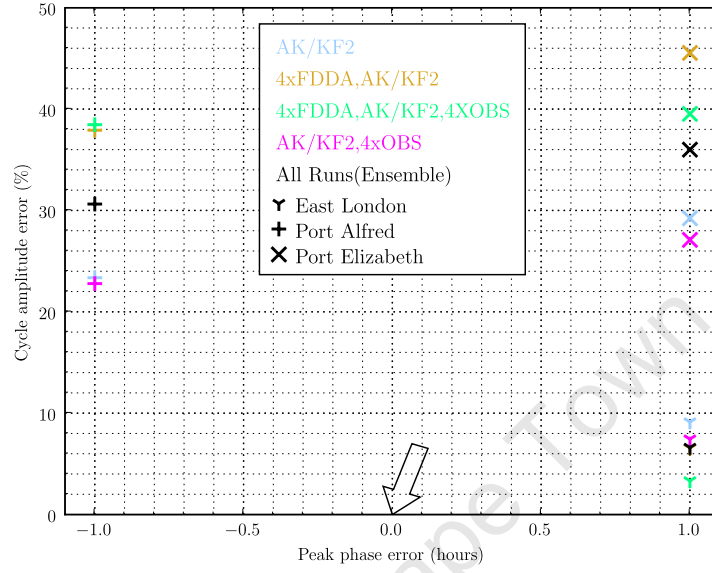
The spatially averaged points (Figure A.1(b)) provide us with a very similar picture. Here both of the runs without relaxation towards NCEP/NCAR Reanalysis (AK/KF2 and AK/KF2 with 4xOBS) are on par in terms of their performance, their distance to the origin is nearly identical.

This is completely in line with our previous analysis and while it does not change any of the conclusions already made it implies that combining model parametrisations that we have found to improve skill individually introduces no regressions.

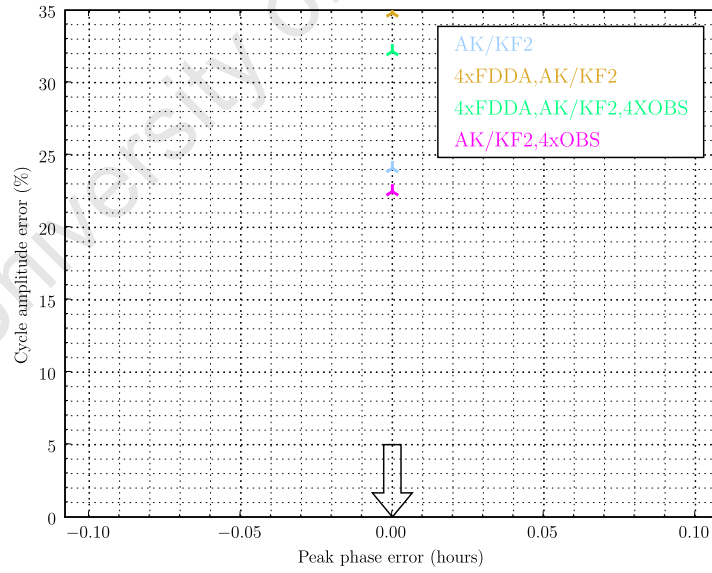
A.2 Daily and Seasonal Cycles

In terms of the diurnal cycle the model's skill also gets augmented as one would expect by these combinations. Figure A.2(b) to the right clearly shows that while the spatial average peak phase error is always 0h, the strength of the cycle certainly varies appreciably. The reader may see that extra observational nudging brings the amplitude closer to reality by 2-3% while enabling grid FDDA, as before, causes a deterioration of the CAE of about 10%.

This trend is even stronger when looking at the individual station results of Figure A.2(a) where the difference lies in the region of 15% for Port Alfred



(a) Individual stations



(b) Spatial Average

Figure A.2: Goal plot of daily cycle amplitude error CAE vs daily peak phase error PPE of wind speed for all four combination runs

and Port Elizabeth, however, for East London the application of NCEP/NCAR relaxation slightly enhances the 24h periodicity.

On average we may then conclude that in terms of the daily cycle the AK/KF2 with 4xOBS combination yields the highest MM5 skill.

Moving on to seasonal performance the goal plot of $MMAE$ vs $DFPD$ is given in Figure A.3 to the right. As usual the $DFPD$ does not actually provide us with more information than the CAE already does but it becomes clear from Figure A.3(a) that AK/KF2 with 4xOBS performs best at both PE and EL, merely at Port Alfred it does not do as well as the reference run.

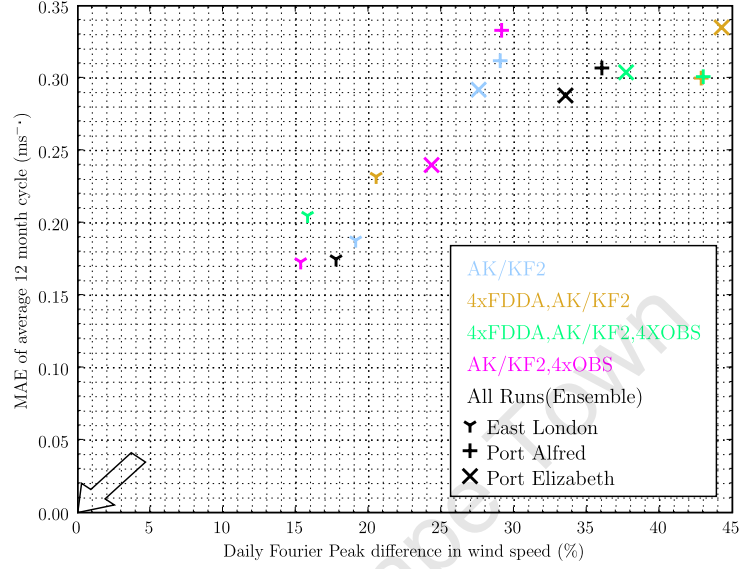
Also for the spatial average statistics of Figure A.3(b) the $MMAE$ values point towards AK/KF2 with 4xOBS coming out on top again, though by a for the $MMAE$ typically insignificant margin of 0.01ms^{-1} with respect to the reference and 0.03ms^{-1} below the worst performer, AK/KF2 with 4xFDDA. The relative rankings of the four runs are the same as for the CAE with grid nudging clearly deteriorating the model's skill.

A.3 Directional Performance

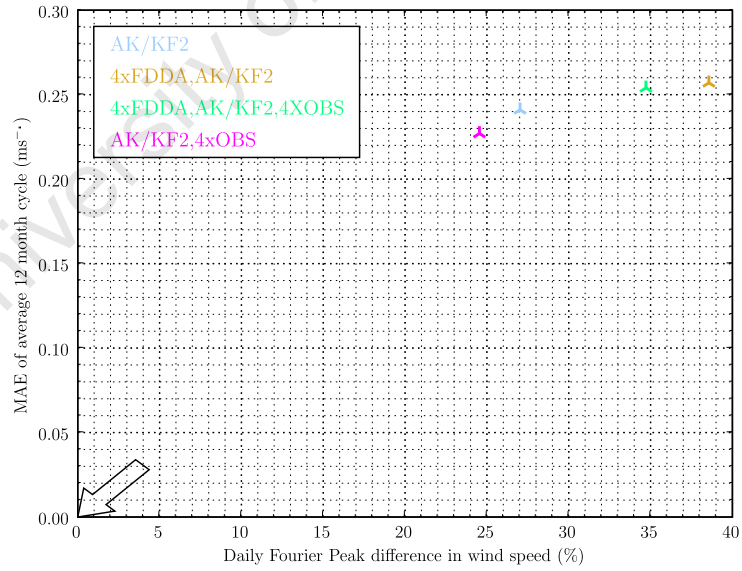
Even with the directional statistics mesh FDDA fails to exhibit any sort of potential. In the spatial mean $WMAESSD$ vs $WMAESSA$ values of Figure A.4(b) on page 165 (analogous to b_V and c_V) AK/KF2 with 4xOBS and 4xFDDA is the best combination but only improves the $WMAESSA$ by about 0.06ms^{-1} and the $WMAESSD$ by a mere 0.01ms^{-1} compared to the reference run, both magnitudes definitely being in the noise region. This is due to the somewhat conflicting individual station performance shown in Figure A.4(a), with grid nudging only showing a slight improvement for Port Alfred and a deterioration otherwise.

Regarding the sector frequencies ($WMAESF$) and power densities ($WMAEPD$) there even appears to be a contradiction between the station specific metrics of Figure A.5(a) on page 166 and the spatially averaged ones in Figure A.5(b). In the former case the addition of grid nudging again only improves the $WMAEPD$ at Port Alfred but has negative consequences for Port Elizabeth and East London. The latter scenario though implies an enhancement across the board by about 8Wm^{-2} in $WMAEPD$ and about 0.5% of $WMAESF$ by about 0.5%. The observations themselves probably speak the truth though.

Further enabling station relaxation only seems to make an appreciable dif-

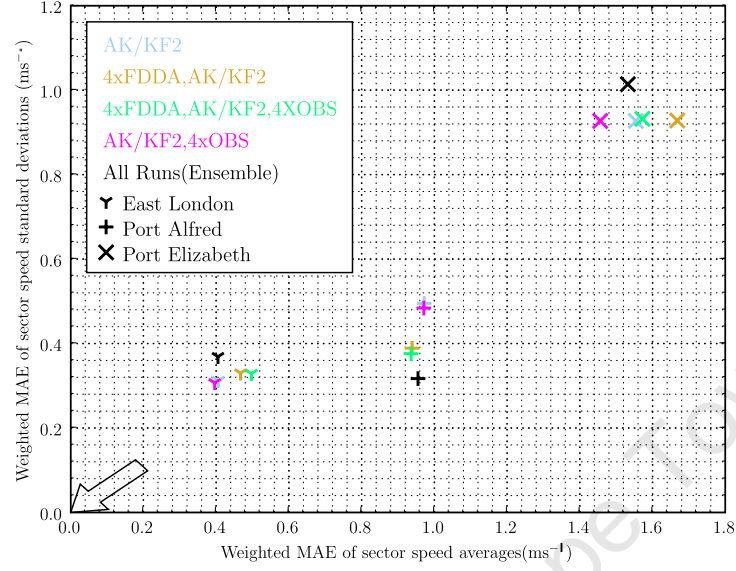


(a) Individual stations

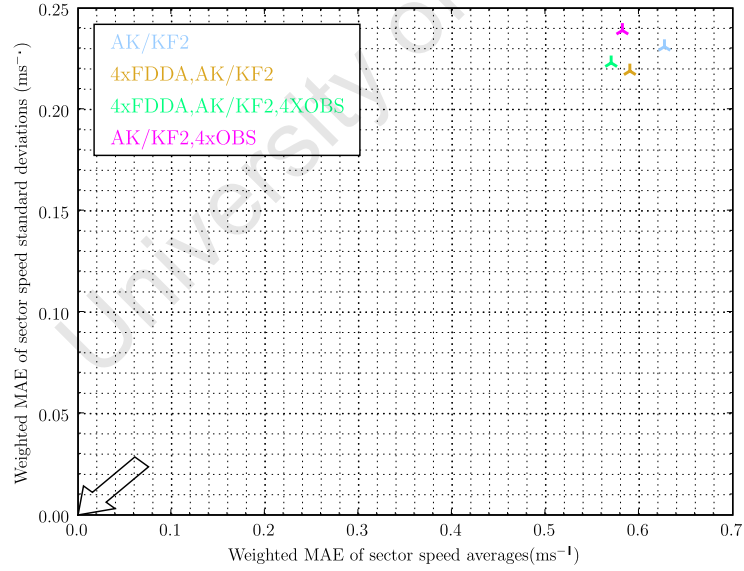


(b) Spatial average

Figure A.3: Goal plot of 12 monthly(seasonal) mean absolute error $MMAE$ vs daily Fourier peak difference $DFPD$ of wind speed for all four combination runs

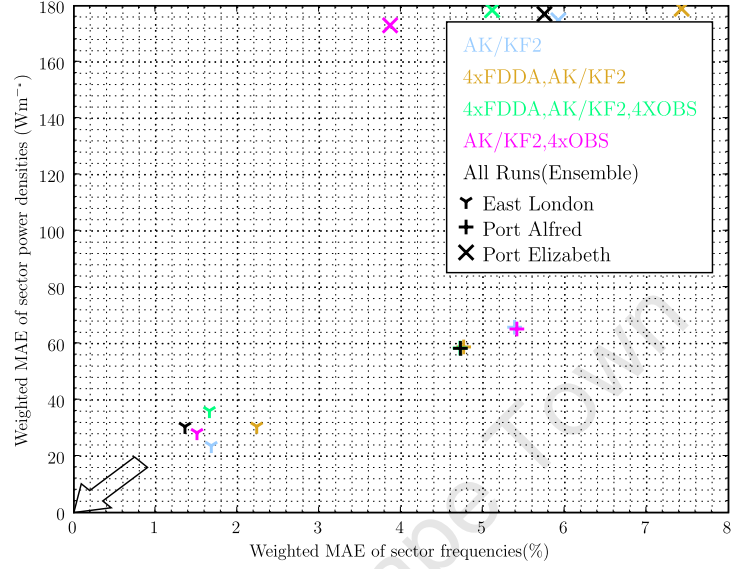


(a) Individual stations

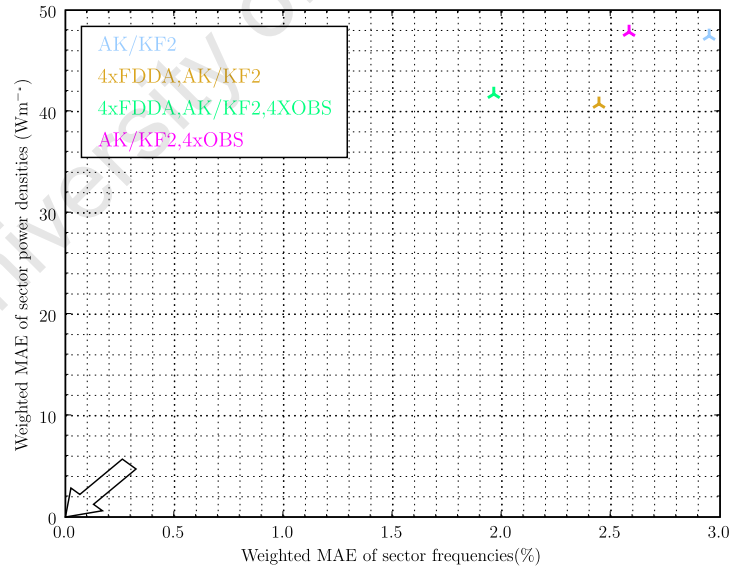


(b) Spatial average

Figure A.4: Goal plot of weighted MAE of sector wind speed averages ($WMAESSA$) and standard deviations ($WMAESSD$) for all four combination runs



(a) Individual stations



(b) Spatial average

Figure A.5: Goal plot of weighted MAE of sector power densities ($WMAEPD$) and frequencies ($WMAESF$) for all four combination runs

ference for Port Elizabeth but appears favourably on the spatial mean too by getting us another 0.5% improvement irrespective of whether grid FDDA is on or not, although it hardly changes the *WMAEPD*.

All in all when it comes to directional performance the optimum MM5 settings are AK/KF2 with 4xOBS, the effect of 4xFDDA is questionable at best.

A.4 Summary

Combining Anthes Kuo/Kain Fritsch 2 with observational nudging and/or grid FDDA yields no surprises over their individual performance. Together they are certainly better than in isolation and can significantly contribute to MM5's skill in terms of wind prediction. The end result is that the strengths and weaknesses of these particular nudging techniques are linearly additive.

University of Cape Town

Appendix B

MM5 Model Configuration

B.1 Sensitivity Runs for the Eastern Cape

B.1.1 TERRAIN

All the domain specific options including those not explicitly discussed in section 4.3.1 on page 66 are given in the terrain.deck file as given below. The code at the end of the user modifiable section has been omitted for brevity.

```
#!/bin/csh -f
# terrain.csh
#
set echo
#
# Set this if you would like to ftp terrain data
#
set ftpdata = false
#
# Set the following for ftp'ing 30 sec elevation data from USGS ftp site
#
set Where30sTer = Data
#set Where30sTer = /your-data-directory
if ( $Where30sTer == ftp) then
#
# Use this if you are ftping from other places
#
set users = Others
#
# Use this if you are ftping from MMM/NCAR
#
set users = MMM
else
set users =
endif
```

```

#
#   Uncomment the following line if using the 30 - 100 cm layer soil file
#
# set BotSoil
#
# -----
#           1. Set up parameter statements
# -----
#
cat > src/parame.incl.tmp << EOF
C       IIMX,JJMX are the maximum size of the domains, NSIZE = IIMX*JJMX
C       PARAMETER (IIMX = 100, JJMX = 100, NSIZE = IIMX*JJMX)
EOF
cat > src/paramed.incl.tmp << EOF
C       ITRH,JTRH are the maximum size of the terrain data.
C       NOBT = ITRH*JTRH, here assuming
C       ITRH= 270 ( 45 deg. in north-south direction, 10 min. resolution)
C       JTRH= 450 ( 75 deg. in north-south direction, 10 min. resolution)
C NOTE:
C   IF USING GLOBAL 30SEC ELEVATION DATASET FROM USGS, NEED TO SET
C   BOTH ITRH AND JTRH BIG. TRY THE COMMENTED PARAMETER LINE FIRST.
C   THIS WILL REQUIRE APPROXI 0.9 GB MEMORY ON A 32-BIT IEEE MACHINE.
C   AN ESTIMATE OF THE DIMENSION SIZE CAN BE MADE FROM Data30s/rdem.out
C   AFTER THE FIRST JOB FAILS. USE (XMAXLAT-XMINLAT)*120 TO ESTIMATE
C   ITRH, AND (XMAXLON-XMINLON)*120 TO ESTIMATE JTRH.
C
C       PARAMETER (ITRH = 500, JTRH = 500, NOBT = ITRH*JTRH)
C       PARAMETER (ITRH = 1500, JTRH = 1800, NOBT = ITRH*JTRH)
EOF
#
# -----
#           2. Set up NAMELIST
# -----
#
if ( -e terrain.namelist ) rm terrain.namelist
cat > terrain.namelist << EOF
&MAPBG
PHIC =  -33.45,      ; CENTRAL LATITUDE (minus for southern hemesphere)
XLONC = 27.05,      ; CENTRAL LONGITUDE (minus for western hemesphere)
IEXP =   .F.,      ; .T. EXPANDED COARSE DOMAIN, .F. NOT EXPANDED.
                ; USEFUL IF RUNNING RAWINS/little_r
AEXP =  360.,      ; APPROX EXPANSION (KM)
;IPROJ = 'LAMCON',  ; LAMBERT-CONFORMAL MAP PROJECTION
;IPROJ = 'POLSTR',  ; POLAR STEREOGRAPHIC MAP PROJECTION
;IPROJ = 'MERCAT',  ; MERCATOR MAP PROJECTION
&END
&DOMAINS
;
MAXNES = 2,        ; NUMBER OF DOMAINS TO PROCESS
NESTIX = 20, 22, 136, 181, 211, 221, ; GRID DIMENSIONS IN Y DIRECTION
NESTJX = 24, 31, 181, 196, 211, 221, ; GRID DIMENSIONS IN X DIRECTION
DIS = 54., 18., 9., 3.0, 1.0, 1.0, ; GRID DISTANCE
NUMNC = 1, 1, 2, 3, 4, 5, ; MOTHER DOMAIN ID
NESTI = 1, 7, 28, 35, 45, 50, ; LOWER LEFT I OF NEST IN MOTHER DOMAIN
NESTJ = 1, 7, 25, 65, 55, 50, ; LOWER LEFT J OF NEST IN MOTHER DOMAIN

```

```

RID      = 1.5, 1.5, 1.5, 3.1, 2.3, 2.3, ; RADIUS OF INFLUENCE IN GRID UNITS (IFANAL=T)
NTYPE    = 4, 6, 4, 6, 6, 6, ; INPUT DATA RESOLUTION
;
;      1: 1 deg (~111 km) global terrain and landuse
;      2: 30 min (~56 km) global terrain and landuse
;      3: 10 min (~19 km) global terrain and landuse
;      4: 5 min (~9 km) global terrain and landuse
;      5: 2 min (~4 km) global terrain and landuse
;      6: 30 sec (~.9 km) global terrain and landuse
;
NSTTYP= 1, 2, 2, 2, 2, 2, ; 1 -- ONE WAY NEST, 2 -- TWO WAY NEST
&END
&OPTN
IFTER    = .TRUE., ; .T.-- TERRAIN, .F.-- PLOT DOMAIN MAPS ONLY
IFANAL   = .F., ; .T.-- OBJECTIVE ANALYSIS, .F.-- INTERPOLATION
ISMTHTR  = 2, ; 1: 1-2-1 smoother, 2: two pass smoother/desmoother
IFEZFUG  = .F., ; .T. USE NCAR GRAPHICS EZMAP WATER BODY INFO TO FUDGE THE LAND USE
; .F. USE LANDWATER MASK DATA
IFTFUG   = .F., ; .T. DON'T DO EZFUDGE WITHIN THE USER-SPECIFIED
; LAT/LON BOXES, need to define namelist fudget
IFFUDG   = .F., ; .T. POINT-BY-POINT FUDGING OF LANDUSE,
; need to define namelist fudge
IPRNTD   = .F., ; PRINT OUT LAT. AND LON. ON THE MESH
IPRTHT   = .T., ; PRINT OUT ALL PROCESSING FIELDS ON THE MESH
IPRINT   = 1, ; = 1: A LOT MORE PRINT OUTPUT IN terrain.print.out
FIN      = 100., 100., 100., 100., 100., 100., ; CONTOUR INTERVAL (meter) FOR TERRAIN HEIGHT PLOT
TRUELAT1=91., ; TRUE LATITUDE 1
TRUELAT2=91., ; TRUE LATITUDE 2, use this if IPROJ='LAMCON'
IFILL    = .TRUE., ; .TRUE. --- color filled plots
LSMDATA  = .TRUE., ; .TRUE. --- Create the data for LSM
VEGTYP   = 1, ; LANDUSE DATA TYPE: =0: old 13 cat; =1: 24 cat USGS; =2: 16 cat SiB
VSPLOT   = .TRUE., ; .TRUE. --- plot Vege., Soil, Vege. Frc. percentages.
IEXTRA   = .TRUE., ; .TRUE. --- Create extra data for Pleim-Xiu LSM
&END
&FUDGE
; USE ONLY IF IFFUDG = .T., POINT-BY-POINT FUDGING OF LANDUSE,
; IFFUG FOR EACH OF THE NESTS: .F. NO FUDGING, .T. FUDGING
IFFUG    = .F.,.F., ; FUDGE FLAGS
; NDFUG : THE NUMBER OF FUDGING POINTS FOR EACH OF NESTS
NDFUG    = 0,0,
; LOCATION (I,J) AND LANDUSE VALUES FOR EACH OF THE NESTS
; NOTE: REGARDLESS OF IFFUG AND NDFUG, 200 VALUES MUST BE GIVEN FOR
; EACH NEST, OR ELSE THE INDEXING WILL GET MESSED UP
; The example below is for two domains. Add more for domain 3 and up
; if needed. Do not remove 0 values for domain 1 and/or 2 even
; they are not used.
;
IFUG(1,1)= 200*0, ; I location for fudge points in domain 1
IFUG(1,2)= 200*0, ; I location for fudge points in domain 2
JFUG(1,1)= 200*0, ; J location for fudge points in domain 1
JFUG(1,2)= 200*0, ; J location for fudge points in domain 2
LNDFUG(1,1)= 200*0, ; land-use value at fudge points for domain 1
LNDFUG(1,2)= 200*0, ; land-use value at fudge points for domain 2
&END
&FUDGET

```

```

; USE ONLY IF IFTFUG=.T., WHICH MEANS TERRAIN WON'T DO EZFUDGE WITHIN
;   THE USER-SPECIFIED LAT/LON BOXES. THIS OPTION IS USED WHEN THERE
;   ARE INLAND BODIES OF WATER THAT ARE DEFINED IN THE LAND USE
;   DATA SET BUT NOT IN THE EZMAP DATA SET. THIS OPTION PREVENTS
;   THOSE BODIES OF WATER FROM BEING WIPE OUT BY EZFUDGE
NFUGBOX = 2          ; NUMBER OF SUBDOMAINS IN WHICH TO
;                   ; TURN OFF EZMAP LAND USE FUDGING
STARTLAT=45.0,44.0,  ; LATITUDES OF LOWER-LEFT CORNERS OF SUBDOMAINS
ENDLAT  =46.5,45.0,  ; LATITUDES OF UPPER-RIGHT CORNERS OF SUBDOMAINS
STARTLON=-95.0,-79.8, ; LONGITUDES OF LOWER-LEFT CORNERS OF SUBDOMAINS
ENDLON  =-92.6,-78.5, ; LONGITUDES OF UPPER-RIGHT CORNERS OF SUBDOMAINS
&END
&EZFUDGE
; USE ONLY IF IFEZFUG=.T., WHICH TURNS ON EZMAP WATER BODY FUDGING OF LANDUSE.
; USERS: FEEL FREE TO ADD ANY MORE LAKE SURFACE HEIGHTS THAT YOU'LL NEED.
; HTPS IS THE HEIGHT IN METERS AND THE INDEX OF HTPS CORRESPONDS TO THE ID
; OF THE 'PS' AREA IN THE FILE ezmap_area_ids.
;
HTPS(441) = -.001   ; Oceans -- Do NOT change this one
HTPS(550) = 183.    ; Lake Superior
HTPS(587) = 177.    ; Lakes Michigan and Huron
HTPS(618) = 176.    ; Lake St. Clair
HTPS(613) = 174.    ; Lake Erie
HTPS(645) = 75.     ; Lake Ontario
HTPS(480) = 1897.   ; Lake Tahoe
HTPS(500) = 1281.   ; Great Salt Lake
&END
EOF
#
# -----
#
#           END OF USER MODIFICATION
#
# -----

```

B.1.2 REGRID

The NCEP/NCAR Reanalysis files have been prepared into the REGRID intermediary format in the “./pregrid” directory by some custom written code which is not listed here. However, REGRID’s namelist.input is given below.

```

&record1
  start_year      = 1996
  start_month     = 03
  start_day       = 15
  start_hour      = 00
  end_year        = 1997
  end_month       = 03
  end_day         = 20
  end_hour        = 00
  interval        = 21600 /
&record2
  ptop_in_Pa      = 1000

```

```

new_levels_in_Pa      = ,
sst_to_ice_threshold  = -9999
linear_interpolation  = .FALSE. /
&record3
root                  = './pregrid/T' './pregrid/HGT' './pregrid/RH'
                    './pregrid/PMSL' './pregrid/U' './pregrid/V' './pregrid/SST' './pregrid/SOILM010'
                    './pregrid/SOILM200' './pregrid/SOILT010' './pregrid/SOILT200' './pregrid/SOILT300'
terrain_file_name     = './../TERRAIN/TERRAIN_DOMAIN2'
constants_full_name   = 'ALMX_FILE' './pregrid/SOILHGT:1997-01-01_00'
                    './pregrid/LANDSEA:1997-01-01_00'/
&record4
print_echo            = .FALSE. ,
print_debug           = .FALSE. ,
print_mask            = .FALSE. ,
print_interp          = .FALSE. ,
print_link_list_store = .FALSE. ,
print_array_store     = .FALSE. ,
print_header          = .FALSE. ,
print_output          = .FALSE. ,
print_file            = .FALSE. ,
print_tc              = .FALSE. ,
print_f77_info        = .FALSE. /
&record5
insert_bogus_storm    = .FALSE.
num_storm              =      1
latc_loc              =     36.0
lonc_loc              =    -35.0
vmax_meters_per_second =     35.0
rmax                  =    90000.0
vmax_ratio            =     0.75 /

```

B.1.3 INTERPF

The following lists the configuration of MM5's INTERPF pre-processor whose main goal is to perform vertical interpolation and thus defines the number (40 in this case) and location of sigma levels in the model.

```

&record0
input_file            = './REGRID/regridder/REGRID_DOMAIN1' / ! pressure-level data file name
&record1
start_year            = 1996                                ! The starting and
start_month           = 03                                  ! ending dates to
start_day             = 15                                  ! process
start_hour            = 00
end_year              = 1997
end_month             = 03
end_day               = 20
end_hour              = 00
interval              = 21600                                ! time difference (s)
less_than_24h         = .FALSE. /                          ! if input is less than 24 h
&record2
sigma_f_bu            = 1.000,0.9975,0.995,0.991,0.986     ! full sigma, bottom-up,
                    0.980,0.973,0.965,0.955,0.945,         ! starting with 1.0, end

```

```

0.935,0.925,0.915,0.900,      ! with 0.0. 40 levels here
0.875,0.850,0.825,0.800,
0.775,0.750,0.725,0.700,
0.675,0.650,0.625,0.600,
0.575,0.550,0.525,0.500,
0.450,0.400,
0.350,0.300,
0.250,0.200,
0.150,0.100,
0.050,0.000

ptop      = 1000                ! top pressure if need to be redefined
isfc      = 0 /                ! # sigma levels to spread
                                ! surface information

&record3
p0         = 1.e5              ! base state sea-level pres (Pa)
t1p        = 50.               ! base state lapse rate d(T)/d(ln P)
ts0        = 285.              ! base state sea-level temp (K)
tiso       = 0./               ! base state isothermal stratospheric temp (K)

&record4
removediv  = .TRUE.            ! T/F remove integrated mean divergence
usesfc     = .TRUE.            ! T/F use surface data
wrth2o     = .TRUE.            ! T/F specific humidity wrt liquid H2O
psfc_method = 0 /              ! 0 => (Tsfc+Tslv)/2 , 1 => diurnally avg temp

&record5
ifdatim    = -1 /              ! # of IC time periods to output

```

B.1.4 MM5

MM5's core configuration is split up into two. Firstly we have the compile time options which mostly select compiler, flags and parametrisations (moisture, cumulus, LSM and PBL) which are specified in a "configure.user" file. Secondly there are the run time options which govern simulation time, timestep, output and save options, FDDA nudging coefficients as well as a few other variables that are given in an "mm5.deck" file.

Both are listed below in their generic forms where the string "<variable>" indicates that some of the runs used different configurations here in order to study their skill. The other settings were kept constant throughout all sensitivity runs as shown.

configure.user (MMP options, commented out sections and general Makefile compile constructs omitted):

```

#-----
# 1. System Variables
#-----
SHELL =      /bin/sh
RANLIB =      echo
.SUFFIXES: .F .i .o .f .c
#-----

```

```

# 3. Fortran options
#   Uncomment the ones you need, including RUNTIME_SYSTEM
#-----
LIBINCLUDE = $(DEVTOP)/include
#-----
#   3i1. PC_PGF77 (LINUX/Portland Group Inc.)
#       pgf77 version 1.6 and above
#       May use pgf90 if the version is 3.1-4
#-----
RUNTIME_SYSTEM = "linux"
FC = pgf90
FCFLAGS = -I$(LIBINCLUDE) -O2 -Mcray=pointer -tp p6 -pc 32 -Mnoframe -byteswapio -mp \
-Mnosgimp
CPP = /lib/cpp
CFLAGS = -O
CPPFLAGS = -I$(LIBINCLUDE)
LDOPTIONS = -O2 -Mcray=pointer -tp p6 -pc 32 -Mnoframe -byteswapio -mp
LOCAL_LIBRARIES =
MAKE = make -r
#-----
#   3i2. PC_INTEL (LINUX/INTEL)
#-----
#RUNTIME_SYSTEM = "linux"
#FC = ifort
#FCFLAGS = -I$(LIBINCLUDE) -O2 -tp p6 -pc 32 -convert big_endian
#CPP = /lib/cpp
#CFLAGS = -O
#CPPFLAGS = -I$(LIBINCLUDE)
#LDOPTIONS = -O2 -tp p6 -pc 32 -convert big_endian
#LOCAL_LIBRARIES =
#MAKE = make -i -r
#-----
# 4. General commands
#-----
AR = ar ru
RM = rm -f
RM_CMD = $(RM) *.CKP *.ln *.BAK *.bak *.o *.i core errs ,* ~* .a \
.emacs_* tags TAGS make.log MakeOut *.f !
GREP = grep -s
CC = cc
#-----
# 5. Options for making ./include/parame.incl
#-----
#
# FDDAGD (integer)           - "1" -> FDDA gridded run
FDDAGD = <variable>
#
# FDDAOBS (integer)         - "1" -> FDDA obs run
FDDAOBS = <variable>
#
# MAXNES (integer)          - Max Number of Domains in simulation
MAXNES = 2
#
# MIX,MJX (integer)         - Maximum Dimensions of any Domain
MIX = 22

```

```

MJX = 31
# MKX (integer)                - Number of half sigma levels in model
MKX = 39
#-----
# 6. Physics Options
#   The first MAXNES values in the list will be used for the corresponding
#   model nests; the rest in the list can be used to compile other options.
#   The exception is FRAD, of which only the first value is used in the model,
#   (i.e., only one radiation option is used for all nests). The rest allow
#   other options to be compiled.
#   Compilation of Arakawa-Schubert cumulus scheme requires imsl.
#-----
# IMPHYS - for explicit moisture schemes (array,integer)
IMPHYS = "4,4,1,1,1,1,1,1,1"
#
#   - Dry,stable,warm rain,simple ice,mix phase,
#   - 1 ,2 ,3 ,4 ,5
#   - graupel(gsfc),graupel(reisner2),schultz
#   - ,6 ,7 ,8
MPHYSTBL = 1
#
#   - 0=do not use look-up tables for moist
#   physics
#   - 1=use look-up tables for moist physics
#   (currently only simple ice and mix phase
#   are available)
#   - 2=optimized exmoir routine (need vslib, if not
#   available set -DvsLIB in compile flags)
#
# ICUPA - for cumulus schemes (array,integer)
#   - None,Kuo,Grell,AS,FC,KF,BM,KF2 - 1,2,3,4,5,6,7,8
ICUPA = "<variable>,<variable>,1,1,1,1,1,1,1"
#
# IBLTYP - for planetary boundary layer (array,integer)
#   - 0=no PBL fluxes,1=bulk,2=Blackadar,
#   3=Burk-Thompson,4=Eta M-Y,5=MRF,
#   6=Gayno-Seaman,7=Pleim-Xiu
IBLTYP = "<variable>,<variable>,0,0,0,0,0,0,0"
#
# FRAD - for atmospheric radiation (integer)
#   - Radiation cooling of atmosphere
#   0=none,1=simple,2=cloud,3=ccm2,rrtm=4
FRAD = "4,4,0,0,0"
#
# IPOLAR - (integer) for polar model used only if ISOIL=1
#   0=not polar (5-layer soil model)
#   1=polar (7-layer snow/soil model)
IPOLAR = 0
#
# ISOIL - for multi-layer soil temperature model (integer)
#   - 0=no,1=yes (only works with IBLTYP=2,4,5,6)
#   2=Noah land-surface scheme (IBLTYP=4,5 only)
#   3=Pleim-Xiu LSM (IBLTYP=7 only)
ISOIL = <variable>
#
# ISHALLO (array,integer)      - Shallow Convection Option
#   1=shallow convection,0=No shallow convection

```

```
ISHALLO = "1,1,0,0,0,0,0,0,0,0"
```

mm5.deck:

```
#!/bin/sh
#
# Version 3 of mm5 job deck
#
# The mm5 executable (mm5.exe) expects to find the following files
#   in the Run/ directory:
#   MMINPUT_DOMAIN1  -|
#   BDYOUT_DOMAIN1   | --> output files from Interpf
#   LOWBDY_DOMAIN1   -|
#   TERRAIN_DOMAIN[2,3..] if running nests --> output from Terrain
#
# If it is a restart run:
#   RESTART_DOMAIN1[2,3..] --> output from MM5 run: renamed from
#                           SAVE_DOMAIN1[2,3...]
#
# If it is gridded FDDA run with surface analysis nudging:
#   SFCFDDA_DOMAIN1[2,3,...]
#
# If it is observational nudging run:
#   MM5OBS_DOMAIN1[2,3..] --> user-created observation files
#
# Output from a MM5 run:
#   If IFTAPE = 1
#   MMOUT_DOMAIN1[2,3...] --> one output for each domain
#   If IFSAVE = TRUE
#   SAVE_DOMAIN1[2,3...]
#
# temp files should be accessible
umask 022
#
# Sections
# 1. Options for namelist ("mmlif")
# 2. Running...
#
BLOCKEND='&END'
#-----
# 1. Options for namelist ("mmlif")
#-----
#
#   The first dimension (column) of the arrays denotes the domain
#   identifier.
#
#       Col 1 = Domain #1, Col 2 = Dom #2, etc.
#
cat > ./Run/oparam << EOF
&OPARAM
;
;   ***** FORECAST TIME AND TIME STEP *****
;
TIMAX = 532800.,           ; forecast length in minutes
TISTEP = 90.,             ; coarse domain DT in model, use 3*DX
```

178

179

180

```

&FPARAM
;
; ***** 4DDA OPTIONS *****
;
; THE FIRST DIMENSION (COLUMN) IS THE DOMAIN IDENTIFIER:
; COLUMN 1 = DOMAIN #1, COLUMN 2 = DOMAIN #2, ETC.
;
; START TIME FOR FDDA (ANALYSIS OR OBS) FOR EACH DOMAIN
; (IN MINUTES RELATIVE TO MODEL INITIAL TIME)
FDASTA=0.,0.,0.,0.,0.,0.,0.,0.,0.,0.
;
; ENDING TIME FOR FDDA (ANALYSIS OR OBS) FOR EACH DOMAIN
; (IN MINUTES RELATIVE TO MODEL INITIAL TIME)
FDAEND=532800.,532800.,0.,0.,0.,0.,0.,0.,0.,0.
;
; ***** ANALYSIS NUDGING *****
;
; THE FIRST DIMENSION (COLUMN) OF THE ARRAYS DENOTES THE
; DOMAIN IDENTIFIER:
; COLUMN 1 = DOMAIN #1, COLUMN 2 = DOMAIN #2, ETC.
; THE SECOND DIMENSION (ROW OR LINE) EITHER REFERS TO THE 3D VS
; SFC ANALYSIS OR WHICH VARIABLE IS ACCESSED:
; LINE 1 = 3D, LINE 2 = SFC OR
; LINE 1 = U, LINE 2 = V, LINE 3 = T, LINE 4 = Q
;
; IS THIS A GRID 4DDA RUN? 0 = NO; 1 = YES
I4D= <variable>,0,0,0,0,0,0,0,0,0,0,
    0,0,0,0,0,0,0,0,0,0,0,
;
; SPECIFY THE TIME IN MINUTES BETWEEN THE INPUT (USUALLY
; FROM INTERP) USED FOR GRID FDDA
DIFTIM=360.,360.,0.,0.,0.,0.,0.,0.,0.,0.,    ; 3D ANALYSIS NUDGING
    180.,180.,0.,0.,0.,0.,0.,0.,0.,0.,    ; SFC ANALYSIS NUDGING
;
; GRID NUDGE THE WIND FIELD? 0 = NO; 1 = YES
IWIND=1,0,0,0,0,0,0,0,0,0,    ; 3D ANALYSIS NUDGING
    0,0,0,0,0,0,0,0,0,0,    ; SFC ANALYSIS NUDGING
;
; NUDGING COEFFICIENT FOR WINDS ANALYSES
GV=<variable>,1.0E-4,0.,0.,0.,0.,0.,0.,0.,    ; 3D ANALYSIS NUDGING
    2.5E-4,1.0E-4,0.,0.,0.,0.,0.,0.,0.,    ; SFC ANALYSIS NUDGING
;
; GRID NUDGE THE TEMPERATURE FIELD? 0 = NO; 1 = YES
ITEMP=1,0,0,0,0,0,0,0,0,0,    ; 3D ANALYSIS NUDGING
    0,0,0,0,0,0,0,0,0,0,    ; SFC ANALYSIS NUDGING
;
; NUDGING COEFFICIENT FOR TEMPERATURE ANALYSES
GT=<variable>,1.0E-4,0.,0.,0.,0.,0.,0.,0.,    ; 3D ANALYSIS NUDGING
    2.5E-4,1.0E-4,0.,0.,0.,0.,0.,0.,0.,    ; SFC ANALYSIS NUDGING
;
; IMOIS=1,0,0,0,0,0,0,0,0,0,    ; 3D ANALYSIS NUDGING
    0,0,0,0,0,0,0,0,0,0,    ; SFC ANALYSIS NUDGING
;
; NUDGING COEFFICIENT FOR THE MIXING RATIO ANALYSES
GQ=<variable>,1.E-5,0.,0.,0.,0.,0.,0.,0.,    ; 3D ANALYSIS NUDGING
    1.E-5,1.E-5,0.,0.,0.,0.,0.,0.,0.,    ; SFC ANALYSIS NUDGING

```

```

;
;      GRID NUDGE THE ROTATIONAL WIND FIELD? 0 = NO; 1 = YES
IROT=0,0,0,0,0,0,0,0,0,0,      ; 3D ANALYSIS NUDGING
;
;      NUDGING COEFFICIENT FOR THE ROTATIONAL COMPONENT OF THE WINDS
GR=5.E6,5.E6,0.,0.,0.,0.,0.,0.,0.,      ; 3D ANALYSIS NUDGING
;
;      IF GRID NUDGING (I4D(1,1)=1) AND YOU WISH TO EXCLUDE THE
;      BOUNDARY LAYER FROM FDDA OF COARSE GRID THREE DIMENSIONAL
;      DATA (USUALLY FROM INTERP),
;      0 = NO,   INCLUDE BOUNDARY LAYER NUDGING
;      1 = YES,  EXCLUDE BOUNDARY LAYER NUDGING
INONBL =1,1,0,0,0,0,0,0,0,0,      ; U WIND
        1,1,0,0,0,0,0,0,0,0,      ; V WIND
        1,1,1,1,1,1,1,1,1,1,      ; TEMPERATURE
        1,1,1,1,1,1,1,1,1,1,      ; MIXING RATIO
;
;      RADIUS OF INFLUENCE FOR SURFACE ANALYSIS (KM).
;      IF I4D(2,1)=1 OR I4D(2,2)=1, ETC, DEFINE RINBLW (KM) USED
;      IN SUBROUTINE BLW TO DETERMINE THE HORIZONTAL VARIABILITY
;      OF THE SURFACE-ANALYSIS NUDGING AS A FUNCTION OF SURFACE
;      DATA DENSITY. OVER LAND, THE STRENGTH OF THE SURFACE-
;      ANALYSIS NUDGING IS LINEARLY DECREASED BY 80 PERCENT AT
;      THOSE GRID POINTS GREATER THAN RINBLW FROM AN OBSERVATION
;      TO ACCOUNT FOR DECREASED CONFIDENCE IN THE ANALYSIS
;      IN REGIONS NOT NEAR ANY OBSERVATIONS.
RINBLW=250.,
;
;      SET THE NUDGING PRINT FREQUENCY FOR SELECTED DIAGNOSTIC
;      PRINTS IN THE GRID (ANALYSIS) NUDGING CODE (IN CGM
;      TIMESTEPS)
NPFG=960,
;
;      ***** OBSERVATION NUDGING *****
;
;      INDIVIDUAL OBSERVATION NUDGING. VARIABLES THAT ARE ARRAYS
;      USE THE FIRST DIMENSION (COLUMN) AS THE DOMAIN IDENTIFIER:
;      COLUMN 1 = DOMAIN #1, COLUMN 2 = DOMAIN #2, ETC.
;
;      IS THIS INDIVIDUAL OBSERVATION NUDGING? 0 = NO; 1 = YES
I4DI  =0,<variable>,0,0,0,0,0,0,0,0,
;
;      OBS NUDGE THE WIND FIELD FROM STATION DATA? 0 = NO; 1 = YES
ISWIND =0,1,0,0,0,0,0,0,0,0,
;
;      NUDGING COEFFICIENT FOR WINDS FROM STATION DATA
GIV  =4.E-4,<variable>,0.,0.,0.,0.,0.,0.,0.,
;
;      OBS NUDGE THE TEMPERATURE FIELD FROM STATION DATA? 0 = NO; 1 = YES
ISTEMP=0,0,0,0,0,0,0,0,0,0,
;
;      NUDGING COEFFICIENT FOR TEMPERATURES FROM STATION DATA
GIT  =4.E-4,4.E-4,0.,0.,0.,0.,0.,0.,0.,0.,
;

```

183

```
#-----
#
#      run MM5
#
echo This version of mm5.deck stops after creating namelist file mmlif.
echo Please run code manually.
date
```

B.2 Wind Atlas Runs for South Africa

B.2.1 TERRAIN

The below terrain.deck file configures both mother domain at 54km and nest at 18km resolution that encompass the whole of South Africa for the two wind atlas simulations. This time an abridged version is given with most comments removed, also note that only the user modifiable parts are reproduced.

```
#!/bin/csh -f
set echo
set ftpdata = false
set Where30sTer = .
if ( $Where30sTer == ftp) then
    set users = MMM
else
    set users =
endif
cat > src/parame.incl.tmp << EOF
    PARAMETER (IIMX = 300, JJMX = 300, NSIZE = IIMX*JJMX)
EOF
cat > src/paramed.incl.tmp << EOF
    PARAMETER (ITRH = 2130, JTRH = 2631, NOBT = ITRH*JTRH)
EOF
if ( -e terrain.namelist ) rm terrain.namelist
cat > terrain.namelist << EOF
&MAPBG
PHIC = -28.5,      ; CENTRAL LATITUDE (minus for southern hemesphere)
XLONC = 24.8,      ; CENTRAL LONGITUDE (minus for western hemesphere)
IEXP = .F.,        ; .T. EXPANDED COARSE DOMAIN, .F. NOT EXPANDED.
                ; USEFUL IF RUNNING RAWINS/little_r
AEXP = 360.,       ; APPROX EXPANSION (KM)
;IPROJ = 'LANCON',  ; LAMBERT-CONFORMAL MAP PROJECTION
;IPROJ = 'POLSTR',  ; POLAR STEREOGRAPHIC MAP PROJECTION
IPROJ = 'MERCAT',   ; MERCATOR MAP PROJECTION
&END
&DOMAINS
;
MAXNES = 2,         ; NUMBER OF DOMAINS TO PROCESS
NESTIX = 51, 112, 136, 181, 211, 221, ; GRID DIMENSIONS IN Y DIRECTION
NESTJX = 56, 124, 181, 196, 211, 221, ; GRID DIMENSIONS IN X DIRECTION
DIS = 54., 18., 9., 3.0, 1.0, 1.0, ; GRID DISTANCE
NUMNC = 1, 1, 2, 3, 4, 5, ; MOTHER DOMAIN ID
```

```

NESTI = 1, 7, 28, 35, 45, 50, ; LOWER LEFT I OF NEST IN MOTHER DOMAIN
NESTJ = 1, 8, 25, 65, 55, 50, ; LOWER LEFT J OF NEST IN MOTHER DOMAIN
RID = 1.5, 1.5, 1.5, 3.1, 2.3, 2.3, ; RADIUS OF INFLUENCE IN GRID UNITS (IFANAL=T)
NTYPE = 4, 6, 4, 6, 6, 6, ; INPUT DATA RESOLUTION
;
NSTTYP= 1, 2, 2, 2, 2, 2, ; 1 -- ONE WAY NEST, 2 -- TWO WAY NEST
&END
&OPTN
IFTER = .TRUE., ; .T.-- TERRAIN, .F.-- PLOT DOMAIN MAPS ONLY
IFANAL = .F., ; .T.-- OBJECTIVE ANALYSIS, .F.-- INTERPOLATION
ISMTHR = 2, ; 1: 1-2-1 smoother, 2: two pass smoother/desmoother
IFEZFUG = .F., ; .T. USE NCAR GRAPHICS EZMAP WATER BODY INFO TO FUDGE THE LAND USE
; .F. USE LANDWATER MASK DATA
IFTFUG = .F., ; .T. DON'T DO EZFUDGE WITHIN THE USER-SPECIFIED
; LAT/LON BOXES, need to define namelist fudget
IFFUG = .F., ; .T. POINT-BY-POINT FUDGING OF LANDUSE,
; need to define namelist fudge
IPRNTD = .F., ; PRINT OUT LAT. AND LON. ON THE MESH
IPRHT = .T., ; PRINT OUT ALL PROCESSING FIELDS ON THE MESH
IPRINT = 0, ; = 1: A LOT MORE PRINT OUTPUT IN terrain.print.out
FIN = 100., 100., 100., 100., 100., 100., ; CONTOUR INTERVAL (meter) FOR TERRAIN HEIGHT PLOT
;TRUELAT1=91., ; TRUE LATITUDE 1
;TRUELAT2=91., ; TRUE LATITUDE 2, use this if IPROJ='LAMCON'
IFILL = .TRUE., ; .TRUE. --- color filled plots
LSMDATA = .TRUE., ; .TRUE. --- Create the data for LSM
VEGTYPE = 1, ; LANDUSE DATA TYPE: =0: old 13 cat; =1: 24 cat USGS; =2: 16 cat SiB
VSPLOT = .TRUE., ; .TRUE. --- plot Vege., Soil, Vege. Frc. percentages.
IEXTRA = .TRUE., ; .TRUE. --- Create extra data for Pleim-Xiu LSM
&END
&FUDGE
IFFUG = .F.,.F., ; FUDGE FLAGS
NDFUG = 0,0,
IFUG(1,1)= 200*0, ; I location for fudge points in domain 1
IFUG(1,2)= 200*0, ; I location for fudge points in domain 2
JFUG(1,1)= 200*0, ; J location for fudge points in domain 1
JFUG(1,2)= 200*0, ; J location for fudge points in domain 2
LNDFUG(1,1)= 200*0, ; land-use value at fudge points for domain 1
LNDFUG(1,2)= 200*0, ; land-use value at fudge points for domain 2
&END
&FUDGET
NFUGBOX = 2 ; NUMBER OF SUBDOMAINS IN WHICH TO
STARTLAT=45.0,44.0, ; LATITUDES OF LOWER-LEFT CORNERS OF SUBDOMAINS
ENDLAT =46.5,45.0, ; LATITUDES OF UPPER-RIGHT CORNERS OF SUBDOMAINS
STARTLON=-95.0,-79.8, ; LONGITUDES OF LOWER-LEFT CORNERS OF SUBDOMAINS
ENDLON =-92.6,-78.5, ; LONGITUDES OF UPPER-RIGHT CORNERS OF SUBDOMAINS
&END
&EZFUDGE
HTPS(441) = -.001 ; Oceans -- Do NOT change this one
HTPS(550) = 183. ; Lake Superior
HTPS(587) = 177. ; Lakes Michigan and Huron
HTPS(618) = 176. ; Lake St. Clair
HTPS(613) = 174. ; Lake Erie
HTPS(645) = 75. ; Lake Ontario
HTPS(480) = 1897. ; Lake Tahoe
HTPS(500) = 1281. ; Great Salt Lake

```

```
&END
EOF
```

B.2.2 REGRID

The regridding of NCEP/NCAR Reanalysis data for input is handled identically to subsection B.1.2 on page 172.

B.2.3 INTERPF

Vertical interpolation and the setup of sigma levels is also configured in the same way as in subsection B.1.3 on page 173.

B.2.4 MM5

The configuration of the core atmospheric model is again very similar to that outlined in detail in subsection B.1.4 on page 174. For completeness sake below are both the significant parts of `configure.user` and `mm5.deck` with comments and irrelevant bits removed.

`configure.user` (customised for the IBM P690 32 CPU SMP machine at CHPC) which both runs use:

```
#### SMP VERSION ####
RUNTIME_SYSTEM = "IBM"
FC = xlf_r
FCFLAGS = -I$(LIBINCLUDE) -O2 -qarch=auto -qmaxmem=-1 -qsmp=omp:noauto -qnosave \
-qstrict -qnocclines
CPP = cpp
CFLAGS = -O
# we REQUIRE the -traditional-cpp because otherwise condensed whitespace messes up
# our preprocessed result
# -w is for suppressing the warnings generated by the trailing
# fixed form line numbering
CPPFLAGS = -I$(LIBINCLUDE) -C -P -traditional-cpp -w
LDOPTIONS = -qmaxmem=-1 -O2 -qarch=auto
LOCAL_LIBRARIES = -lmass -lxlsm
MAKE = make
AR = ar ru
RM = rm -f
RM_CMD = $(RM) *.CKP *.ln *.BAK *.bak *.o *.i core errs ,* *~ *.a \
.emacs_* tags TAGS make.log MakeOut *.f !
GREP = grep -s
CC = cc
FDDAGD = 1
FDDA0BS = 1
MAXNES = 2
MIX = 112
```

```

MJX = 124
MKX = 39
IMPHYS = "4,4,1,1,1,1,1,1,1"
MPHYSTBL = 1
ICUPA = "2,8,1,1,1,1,1,1,1"
IBLTYP = "7,7,0,0,0,0,0,0,0"
FRAD = "4,4,0,0,0"
IPOLAR = 0
ISOIL = 3
ISHALLO = "1,1,0,0,0,0,0,0,0"

```

The mm5.deck has been kept generic too with only one single change between the run utilising grid FDDA and the one without, indicated by “<variable>”:

```

cat > ./Run/oparam << EOF
&OPARAM
; ***** FORECAST TIME AND TIME STEP *****
TIMAX = 532800., ; forecast length in minutes
TISTEP = 90., ; coarse domain DT in model, use 3*DX
; ***** OUTPUT/RESTART OPTIONS *****
IFREST = .FALSE., ; whether this is a restart
IXTIMR = 9999, ; restart time in minutes
IFSAVE = .TRUE., ; save data for restart
SVLAST = .TRUE., ; T: only save the last file for restart
SAVFRQ = 360., ; how frequently to save data (in minutes)
IFTAPE = 1, ; model output: 0,1
TAPFRQ = 60., ; how frequently to output model results (in minutes)
BUFFRQ = 0., ; how frequently to split model output files (in minutes)
INCTAP = 6,1,1,1,1,1,1,1,1, ; multipliers of TAPFRQ for outputting
IFRSFA = .FALSE., ; IF this is a RESTART run, AND IF FDDA is ON,
; AND IF multiple input FILES are used, set this to .TRUE.
; set CDATEST to the INITIAL time of the first run
IFSKIP = .FALSE., ; whether to skip input files - DO NOT use this for restart
CDATEST = '1993-03-13_00:00:00', ; IF IFSKIP=.TRUE. this is the code start date
IFPRT = 0, ; sample print out: =1, a lot of print
PRTFRQ = 720., ; Print frequency for sample output (in minutes)
MASCHK = 99999, ; mass conservation check (KTAU or no. of time steps)
IFTSOUT = .FALSE., ; output time series (default 30 points)
TSLAT = 0.0,0.0,0.0,0.0,0.0, ; latitudes of time series points (S is negative)
TSLON = 0.0,0.0,0.0,0.0,0.0, ; longitudes of time series points (W is negative)
/
EOF
cat > ./Run/lparam << EOF
&LPARAM
RADFRQ = 30., ;atmospheric radiation calculation frequency (in minutes)
IMVDIF = 0, ;moist vertical diffusion in clouds - 0, 1 (IBLTYP=2,5 only)
IVQADV = 1, ;vertical moisture advection uses log interpolation - 0, linear - 1
IVTADV = 1, ;vertical temperature advection uses theta interpolation - 0, linear - 1
ITHADV = 1, ;advection of temperature uses potential temperature - 1, standard - 0
ITPDIF = 1, ;sigma-diffusion using temperature - 0, sigma-diffusion using perturbation
TDKORR = 2, ;temperature gradient correction for z-diffusion at ground level
ICOR3D = 1, ;3D Coriolis force - 0, 1
IEXSI = 0, ;initial sea-ice - 0, 1(base on SST), 2(read in) (ISOIL=1 only)
IFUPR = 1, ;upper radiative boundary condition - 0, 1

```

188

```

EOF
cat > ./Run/pparam << EOF
&PPARAM
ZZLND = 0.1,           ; roughness length over land in meters
ZZWTR = 0.0001,        ; roughness length over water in meters
ALBLND = 0.15,         ; albedo
THINLD = 0.04,         ; surface thermal inertia
XMAVA = 0.3,           ; moisture availability over land as a decimal fraction of one
CONF = 1.0,            ; non-convective precipitation saturation threshold (=1: 100%)
/
EOF
cat > ./Run/fparam << EOF
&FPARAM
; ***** 4DDA OPTIONS *****
FDATA=0.,0.,0.,0.,0.,0.,0.,0.,0.,0.,
FDAEND=532800.,532800.,0.,0.,0.,0.,0.,0.,0.,0.,
; ***** ANALYSIS NUDGING *****
I4D= <variable>,0,0,0,0,0,0,0,0,0,
      0,0,0,0,0,0,0,0,0,
DIFTIM=360.,360.,0.,0.,0.,0.,0.,0.,0.,0., ; 3D ANALYSIS NUDGING
      180.,180.,0.,0.,0.,0.,0.,0.,0.,0., ; SFC ANALYSIS NUDGING
IWIND=1,0,0,0,0,0,0,0,0,0, ; 3D ANALYSIS NUDGING
      0,0,0,0,0,0,0,0,0,0, ; SFC ANALYSIS NUDGING
GV=1.0E-3,1.0E-4,0.,0.,0.,0.,0.,0.,0.,0., ; 3D ANALYSIS NUDGING
      2.5E-4,1.0E-4,0.,0.,0.,0.,0.,0.,0.,0., ; SFC ANALYSIS NUDGING
ITEMP=1,0,0,0,0,0,0,0,0,0, ; 3D ANALYSIS NUDGING
      0,0,0,0,0,0,0,0,0,0, ; SFC ANALYSIS NUDGING
GT=1.0E-3,1.0E-4,0.,0.,0.,0.,0.,0.,0.,0., ; 3D ANALYSIS NUDGING
      2.5E-4,1.0E-4,0.,0.,0.,0.,0.,0.,0.,0., ; SFC ANALYSIS NUDGING
IMOIS=1,0,0,0,0,0,0,0,0,0, ; 3D ANALYSIS NUDGING
      0,0,0,0,0,0,0,0,0,0, ; SFC ANALYSIS NUDGING
GQ=4.E-5,1.E-5,0.,0.,0.,0.,0.,0.,0.,0., ; 3D ANALYSIS NUDGING
      1.E-5,1.E-5,0.,0.,0.,0.,0.,0.,0.,0., ; SFC ANALYSIS NUDGING
IRDT=0,0,0,0,0,0,0,0,0,0, ; 3D ANALYSIS NUDGING
GR=5.E6,5.E6,0.,0.,0.,0.,0.,0.,0.,0., ; 3D ANALYSIS NUDGING
INONBL =1,1,0,0,0,0,0,0,0,0, ; U WIND
      1,1,0,0,0,0,0,0,0,0, ; V WIND
      1,1,1,1,1,1,1,1,1,1, ; TEMPERATURE
      1,1,1,1,1,1,1,1,1,1, ; MIXING RATIO
RINBLW=250.,
NPFG=960,
; ***** OBSERVATION NUDGING *****
I4DI =0,1,0,0,0,0,0,0,0,0,
ISWIND =0,1,0,0,0,0,0,0,0,0,
GIV =4.E-4,1.6E-3,0.,0.,0.,0.,0.,0.,0.,0.,
ISTEMP=0,0,0,0,0,0,0,0,0,0,
GIT =4.E-4,4.E-4,0.,0.,0.,0.,0.,0.,0.,0.,
ISHOIS=0,0,0,0,0,0,0,0,0,0,
GIQ =4.E-4,4.E-4,0.,0.,0.,0.,0.,0.,0.,0.,
RINXY=20.,
RINSIG=0.001,
TWINDU=30.0,
NPFI=40,
IONF=2,
IDYNIN=0, ;for dynamic initialization using a ramp-down function to gradually

```

```
DTRAMP=60.;;the time period in minutes over which the  
/  
EOF
```

University of Cape Town

Appendix C

MM5 Source Code Modifications

C.1 Pleim Xiu LSM Roughness Length Output

The following patch will, upon model initialisation, print to stdout what PX's logarithmic averaging has produced in terms of roughness lengths z_0 per grid cell, starting with I=0 and J=0, then I=0, J=1 etc. This is the exact grid order that GrADS also uses as convention. The patch also produces a roughness_nxm.dat file for each domain (n and m being the grid dimensions) which will serve as a GrADS .dat binary data file. Then all we require is a .ctl descriptor file from any similar surface field obtained through MM5toGrADS, change the variable name and description and the roughness lengths can readily be displayed.

```
--- MM5/physics/pbl_sfc/pxpbl/vegeland.F      2004-12-23 00:38:09.000000000 +0200
+++ vegeland.F  2008-02-10 21:48:03.000000000 +0200
@@ -59,6 +59,7 @@
     REAL SIG,ZNOTC,CLAI                                VEGELAND.55
     REAL VEGF,FAREA,FWAT,FDATE                          VEGELAND.56
     INTEGER ILX,JLX                                      VEGELAND.57
+    CHARACTER*25 RFNAME                                  MYCHANGE.01
C*****
C USGS LU characterization                              VEGELAND.59
c-----                                                VEGELAND.60
@@ -163,6 +164,8 @@
     FDATE = FLOAT(JDATE-JM0(IM01))/FLOAT(JM0(IM02)-JM0(IM01))  VEGELAND.159
     ENDIF                                                    VEGELAND.160
     ENDIF                                                    VEGELAND.161
```

```

+      WRITE (RFNAME, "('pxroughness_',I4.4,'x',I4.4,'.dat')") ILX, JLX      MYCHANGE.002
+      OPEN (UNIT=69, FILE=RFNAME, ACCESS='DIRECT', RECL=4)                  MYCHANGE.003
      C...START OF GRID CELL BY GRID CELL PROCESSING                        VEGELAND.162
                                                                              VEGELAND.163
      DO 99 IRT=1,ILX                                                        VEGELAND.164
@@ -287,6 +290,8 @@
          XVEGMN(IRT,JCT) = SUMVMN/ALAI                                     VEGELAND.281
                                                                              VEGELAND.282
          ENDIF                                                              VEGELAND.283
          ZNT(IRT,JCT) = ZNT(IRT,JCT)*0.01                                !CONVERT TO M
+      PRINT *, 'RLENGTH AT I=', IRT, ' J=', JCT, ' IS ', ZNT(IRT,JCT)      MYCHANGE.004
+      WRITE (69, REC=(IRT-1)*JLX+JCT) ZNT(IRT,JCT)                        MYCHANGE.005
          XVEG(IRT,JCT) = XVEG(IRT,JCT)*0.01                                !CONVERT TO FRAC
          XVEGMN(IRT,JCT) = XVEGMN(IRT,JCT)*0.01                          VEGELAND.285
                                                                              VEGELAND.286

@@ -295,6 +300,7 @@
      C      & AVGRSMN, AVGZO, AVGVEG,AVGVMN, ALAI                          VEGELAND.289
                                                                              VEGELAND.290
      99      CONTINUE                                                        VEGELAND.291
+      CLOSE (69)                                                            MYCHANGE.006
      C805      FORMAT (2I4, 3F15.2,F15.4,3F15.2)                          VEGELAND.292
      RETURN                                                                  VEGELAND.293
      END                                                                    VEGELAND.294

```

C.2 Observational Nudging Patch

This patch addresses the limitations of the MM5 code which disallows observational nudging to span year boundaries. It involves a different and more flexible date and time format which allows for full year specification. Detailed explanations about this new format and how it works can be found online at <http://www.mmm.ucar.edu/mm5/mm5v3/util.html> (where this patch may be downloaded) as well as in the MM5 mailing list. The nature of the patch is summarised in the following note

```

Now works with all sorts of year transitions, leap year or not, all based on the starting
year of the simulation and a new YYYYdddhf specification of TIMEOB in the MM5OBS_DOMAINx
files. Involves at its core a new 'global' variable MYEAR(start year of simulation) which
is already properly initialised in domain/initial/param.F but then also becomes visible
in in4dob.F by inclusion of various.incl. That in turn requires IN4DOB(...) to be
rewritten to IN4DOB(INEST) in all invocations. Affected source files (diff for details):

fdda/util/setfd.F: Change of IN4DOB invocation to IN4DOB(INEST)
domain/nest/initnest.F: Change of IN4DOB invocation to IN4DOB(INEST)
domain/initial/init.F: Change of IN4DOB invocation to IN4DOB(INEST)
domain/initial.param.F: Removal of MYEAR declaration
include/various.incl: Declaration of INTEGER MYEAR with COMMON block
include/parfdda.incl: Removal of RTLAST, now an INTEGER ITLAST
fdda/obs/in4dob.F: Changes too numerous to list, diff for details

```

Here is the patch:

```

diff -X source_excludes -Nur MM5/domain/initial/init.F MM5_n2/domain/initial/init.F
--- MM5/domain/initial/init.F 2004-12-23 00:52:45.000000000 +0200
+++ MM5_n2/domain/initial/init.F 2006-10-13 12:34:53.000000000 +0200
@@ -886,7 +886,7 @@
C IF OBS NUDGING ON ANY MESH, IN4DOB MUST BE CALLED HERE... INIT.794
C INIT.795
#ifdef FDDA0BS1 INIT.796
- IF(I4DITOT.GE.1)CALL IN4DOB(1,XTIME,KTAU,KTAUR,DTMIN,JULDAY,GMT) INIT.797
+ IF(I4DITOT.GE.1)CALL IN4DOB(1) INIT.797
#endif INIT.798
#ifdef FDDAGD1 INIT.799
IF(I4D(1,1).EQ.1)CALL IN4DGD(1,IL,JL,KL) INIT.800
diff -X source_excludes -Nur MM5/domain/initial/param.F MM5_n2/domain/initial/param.F
--- MM5/domain/initial/param.F 2005-03-18 21:31:36.000000000 +0200
+++ MM5_n2/domain/initial/param.F 2006-10-24 18:34:33.000000000 +0200
@@ -60,7 +60,7 @@
INTEGER LI,LS,LC PARAM.53
INTEGER NN2,NN3,N,NN,NNL,IN,NV 24SEP99.31
INTEGER NOC,NOV PARAM.55
- INTEGER MYEAR,MMON,MDAY,MHOUR,MMIN,MSEC PARAM.56
+ INTEGER MMON,MDAY,MHOUR,MMIN,MSEC PARAM.56
INTEGER MY1,MY2,MY3,MONSS PARAM.57
REAL EQOFT PARAM.58
REAL AA,CC,SSUM,XX,XTOP,XBOT,VQMAX,YY,WK,QK,WKP1,QKP1 19DEC02.211
@@ -2298,7 +2298,7 @@
# include <mpp_param_25.incl> PARAM.1706
#endif PARAM.1707
C PARAM.1708
-C-----IF ITGFLG=1, SPECIFY THE JULIAN DATE AND GMT OF THE INITIAL DATA. PARAM.1709
+C-----IF ITGFLG=1, SPECIFY THE YEAR,JULIAN DATE AND GMT OF THE INITIAL DATA. PARAM.1709
C DECTIM : IS THE TIME IN MINUTES AFTER WHICH THE SOLAR DECLINATION PARAM.1710
C ANGLE MUST BE RECALCULATED. PARAM.1711
C PARAM.1712
diff -X source_excludes -Nur MM5/domain/nest/initnest.F MM5_n2/domain/nest/initnest.F
--- MM5/domain/nest/initnest.F 2004-12-23 00:53:48.000000000 +0200
+++ MM5_n2/domain/nest/initnest.F 2006-10-13 12:33:27.000000000 +0200
@@ -2159,7 +2159,7 @@
900 CONTINUE INITNEST.1855
#ifdef FDDA0BS1 INITNEST.1856
IF(I4DI(NUMNES).EQ.1)THEN INITNEST.1857
- CALL IN4DOB(NUMNES,XTIME,KTAU,KTAUR,DTMIN,JULDAY,GMT) INITNEST.1858
+ CALL IN4DOB(NUMNES) INITNEST.1858
PRINT *, ' CALLING IN4DOB FOR DOMAIN ',NUMNES INITNEST.1859
ENDIF INITNEST.1860
#endif INITNEST.1861
diff -X source_excludes -Nur MM5/fdda/obs/in4dob.F MM5_n2/fdda/obs/in4dob.F
--- MM5/fdda/obs/in4dob.F 1999-08-01 01:19:25.000000000 +0200
+++ MM5_n2/fdda/obs/in4dob.F 2006-10-26 14:41:14.000000000 +0200
@@ -1,6 +1,6 @@
C IN4DOB.1
C IN4DOB.2
- SUBROUTINE IN4DOB(INEST,XTIME,KTAU,KTAUR,DTMIN,JULDAY,GMT) IN4DOB.3
+ SUBROUTINE IN4DOB(INEST) IN4DOB.3
C THIS IS SUBROUTINE READS AN OBSERVATION DATA FILE AND IN4DOB.4

```

```

C SELECTS ONLY THOSE VALUES OBSERVED AT TIMES THAT FALL
C WITHIN A TIME WINDOW (TWINDO) CENTERED ABOUT THE CURRENT
C
C -14,14 +14,22 C
C      IVAR=3   TOBS
C      IVAR=4   QOBS
C      IVAR=5   PSOBS (CROSS)
+C Note also: Kilian Hagemann modified this source code so that it works
+C correctly over year boundaries
#      include <parame.incl>
#      include <param2.incl>
#      include <parfdda.incl>
#      include <param3.incl>
#      include <fddaob.incl>
+#      include <various.incl>
+      INTEGER ITIMOB
+      REAL RT0, RTIMOB
+      LOGICAL OPENED
+      DATA NMOVE/0/, IEOF/0/, NVOLA/61/
+      IF(I4DI(INEST).NE.1)RETURN
+C Debug statements if desired
+C      PRINT *, 'Calling IN4DOB with INEST', INEST, 'XTIME', XTIME,
+C      & 'KTAU', KTAU, 'KTAUR', KTAUR, 'GMT', GMT
+      IF(KTAU.EQ.0.OR.KTAU.EQ.KTAUR)THEN
+          DO 5 N=1,NIOBF
+              TIMEOB(N)=99999.
C
C -38,9 +46,12 C
+      10 CONTINUE
+      TBACK=XHOUR-WINDOW
+      TFORWD=XHOUR+WINDOW
+C Debug statements if desired
+C      PRINT *, 'Proceeding in IN4DOB with TBACK', TBACK, 'TFORWD',
+C      & TFORWD, 'XHOUR', XHOUR, 'NSTA', NSTA
C CHECK FOR OLD OBS THAT CAN BE DISCARDED
C      IF(KTAU.EQ.0.OR.KTAU.EQ.KTAUR)THEN
-      RTLAST(INEST)=-999.
+      ITLAST(INEST)=-999
+      INQUIRE (NVOLA+INEST-1,OPENED=OPENED)
+      IF (.NOT. OPENED) THEN
+          OPEN(NVOLA+INEST-1,
C
C -48,23 +59,22 C
+      &      FORM='UNFORMATTED',STATUS='OLD')
+      ENDIF
+      #ifndef MPP1
-      11 READ (NVOLA+INEST-1,END=12)RTIMOB,RIO(1),RJO(1),RKO(1),(
+      11 READ (NVOLA+INEST-1,END=12)ITIMOB,RIO(1),RJO(1),RKO(1),(
+          +          VAROBS(IVAR,1),IVAR=1,5)
+C      PRINT *, 'FIRST IN4DOB READ: ITIMOB', ITIMOB, 'RIO', RIO(1),
+C      +      'RJO', RJO(1), 'RKO', RKO(1)
+      #else
+      11 CONTINUE
+      # include "mpp_in4dob_10.incl"
+      #endif
-      IF(RTIMOB.LT.RTLAST(INEST))THEN
+      IF(ITIMOB.LT.ITLAST(INEST))THEN
+      PRINT *, 'OBS ARE NOT IN CHRONOLOGICAL ORDER'

```

```

        PRINT *, 'NEW YEAR?'
        STOP 11
    ELSE
-       RTLAST(INEST)=RTIMOB
+       ITLAST(INEST)=ITIMOB
    ENDIF
-       JULOB=RTIMOB/100.+0.000001
-       RJULOB=FLOAT(JULOB)*100.
-       HOUROB=RTIMOB-RJULOB
-       RTIMOB=FLOAT(JULOB-JULDAY)*24.-GMT+HOUROB
+       RTIMOB=RTO(ITIMOB,JULDAY,GMT,MYEAR)
        TIMEOB(1)=RTIMOB
        IF(RTIMOB.LT.TBACK-WINDOW)GOTO 11
    ENDIF
@@ -138,27 +148,24 @@
        NVOL=NVOLA+INEST-1
        IF(IEODI(INEST).EQ.1)GOTO 111
    #ifndef MPP1
-       READ (NVOL,END=111)TIMEOB(N),RIO(N),RJO(N),RKO(N),(VAROBS(IVAR,N),
+       READ (NVOL,END=111)ITIMOB,RIO(N),RJO(N),RKO(N),(VAROBS(IVAR,N),
+           +           IVAR=1,5)
+C     PRINT *, 'SECOND IN4DOB READ: ITIMOB', ITIMOB, 'RIO(N)', RIO(N),
+C     +           'RJO(N)', RJO(N), 'RKO(N)', RKO(N)
    #else
        #include "mpp_in4dob_20.incl"
    #endif
-       IF(TIMEOB(N).LT.RTLAST(INEST))THEN
+       IF(ITIMOB.LT.ITLAST(INEST))THEN
        PRINT *, 'OBS ARE NOT IN CHRONOLOGICAL ORDER'
        PRINT *, 'NEW YEAR?'
        STOP 111
    ELSE
-       RTLAST(INEST)=TIMEOB(N)
+       ITLAST(INEST)=ITIMOB
    ENDIF
    CCCCCC DATA MUST BE WITHIN VERTICAL DOMAIN CCCCCCCCCC
        IF(RKO(N).GT.FLOAT(KX).OR.RKO(N).LT.1.)STOP 70
- C CONVERT TIMEOB FROM JULIAN DATE AND GMT FORM TO FORECAST
- C TIME IN HOURS (EX. TIMEOB=13002.4 REPRESENTS JULDAY 130
- C AND GMT (HOUR) = 2.4)
-       JULOB=TIMEOB(N)/100.+0.000001
-       RJULOB=FLOAT(JULOB)*100.
-       HOUROB=TIMEOB(N)-RJULOB
-       TIMEOB(N)=FLOAT(JULOB-JULDAY)*24.-GMT+HOUROB
+       RTIMOB=RTO(ITIMOB,JULDAY,GMT,MYEAR)
+       TIMEOB(N)=RTIMOB
        IF(IDYNIN.EQ.1.AND.TIMEOB(N)*60..GT.DATEND(INEST))THEN
        PRINT*, ' IN4DOB: FOR INEST = ',INEST,' AT XTIME = ',XTIME,
1' TIMEOB = ',TIMEOB(N)*60.,' AND DATEND = ',DATEND(INEST),' : '
@@ -317,3 +324,53 @@
        RETURN
        END
+
+     REAL FUNCTION RTO(ITIMOB,JULDAY,GMT,MYEAR)

```

```

+C This is my own new way with obs date as YYYYdddhf INTEGER
+C CONVERT OBS TIME FROM YEAR, JULIAN DATE AND GMT YYYYdddhf INTEGER FORM
+C TO FORECAST TIME IN HOURS (EX. TIMEOB=1995130024 REPRESENTS YEAR 1995
+C JULDAY 130 AND GMT (HOUR) = 2.4)
+     INTEGER ITIMOB,JULDAY,GMT,YEAR, JULOB, IJULOB, DAYS, Y
+     REAL HOUROB
+     LOGICAL DATBAD
+     EXTERNAL NFEB
+C Extract year, julian day and hour of observation
+     YEAR=ITIMOB/1000000
+     JULOB=MOD(ITIMOB,1000000)/1000
+     IJULOB=JULOB*1000
+     HOUROB=FLOAT(ITIMOB-IJULOB-YEAR*1000000)/10.
+C Consistency check
+     DATBAD=.FALSE.
+C Maximum 4 byte signed int is 2147483647, min year is arbitrary
+     IF ((YEAR.LT.1000).OR.(YEAR.GT.2146)) DATBAD=.TRUE.
+     IF ((JULOB.LT.1).OR.(JULOB.GT.(337+NFEB(YEAR)))) DATBAD=.TRUE.
+     IF (HOUROB.GT.23.91) DATBAD=.TRUE.
+     IF (DATBAD) THEN
+         PRINT *, 'INVALID DATETIME SPECIFICATION: ', ITIMOB
+         PRINT *, 'ENSURE OBS TIME IS IN YYYYdddhf INTEGER FORMAT'
+         PRINT *, 'AND THAT DATETIME FIELDS ARE WITHIN PROPER BOUNDS'
+         STOP 69
+     ENDIF
+C Now figure out the model time in hours from this information
+     IF (YEAR.EQ.MYEAR) THEN
+         DAYS=JULOB-JULDAY
+     ELSE
+C Completed days in starting year
+         DAYS=337+NFEB(MYEAR)-JULDAY
+         DO 10 Y=MYEAR+1, YEAR-1, 1
+C Plus all the days in the years between
+         DAYS=DAYS+337+NFEB(Y)
+     10     CONTINUE
+C Plus the current years' julian day
+         DAYS=DAYS+JULOB
+     ENDIF
+     RTO=FLOAT(DAYS)*24.-GMT+HOUROB
+C Debug statements if desired
+C     PRINT *, 'ITIMOB READ IN:', ITIMOB
+C     PRINT *, 'FIELDS EXTRACTED: YEAR=', YEAR, ' JULIAN DAY=', JULOB,
+C     &         'HOUR=', HOUROB
+C     PRINT *, 'DAYS SINCE MODEL START = ', DAYS
+C     PRINT *, 'HOURS SINCE MODEL START =', RTO
+C     PRINT *, 'INITIAL STARTING YEAR = ', MYEAR
+     RETURN
+     END
diff -X source_excludes -Nur MM5/fdda/util/setfd.F MM5_n2/fdda/util/setfd.F
--- MM5/fdda/util/setfd.F      2000-11-07 19:45:50.000000000 +0200
+++ MM5_n2/fdda/util/setfd.F   2006-10-13 12:31:36.000000000 +0200
@@ -108,7 +108,7 @@
ENDIF
INFR=IONF*(IRATIO**LEVIDN(INEST))
IF((KTAU.GT.KTAUR.AND.MOD(KTAU,INFR).EQ.0).OR.(KTAU.EQ.KTAUR.AND.
SETFD.108
SETFD.109
SETFD.110

```

```

- + IFREST))CALL IN4DOB(INEST,XTIME,KTAU,KTAUR,DTMIN,JULDAY,GMT) SETFD.111
+ + IFREST))CALL IN4DOB(INEST) SETFD.111
C SETFD.112
    IF(IFREST.AND.KTAU.EQ.KTAUR.AND.I4DI(INEST).EQ.1.AND.NSTA.GT.0) SETFD.113
+ THEN SETFD.114

diff -X source_excludes -Nur MM5/include/parfdda.incl MM5_n2/include/parfdda.incl
--- MM5/include/parfdda.incl 1999-08-01 01:19:51.000000000 +0200
+++ MM5_n2/include/parfdda.incl 2006-10-10 17:01:20.000000000 +0200
@@ -1,15 +1,15 @@
    REAL DATEND,DCON,DIFTIM,DIFZ,DPSMX,DTRAMP,FDAEND,FDASTA,GIP,GIQ, PARFDDA.1
- 1 GIT,GIV,PFREE,RINBLW,RINFMN,RINFMX,RINP,RINSIG,RINXY,RTLAST, PARFDDA.2
+ 1 GIT,GIV,PFREE,RINBLW,RINFMN,RINFMX,RINP,RINSIG,RINXY, PARFDDA.2
 2 TFACI,TIMANL,TIMOBP,TWINDO,WNDFAC,XTIM1,XTIM2,XTIMT,XTNEX PARFDDA.3
  INTEGER I4D,I4DI,I4DITOT,IBLNOP,IDQ,IDQN,IDYNIN,IEOD,IEODI,IMOIS, PARFDDA.4
 1 IMOISP,INDT,INDXY,INDZ,INONBL,INOPRO,INT4,IONF,IPSTR,IROT,ISMOIS, PARFDDA.5
- 2 ISPSTR,ISTEMP,ISWIND,ITEMP,IUN1,IUN2,IWIND,IWINDS,IWTSIG,NPF, PARFDDA.6
- 3 NPGF,NPFI,NPFV,NSTAT PARFDDA.7
+ 2 ISPSTR,ISTEMP,ISWIND,ITEMP,ITLAST,IUN1,IUN2,IWIND,IWINDS,IWTSIG, PARFDDA.6
+ 3 NPF,NPGF,NPFI,NPFV,NSTAT PARFDDA.7
  COMMON/PARFDDA/ I4D,IWIND,ITEMP,IMOIS,IPSTR, PARFDDA.8
 3 INT4,INDXY,INDT,INDZ,INONBL, PARFDDA.9
 4 IUN1,IUN2,XTIMT,XTIM1,XTIM2,XTNEX, PARFDDA.10
 5 ISWIND,ISMOIS,ISTEMP,ISPSTR, PARFDDA.11
- 6 GIV,GIT,GIQ,GIP,RINXY,RINSIG,I4DI,TWINDO,RTLAST, PARFDDA.12
+ 6 GIV,GIT,GIQ,GIP,RINXY,RINSIG,I4DI,TWINDO,ITLAST, PARFDDA.12
 7 TIMANL,IROT,IMOISP,RINP,IDQ,IDQN,TIMOBP, PARFDDA.13
 8 INOPRO,DIFTIM,IBLNOP,NPF,NPGF,NPFI,RINBLW, PARFDDA.14
 9 IWINDS,NSTAT,NPFV,IWTSIG,RINFMX,RINFMN,DPSMX, PARFDDA.15
@@ -22,7 +22,7 @@
 2 XTIMT(MAXSES),XTIM1(MAXSES),XTIM2(MAXSES),ISWIND(MAXSES), PARFDDA.22
 4 ISMOIS(MAXSES),ISTEMP(MAXSES),ISPSTR(MAXSES),GIV(MAXSES), PARFDDA.23
 5 GIQ(MAXSES),GIT(MAXSES),GIP(MAXSES),I4DI(MAXSES), PARFDDA.24
- 5 RTLAST(MAXSES),TIMANL(MAXSES),FDASTA(MAXSES),FDAEND(MAXSES), PARFDDA.25
+ 5 ITLAST(MAXSES),TIMANL(MAXSES),FDASTA(MAXSES),FDAEND(MAXSES), PARFDDA.25
 6 ISTAT(MAXSES),IROT(MAXSES),IMOISP(MAXSES),TIMOBP(MAXSES), PARFDDA.26
 7 INOPRO(MAXSES,2),IBLNOP(MAXSES,4),DIFTIM(MAXSES,2), PARFDDA.27
 8 INONBL(MAXSES,4),IWINDS(MAXSES,2),NSTAT(MAXSES),IWTSIG(MAXSES), PARFDDA.28

diff -X source_excludes -Nur MM5/include/various.incl MM5_n2/include/various.incl
--- MM5/include/various.incl 1999-08-01 01:20:06.000000000 +0200
+++ MM5_n2/include/various.incl 2006-10-24 19:14:36.000000000 +0200
@@ -31,3 +31,5 @@
 4 IADMI,IADMA,IAQMI,IAQMA,IAQME, VARIOUS.31
 5 IAQMR,IAFNG,IAGNG VARIOUS.32
C VARIOUS.33
+ INTEGER MYEAR
+ COMMON/OBYEAR/ MYEAR

```

It turns out that there is another bug in the standard vtran.F code which consistently caused a segmentation fault during the sensitivity simulations of chapter 4 on page 57 when observational nudging was switched on. It seems to be caused by the fact that the number of vertical levels exceeded the horizontal dimensions for both domains. The following patch fixes this problem, more informa-

tion is in the mailing list post (see <http://mailman.ucar.edu/pipermail/mm5-users/2006/001050.html>).

```

--- MM5/domain/io/vtran.F      1999-08-01 01:18:39.000000000 +0200
+++ MM5_n2/domain/io/vtran.F   2008-03-02 12:18:29.000000000 +0200
@@ -4,7 +4,7 @@
C      PURPOSE  GET RID OF LIB=INSLIB ON THE SEGLDR CARD, THIS DOES A      VTRAN.4
C                  SIMPLE VECTOR SLAB TRANSPOSE.                          VTRAN.5
C                                                                VTRAN.6
-   DIMENSION FXY(MIX*MJX) ,FXY(IMAX*JMAX)                             VTRAN.7
+   DIMENSION FXY(IMAX*JMAX) ,FXY(IMAX*JMAX)                             VTRAN.7
C                                                                VTRAN.8
C      ICOUNT=0                                                         VTRAN.9
C      DO 500 I=1,IMAX                                                  VTRAN.10

```

C.3 Output Reduction Patch

In order to save on disk space and post-processing time the author applied the following patch to the MM5 source code which limits the model output to only those variables which are absolutely necessary in the analysis, namely *u*, *v*, temperature and selected terrain and surface fields. This turned out to reduce the MMOUT* files by a factor of three in size which is a significant advantage when performing dozens of simulations.

The modification itself is fairly trivial and mostly consists of removing output statements for fields which are not strictly required for wind evaluation.

```

--- MM5/domain/io/outtapi.F     2005-05-12 21:35:46.000000000 +0200
+++ MM5_n2/domain/io/outtapi.F  2008-06-08 16:20:15.000000000 +0200
@@ -766,259 +766,6 @@
+   +   'TEMPERATURE',                                                    OUTTAP.685
+   +   IX,JX,KZZ,MIX,MJX,MXX )                                           OUTTAP.686
C                                                                OUTTAP.687
-C----- OPTIONAL 3D MOISTURE VARIABLES                                OUTTAP.688
-C                                                                OUTTAP.689
-C      1) QV                                                                OUTTAP.690
-C      2) CLOUD WATER                                                    OUTTAP.691
-C      3) RAIN WATER                                                    OUTTAP.692
-C      4) SNOW                                                            OUTTAP.693
-C      5) ICE                                                            OUTTAP.694
-C      6) GRAUPEL                                                        OUTTAP.695
-C      7) ICE NUMBER CONCENTRATION                                       OUTTAP.696
-C                                                                OUTTAP.697
-   IF (IMPHYS(NESCOU).GE.2) THEN                                         OUTTAP.698
-                                                                OUTTAP.699
-   CALL FILLCRS(QVA,MIX,MJX,MXX,IL,JL,KL)                               07NOV00.419
-   CALL DECOUPLE(QVA,MIX,MJX,MXX,PSA,HDOH,IL,JL,KL)                     OUTTAP.700
-   CALL WRITE_FIELDREC( iutl,D3,nescou,HDOH,XTIME,                      OUTTAP.701
-   +   'Q',CURRENT_DATE,                                                 OUTTAP.702
-   +   'C','YXS','kg/kg',                                               OUTTAP.703

```

```

- +          'MIXING RATIO',                                OUTTAP.704
- +          IX,JX,KZZ,MIX,MJX,MKX )                        OUTTAP.705
-                                                         OUTTAP.706
-C                                                         OUTTAP.707
-          IF (IMPHYS(NESCOU).GE.3)THEN                     OUTTAP.708
-C                                                         OUTTAP.709
-          CALL FILLCRS(QCA,MIX,MJX,MKX,IL,JL,KL)           07NOV00.420
-          CALL DECOUPLE(QCA,MIX,MJX,MKX,PSA,HDOH,IL,JL,KL) OUTTAP.710
-          CALL WRITE_FIELDREC( iut1,D3,nescou,HDOH,XTIME,   OUTTAP.711
- +          'CLW',CURRENT_DATE,                             OUTTAP.712
- +          'C','YXS','kg/kg',                               OUTTAP.713
- +          'CLOUD WATER MIXING RATIO',                     OUTTAP.714
- +          IX,JX,KZZ,MIX,MJX,MKX )                         OUTTAP.715
-                                                         OUTTAP.716
-C                                                         OUTTAP.717
-          CALL FILLCRS(QRA,MIX,MJX,MKX,IL,JL,KL)           07NOV00.421
-          CALL DECOUPLE(QRA,MIX,MJX,MKX,PSA,HDOH,IL,JL,KL) OUTTAP.718
-          CALL WRITE_FIELDREC( iut1,D3,nescou,HDOH,XTIME,   OUTTAP.719
- +          'RNW',CURRENT_DATE,                             OUTTAP.720
- +          'C','YXS','kg/kg',                               OUTTAP.721
- +          'RAIN WATER MIXING RATIO',                     OUTTAP.722
- +          IX,JX,KZZ,MIX,MJX,MKX )                         OUTTAP.723
-                                                         OUTTAP.724
-C                                                         OUTTAP.725
-          IF (IMPHYS(NESCOU).GE.5)THEN                     OUTTAP.726
-                                                         OUTTAP.727
-          CALL FILLCRS(QIA,MIX,MJX,MKX,IL,JL,KL)           07NOV00.422
-          CALL DECOUPLE(QIA,MIX,MJX,MKX,PSA,HDOH,IL,JL,KL) OUTTAP.728
-          CALL WRITE_FIELDREC( iut1,D3,nescou,HDOH,XTIME,   OUTTAP.729
- +          'ICE',CURRENT_DATE,                             OUTTAP.730
- +          'C','YXS','kg/kg',                               OUTTAP.731
- +          'CLOUD ICE MIXING RATIO',                       07NOV00.423
- +          IX,JX,KZZ,MIX,MJX,MKX )                         OUTTAP.733
-                                                         OUTTAP.734
-C                                                         OUTTAP.735
-          CALL FILLCRS(QNIA,MIX,MJX,MKX,IL,JL,KL)          07NOV00.424
-          CALL DECOUPLE(QNIA,MIX,MJX,MKX,PSA,HDOH,IL,JL,KL) OUTTAP.736
-          CALL WRITE_FIELDREC( iut1,D3,nescou,HDOH,XTIME,   OUTTAP.737
- +          'SNOW',CURRENT_DATE,                             OUTTAP.738
- +          'C','YXS','kg/kg',                               OUTTAP.739
- +          'SNOW MIXING RATIO',                             OUTTAP.740
- +          IX,JX,KZZ,MIX,MJX,MKX )                         OUTTAP.741
-                                                         OUTTAP.742
-          ENDIF                                             OUTTAP.743
-C                                                         OUTTAP.744
-          IF (IMPHYS(NESCOU).GE.6)THEN                     OUTTAP.745
-                                                         OUTTAP.746
-          CALL FILLCRS(QGA,MIX,MJX,MKX,IL,JL,KL)           07NOV00.425
-          CALL DECOUPLE(QGA,MIX,MJX,MKX,PSA,HDOH,IL,JL,KL) OUTTAP.747
-          CALL WRITE_FIELDREC( iut1,D3,nescou,HDOH,XTIME,   OUTTAP.748
- +          'GRAUPEL',CURRENT_DATE,                         OUTTAP.749
- +          'C','YXS','kg/kg',                               OUTTAP.750
- +          'GRAUPEL MIXING RATIO',                         OUTTAP.751
- +          IX,JX,KZZ,MIX,MJX,MKX )                         OUTTAP.752
-                                                         OUTTAP.753

```

```

-C
-      IF (IMPHYS(NESCOU).EQ.7) THEN
-          CALL FILLCRS(QNCA,MIX,MJX,MKX,IL,JL,KL)
-          CALL DECOUPLE(QNCA,MIX,MJX,MKX,PSA,HDOTH2,IL,JL,KL)
-          CALL RHO_MLT(HDOTH2,IL,JL,KL,PSB,HDOTH,IL,JL,KL)
-          CALL WRITE_FIELDREC( iutl,D3,nescou,HDOTH,XTIME,
-      +          'NCI',CURRENT_DATE,
-      +          'C','YXS','J/m^3',
-      +          'NUMBER CONCENTRATION OF ICE',
-      +          IX,JX,KZZ,MIX,MJX,MKX )
-
-      ENDIF
-      ENDIF
-      ENDIF
-      ENDIF
-C
-C----- OPTIONAL NAVY PBL SCHEME 3D FORECAST VARIABLES
-C
-C      1) TURBULENT KINETIC ENERGY
-C
-      IF (IBLTYP(NESCOU).EQ.3.OR.IBLTYP(NESCOU).EQ.4) THEN
-
-          CALL FILLCRS(TURKE,MIX,MJX,MKX,IL,JL,KL)
-          CALL EQUATO(TURKE,MIX,MJX,MKX,HDOTH,IL,JL,KL)
-          CALL WRITE_FIELDREC( iutl,D3,nescou,HDOTH,XTIME,
-      +          'TKE',CURRENT_DATE,
-      +          'C','YXS','J/kg',
-      +          'TURBULENT KINETIC ENERGY',
-      +          IX,JX,KZZ,MIX,MJX,MKX )
-      ENDIF
-
-      IF (IBLTYP(NESCOU).EQ.6) THEN
-
-          CALL FILLCRS(TKEA,MIX,MJX,MKX+1,IL,JL,KL+1)
-          CALL EQUATO(TKEA,MIX,MJX,MKX+1,HDOTH,IL,JL,KL+1)
-          CALL WRITE_FIELDREC( iutl,D3,nescou,HDOTH,XTIME,
-      +          'TKE',CURRENT_DATE,
-      +          'C','YXW','J/kg',
-      +          'TURBULENT KINETIC ENERGY',
-      +          IX,JX,KZZ+1,MIX,MJX,MKX+1 )
-      ENDIF
-C
-C----- OPTIONAL ATMOSPHERIC RADIATION VARIABLE
-C
-C      1) ATMOSPHERIC RADIATION TENDENCY
-C
-      IF (IFRAD.GE.2) THEN
-          CALL FILLCRS(RTTEN,MIX,MJX,MKX,IL,JL,KL)
-          CALL EQUATO(RTTEN,MIX,MJX,MKX,HDOTH,IL,JL,KL)
-#ifndef MPP1
-      DO I=1,IL*JL*KL
-          HDOTH(I)=HDOTH(I)*86400.
-      ENDDO
-#else
-      DO I=1,MIX*MJX*MKX

```

```

-          HDOTH(I)=HDOTH(I)*86400.                                OUTTAP.802
-          ENDDO                                                    OUTTAP.803
-#endif                                                            OUTTAP.804
-          CALL WRITE_FIELDREC( iutl,D3,nescou,HDOTH,XTIME,        OUTTAP.805
-          +          'RAD TEND',CURRENT_DATE,                    OUTTAP.806
-          +          'C','YXS','K/DAY',                          OUTTAP.807
-          +          'ATMOSPHERIC RADIATION TENDENCY',            OUTTAP.808
-          +          IX,JX,KZZ,MIX,MJX,MKX )                     OUTTAP.809
-          ENDIF                                                    OUTTAP.810
-C                                                                    OUTTAP.811
-C----- NON-HYDROSTATIC 3D FORECAST VARIABLES                  OUTTAP.812
-C                                                                    OUTTAP.813
-C          1) VERTICAL WIND COMPONENT                             OUTTAP.814
-C          2) PRESSURE PERTURBATION                              OUTTAP.815
-C                                                                    OUTTAP.816
-          CALL FILLCRS(WA,MIX,MJX,MKX+1,IL,JL,KL+1)              07NOV00.430
-          CALL DECOUPLE(WA,MIX,MJX,MKX+1,PSA,HDOTH,IL,JL,KL+1)    OUTTAP.817
-          CALL WRITE_FIELDREC( iutl,D3,nescou,HDOTH,XTIME,        OUTTAP.818
-          +          'W',CURRENT_DATE,                            OUTTAP.819
-          +          'C','YXW','m/s',                             OUTTAP.820
-          +          'VERTICAL WIND COMPONENT',                   OUTTAP.821
-          +          IX,JX,KZZ+1,MIX,MJX,MKX+1 )                  OUTTAP.822
-                                                                    OUTTAP.823
-C                                                                    OUTTAP.824
-          CALL FILLCRS(PPA,MIX,MJX,MKX,IL,JL,KL)                 07NOV00.431
-          CALL DECOUPLE(PPA,MIX,MJX,MKX,PSA,HDOTH,IL,JL,KL)        OUTTAP.825
-          CALL WRITE_FIELDREC( iutl,D3,nescou,HDOTH,XTIME,        OUTTAP.826
-          +          'PP',CURRENT_DATE,                           OUTTAP.827
-          +          'C','YXS','Pa',                               OUTTAP.828
-          +          'PRESSURE PERTURBATION',                     OUTTAP.829
-          +          IX,JX,KZZ,MIX,MJX,MKX )                       OUTTAP.830
-                                                                    OUTTAP.831
-C----- 2D FIELDS                                              OUTTAP.832
-C                                                                    OUTTAP.833
-C          1) PSTAR (in Pascal)                                   OUTTAP.834
-C                                                                    OUTTAP.835
-          DO I=1,ILX                                              OUTTAP.836
-          DO J=1,JLX                                              OUTTAP.837
-          HSCRL(I,J)=PSA(I,J)*1000.                               OUTTAP.838
-          ENDDO                                                    OUTTAP.839
-          ENDDO                                                    OUTTAP.840
-          CALL EQUATO(HSCRL,MIX,MJX,1,HDOTH,IL,JL,1)              24SEP99.148
-          CALL WRITE_FIELDREC( iutl,D2,nescou,HDOTH,XTIME,        OUTTAP.842
-          +          'PSTARCRS',CURRENT_DATE,                     OUTTAP.843
-          +          'C','YX','Pa',                                OUTTAP.844
-          +          '(REFERENCE) SURFACE PRESSURE MINUS PTOP',    OUTTAP.845
-          +          IX,JX,1,MIX,MJX,1 )                           OUTTAP.846
-                                                                    OUTTAP.847
-C----- 2D FORECAST VARIABLES                                  OUTTAP.848
-C                                                                    OUTTAP.849
-C          1) GROUND TEMPERATURE                                 OUTTAP.850
-C                                                                    OUTTAP.851
-          CALL EQUATO(TGA,MIX,MJX,1,HDOTH,IL,JL,1)                24SEP99.149
-          CALL WRITE_FIELDREC( iutl,D2,nescou,HDOTH,XTIME,        OUTTAP.853
-          +          'GROUND T',CURRENT_DATE,                     OUTTAP.854

```

```

-      +                      'C','YX','K',                      OUTTAP.855
-      +                      'GROUND TEMPERATURE',             OUTTAP.856
-      +                      IX,JX,1,MIX,MJX,1 )                 OUTTAP.857
-                                                                OUTTAP.858
-C                                                                OUTTAP.859
-C----- 2D FORECAST VARIABLES                                  OUTTAP.860
-C                                                                OUTTAP.861
-C          1) ACCUMULATED CONVECTIVE PRECIPITATION              OUTTAP.862
-C          2) ACCUMULATED RESOLVED PRECIPITATION                OUTTAP.863
-C                                                                OUTTAP.864
-      CALL EQUATO(RAINC,MIX,MJX,1,HDOH,IL,JL,1)                24SEP99.150
-      CALL WRITE_FIELDREC( iutl,D2,nescou,HDOH,XTIME,          OUTTAP.866
-      +                      'RAIN CON',CURRENT_DATE,           OUTTAP.867
-      +                      'C','YX','cm',                      OUTTAP.868
-      +                      'ACCUMULATED CONVECTIVE PRECIPITATION', OUTTAP.869
-      +                      IX,JX,1,MIX,MJX,1 )                 OUTTAP.870
-                                                                OUTTAP.871
-C                                                                OUTTAP.872
-      CALL EQUATO(RAINNC,MIX,MJX,1,HDOH,IL,JL,1)                24SEP99.151
-      CALL WRITE_FIELDREC( iutl,D2,nescou,HDOH,XTIME,          OUTTAP.874
-      +                      'RAIN NON',CURRENT_DATE,           OUTTAP.875
-      +                      'C','YX','cm',                      OUTTAP.876
-      +                      'ACCUMULATED NONCONVECTIVE PRECIPITATION', OUTTAP.877
-      +                      IX,JX,1,MIX,MJX,1 )                 OUTTAP.878
-C                                                                OUTTAP.879
-C----- OPTIONAL 2D FORECAST VARIABLE                          OUTTAP.880
-C                                                                OUTTAP.881
-C          1) VISIBILITY                                         OUTTAP.882
-C          2) FOG DEPTH                                          OUTTAP.883
-C                                                                OUTTAP.884
-      IF(IFOGSW(NESCOU).EQ.1)THEN                                OUTTAP.885
-      CALL EQUATO(VISB(NESCOU,1,1),MIX,MJX,1,HDOH,IL,JL,1)    24SEP99.152
-      CALL WRITE_FIELDREC( iutl,D2,nescou,HDOH,XTIME,          OUTTAP.887
-      +                      'VISB',CURRENT_DATE,                OUTTAP.888
-      +                      'C','YX','m',                        OUTTAP.889
-      +                      'VISIBILITY',                        OUTTAP.890
-      +                      IX,JX,1,MIX,MJX,1 )                 OUTTAP.891
-                                                                OUTTAP.892
-      CALL EQUATO(FOGDEP(NESCOU,1,1),MIX,MJX,1,HDOH,IL,JL,1)  24SEP99.153
-      CALL WRITE_FIELDREC( iutl,D2,nescou,HDOH,XTIME,          OUTTAP.894
-      +                      'FOGDEPTH',CURRENT_DATE,           OUTTAP.895
-      +                      'C','YX','m',                        OUTTAP.896
-      +                      'FOG DEPTH',                        OUTTAP.897
-      +                      IX,JX,1,MIX,MJX,1 )                 OUTTAP.898
-      ENDIF                                                      OUTTAP.899
-C                                                                05DEC01.250
-C OPTIONAL KF-2 VARIABLES                                       05DEC01.251
-C                                                                05DEC01.252
-      IF(ICUPA(NESCOU).EQ.8)THEN                                05DEC01.253
-      CALL EQUATO(SUMFB,MIX,MJX,1,HDOH,IL,JL,1)                05DEC01.254
-      CALL WRITE_FIELDREC( iutl,D2,nescou,HDOH,XTIME,          05DEC01.255
-      +                      'SUMFB',CURRENT_DATE,                05DEC01.256
-      +                      'C','YX','m',                        05DEC01.257
-      +                      'KF MASS FLUX UPDRAFT   MAX VALUE IN PREVIOUS 05DEC01.258
-      +                      & HOUR',                             05DEC01.259

```

```

-      +                      IX,JX,1,MIX,MJX,1 )                      05DEC01.260
-                                                                    05DEC01.261
-      CALL EQUATO(SPSRC,MIX,MJX,1,HDOTH,IL,JL,1)                    05DEC01.262
-      CALL WRITE_FIELDREC( iutl,D2,nescou,HDOTH,XTIME,              05DEC01.263
-      +                      'SPSRC',CURRENT_DATE,                  05DEC01.264
-      +                      'C','YX','m',                          05DEC01.265
-      +                      'KF SOURCE LYR UPDRAFT  MAX VALUE IN PREVIOUS  05DEC01.266
-      & HOUR',                                                       05DEC01.267
-      +                      IX,JX,1,MIX,MJX,1 )                      05DEC01.268
-      DO 773 I = 1,MIX                                              05DEC01.269
-      DO 773 J = 1,MJX                                              05DEC01.270
-      SUMFB(I,J) = 0.                                              05DEC01.271
-      SPSRC(I,J) = 0.                                              05DEC01.272
- 773    CONTINUE                                                  05DEC01.273
-      ENDIF                                                         05DEC01.274
-C                                                                    OUTTAP.900
C----- NON-FORECAST (TERRESTRIAL) 2D ARRAYS                       OUTTAP.901
C                                                                    OUTTAP.902
C      1) TERRAIN                                                    OUTTAP.903
@@ -1068,20 +815,6 @@
-      +                      'D','YX','(DIMENSIONLESS)',            OUTTAP.946
-      +                      'MAP SCALE FACTOR',                    OUTTAP.947
-      +                      IX,JX,1,MIX,MJX,1 )                      OUTTAP.948
-C                                                                    OUTTAP.949
-      CALL EQUATO(F,MIX,MJX,1,HDOTH,IL,JL,1)                        24SEP99.157
-      CALL WRITE_FIELDREC( iutl,D2,nescou,HDOTH,XTIME,              OUTTAP.951
-      +                      'CORIOLIS',CURRENT_DATE,              OUTTAP.952
-      +                      'D','YX','1/s',                        OUTTAP.953
-      +                      'CORIOLIS PARAMETER',                  OUTTAP.954
-      +                      IX,JX,1,MIX,MJX,1 )                      OUTTAP.955
-C                                                                    OUTTAP.956
-      CALL EQUATO(TMN,MIX,MJX,1,HDOTH,IL,JL,1)                      24SEP99.158
-      CALL WRITE_FIELDREC( iutl,D2,nescou,HDOTH,XTIME,              OUTTAP.958
-      +                      'RES TEMP',CURRENT_DATE,              OUTTAP.959
-      +                      'C','YX','K',                          OUTTAP.960
-      +                      'INFINITE RESERVOIR SLAB TEMPERATURE',  OUTTAP.961
-      +                      IX,JX,1,MIX,MJX,1 )                      OUTTAP.962
C                                                                    OUTTAP.963
-      CALL EQUATO(XLAT,MIX,MJX,1,HDOTH,IL,JL,1)                    24SEP99.159
-      CALL WRITE_FIELDREC( iutl,D2,nescou,HDOTH,XTIME,              OUTTAP.965
@@ -1104,37 +837,6 @@
-      +                      'LANDUSE CATEGORY',                    OUTTAP.982
-      +                      IX,JX,1,MIX,MJX,1 )                      OUTTAP.983
C                                                                    19DEC02.562
-      IF (IFSNOU(NESCOU).EQ.1.AND.ISOIL.NE.2)THEN                  19DEC02.563
-      CALL EQUATO(SNOWC,MIX,MJX,1,HDOTH,IL,JL,1)                    05DEC01.276
-      CALL WRITE_FIELDREC( iutl,D2,nescou,HDOTH,XTIME,              05DEC01.277
-      +                      'SNOWCOVR',CURRENT_DATE,              OUTTAP.987
-      +                      'C','YX','(DIMENSIONLESS)',            OUTTAP.988
-      +                      'SNOW COVER FLAGS',                    OUTTAP.989
-      +                      IX,JX,1,MIX,MJX,1 )                      OUTTAP.990
-C                                                                    19DEC02.564
-      ELSE IF (IFSNOU(NESCOU).EQ.2)THEN                            05DEC01.278
-      CALL EQUATO(SNOWC,MIX,MJX,1,HDOTH,IL,JL,1)                    05DEC01.279
-      CALL WRITE_FIELDREC( iutl,D2,nescou,HDOTH,XTIME,              05DEC01.280

```

```

-      +                      'WEASD' , CURRENT_DATE,          19DEC02.565
-      +                      'C' , 'YX' , 'mm' ,                05DEC01.282
-      +                      'SNOW COVER LIQUID EQUIVALENT' ,   05DEC01.283
-      +                      IX,JX,1,MIX,MJX,1 )                05DEC01.284
-      ENDIF                                                    05DEC01.285
-      IF(IPOLAR.EQ.1)THEN                                       19DEC02.566
-      CALL EQUATO(SEAICE,MIX,MJX,1,HDOH,IL,JL,1)               19DEC02.567
-      CALL WRITE_FIELDREC( iut1,D2,nescou,HDOH,XTIME,          19DEC02.568
-      +                      'SEAICEFR',CURRENT_DATE,           19DEC02.569
-      +                      'C' , 'YX' , '(DIMENSIONLESS)' ,   19DEC02.570
-      +                      'SEA ICE FRACTION' ,               19DEC02.571
-      +                      IX,JX,1,MIX,MJX,1 )                19DEC02.572
-      CALL EQUATO(TGSI,MIX,MJX,1,HDOH,IL,JL,1)               23DEC04.233
-      CALL WRITE_FIELDREC( iut1,D2,nescou,HDOH,XTIME,          23DEC04.234
-      +                      'TGSI' , CURRENT_DATE,             23DEC04.235
-      +                      'C' , 'YX' , 'K' ,                 23DEC04.236
-      +                      'SEA ICE TEMPERATURE' ,           23DEC04.237
-      +                      IX,JX,1,MIX,MJX,1 )                23DEC04.238
-      ENDIF                                                    19DEC02.573
-C                                                                24SEP99.163
-      CALL EQUATO(TSS,MIX,MJX,1,HDOH,IL,JL,1)                 24SEP99.164
-      CALL WRITE_FIELDREC( iut1,D2,nescou,HDOH,XTIME,          24SEP99.165
-      +                      'TSEASFC' , CURRENT_DATE,          24SEP99.166
-      +                      'SEA SURFACE TEMPERATURE' ,        24SEP99.168
-      +                      IX,JX,1,MIX,MJX,1 )                24SEP99.169
-      C                                                         OUTTAP.991
-C----- OPTIONAL 2D FORECAST VARIABLES (SINCE V2 RELEASE 2-7)  OUTTAP.992
-C                                                         OUTTAP.993
-C      1) PBL HEIGHT (M)                                       OUTTAP.994
-C      2) PBL REGIME (BULK RI (BLTYP=1) OR 1-4 (IBLTYP=2,5))  OUTTAP.995
-C      NOTE: ABOVE PBL VARIABLES ONLY DEFINED FOR IBLTYP=1,2,5 OUTTAP.996
-C      3) SURFACE SENSIBLE HEAT FLUX (W/M^2)                  OUTTAP.997
-C      4) SURFACE LATENT HEAT FLUX (W/M^2)                    OUTTAP.998
-C      5) FRICTIONAL VELOCITY (M/S)                           OUTTAP.999
-C      6) SHORT WAVE RADIATION (W/M^2)                         OUTTAP.1000
-C      7) LONG WAVE RADIATION (W/M^2)                          OUTTAP.1001
-C      8-MLX) SOIL TEMPERATURE: LAYERS 1-MLX                 OUTTAP.1002
-C                                                         OUTTAP.1003
-      CALL EQUATO(PBL,MIX,MJX,1,HDOH,IL,JL,1)                 24SEP99.170
-      CALL WRITE_FIELDREC( iut1,D2,nescou,HDOH,XTIME,          OUTTAP.1005
-      +                      'PBL HGT' , CURRENT_DATE,          OUTTAP.1006
-      +                      'C' , 'YX' , 'm' ,                 OUTTAP.1007
-      +                      'PBL HEIGHT' ,                     OUTTAP.1008
-      +                      IX,JX,1,MIX,MJX,1 )                OUTTAP.1009
-C                                                         OUTTAP.1010
-      CALL EQUATO(REGIME,MIX,MJX,1,HDOH,IL,JL,1)              24SEP99.171
-      CALL WRITE_FIELDREC( iut1,D2,nescou,HDOH,XTIME,          OUTTAP.1012
-      +                      'REGIME' , CURRENT_DATE,           OUTTAP.1013
-      +                      'C' , 'YX' , '(DIMENSIONLESS)' ,   OUTTAP.1014
-      +                      'PBL REGIME' ,                     OUTTAP.1015
-      +                      IX,JX,1,MIX,MJX,1 )                OUTTAP.1016
-C                                                         OUTTAP.1017
-      CALL EQUATO(HFX,MIX,MJX,1,HDOH,IL,JL,1)                 24SEP99.172
-      CALL WRITE_FIELDREC( iut1,D2,nescou,HDOH,XTIME,          OUTTAP.1019

```

```

- +          'SHFLUX',CURRENT_DATE,                OUTTAP.1020
- +          'C','YX','W/m^2',                      OUTTAP.1021
- +          'SENSIBLE HEAT FLUX',                  OUTTAP.1022
- +          IX,JX,1,MIX,MJX,1 )                    OUTTAP.1023
-C                                                  OUTTAP.1024
- DO I = 1,ILX                                     OUTTAP.1025
-   DO J = 1,JLX                                   07NOV00.432
-     IF (IBLTYP(NESCOU).EQ.7) THEN                07NOV00.433
-       HSCRL(I,J) = QFX(I,J) ! in PX QFX already includes Lv 07NOV00.434
-     ELSE                                          07NOV00.435
-       HSCRL(I,J) = QFX(I,J)*XLV                 07NOV00.436
-     ENDIF                                         07NOV00.437
-   END DO                                         OUTTAP.1028
- END DO                                           OUTTAP.1029
- CALL EQUATO(HSCRL,MIX,MJX,1,HDOH,IL,JL,1)       24SEP99.173
- CALL WRITE_FIELDREC( iutl,D2,nescou,HDOH,XTIME, OUTTAP.1031
- +          'LHFLUX',CURRENT_DATE,                OUTTAP.1032
- +          'C','YX','W/m^2',                      OUTTAP.1033
- +          'LATENT HEAT FLUX',                    OUTTAP.1034
- +          IX,JX,1,MIX,MJX,1 )                    OUTTAP.1035
-C                                                  OUTTAP.1036
- CALL EQUATO(UST,MIX,MJX,1,HDOH,IL,JL,1)         24SEP99.174
- CALL WRITE_FIELDREC( iutl,D2,nescou,HDOH,XTIME, OUTTAP.1038
- +          'UST',CURRENT_DATE,                    OUTTAP.1039
- +          'C','YX','m/s',                        OUTTAP.1040
- +          'FRICTIONAL VELOCITY',                 OUTTAP.1041
- +          IX,JX,1,MIX,MJX,1 )                    OUTTAP.1042
-C                                                  OUTTAP.1043
- DO I = 1,ILX                                     OUTTAP.1044
-   DO J = 1,JLX                                   OUTTAP.1045
-     IF (IBLTYP(NESCOU).EQ.7) THEN                07NOV00.438
-       HSCRL(I,J) = GSW(I,J)/(1.-ALBEDO(I,J))      07NOV00.439
-     ELSE                                          07NOV00.440
-       HSCRL(I,J) = GSW(I,J)/(1.-ALB(I,J))         07NOV00.441
-     ENDIF                                         07NOV00.442
-   END DO                                         OUTTAP.1047
- END DO                                           OUTTAP.1048
- CALL EQUATO(HSCRL,MIX,MJX,1,HDOH,IL,JL,1)       24SEP99.175
- CALL WRITE_FIELDREC( iutl,D2,nescou,HDOH,XTIME, OUTTAP.1050
- +          'SWDOWN',CURRENT_DATE,                 OUTTAP.1051
- +          'C','YX','W/m^2',                      OUTTAP.1052
- +          'SURFACE DOWNWARD SHORTWAVE RADIATION', OUTTAP.1053
- +          IX,JX,1,MIX,MJX,1 )                    OUTTAP.1054
-C                                                  OUTTAP.1055
- CALL EQUATO(GLW,MIX,MJX,1,HDOH,IL,JL,1)         24SEP99.176
- CALL WRITE_FIELDREC( iutl,D2,nescou,HDOH,XTIME, OUTTAP.1057
- +          'LWDOWN',CURRENT_DATE,                 OUTTAP.1058
- +          'C','YX','W/m^2',                      OUTTAP.1059
- +          'SURFACE DOWNWARD LONGWAVE RADIATION', OUTTAP.1060
- +          IX,JX,1,MIX,MJX,1 )                    OUTTAP.1061
- IF (IFRAD.GE.2) THEN                             07NOV00.443
-   CALL EQUATO(OSW,MIX,MJX,1,HDOH,IL,JL,1)       07NOV00.444
-   CALL WRITE_FIELDREC( iutl,D2,nescou,HDOH,XTIME, OUTTAP.1061
- +          'SWOUT',CURRENT_DATE,                  OUTTAP.1062
- +          'C','YX','W/m^2',                      OUTTAP.1063

```

```

-      +                      'TOP OUTGOING SHORTWAVE RADIATION',
-      +                      IX,JX,1,MIX,MJX,1 )
-C
-      CALL EQUATO(OLW,MIX,MJX,1,HDOH,IL,JL,1)
-      CALL WRITE_FIELDREC( iutl,D2,nescou,HDOH,XTIME,
-      +                      'LWOUT',CURRENT_DATE,
-      +                      'C','YX','W/m^2',
-      +                      'TOP OUTGOING LONGWAVE RADIATION',
-      +                      IX,JX,1,MIX,MJX,1 )
-      ENDIF
-C
-      IF(ISOIL.EQ.1) THEN
-      DO L = 1,MLX
-#ifndef MPP1
-      CALL EQUATO(TSLA(1,1,L),MIX,MJX,1,HDOH,IL,JL,1)
-#else
-      CALL EQUATO(TSLA(NOFLIC(1),NOFLIC(1),L),
-      +                      MIX,MJX,1,HDOH,IL,JL,1)
-#endif
-      WRITE (NAME(1:8),FMT=350)L
-      WRITE (DESCRIPTION(1:28),FMT=360)L
-      CALL WRITE_FIELDREC( iutl,D2,nescou,HDOH,XTIME,
-      +                      NAME,CURRENT_DATE,
-      +                      'C','YX','K',
-      +                      DESCRIPTION,
-      +                      IX,JX,1,MIX,MJX,1 )
-      END DO
-      END IF
-350  FORMAT('SOIL T',I2)
-360  FORMAT('SOIL TEMPERATURE IN LAYER ',I2)
-
-      IF(ISOIL.GE.2) THEN
-      DO L = 1,MLX
-#ifndef MPP1
-      CALL EQUATO(STCA(1,1,L),MIX,MJX,1,HDOH,IL,JL,1)
-#else
-      CALL EQUATO(STCA(NOFLIC(1),NOFLIC(1),L),
-      +                      MIX,MJX,1,HDOH,IL,JL,1)
-#endif
-      WRITE (NAME(1:8),FMT=350)L
-      WRITE (DESCRIPTION(1:28),FMT=360)L
-      CALL WRITE_FIELDREC( iutl,D2,nescou,HDOH,XTIME,
-      +                      NAME,CURRENT_DATE,
-      +                      'C','YX','K',
-      +                      DESCRIPTION,
-      +                      IX,JX,1,MIX,MJX,1 )
-      END DO
-C
-      DO L = 1,MLX
-
-#ifndef MPP1
-      CALL EQUATO(SMCA(1,1,L),MIX,MJX,1,HDOH,IL,JL,1)
-#else
-      CALL EQUATO(SMCA(NOFLIC(1),NOFLIC(1),L),
-      +                      MIX,MJX,1,HDOH,IL,JL,1)

```

```

-#endif
-      WRITE (NAME(1:8),FMT=370)L
-      WRITE (DESCRIPTION(1:25),FMT=380)L
-      CALL WRITE_FIELDREC( iut1,D2,nescou,HDOH,XTIME,
-      +      NAME,CURRENT_DATE,
-      +      'C','YX','m^3/m^3',
-      +      DESCRIPTION,
-      +      IX,JX,1,MIX,MJX,1 )
-
-      END DO
-370  FORMAT('SOIL M',I2)
-380  FORMAT('TOTAL SOIL MOIS IN LYR ',I2)
-C
-      DO L = 1,MLX
-
-#ifndef MPP1
-      CALL EQUATO(SWCA(1,1,L),MIX,MJX,1,HDOH,IL,JL,1)
-#else
-      CALL EQUATO(SWCA(NOFLOC(1),NOFLOC(1),L),
-      +      MIX,MJX,1,HDOH,IL,JL,1)
-#endif
-      WRITE (NAME(1:8),FMT=371)L
-      WRITE (DESCRIPTION(1:25),FMT=381)L
-      CALL WRITE_FIELDREC( iut1,D2,nescou,HDOH,XTIME,
-      +      NAME,CURRENT_DATE,
-      +      'C','YX','m^3/m^3',
-      +      DESCRIPTION,
-      +      IX,JX,1,MIX,MJX,1 )
-
-      END DO
-371  FORMAT('SOIL W',I2)
-381  FORMAT('SOIL LQD WATER IN LYR ',I2)
-C
-
-      CALL EQUATO(CMCA,MIX,MJX,1,HDOH,IL,JL,1)
-      CALL WRITE_FIELDREC( iut1,D2,nescou,HDOH,XTIME,
-      +      'CANOPYM ',CURRENT_DATE,
-      +      'C','YX','m',
-      +      'CANOPY MOISTUR E CONTENT',
-      +      IX,JX,1,MIX,MJX,1 )
-C
-      IF(ISOIL.EQ.2) THEN
-      CALL EQUATO(SNOA,MIX,MJX,1,HDOH,IL,JL,1)
-      CALL WRITE_FIELDREC( iut1,D2,nescou,HDOH,XTIME,
-      +      'WEASD ',CURRENT_DATE,
-      +      'C','YX','mm',
-      +      'WATER EQUIVALENT SNOW DEPTH',
-      +      IX,JX,1,MIX,MJX,1 )
-C
-C CONVERT SNOWH TO METERS FOR OUTPUT
-      DO I=1,ILX
-      DO J=1,JLX
-      HSCRL(I,J)=SNOWHA(I,J)*0.001
-      ENDDO
-      ENDDO

```

```

24SEP99.194
OUTTAP.1093
OUTTAP.1094
OUTTAP.1095
OUTTAP.1096
OUTTAP.1097
OUTTAP.1098
OUTTAP.1099
OUTTAP.1100
OUTTAP.1101
OUTTAP.1102
19DEC02.574
19DEC02.575
19DEC02.576
19DEC02.577
19DEC02.578
19DEC02.579
19DEC02.580
19DEC02.581
19DEC02.582
19DEC02.583
19DEC02.584
19DEC02.585
19DEC02.586
19DEC02.587
19DEC02.588
19DEC02.589
19DEC02.590
19DEC02.591
19DEC02.592
19DEC02.593
19DEC02.594
OUTTAP.1104
OUTTAP.1105
24SEP99.195
OUTTAP.1107
OUTTAP.1108
OUTTAP.1109
OUTTAP.1110
OUTTAP.1111
OUTTAP.1112
07NOV00.459
24SEP99.196
OUTTAP.1114
19DEC02.595
OUTTAP.1116
OUTTAP.1117
OUTTAP.1118
24SEP99.197
19DEC02.596
19DEC02.597
19DEC02.598
19DEC02.599
19DEC02.600
19DEC02.601

```

```

-      CALL EQUATO(HSCRL,MIX,MJX,1,HDOH,IL,JL,1) 19DEC02.602
-      CALL WRITE_FIELDREC( iutl,D2,nescou,HDOH,XTIME, 19DEC02.603
-      +      'SNOWH ',CURRENT_DATE, 19DEC02.604
-      +      'C','YX','m', 19DEC02.605
-      +      'PHYSICAL SNOW DEPTH', 19DEC02.606
-      +      IX,JX,1,MIX,MJX,1 ) 19DEC02.607
-C 19DEC02.608
-      CALL EQUATO(SNOWC,MIX,MJX,1,HDOH,IL,JL,1) 19DEC02.609
-      CALL WRITE_FIELDREC( iutl,D2,nescou,HDOH,XTIME, 19DEC02.610
-      +      'SNOWCOVR',CURRENT_DATE, 19DEC02.611
-      +      'C','YX','fraction', 19DEC02.612
-      +      'FRACTIONAL SNOW COVER', 19DEC02.613
-      +      IX,JX,1,MIX,MJX,1 ) 19DEC02.614
-C 19DEC02.615
-      CALL EQUATO(ALB,MIX,MJX,1,HDOH,IL,JL,1) 19DEC02.616
-      CALL WRITE_FIELDREC( iutl,D2,nescou,HDOH,XTIME, 19DEC02.617
-      +      'ALB ',CURRENT_DATE, 19DEC02.618
-      +      'C','YX','fraction', 19DEC02.619
-      +      'ALBEDO', 19DEC02.620
-      +      IX,JX,1,MIX,MJX,1 ) 19DEC02.621
-C 19DEC02.622
-      CALL EQUATO(GRNFLX,MIX,MJX,1,HDOH,IL,JL,1) 19DEC02.623
-      CALL WRITE_FIELDREC( iutl,D2,nescou,HDOH,XTIME, 19DEC02.624
-      +      'GRNFLX ',CURRENT_DATE, 19DEC02.625
-      +      'C','YX','W m{-2}', 19DEC02.626
-      +      'GROUND HEAT FLUX', 19DEC02.627
-      +      IX,JX,1,MIX,MJX,1 ) 19DEC02.628
-C 19DEC02.629
-      CALL EQUATO(VEGFRC,MIX,MJX,1,HDOH,IL,JL,1) 19DEC02.630
-      CALL WRITE_FIELDREC( iutl,D2,nescou,HDOH,XTIME, 19DEC02.631
-      +      'VEGFRC',CURRENT_DATE, 19DEC02.632
-      +      'C','YX','fraction', 19DEC02.633
-      +      'VEGETATION COVERAGE', 19DEC02.634
-      +      IX,JX,1,MIX,MJX,1 ) 19DEC02.635
-      19DEC02.636
-C 19DEC02.637
-      CALL EQUATO(XICE,MIX,MJX,1,HDOH,IL,JL,1) 24SEP99.198
-      CALL WRITE_FIELDREC( iutl,D2,nescou,HDOH,XTIME, 24SEP99.199
-      +      'SEAICE ',CURRENT_DATE, 24SEP99.200
-      +      'C','YX','(DIMENSIONLESS)', 24SEP99.201
-      +      'SEA ICE FLAG', 24SEP99.202
-      +      IX,JX,1,MIX,MJX,1 ) 24SEP99.203
-C 19DEC02.637
-      CALL EQUATO(SSROFF,MIX,MJX,1,HDOH,IL,JL,1) 24SEP99.204
-      CALL WRITE_FIELDREC( iutl,D2,nescou,HDOH,XTIME, 19DEC02.637
-      +      'SFCRNOFF ',CURRENT_DATE, 19DEC02.638
-      +      'C','YX','mm', 19DEC02.639
-      +      'SURFACE RUNOFF', 19DEC02.640
-      +      IX,JX,1,MIX,MJX,1 ) 19DEC02.641
-C 19DEC02.642
-      CALL EQUATO(BGROFF,MIX,MJX,1,HDOH,IL,JL,1) 24SEP99.205
-      CALL WRITE_FIELDREC( iutl,D2,nescou,HDOH,XTIME, 19DEC02.642
-      +      'UGDRNOFF ',CURRENT_DATE, 19DEC02.643
-      +      'C','YX','mm', 19DEC02.644
-      +      'UNDERGROUND RUNOFF', 19DEC02.645

```

```

-      +                      IX,JX,1,MIX,MJX,1 )                      OUTTAP.1132
-C                                                                OUTTAP.1133
-      IF(RDBRDALB)THEN                                           19DEC02.638
-      DO MONTHLOOP=1,12                                           19DEC02.639
-      WRITE(MONALBNAMECHAR,FMT='("MONALB",I2.2)') MONTHLOOP      19DEC02.640
-#ifndef MPP1                                                       19DEC02.641
-      CALL EQUATO(ALBMON(1,1,MONTHLOOP),MIX,MJX,1,HDOETH,IL,JL,1) 19DEC02.642
-#else                                                                19DEC02.643
-      CALL EQUATO(ALBMON(NOFLIC(1),NOFLIC(1),MONTHLOOP),          19DEC02.644
-      +                      MIX,MJX,1,HDOETH,IL,JL,1)            19DEC02.645
-#endif                                                             19DEC02.646
-      CALL WRITE_FIELDREC( iut1,D2,nescou,HDOETH,XTIME,          19DEC02.647
-      +                      MONALBNAMECHAR,CURRENT_DATE,         19DEC02.648
-      +                      'C','YX','PERCENT',                  19DEC02.649
-      +                      'MONTHLY ALBEDO FOR ' // MONTHNAME(MONTHLOOP), 19DEC02.650
-      +                      IX,JX,1,MIX,MJX,1 )                  19DEC02.651
-      END DO                                                       19DEC02.652
-      CALL EQUATO(ALBEDO,MIX,MJX,1,HDOETH,IL,JL,1)              19DEC02.653
-      CALL WRITE_FIELDREC( iut1,D2,nescou,HDOETH,XTIME,          19DEC02.654
-      +                      'ALBEDO ',CURRENT_DATE,              19DEC02.655
-      +                      'C','YX','PERCENT',                  19DEC02.656
-      +                      'BACKGROUND ALBEDO',                 19DEC02.657
-      +                      IX,JX,1,MIX,MJX,1 )                  19DEC02.658
-C                                                                19DEC02.659
-                                                                19DEC02.660
-      ENDIF                                                         19DEC02.661
-C                                                                19DEC02.662
-      IF(RDMAXALB)THEN                                           19DEC02.663
-      CALL EQUATO(ALBMAX,MIX,MJX,1,HDOETH,IL,JL,1)              19DEC02.664
-      CALL WRITE_FIELDREC( iut1,D2,nescou,HDOETH,XTIME,          19DEC02.665
-      +                      'ALBSNOMX ',CURRENT_DATE,            19DEC02.666
-      +                      'C','YX','PERCENT',                  19DEC02.667
-      +                      'MAXIMUM SNOW ALBEDO',                19DEC02.668
-      +                      IX,JX,1,MIX,MJX,1 )                  19DEC02.669
-      ENDIF                                                         19DEC02.670
-C                                                                19DEC02.671
-      ENDIF                                                         07NOV00.460
-      END IF                                                         OUTTAP.1134
-C                                                                07NOV00.461
C----- EXTRA OUTPUT FOR PBL:  T2, Q2, U10, and V10:             05DEC01.286
C                                                                07NOV00.463
      IF (IBLTYP(NESCOU).EQ.2 .OR. IBLTYP(NESCOU).EQ.4 .OR.
+      IBLTYP(NESCOU).EQ.5) THEN                                05DEC01.287
-      CALL EQUATO(TA2,MIX,MJX,1,HDOETH,IL,JL,1)                05DEC01.288
-      CALL WRITE_FIELDREC( iut1,D2,nescou,HDOETH,XTIME,          07NOV00.465
-      +                      'T2',CURRENT_DATE,                   07NOV00.466
-      +                      'C','YX','K',                         07NOV00.467
-      +                      '2-meter Temperature',               07NOV00.468
-      +                      IX,JX,1,MIX,MJX,1 )                  07NOV00.469
-C                                                                07NOV00.470
-      CALL EQUATO(QA2,MIX,MJX,1,HDOETH,IL,JL,1)                07NOV00.471
-      CALL WRITE_FIELDREC( iut1,D2,nescou,HDOETH,XTIME,          07NOV00.472
-      +                      'Q2',CURRENT_DATE,                   07NOV00.473
-      +                      'C','YX','kg kg{-1}',                07NOV00.474
-      +                      '2-meter Mixing Ratio',              07NOV00.475

```

```

-      +                      IX,JX,1,MIX,MJX,1 )                      07NOV00.477
-C                                                                07NOV00.478
-C      MAKE USE OF FILLCRS ROUTINE TO FILL THE LAST ROW AND COLUMN  07NOV00.479
-C      OF U10 AND V10                                              07NOV00.480
-C                                                                07NOV00.481
-      CALL FILLCRS(UA10,MIX,MJX,1,IL,JL,1)                        07NOV00.482
-      CALL EQUATO(UA10,MIX,MJX,1,HDOTH,IL,JL,1)                   07NOV00.483
-      CALL WRITE_FIELDREC( iutl,D2,nescou,HDOTH,XTIME,            07NOV00.484
@@ -1467,41 +865,8 @@
-      +                      IX,JX,1,MIX,MJX,1 )                      07NOV00.496
-      ENDIF                                                         07NOV00.497
-      C                                                             07NOV00.498
-C      MOISTURE AVAILABILITY IF VARYING IN TIME                    07NOV00.499
-C                                                                07NOV00.500
-      IF (IMOIAV(NESCOU) .GE. 1) THEN                               07NOV00.501
-      CALL EQUATO(MAVAIL,MIX,MJX,1,HDOTH,IL,JL,1)                 07NOV00.502
-      CALL WRITE_FIELDREC( iutl,D2,nescou,HDOTH,XTIME,            07NOV00.503
-      +                      'MAVAIL',CURRENT_DATE,                07NOV00.504
-      +                      'C','YX','dimensionless',              07NOV00.505
-      +                      'MOISTURE AVAILABILITY',               07NOV00.506
-      +                      IX,JX,1,MIX,MJX,1 )                    07NOV00.507
-      ENDIF                                                         07NOV00.508
-C                                                                OUTTAP.1135
-C----- PHYSICAL PARAMETERS FROM LANDUSE.TBL                     OUTTAP.1136
-      C                                                             OUTTAP.1137
-      IF (ISFPAR(NESCOU).EQ.1) THEN                                07NOV00.509
-      CALL EQUATO(ALBD,NLU,NSEAS,1,HDOTH,LUN,NSEAS,1)             24SEP99.206
-      CALL WRITE_FIELDREC( iutl,D2,nescou,HDOTH,XTIME,            OUTTAP.1139
-      +                      'ALBD ',CURRENT_DATE,                  OUTTAP.1140
-      +                      ' ','CA','PERCENT',                     OUTTAP.1141
-      +                      'SURFACE ALBEDO',                       OUTTAP.1142
-      +                      LUN,NSEAS,1,NLU,NSEAS,1 )              07NOV00.510
-C                                                                OUTTAP.1144
-      CALL EQUATO(SLMO,NLU,NSEAS,1,HDOTH,LUN,NSEAS,1)             24SEP99.207
-      CALL WRITE_FIELDREC( iutl,D2,nescou,HDOTH,XTIME,            OUTTAP.1146
-      +                      'SLMO ',CURRENT_DATE,                  OUTTAP.1147
-      +                      ' ','CA','fraction',                     OUTTAP.1148
-      +                      'SURFACE MOISTURE AVAILABILITY',        OUTTAP.1149
-      +                      LUN,NSEAS,1,NLU,NSEAS,1 )              07NOV00.511
-C                                                                OUTTAP.1151
-      CALL EQUATO(SFEM,NLU,NSEAS,1,HDOTH,LUN,NSEAS,1)             24SEP99.208
-      CALL WRITE_FIELDREC( iutl,D2,nescou,HDOTH,XTIME,            OUTTAP.1153
-      +                      'SFEM ',CURRENT_DATE,                  OUTTAP.1154
-      +                      ' ','CA','fraction',                     OUTTAP.1155
-      +                      'SURFACE EMISSIVITY AT 9 um',           OUTTAP.1156
-      +                      LUN,NSEAS,1,NLU,NSEAS,1 )              07NOV00.512
-C                                                                OUTTAP.1158
-      CALL EQUATO(SFZO,NLU,NSEAS,1,HDOTH,LUN,NSEAS,1)             24SEP99.209
-      CALL WRITE_FIELDREC( iutl,D2,nescou,HDOTH,XTIME,            OUTTAP.1160
-      +                      'SFZO ',CURRENT_DATE,                  OUTTAP.1161
@@ -1509,111 +874,10 @@
-      +                      'SURFACE ROUGHNESS LENGTH',             OUTTAP.1163
-      +                      LUN,NSEAS,1,NLU,NSEAS,1 )              07NOV00.513
-      C                                                             OUTTAP.1165
-      CALL EQUATO(THERIN,NLU,NSEAS,1,HDOTH,LUN,NSEAS,1)           24SEP99.210

```

```

-      CALL WRITE_FIELDREC( iutl,D2,nescou,HDOH,XTIME,          OUTTAP.1167
-      +      'THERIN ',CURRENT_DATE,                          OUTTAP.1168
-      +      ' ', 'CA', '100*cal cm^-2 K^-1 s^1/2',           OUTTAP.1169
-      +      'SURFACE THERMAL INERTIA',                       OUTTAP.1170
-      +      LUN,NSEAS,1,NLU,NSEAS,1 )                       07NOV00.514
-C                                                                OUTTAP.1172
-      CALL EQUATO(SFHC,NLU,NSEAS,1,HDOH,LUN,NSEAS,1)          24SEP99.211
-      CALL WRITE_FIELDREC( iutl,D2,nescou,HDOH,XTIME,          OUTTAP.1174
-      +      'SFHC ',CURRENT_DATE,                             OUTTAP.1175
-      +      ' ', 'CA', 'J m^-3 K^-1',                        OUTTAP.1176
-      +      'SOIL HEAT CAPACITY',                            OUTTAP.1177
-      +      LUN,NSEAS,1,NLU,NSEAS,1 )                       07NOV00.515
-C                                                                OUTTAP.1179
-      CALL EQUATO(SCFX,NLU,1,1,HDOH,LUN,1,1)                  24SEP99.212
-      CALL WRITE_FIELDREC( iutl,D1,nescou,HDOH,XTIME,          OUTTAP.1181
-      +      'SCFX ',CURRENT_DATE,                             OUTTAP.1182
-      +      ' ', 'CA', 'fraction',                           OUTTAP.1183
-      +      'SNOW COVER EFFECT',                             OUTTAP.1184
-      +      LUN,1,1,NLU,1,1 )                                07NOV00.516
-      END IF                                                    07NOV00.517
-C                                                                07NOV00.518
C---- Output 2-d variables if using the PX land surface/PBL model 07NOV00.519
C                                                                07NOV00.520
      IF (IBLTYP(NESCOU) .EQ. 7) THEN                            07NOV00.521
C                                                                07NOV00.522
-      CALL EQUATO(TA2,MIX,MJX,1,HDOH,IL,JL,1)                 19DEC02.672
-      CALL WRITE_FIELDREC( iutl,D2,nescou,HDOH,XTIME,          19DEC02.673
-      +      'T2',CURRENT_DATE,                                19DEC02.674
-      +      'C', 'YX', 'K',                                   19DEC02.675
-      +      '1.5-meter Temperature',                          19DEC02.676
-      +      IX,JX,1,MIX,MJX,1 )                               19DEC02.677
-C                                                                19DEC02.678
-      CALL EQUATO(MOL,MIX,MJX,1,HDOH,IL,JL,1)                 07NOV00.523
-      CALL WRITE_FIELDREC( iutl,D2,nescou,HDOH,XTIME,          07NOV00.524
-      +      'M-0 LENG',CURRENT_DATE,                          07NOV00.525
-      +      'C', 'YX', 'm',                                   07NOV00.526
-      +      'MONIN-OBUKOV LENGTH',                            07NOV00.527
-      +      IX,JX,1,MIX,MJX,1 )                               07NOV00.528
-C                                                                07NOV00.529
-      CALL EQUATO(RADNET,MIX,MJX,1,HDOH,IL,JL,1)              07NOV00.530
-      CALL WRITE_FIELDREC( iutl,D2,nescou,HDOH,XTIME,          07NOV00.531
-      +      'NET RAD',CURRENT_DATE,                           07NOV00.532
-      +      'C', 'YX', 'W/m^2',                               07NOV00.533
-      +      'SURFACE NET RADIATION',                          07NOV00.534
-      +      IX,JX,1,MIX,MJX,1 )                               07NOV00.535
-C                                                                07NOV00.536
-      CALL EQUATO(ALBEDO,MIX,MJX,1,HDOH,IL,JL,1)              07NOV00.537
-      CALL WRITE_FIELDREC( iutl,D2,nescou,HDOH,XTIME,          07NOV00.538
-      +      'ALBEDO',CURRENT_DATE,                            07NOV00.539
-      +      'C', 'YX', 'fraction',                             07NOV00.540
-      +      'SURFACE ALBEDO',                                 07NOV00.541
-      +      IX,JX,1,MIX,MJX,1 )                               07NOV00.542
-C                                                                07NOV00.543
-      CALL EQUATO(GRNFLX,MIX,MJX,1,HDOH,IL,JL,1)              07NOV00.544
-      CALL WRITE_FIELDREC( iutl,D2,nescou,HDOH,XTIME,          07NOV00.545

```

212

```

-      DO I=1,IL                                07NOV00.601
-      DO J=1,JL                                07NOV00.602
-      HSCRL(I,J)=FLOAT(ISLTYP(I,J))            07NOV00.603
-      ENDDO                                    07NOV00.604
-      ENDDO                                    07NOV00.605
-      CALL EQUATO(HSCRL,MIX,MJX,1,HDOH,IL,JL,1) 07NOV00.606
-      CALL WRITE_FIELDREC( iutl,D2,nescou,HDOH,XTIME, 07NOV00.607
-      +      'ISLTYP',CURRENT_DATE,            07NOV00.608
-      +      'C','YX','USDA cat',              07NOV00.609
-      +      'SOIL TEXTURE TYPE',              07NOV00.610
-      +      IX,JX,1,MIX,MJX,1 )               07NOV00.611
-C      ENDDO                                    07NOV00.612
-      ENDDO                                    07NOV00.613
-      ENDDO                                    07NOV00.614
C      OUTTAP.1186
C---- 1D FIELD                                OUTTAP.1187
@@ -1656,8 +908,8 @@
C---- CHECK TO SEE IF OUTPUT FILES NEED TO BE CLOSED
C
      IF (BUFFRQ.GE.TAPFRQ) THEN                OUTTAP.1205
-      IF (MOD(NINT(XTIME/TAPFRQ),NINT(BUFFRQ/TAPFRQ+0.01)).EQ.0  OUTTAP.1206
-      *      .OR. (BUFFRQ.EQ.TAPFRQ .AND. XTIME.GT.O.)) THEN    OUTTAP.1207
+      IF ((MOD(NINT(XTIME/TAPFRQ),NINT(BUFFRQ/TAPFRQ+0.01)).EQ.0 24SEP99.214
+      *      .OR. BUFFRQ.EQ.TAPFRQ ) .AND. XTIME.GT.O.) THEN    24SEP99.215
+      *      .OR. BUFFRQ.EQ.TAPFRQ ) .AND. XTIME.GT.O.) THEN    23AUG05
      IF ( DM_IONODE() ) CLOSE (IUTL)          23AUG05
C      PRINT 940, IUTL                          OUTTAP.1210
940      FORMAT(2X,'OUTPUT UNIT ',I2,' IS CLOSED.') OUTTAP.1211

```

University of Cape Town

Bibliography

- G. Adrian and F. Fiedler. Simulation of Unstationary Wind and Temperature Fields Over Complex Terrain and Comparison with Observations. *Beitraege zur Physik der Atmosphaere*, 64:27–48, 1991.
- C. Ambroise, G. Seze, F. Badran, and S. Thiria. Hierarchical Clustering of Self-Organizing Maps for Cloud Classification. *Neurocomputing*, 30(1):47–52, 2000.
- J. Anderson, E. Hardy, J. Roach, and R. Witmer. A Land Use and Land Cover Classification System for Use with Remote Sensor Data. Professional Paper 964, U.S. Geological Survey, 1976.
- R. Anthes and T. Warner. Development of Hydrodynamic Models Suitable for Air Pollution and Other Mesometeorological Studies. *Monthly Weather Review*, 106(8):1045–1078, 1978.
- R. Anthes, E. Hsieh, and Y. Kuo. *Description of the Penn State/NCAR Mesoscale Model Version 4 (MM4)*. National Center for Atmospheric Research (US and Acid Deposition Modeling Project and Atmospheric Analysis and Prediction Division), 1987.
- I. Antoniou, H. Jorgensen, T. Mikkelsen, T. Pedersen, G. Warmbier, and D. Smith. Comparison of Wind Speed and Power Curve Measurements Using a Cup Anemometer, a LIDAR, and a SODAR. In *EWEC Wind Energy Conference, London*, 2004.
- G. Austin, A. Williams, G. Morris, R. Spalding-Fecher, and R. Worthington. Employment Potential of Renewable Energy in South Africa. Technical report, Agama Energy, 2003.

- K. Ayotte, R. Davy, and P. Coppin. A Simple Temporal and Spatial Analysis of Flow in Complex Terrain in the Context of Wind Energy Modelling. *Boundary-Layer Meteorology*, 98(2):275–295, 2001.
- D. Banks and J. Schäffler. The Potential Contribution of Renewable Energy in South Africa. Technical report, Sustainable Energy & Climate Change Project (SECCP), Earthlife Africa Jhb, February 2006. Draft Update Report.
- D. Banks and J. Schäffler. A Proposed Renewable Energy Plan of Action for the Western Cape. Technical report, Western Cape Department of Environmental Affairs and Development Planning, May 2007. Final Draft for Approval.
- A. Betz. *Introduction to the Theory of Flow Machines*. Oxford: Pergamon Press, 1966.
- A. Blackadar. High Resolution Models of the Planetary Boundary Layer. *Advances in Environmental Science and Engineering*, 1(1):50–85, 1979.
- A. Brebøl, T. Henriksen, and E. James-Smith. Electricity from Renewable Resources in the Western Cape. Master’s thesis, Roskilde University, 2003.
- J. Breckling. *The Analysis of Directional Time Series: Applications to Wind Speed and Direction*. Springer-Verlag, 1989.
- D. Bruce Turner. Comparison of Three Methods for Calculating the Standard Deviation of the Wind Direction. *Journal of Applied Meteorology*, 25(5)(5):703–707, 1986.
- BTM Consult. International Wind Energy Development - World Market Update 2006. Technical report, BTM Consult ApS, Ringkøbing, Denmark, March 2007.
- S. Burk and W. Thompson. A Vertically Nested Regional Numerical Weather Prediction Model with Second-Order Closure Physics. *Monthly Weather Review*, 117(11):2305–2324, 1989.
- D. Burridge, E. Källén, and C. Pastre. Evaluation of the International HIRLAM Project and HIRLAM Operational Applications. Technical report, HIRLAM, January 2005.
- T. Burton. *Wind Energy Handbook*. Wiley, 2001.

- W. Canter. Wind Turbine Data Summary Tables. *Windstats Newsletter*, 19(2):8, 2006.
- T. Cavazos. Using Self-Organizing Maps to Investigate Extreme Climate Events: An Application to Wintertime Precipitation in the Balkans. *Journal of Climate*, 13(10):1718–1732, 2000.
- T. Cavazos and B. Hewitson. Relative Performance of Empirical Predictors of Daily Precipitation. *Conference Proceedings of the International Environmental Modelling and Software Society*, 2:349–354, 2002.
- A. Chandrasekar, C. Philbrick, B. Doddridge, R. Clark, and P. Georgopoulos. A Comparative Study of Prognostic MM5 Meteorological Model with Aircraft, Wind Profiler, Lidar, Tethered Balloon and RASS Data Over Philadelphia During a 1999 Summer Episode. *Environmental Fluid Mechanics*, 4(4):339–365, 2005.
- F. Chen and J. Dudhia. Coupling an Advanced Land Surface-Hydrology Model with the Penn State-NCAR MM5 Modeling System. Part I: Model Implementation and Sensitivity. *Monthly Weather Review*, 129(4):569–585, 2001.
- R. Crane and B. Hewitson. Clustering and Upscaling of Station Precipitation Records to Regional Patterns Using Self-Organizing Maps (SOMs). *CLIMATE RESEARCH*, 25(2):95–107, 2003.
- P. D’Abreton and E. Law. South African Renewable Energy Resource Database. Technical report, CSIR and Eskom, 1999.
- J. Davis. *Statistics and Data Analysis in Geology*. John Wiley & Sons, 1973.
- D. Delaunay. The Navier-Stokes Solver Meteodyn NS. Online. <http://www.meteodyn.com>, 2007.
- R. Diab. Wind Energy Potential over South Africa. Technical report, CSIR, 1979.
- R. Diab. *Wind Atlas of South Africa*. Department of Mineral and Energy Affairs, Pretoria, 1995.
- R. Diab. Estimating Power from Wind. *South African Journal of Science*, 77:126–129, 1981.

- R. Diab. Contribution of Wind Power to South Africa's Energy Needs. *South African Journal of Science*, 81:460–463, 1985.
- J. Dudhia. Numerical Study of Convection Observed During the Winter Monsoon Experiment Using a Mesoscale Two-Dimensional Model. *Journal of the Atmospheric Sciences*, 46(20):3077–3107, 1989.
- J. Dudhia. A Multi-layer Soil Temperature Model for MM5. In *Preprints, The Sixth PSU/NCAR Mesoscale Model Users' Workshop*, pages 49–50, July 1996.
- J. Dudhia, D. Gill, Y. Guo, K. Manning, and W. Wang. *PSU/NCAR Mesoscale Modeling System Tutorial Class Notes and User's Guide: MM5 Modeling System Version 3*. National Center for Atmospheric Research, <http://www.mmm.ucar.edu/mm5/documents/tutorial-v3-notes.html>, January 2004. Online.
- R. Dutkiewicz and J. Denison. *Wind Energy Monitoring*. Energy Research Institute, University of Cape Town, 1991.
- EDRC. Policies and Measures for Renewable Energy and Energy Efficiency in South Africa. Technical report, Energy & Development Research Centre, University of Cape Town, 2003. Prepared for Sustainable Energy and Climate Change Partnership.
- D. Elliott, C. Holladay, W. Barchet, H. Foote, and W. Sandusky. Wind Energy Resource Atlas of the United States. Technical Report DOE/CH 10093-4, US Department of Energy, October 1987.
- K. Emanuel and D. Raymond. *Convective Parameterization for Mesoscale Models: The Kain-Fritsch scheme*, volume 24 of *Meteorological Monographs*. American Meteorological Society, 1993.
- M. Evans, N. Hastings, and B. Peacock. *Statistical Distributions*. Wiley Publishing, 2000.
- F. Castellani and G. Franceschini. The Umbria Region Wind Atlas. In *43rd AIAA Aerospace Sciences Meeting and Exhibit*, 2005.
- E. Feitosa, A. Pereira, and D. Veleza. Brazilian Wind Atlas Project. Technical report, Brazilian Wind Energy Centre (CBEE), 2001.

- A. Fellows and G. Gow. Wind Resource Estimation in South Africa. In *Proceedings 31st Meeting of Experts, State of the Art on Wind Resource Estimation, Publication IEA Annex XI*, 1999.
- N. Fisher. *Statistical Analysis of Circular Data*. Cambridge University Press, 1993.
- N. Fisher. Problems with the Current Definitions of the Standard Deviation of Wind Direction. *Journal of Applied Meteorology*, 26(11)(11):1522–1529, 1987.
- H. Frank and L. Landberg. Modelling the Wind Climate of Ireland. *Boundary-Layer Meteorology*, 85(2):359–377, December 1997.
- H. P. Frank, O. Rathmann, N. G. Mortensen, and L. Landberg. *The Numerical Wind Atlas - the KAMM/WAsP Method*. Risø National Laboratory, 2001.
- J. Fritsch and C. Chappell. Numerical Prediction of Convectively Driven Mesoscale Pressure Systems. Part I: Convective Parameterization. *Journal of the Atmospheric Sciences*, 37(8):1722–1733, 1980.
- G. Gaile and J. Burt. *Directional Statistics*. Geo Abstracts Ltd., 1980.
- D. Gesch, K. Verdin, and S. Greenlee. New Land Surface Digital Elevation Model Covers the Earth. *Eos, Transactions, American Geophysical Union*, 80(6):69–70, 1999.
- G. Giebel, R. Brownsword, and G. Kariniotakis. The State-Of-The-Art in Short-Term Prediction of Wind Power. A Literature Overview. Technical Report Version 1.1, Risø National Laboratory, 2003. Project ANEMOS.
- A. Gravidahl. Norwegian Wind Atlas. Technical report, Vector AS, NVE/ENOVA, 2003. URL http://windsim.com/wind_energy/wind_atlas.
- A. Gravidahl and K. Harstveit. Windsim - Flow Simulations in Complex Terrain Assessments of Wind Resources Along the Norwegian Coast, 2000.
- G. A. Grell, J. Dudhia, and D. R. Stauffer. A Description of the Fifth-Generation Penn State/ NCAR Mesoscale Model (MM5). Technical Note NCAR/TN-398+STR, NCAR, 1995.
- G. Guyot. *Physics of the Environment and Climate*. John Wiley & Sons. Praxis, Chichester. GB, 1998. 138pp, 226pp.

- K. Hagemann. Validation of Mesoscale Modelling Data for Wind Resource Assessment. BSc Honours Thesis, September 2003. Department of Environmental and Geographical Science, University of Cape Town.
- R. Hammouche. *Atlas Vent de l'Algérie (Wind Atlas of Algeria)*. Office National de la Météorologie, 1991.
- B. Hewitson and R. Crane. Self-Organizing Maps: Applications to Synoptic Climatology. *Climate Research*, 22(1):13–26, 2002.
- S. Hong and H. Pan. Nonlocal Boundary Layer Vertical Diffusion in a Medium-Range Forecast Model. *Monthly Weather Review*, 124(10):2322–2339, 1996.
- D. Hudson. *Antarctic Sea Ice Extent, Southern Hemisphere Circulation and South African Rainfall*. PhD thesis, University of Cape Town, 1998.
- A. Ilinca, E. McCarthy, J.-L. Chaumel, and J.-L. Retiveau. Wind Potential Assessment of Quebec Province. *Renewable Energy*, 28:16, 2003.
- P. Jackson and J. Hunt. Turbulent Wind Flow Over a Low Hill. *Quarterly Journal of the Royal Meteorological Society*, 101(430):929–955, 1975.
- Z. Janjić. The Step-Mountain Eta Coordinate Model: Further Developments of the Convection, Viscous Sublayer and Turbulence Closure Schemes. *Monthly Weather Review*, 122(5):927–945, 1994.
- Y. Jia and R. Morris. Meteorological Modeling of the August 1999 Episode to Support the 8-Hour Ozone Early Action Compacts for Tulsa and Oklahoma City. Technical report, Prepared for Oklahoma Department of Environmental Quality, September 2003. Revised Draft Final Report.
- B. Jiménez, F. Durante, B. Lange, T. Kreutzer, and J. Tambke. Offshore Wind Resource Assessment in the German Part of the North Sea: Comparative Study Between MM5 and WAsP. *Geophysical Research Abstracts*, 8:00555, 2006.
- C. Justus. *Winds and Wind System Performance*. Franklin Institute Press, 1978. Research supported by the National Science Foundation and Energy Research and Development Administration.
- C. G. Justus and A. Mikhail. Height Variations of Wind Speed and Wind Distribution Statistics. *Geophysical Research Letters*, 3:261–264, 1976.

- J. Kain. The Kain-Fritsch Convective Parameterization: An update. *Journal of Applied Meteorology*, 43(3):170–181, 2004.
- P. Kallberg, S. Uppala, A. Hernandez, A. Nomura, and E. Serrano. ERA description. ECMWF Re-analysis Project Report Series 1. Technical report, European Centre for Medium Range Weather Forecasts, 1997. 872 pp.
- E. Kalnay et al. The NCEP/NCAR 40-year Reanalysis Project. *Bulletin of the American Meteorological Society*, 77:437:471, 1996.
- M. Kanamitsu, W. Ebisuzaki, J. Woollen, S. Yang, J. Hnilo, M. Fiorino, and G. Potter. NCEP–DOE AMIP-II Reanalysis (R-2). *Bulletin of the American Meteorological Society*, 83(11):1631–1643, 2002.
- T. Kohonen. *Self-organization and Associative Memory*. Springer-Verlag New York, Inc. New York, NY, USA, 1989.
- T. Kohonen. *Self-Organizing Maps*, volume 30. Springer, 3rd edition edition, 2001.
- T. Kohonen, J. Hynninen, J. Kangas, and J. Laaksonen. SOMPAK: The Self-Organizing Map Program Package. Technical Report Report A31, Helsinki University of Technology, Laboratory of Computer and Information Science, 1996.
- I. Kotzé. Photovoltaics and Rural Electrification in South Africa - Problems and Prospects. In *Conference of the International Solar Energy Society*, 1998.
- R. Krieg. *Vindatlas för Sverige (Wind Atlas for Sweden)*. SMHI, Norrköping, 1992.
- B. Källstrand, H. Bergström, J. Højstrup, and A.-S. Smedman. Mesoscale Wind Field Modifications Over the Baltic Sea. *Boundary-Layer Meteorology*, 95(2): 161–188, May 2000.
- L. Landberg, L. Myllerup, O. Rathmann, E. Petersen, B. Joergensen, J. Badger, and N. Mortensen. Wind Resource Estimation - An Overview. *Wind Energy*, 6(3):261–271, 2003.
- F. Langhans. Capacity Building in Energy Efficiency and Renewable Energy - Review of Wind Energy Resource Studies in South Africa. Technical report, Department of Minerals and Energy and Tripod Wind Energy ApS, 2003.

- D. Legates and G. McCabe Jr. Evaluating the Use of "Goodness-of-Fit" Measures in Hydrologic and Hydroclimatic Model Validation. *Water Resources Research*, 35(1):233–241, 1999.
- A. Llombart, A. Talayero, A. Mallet, and E. Telmo. Performance Analysis of Wind Resource Assessment Programs in Complex Terrain. In *International Conference on Renewable Energies and Power Quality*, 2006.
- T. Loveland, B. Reed, J. Brown, D. Ohlen, Z. Zhu, L. Yang, and J. Merchant. Development of a Global Land Cover Characteristics Database and IGBP DISCover from 1 km AVHRR Data. *International Journal of Remote Sensing*, 21(6):1303–1330, 2000.
- D. Loy. Energy-policy Framework Conditions for Electricity Markets and Renewable Energies - 21 Country Analyses. Technical report, Deutsche Gesellschaft für Technische Zusammenarbeit (GTZ) GmbH, Division Environment and Infrastructure, June 2004. Page 91.
- B. Malmgren and A. Winter. Climate Zonation in Puerto Rico Based on Principal Components Analysis and an Artificial Neural Network. *Journal of Climate*, 12(4):977–985, April 1999.
- T. Malyshev, D. Justus, L. P. Alisse, and T. Gül. Renewable Energy Market & Policy Trends in IEA Countries. Technical report, International Energy Agency (IEA), 2004.
- K. Mardia. *Statistics of Directional Data*. JSTOR, 1975.
- J. Maroga. Partnering for Growth - Eskom Annual Report. Technical report, Eskom, 2007.
- J. W. Martens, T. de Lange, J. Cloin, S. Szewczuk, R. Morris, and J. Zak. Accelerating the Market Penetration of Renewable Energy Technologies in South Africa. Technical report, ECN, CSIR, Riso, Richard Morris and Associates, 2001.
- H.-T. Mengelkamp. Wind Climate Simulation over Complex Terrain and Wind Turbine Energy Output Estimation. *Theoretical and Applied Climatology*, 63 (3 - 4):129–139, August 1999.

- E. Mlawer, S. Taubman, P. Brown, M. Iacono, and S. Clough. Radiative Transfer for Inhomogeneous Atmospheres: RRTM, a Validated Correlated-k Model for the Longwave. *J. Geophys. Res.*, 102(16):663–16, 1997.
- P. Moreno, A. Gravdahl, and M. Romero. Wind Flow Over Complex Terrain: Application of Linear and CFD Models. In *European Wind Energy Conference and Exhibition*, 2003.
- Y. Mori. Methods for Estimating the Mean and the Standard Deviation of Wind Direction. *Journal of Applied Meteorology*, 26(9):1282–1284, 1987.
- M. New, R. Washington, C. Jack, B. Hewitson, and U. Oxford. Sensitivity of Southern African Climate to Soil Moisture. *CLIVAR Exchanges*, 8:45–47, 2003.
- F. Ni, T. Cavazos, M. Hughes, A. Comrie, and G. Funkhouser. Cool-season Precipitation in the Southwestern USA Since AD 1000: Comparison of Linear and Nonlinear Techniques for Reconstruction. *International Journal of Climatology*, 22(13):1645–1662, 2002.
- H. Nielsen, H. Madsen, T. Nielsen, J. Badger, G. Giebel, L. Landberg, K. Sattler, and H. Feddersen. Wind Power Ensemble Forecasting. In *CD-Proceedings of the 2004 Global Windpower Conference, Chicago, US*, 2004.
- T. Oke. *Initial Guidance to Obtain Representative Meteorological Observations at Urban Sites*. World Meteorological Organization, 2004.
- J. Peixoto and A. Oort. *Physics of Climate*. American Institute of Physics, 1992.
- M. Perera. Shelter Behind Two-Dimensional Solid and Porous Fences. *Journal of wind Engineering and Industrial Aerodynamics*, 8:98–104, 1981.
- L. Pitt, G. C. van Kooten, M. Love, and N. Djilali. Utility-scale Wind Power: Impacts of Increased Penetration. In *Proceedings of the International Green Energy Conference*, Waterloo, Ontario, Canada, 12-16 June 2005.
- J. Pleim and J. Chang. Non-local Closure Model for Vertical Mixing in the Convective Boundary Layer. *Atmospheric Environment*, 26(6):965–981, 1992.
- W. Press, S. Teukolsky, W. Vetterling, and B. Flannery. *Numerical Recipes in C: The Art of Scientific Computing*. Cambridge University Press, 2002.

- S. Pryor, R. Barthelmie, and E. Kjellström. Potential Climate Change Impact on Wind Energy Resources in Northern Europe: Analyses Using a Regional Climate Model. *Climate Dynamics*, 25(7):815–835, 2005.
- R. Ragoonathun, D. van der Westhuizen, and I. Smit. South African Wind Atlas. Technical report, Eskom, CSIR, DME and Danced, 2001.
- D. Reusch, R. Alley, and B. Hewitson. Relative Performance of Self-Organizing Maps and Principal Component Analysis in Pattern Extraction from Synthetic Climatological Data. *Polar Geography*, 29(3):188–212, 2005.
- R. Reynolds, T. Smith, C. Liu, D. Chelton, K. Casey, and M. Schlax. Daily High-Resolution-Blended Analyses for Sea Surface Temperature. *Journal of Climate*, 20(22):5473–5496, 2007. Data available online at <http://www.ncdc.noaa.gov/oa/climate/research/sst/oi-daily.php>.
- J. Sammon. A Nonlinear Mapping for Data Structure Analysis. *IEEE Transactions on Computers*, 18(5):401–409, 1969.
- S. Sandström. Simulations of the Climatological Wind Field in the Baltic Sea Area Using a Mesoscale Higher-Order Closure Model. *Journal of Applied Meteorology*, 36:1541–1552, November 1997.
- J. Schoof and S. Pryor. An Evaluation of Two GCMs: Simulation of North American Teleconnection Indices and Synoptic Phenomena. *INTERNATIONAL JOURNAL OF CLIMATOLOGY*, 26(2):267, 2006.
- J. Schäffler. Wind Energy in the Western Cape: Grid-connected Renewable Distributed Electricity Generation. Master’s thesis, University of Cape Town, Energy and Development Research Centre, 2001.
- P. Shafran, N. Seaman, and G. Gayno. Evaluation of Numerical Predictions of Boundary Layer Structure during the Lake Michigan Ozone Study. *Journal of Applied Meteorology*, 39(3):412–426, 2000.
- J. Simpson. *Sea Breeze and Local Winds*. 1994.
- W. C. Skamarock, J. B. Klemp, J. Dudhia, D. O. Gill, D. M. Barker, W. Wang, and J. G. Powers. A Description of the Advanced Research WRF Version 2. Technical Note NCAR/TN-468+STR, NCAR, June 2005.

- A. Starkov, L. Landberg, P. Bezroukikh, and M. Borisenko. *Russian Wind Atlas*. Russian-Danish Institute for Energy Efficiency, Moscow; Risø National Laboratory, Roskilde, 2000.
- D. Stauffer and N. Seaman. Use of Four-Dimensional Data Assimilation in a Limited-Area Mesoscale Model. Part I: Experiments with Synoptic-Scale Data. *Monthly Weather Review*, 118(6):1250–1277, 1990.
- D. Stauffer, N. Seaman, and F. Binkowski. Use of Four-Dimensional Data Assimilation in a Limited-Area Mesoscale Model. Part II: Effects of Data Assimilation within the Planetary Boundary Layer. *Monthly Weather Review*, 119(3):734–754, 1991.
- M. Stephens. EDF Statistics for Goodness of Fit and Some Comparisons. *Journal of the American Statistical Association*, 69(347):730–737, 1974.
- A. J. Storkey. Generalised Propagation for Fast Fourier Transforms with Partial or Missing Data. In S. Thrun, L. Saul, and B. Schölkopf, editors, *Advances in Neural Information Processing Systems 16*. MIT Press, Cambridge, MA, 2004.
- M. Tadross, W. Gutowski, B. Hewitson, C. Jack, and M. New. MM5 Simulations of Interannual Change and the Diurnal Cycle of Southern African Regional Climate. *Theoretical and Applied Climatology*, 86(1):63–80, 2006.
- E. S. Takle and J. Brown. Note on the Use of Weibull Statistics to Characterize Wind Speed Data. *Journal of Applied Meteorology*, 17:556–559, 1978.
- W. Tennant. An Assessment of Intraseasonal Variability from 13-Yr GCM Simulations. *Monthly Weather Review*, 131(9):1975–1991, 2003.
- W. Tennant. Considerations When Using Pre-1979 NCEP/NCAR Reanalyses in the Southern Hemisphere. *Geophysical Research Letters*, 31(11):11112, 2004.
- O. Teruo, S. Susumu, F. Kazuhito, H. Atsushi, and Y. Takashi. Feasibility of Offshore Wind Power Generation Over Ise Bay Part I: Ise Bay in the Large-Scale Wind Atlas. In *Proceedings of JSES/JWEA Joint Conference*, pages 99–102, 2003.
- G. Tonnesen, Z. Wang, M. Omary, C. Chien, R. Morris, G. Mansell, S. Kemball-Cook, G. Yarwood, Z. Adelman, A. Holland, and K. Hanisak. 2004 Interim Report for the Western Regional Air Partnership (WRAP) Regional Modeling

- Center (RMC). Technical report, Western Regional Air Partnership, 2004. WGA Contract Number: 30203.
- C. Tremback. MM5 Simulations for TexAQs 2000 Episode. Technical report, ATMET LLC, August 2003.
- I. Troen and E. Petersen. *European Wind Atlas*. Risø National Laboratory, Roskilde, 1987.
- I. Troen, N. Mortensen, and E. Petersen. WAsP Wind Atlas Analysis and Application Programme. User's Guide. Technical report, Risø National Laboratory, Roskilde, Denmark, 1987.
- S. E. Tuller and A. C. Brett. The Characteristics of Wind Velocity that Favor the Fitting of a Weibull Distribution in Wind Speed Analysis. *Journal of Climate and Applied Meteorology*, 23:124–134, 1983.
- P. Tyson and R. Preston-Whyte. *The Weather and Climate of Southern Africa*. Oxford University Press, second edition, 2000.
- H. A. van der Linde. Wind Energy in South Africa. In *World Renewable Energy Congress*, 1996.
- C. Willmott. On the Validation of Models. *Physical Geography*, 2(2):184–194, 1981.
- C. Willmott. Some Comments on the Evaluation of Model Performance. *Bulletin of the American Meteorological Society*, 63(11):1309–1313, 1982.
- C. Willmott, R. Davis, J. Feddema, K. Klink, D. Legates, C. Rowe, S. Ackleson, and J. O'Donnell. Statistics for the Evaluation and Comparison of Models. *Journal of Geophysical Research*, 90(C12):8995–9005, 1985.
- N. Wood. Wind Flow Over Complex Terrain: A Historical Perspective and the Prospect for Large-Eddy Modelling. *Boundary-Layer Meteorology*, 96(1): 11–32, 2000.
- A. Xiu and J. Pleim. Development of a Land Surface Model. Part I: Application in a Mesoscale Meteorological Model. *Journal of Applied Meteorology*, 40(2): 192–209, 2000.

- M. Xu, Y. Liu, C. Davis, and T. Warner. Sensitivity Study on Nudging Parameters for a Mesoscale FDDA System. In *Conference on Weather Analysis and Forecasting*, volume 19, pages 127–130. AMS, 2002.
- D. Zhang and W. Zheng. Diurnal Cycles of Surface Winds and Temperatures as Simulated by Five Boundary Layer Parameterizations. *Journal of Applied Meteorology*, 43(1):157–169, January 2004.
- K. Zhang, H. Mao, K. Civerolo, S. Berman, J. Ku, S. Rao, B. Doddridge, C. Philbrick, and R. Clark. Numerical Investigation of Boundary-Layer Evolution and Nocturnal Low-Level Jets: Local versus Non-Local PBL Schemes. *Environmental Fluid Mechanics*, 1(2):171–208, 2001.