

**ENSEMBLE FORECASTING AND DATA ASSIMILATION
IN NUMERICAL WEATHER MODELING FOR EGYPT**

By

HAMADA SULTAN SOLIMAN BADR

A Thesis Submitted to the
Faculty of Engineering, Cairo University
in Partial Fulfillment of the
Requirements for the Degree of

MASTER OF SCIENCE

In

AEROSPACE ENGINEERING

FACULTY OF ENGINEERING, CAIRO UNIVERSITY
GIZA, EGYPT

February 2011

ENSEMBLE FORECASTING AND DATA ASSIMILATION IN NUMERICAL WEATHER MODELING FOR EGYPT

By

HAMADA SULTAN SOLIMAN BADR

A Thesis Submitted to the
Faculty of Engineering, Cairo University
in Partial Fulfillment of the
Requirements for the Degree of

MASTER OF SCIENCE

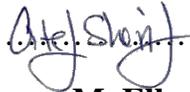
In

AEROSPACE ENGINEERING

Under Supervision of

Atef O. Sherif,

Professor, Department of Aerospace Engineering,
Faculty of Engineering, Cairo University

..........

Basman M. Elhadidi,

Associate Professor, Department of Aerospace Engineering,
Faculty of Engineering, Cairo University

..........

and

Hamdy A. Kandil

Associate Professor, Department of Mechanical Engineering,
Faculty of Engineering, Alexandria University,
Faculty of Engineering and Materials Science, German University in Cairo

..........

FACULTY OF ENGINEERING, CAIRO UNIVERSITY
GIZA, EGYPT

February 2011

ENSEMBLE FORECASTING AND DATA ASSIMILATION IN NUMERICAL WEATHER MODELING FOR EGYPT

By

HAMADA SULTAN SOLIMAN BADR

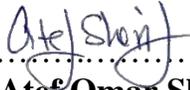
A Thesis Submitted to the
Faculty of Engineering, Cairo University
in Partial Fulfillment of the
Requirements for the Degree of

MASTER OF SCIENCE

In

AEROSPACE ENGINEERING

Approved by the Examining Committee:



.....
Prof. Dr. **Atef Omar Sherif**, Thesis Main Advisor
Professor, Department of Aerospace Engineering, Cairo University



.....
Prof. Dr. **Jimmy Dudhia**, Member
Project Scientist II, National Center for Atmospheric Research, USA



.....
Prof. Dr. **Mohamed Ezz Eldin Elraey**, Member
Former Dean, Institute of Graduate Studies and Research, Alexandria University

FACULTY OF ENGINEERING, CAIRO UNIVERSITY
GIZA, EGYPT

February 2011

ACKNOWLEDGEMENT

In the name of God the Most Gracious the Most Merciful

I would like to thank all my family, especially my mother, friends and colleagues for their continuous love and support which makes me able to face any hard situation in my life and this has a great impact on the performance and accuracy of my work. May God reward them all and keep them to stand by me forever.

I would like to express my respect to my supervisor Professor Atef Sherif. He is not just a supervisor for this work, he is my teacher during my undergraduate study and till now, and I always learn more from his ideas directly and indirectly.

Also, I want to record my gratitude to Dr. Basman Elhadidi, he is not just a supervisor but he is a friend closer than a brother too. And I need to thank Dr. Hamdy Kandil for his revision of this thesis.

At last, I would like to thank the NARSS research team members, Ahmed Abdelmoaty and Mohammed Abdelkader, for their continuous help and support. Also, thanks to NARSS administrators for providing the time, computational facilities and remotely-sensed data required for performing most simulations.

ABSTRACT

In this thesis, the effects of using nudging-based data assimilation on the results of a numerical weather modeling system for Egypt were studied based on the MM5 model. Effects of different nudging options were studied for different weather regimes and at locations with different observation site densities. Effects on accuracy at the observational-data-void areas and at the locations with difficult terrain properties were of prime concern. The collection, processing, and visualization of satellite observational data were also included. It was found that data assimilation improves the accuracy of the model and the computed temperatures correlate better with the observed temperatures, especially near the surface. The accuracy was affected at the observational-data-void areas, near coast lines, and at the locations with complex terrain. This may be fixed by increasing the number and quality of observational stations at these locations.

Then, the effects of changing the initialization, time and method, and the different physics options of the model were studied. It was found that the model is sensitive to the initialization and the physics options. This was utilized for a preliminary ensemble forecasting experiment for Egypt to cope with the uncertainties associated with the initial conditions and the model. Nine different physics options, based on three planetary boundary layer schemes and three radiation schemes, and four sets of initial conditions were used to construct an ensemble of 36 forecasts. Nine forecasts were used as reference forecasts and the other 27 forecasts utilized data assimilation. The results, of the ensemble forecasting experiment, were evaluated based on percentage difference in maximum, minimum and average temperatures with the observations. It was found that the ensemble mean is better and more correlated to observations than a single control forecast. Moreover, the accuracy of the results is proportional to the number and quality of the ensemble members.

TABLE OF CONTENTS

ACKNOWLEDGEMENT	VII
ABSTRACT	II
TABLE OF CONTENTS	II
LIST OF TABLES	VIII
LIST OF FIGURES	X
NOMENCLATURE	XIV
ABBREVIATIONS	XVI
1. INTRODUCTION	1
1.1 NUMERICAL WEATHER MODELING	2
1.1.1 Initial Boundary Value Problem	5
1.1.2 Observational Data	6
1.1.3 Data Assimilation	8
1.1.4 Ensemble Forecasting	9
1.2 MOTIVATION	11
1.3 LITERATURE REVIEW	12
1.3.1 Data Assimilation	13
1.3.2 Ensemble Forecasting	15
1.4 THESIS OBJECTIVES	19
1.5 THESIS ORGANIZATION	20
2. MODELING SYSTEM	23
2.1 INTRODUCTION	23
2.1.1 A Modeling System Framework	23
2.1.2 The MM5 Modeling System	24
2.1.3 The MM5 Forecast Variables	28
2.2 GOVERNING EQUATIONS	29
2.2.1 Pressure Tendency Equation	31
2.2.2 Momentum Equations	32
2.2.3 Temperature Tendency Equation	33
2.3 INITIAL AND BOUNDARY CONDITIONS	33

2.3.1 Initial Conditions	37
2.3.2 Surface Boundary Conditions	38
2.3.3 Lower Boundary Conditions	39
2.3.4 Upper Boundary Conditions	40
2.3.5 Lateral Boundary Conditions	42
2.4 SOLUTION TECHNIQUE	44
2.4.1 Finite Differencing	44
2.4.2 Physical Parameterizations	45
2.4.3 Automation Scripts	46
2.5 MODEL EVALUATION	48
3. OBSERVATIONAL DATA	51
3.1 INTRODUCTION	51
3.2 AVAILABLE DATA	52
3.2.1 Research Data Archive	52
3.2.2 NOAA Satellite Data	53
3.3 DATA PROCESSING	55
3.3.1 Processing of RDA Data	56
3.3.2 Processing of NOAA Data	57
3.3.2.1 Transmission of NOAA Data	60
3.3.2.2 Geometric Correction	61
3.3.2.3 Cloud Masking	62
3.4 DATA ANALYSIS	63
3.4.1 ATOVS Data Retrievals	63
3.4.2 AVHRR Data Retrievals	65
3.5 DATA QUALITY	69
4. DATA ASSIMILATION	71
4.1 INTRODUCTION	71
4.1.1 Observational Increments	71
4.1.2 Observational Errors	72
4.2 ASSIMILATION METHODS	73
4.2.1 Types of Data Assimilation	73
4.2.2 Methods of Data Assimilation	74
4.3 OBJECTIVE ANALYSIS	76
4.3.1 Objective Analysis Techniques	78

4.3.1.1 Cressman Schemes	78
4.3.1.2 Multi-quadric Scheme	79
4.3.2 Quality Control for Observations.....	80
4.3.2.1 Quality Control on Individual Reports	81
4.3.2.2 The ERRMAX Test.....	81
4.3.2.3 The Buddy Test	82
4.4 FOUR-DIMENSIONAL DATA ASSIMILATION	82
4.4.1 Applications of FDDA.....	82
4.4.1.1 Dynamic Initialization.....	83
4.4.1.2 Dynamic Analysis	83
4.4.1.3 Boundary Conditions.....	84
4.4.2 FDDA Methods.....	84
4.4.2.1 Analysis or Grid Nudging	84
4.4.2.2 Station or Observational Nudging	86
4.5 APPLICATION OF DATA ASSIMILATION	87
5. ENSEMBLE FORECASTING	91
5.1 INTRODUCTION.....	91
5.1.1 Atmospheric Predictability	92
5.1.2 Sources of Errors in NWP Model	93
5.1.3 Construction of an Ensemble	94
5.2 ENSEMBLE FORECASTING METHODS	95
5.2.1 Early Methods.....	96
5.2.1.1 Stochastic-Dynamic Forecasting	96
5.2.1.2 Monte Carlo Forecasting	97
5.2.1.3 Lagged Average Forecasting	98
5.2.2 Operational Methods.....	99
5.2.3 Ensemble Kalman Filtering	99
5.3 ENSEMBLE FORECASTING EXPERIMENT	100
6. RESULTS AND DISCUSSIONS.....	103
6.1 USING FDDA IN NWP FOR EGYPT	103
6.1.1 Comparison of Different FDDA Methods	104
6.1.1.1 Cost versus Gain Statistics	104
6.1.1.2 Errors at Different Atmospheric Layers	106
6.1.1.3 Scatter Plots for Different FDDA methods	107
6.1.2 FDDA Results in the Four Seasons	108
6.1.3 Summary of the FDDA Benefits.....	110

6.2 EVALUATION OF FDDA ON NWP FOR EGYPT	113
6.3 EFFECT OF CHANGING MODEL CONFIGURATIONS.....	121
6.3.1 Model Sensitivity to Initialization Time	121
6.3.2 Model Sensitivity to Dynamic Initialization	123
6.3.3 Model Sensitivity to Physics Options	125
6.4 ENSEMBLE FORECASTING EXPERIMENT	127
7. CONCLUSIONS AND FUTURE WORK.....	131
7.1 USING FDDA IN NWP FOR EGYPT	131
7.2 EVALUATION OF FDDA ON NWP FOR EGYPT	132
7.3 EFFECT OF CHANGING MODEL CONFIGURATIONS.....	133
7.4 ENSEMBLE FORECASTING EXPERIMENT	134
7.5 SUGGESTIONS FOR FUTURE WORK.....	134
REFERENCES.....	137
APPENDICES	149
A HISTORICAL OVERVIEW OF NWP	151
A.1. THE EARLY HISTORY OF NWP	151
A.2. THE EVOLUTION OF FORECAST SKILL	153
B THE MM5 PHYSICS OPTIONS	159
B.1. CUMULUS PARAMETERIZATION	159
B.2. PBL AND DIFFUSION SCHEMES	161
B.3. EXPLICIT MOISTURE SCHEMES	163
B.4. RADIATION SCHEMES.....	164
B.5. SURFACE SCHEMES	165
C GLOBAL OBSERVING SYSTEM.....	167
C.1. SURFACE OBSERVATIONS	167
C.2. UPPER-AIR OBSERVATIONS	167
C.3. MARINE OBSERVATIONS	168
C.4. AIRCRAFT OBSERVATIONS	169
C.5. SATELLITE OBSERVATIONS	170
C.6. OTHER OBSERVATION PLATFORMS.....	171
D CRESSMAN ANALYSIS	173
D.1. STANDARD CRESSMAN SCHEME	173

D.2. ELLIPSE CRESSMAN SCHEME	173
D.3. BANANA CRESSMAN SCHEME.....	174
E OPERATIONAL ENSEMBLE FORECASTING	177
E.1. BREEDING	177
E.2. SINGULAR VECTORS	180
E.3. MULTIPLE DATA ASSIMILATION SYSTEMS.....	182
E.4. PERTURBED PHYSICAL PARAMETERIZATION.....	182
E.5. MULTIPLE SYSTEM ENSEMBLES	182
E.6. OTHER METHODS	184
F DETAILED RESULTS	185
F.1. USING FDDA IN NWP FOR EGYPT	185
F.2. EVALUATION OF FDDA ON NWP FOR EGYPT.....	198
F.3. MODEL SENSITIVITY TO PHYSICS OPTIONS	205
F.4. ENSEMBLE FORECASTING EXPERIMENT	241
G AUTOMATION SCRIPTS	259
G.1. INITIALIZATION SCRIPT	259
G.2. RUNNING SCRIPT	272
G.3. VISUALIZATION SCRIPT.....	280

LIST OF TABLES

TABLE 1-1 DIFFERENT MODELING SCALES, GRID LENGTHS AND LOCALES	2
TABLE 1-2 DIFFERENT CLIMATE AND GLOBAL/REGIONAL WEATHER MODELS	4
TABLE 1-3 A TYPICAL NUMBER OF THE MODEL DOFS AND AVAILABLE OBSERVATIONS	5
TABLE 2-1 HISTORY OF THE MM5 AND WRF MODELS UPDATED FROM [63].....	25
TABLE 2-2 SUMMARY OF THE FUNCTIONS, INPUT AND OUTPUT OF MM5 PROGRAMS [17].....	27
TABLE 2-3 THE MM5 FORECAST VARIABLES [17].....	28
TABLE 3-1 RESEARCH DATA ARCHIVE (RDA) DATASETS.....	53
TABLE 3-2 INSTRUMENTATION ONBOARD NOAA SATELLITES [74]	54
TABLE 3-3 NOAA/AVHRR AND NOAA/ATOVS SATELLITE DATA	55
TABLE 3-4 CALIBRATION OF AVHRR VISIBLE AND NEAR-INFRARED CHANNELS	66
TABLE 3-5 CALIBRATION OF AVHRR MIDDLE- AND THERMAL- CHANNELS	67
TABLE 3-6 SAMPLE RETRIEVALS OF THE NOAA/AVHRR DATA AT NARSS	68
TABLE 4-1 EMPIRICAL, STATISTICAL AND ADVANCED DA METHODS	75
TABLE 4-2 COMPARISON BETWEEN EMPIRICAL AND VARIATIONAL ASSIMILATION METHODS	76
TABLE 4-3 DIFFERENT CRESSMAN SCHEMES AND THE CORRESPONDING WEIGHTING FUNCTIONS [58].....	79
TABLE 4-4 SELECTED TIME PERIODS FOR TESTING FDDA ENHANCEMENT OF THE MM5 SIMULATIONS	87
TABLE 4-5 DESCRIPTION OF DIFFERENT SETS OF DA SIMULATIONS	88
TABLE 4-6 GRID AND PHYSICS OPTIONS FOR THE THREE DOMAINS USED IN DA SIMULATIONS.....	88
TABLE 4-7 DESCRIPTION OF THE SIMULATIONS PERFORMED FOR THE EVALUATION OF FDDA	90
TABLE 5-1 DIFFERENT SOURCES OF NWP ERRORS [85]	93
TABLE 5-2 SELECTED TIME PERIODS FOR TESTING MODEL SENSITIVITY TO DIFFERENT PHYSICS OPTIONS	100
TABLE 5-3 ENSEMBLE MEMBERS USED IN THE EF EXPERIMENT FOR EGYPT	101
TABLE 6-1 DESCRIPTION OF THE PERFORMED SIMULATIONS FOR EVALUATING FDDA BENEFITS	103
TABLE 6-2 RMSE AND MABE AND ASSOCIATED GAINS AT DIFFERENT ATMOSPHERIC LAYERS.....	106

TABLE 6-3 STATISTICAL SUMMARY OF THE FDDA BENEFITS IN THE FOUR SEASONS AT ALL STATIONS112

TABLE 6-4 DESCRIPTION OF THE PERFORMED SIMULATIONS113

TABLE 6-5 SUMMARY OF RMSE FOR ALL OBSERVATION SETS AND ALL SIMULATIONS119

TABLE 6-6 SUMMERY OF THE MODEL SENSITIVITY TO INITIALIZATION TIME122

TABLE 6-7 SUMMARY OF THE RESULTS OF TESTING THE MODEL SENSITIVITY TO DYNAMIC INITIALIZATION 124

LIST OF FIGURES

FIGURE 1-1 THE RELATION BETWEEN GLOBAL AND REGIONAL WEATHER MODELS	3
FIGURE 1-2 WMO OPERATIONAL NETWORK [23].....	6
FIGURE 1-3 TYPICAL 6-HOUR ANALYSIS CYCLE [21]	8
FIGURE 1-4 CYCLIC DEPENDENCY OF EF, DA AND NWP	10
FIGURE 2-1 FRAMEWORK OF THE NUMERICAL WEATHER MODELING SYSTEM AT NARSS.....	24
FIGURE 2-2 THE MM5 FLOW CHART [17; ON LEFT] AND CURRENT IMPLEMENTATION [ON RIGHT]	26
FIGURE 2-3 SCHEMATIC REPRESENTATION OF THE HORIZONTAL GRID STRUCTURE [17]	30
FIGURE 2-4 SCHEMATIC REPRESENTATION OF THE VERTICAL GRID STRUCTURE [17]	31
FIGURE 2-5 NESTED DOMAINS USED IN MM5 FOR EGYPT	34
FIGURE 2-6 DEMONSTRATION FOR THE MERCATOR PROJECTION IN THE MM5 MODEL	35
FIGURE 2-7 THE MM5 HORIZONTAL GRID FOR EGYPT DOMAIN	36
FIGURE 2-8 THE VERTICAL GRID IN THE MM5 MODEL FOR EGYPT DOMAIN	37
FIGURE 2-9 THE INITIAL CONDITIONS OF TEMPERATURE FOR EGYPT DOMAIN.....	38
FIGURE 2-10 LEVELS OF SURFACE BOUNDARY CONDITIONS IN THE MM5 MODEL	39
FIGURE 2-11 THE LOWER TEMPERATURE BOUNDARY CONDITIONS FOR EGYPT DOMAIN	40
FIGURE 2-12 TEMPERATURE DISTRIBUTION AT THE UPPER BOUNDARY FOR EGYPT DOMAIN.....	41
FIGURE 2-13 LATERAL BOUNDARY CONDITIONS FOR TEMPERATURE	43
FIGURE 2-14 TEMPORAL FINITE DIFFERENCING IN THE MM5 MODEL [18].....	44
FIGURE 2-15 INTERACTIONS OF PARAMETERIZATIONS IN THE MM5 MODEL [61].....	45
FIGURE 2-16 THE NAMING CONVENTION FOR THE MM5 RUN DIRECTORY	46
FIGURE 2-17 SCHEMATIC FLOWCHART FOR THE INITIALIZATION SCRIPT	47
FIGURE 2-18 THE LOCATIONS OF WMO STATIONS USED FOR THE MODEL EVALUATION	49
FIGURE 3-1 WMO GLOBAL OBSERVING SYSTEM (GOS) [71].....	51
FIGURE 3-2 PROCESSING OF THE RDA CONVENTIONAL DATA AT NARSS	57

FIGURE 3-3 PROCESSING OF NOAA DATA AT NARSS	58
FIGURE 3-4 DATA TRANSMISSION IN HRPT FORMAT [74]	60
FIGURE 3-5 GEOMETRICALLY UNCORRECTED [ON LEFT] AND CORRECTED [ON RIGHT] IMAGES	61
FIGURE 3-6 AVHRR TRUE-COLOR IMAGE [ON LEFT] AND THE CLOUD MASK [ON RIGHT].....	62
FIGURE 3-7 THREE-DIMENSIONAL RETRIEVALS FROM NOAA/ATOVS DATA AT NARSS	64
FIGURE 3-8 TWO-DIMENSIONAL RETRIEVALS FROM NOAA/ATOVS DATA AT NARSS	64
FIGURE 3-9 AVHRR TRUE-COLOR IMAGE AND A SPATIAL SUBSET FOR THE REGION OF INTEREST	65
FIGURE 3-10 VALIDATION OF ATOVS RETRIEVALS AT HELWAN STATION AT 00Z 01 OCT 2005	69
FIGURE 4-1 SCHEMATIC OF GRID POINTS AND IRREGULARLY-DISTRIBUTED OBSERVATIONS [21].....	77
FIGURE 4-2 OBJECTIVE ANALYSIS AND CYCLIC DEPENDENCY	77
FIGURE 4-3 USING FDDA FOR DYNAMIC INITIALIZATION	83
FIGURE 4-4 USING FDDA FOR DYNAMIC ANALYSIS.....	83
FIGURE 4-5 USING FDDA FOR BOUNDARY CONDITIONS	84
FIGURE 5-1 EVOLUTION OF FORECASTING SKILLS TO PROBABILISTIC FORECASTING	91
FIGURE 5-2 “IDENTICAL TWINS” EXPERIMENT WITH LORENZ THREE-VARIABLE MODEL	92
FIGURE 5-3 INITIAL CONDITIONS AND MODEL-RELATED UNCERTAINTIES IN EF	94
FIGURE 5-4 SCHEMATIC OF THE ESSENTIAL COMPONENTS OF AN ENSEMBLE OF FORECASTS [87].....	94
FIGURE 5-5 SCHEMATIC TIME EVOLUTION OF MCF [87]	97
FIGURE 5-6 SCHEMATIC TIME EVOLUTION OF LAF [87]	98
FIGURE 6-1 COST VERSUS GAIN STATISTICS OF DIFFERENT FDDA TECHNIQUES	104
FIGURE 6-2 VERTICAL TEMPERATURE PROFILES FOR DIFFERENT FDDA TECHNIQUES	105
FIGURE 6-3 RMSE AND MABE OF TEMPERATURE AT DIFFERENT ATMOSPHERIC LAYERS	106
FIGURE 6-4 SCATTER PLOTS OF NEAR-SURFACE TEMPERATURE FOR DIFFERENT FDDA TECHNIQUES	107
FIGURE 6-5 RMSE AND MABE OF NEAR-SURFACE TEMPERATURE FOR ALL SEASONS IN ALEXANDRIA	108
FIGURE 6-6 NEAR-SURFACE TEMPERATURES FOR THE FOUR SEASONS IN ALEXANDRIA	109
FIGURE 6-7 DAILY RMSE AND MABE OF NEAR-SURFACE TEMPERATURE IN ALEXANDRIA	110

FIGURE 6-8 TEMPERATURE DIFFERENCE BETWEEN FDDA1 AND FDDA0 SIMULATIONS	111
FIGURE 6-9 STATIONS MAP FOR THE STATISTICAL SUMMARY OF THE FDDA BENEFITS.....	112
FIGURE 6-10 SAMPLE RESULTS FOR THE DIFFERENT SIMULATIONS AND OBSERVATION SETS IN SUMMER	114
FIGURE 6-11 SCATTER PLOTS OF NEAR-SURFACE TEMPERATURE FOR THE DIFFERENT OBSERVATION SETS ..	115
FIGURE 6-12 CORRELATIONS OF ALL SIMULATIONS FOR THE FOUR OBSERVATION SETS	116
FIGURE 6-13 SCATTER PLOTS FOR THE INTERPOLATION SET WITH AND WITHOUT COASTAL STATIONS.....	117
FIGURE 6-14 AVERAGE RMSE IN SUMMER AND WINTER	118
FIGURE 6-15 TEMPERATURE DIFFERENCE BETWEEN REFERENCE AND “DA MQ” SIMULATIONS	120
FIGURE 6-16 NEAR-SURFACE TEMPERATURE AND ABE FOR DIFFERENT INITIALIZATION TIMES	122
FIGURE 6-17 RMSE AND MABE FOR DIFFERENT INITIALIZATION TIMES	122
FIGURE 6-18 TOTAL AND DAILY RMSE WHEN USING DYNAMIC INITIALIZATION	123
FIGURE 6-19 TOTAL AND DAILY MABE WHEN USING DYNAMIC INITIALIZATION.....	123
FIGURE 6-20 NEAR-SURFACE TEMPERATURE AND ABE WHEN USING DYNAMIC INITIALIZATION	124
FIGURE 6-21 NEAR-SURFACE TEMPERATURE AT LUXOR STATION IN AUTUMN	126
FIGURE 6-22 CHANGE IN AVERAGE TEMPERATURE AT LUXOR STATION IN AUTUMN	126
FIGURE 6-23 NEAR-SURFACE TEMPERATURE AT LUXOR STATION IN SUMMER	127
FIGURE 6-24 MIN, MAX AND MEAN TEMPERATURE AT LUXOR STATION IN SUMMER.....	128
FIGURE 6-25 CHANGE IN AVERAGE TEMPERATURE AT LUXOR STATION IN SUMMER	128

NOMENCLATURE

c_p	Air Specific Heat at constant pressure
ds	Grid Spacing
E	Latent Heat Flux
F	Force
g	Earth Gravity
G	Soil Heat Flux, Nudging Factor
H	Sensible Heat Flux, Forward Operator
L	Net longwave Radiation
M	Map scale factor
N	Total Number of Observations
p	Pressure
Q	Water vapor mixing ratio
r_{ah}	Aerodynamic Resistance
R	Radius of Influence
R_e	Radius of the Earth's
R_{iB}	Richardson Number
S	Net Solar Radiation
t	Time
T	Temperature
T_g	Ground Temperature
U	Horizontal velocity component in the West-East direction
V	Horizontal velocity component in the South-North direction
W	Vertical velocity component
x	Spatial Vector
x	State Vector
z	Elevation from ground level

Greek

ρ	Air Density
σ	Sigma levels, Stefan Boltzmann Constant
λ	Latitude
ϕ	Longitude, Probability Density Function
ψ	Stability Parameters
α	Absorptivity
ε	Ground emissivity
θ	Potential Temperature
γ	Observational Quality Factor

Subscripts

a	Analysis
b	Background
e	Earth
g	Ground
Surface	Model Surface Layer
Top	Model Top Layer

Superscripts

a	Analysis
b	Background
o	Observation

ABBREVIATIONS

3DDA	Three-Dimensional Data Assimilation
3DVar	Three-Dimensional Variational Data Assimilation
4DDA	Four-Dimensional Data Assimilation
4DVar	Four-Dimensional Variational Data Assimilation
AMDAR	Aircraft Meteorological Data Relay
ASAP	Automated Shipboard Aerological Programme
AVHRR	Advanced Very High Resolution Radiometric Rader
ADP	Automated Data Processing
ATOVS	Advanced TIROS Operational Vertical Sounder
BV	Bred Vector
CBL	Convective Boundary Layer
CFL	Courent-Friedrichs-Lewy number
CHAMP	CHALLENGING Minisatellite Payload
CISL	Computational and Information Systems Laboratory
CLASS	Comprehensive Large Array-data Stewardship System
CLT	Cairo Local Time
CM	Cressman Interpolation
DA	Data Assimilation
DALR	Dry Adiabatic Lapse Rate
DCS	Data Collection System
DF	Deterministic Forecasting
DOF	Degrees of Freedom
DSS	Data Support Section
ECMWF	European Centre for Medium-Range Weather Forecasts
EEAA	Egyptian Environmental Affairs Agency
EF	Ensemble Forecasting

EnKF	Ensemble Kalman Filter
ETKF	Ensemble Transform Kalman Filter
ENIAC	Electronic Numerical Integrator and Computer
FVLab	Flow Visualization Lab
FDDA	Four-Dimensional Data Assimilation
FNL	NCEP Final Reanalysis
GCM	General Circulation Models
GCOS	Global Climate Observing System
GDPFS	Global Data Processing and Forecasting System
GFS	Global Forecast System
GOS	Global Observing System
GSN	Global Surface Network
GTS	Global Telecommunication System
GUAN	GCOS Upper-air Network
HPC	High Performance Computing
HRPT	High Resolution Picture Transmission
KF	Kalman Filtering
LAF	Lagged Averaged Forecasting
LAM	Limited-Area Models
LIDAR	Light Detection And Ranging
LLV	Leading Lyapunov Vector
LPW	Layer Precipitable Water
LSM	Land Surface Model
LSP	Land Surface Processes
LST	Land Surface Temperature
LV	Lyapunov Vector
MABE	Maximum Absolute Bias Error
MCF	Monte Carlo Forecasting

MM5	Mesoscale Model Version 5
MODIS	Moderate Resolution Imaging Spectroradiometer
MPI	Message Pass Interface
MPP	Massively Parallel Processor
MSVLab	Modeling-Simulation and Visualization Lab
MQ	Multi-quadric Interpolation
NARSS	National Authority for Remote Sensing and Space Sciences
NASA	National Aeronautic and Space Administration
NCAR	National Center for Atmospheric Research
NCEP	National Centers for Environmental Prediction
NCL	NCAR Command Language
NDVI	Normalized Difference Vegetation Index
NMHS	<i>National</i> Meteorological and Hydrological Services
NOAA	National Oceanic and atmospheric Administration
NOAH	National Centers for Environmental Prediction, Oregon State University, Air Force , Hydrological Research Lab.
NSF	National Science Foundation
NWM	Numerical Weather Modeling
NWP	Numerical Weather Prediction
OA	Objective Analysis
OSU	Oregon State University
PBL	Planetary Boundary Layer
PM	Particulate Matter
PSU	Pennsylvania State University
R&D	Research and Development
RBSN	Regional Basic Synoptic Network
RCM	Regional Climate Model
RDA	Research Data Archive

RMSE	Root Mean Square Error
RSMC	Regional Specialized Meteorological Centers
SCM	Successive Correction Method
SDF	Stochastic-Dynamic Forecasting
SLAF	Scaled Lagged Averaged Forecasting
SLP	Sea Level Pressure
SREF	Short-Range Ensemble Forecasting
SST	Sea Surface Temperature
SV	Singular Vectors
T-RANS	Time-dependant Reynolds Averaged Navier-Stokes.
TLM	Tangent Linear Model
TPW	Total Precipitable Water
TSHF	Total Surface Heat Flux
TSHF	Total Surface Heat Flux
UCAR	University Corporation for Atmospheric Research
UHI	Urban Heat Island
USGS	United States Geological Survey
UTC	Universal Time Clock
Var	Variational Data Assimilation
VOS	Voluntary Observing Ship
WALR	Wet Adiabatic Lapse Rate
WCS	Weather and Climate Station
WMC	World Meteorological Centers
WMO	World Meteorological Organization
WRF	Weather Research and Forecast

CHAPTER ONE

INTRODUCTION

In 1998, the Flow Visualization Lab (FVLab) was established in the department of aerospace engineering in Cairo University. The lab started an extensive research work in the area of weather modeling in the department. The main objectives of FVLab were the numerical modeling of weather and climate-related issues for Egypt, the meteorological assessment of air pollution problem, modeling long-term local meteorological climate changes and the investigation of wind and solar energy potentials in Egypt. The cooperation of FVLab with the Egyptian Environmental Affairs Agency (EEAA) started in 2000 using the Eta model [1] for weather forecasting.

The Modeling-Simulation and Visualization Lab (MSVLab) was established at the National Authority for Remote Sensing and Space Sciences of Egypt (NARSS) in 2003 to continue building the aerodynamic weather modeling system for Egypt. In 2003, the first weather prediction model for Egypt was built in MSVLab using the MM5 solver on a cluster of three dual-processor 64-bit machines [2]. Applications of this model started along two lines; first to study the possible meteorological changes due to aggressive land use/land cover changes and the formation of artificial water bodies, see for example [3, 4], and the second is to investigate the extreme air pollution events and temperature inversion over Egypt, see for example, see for example [5, 6, 6, 8, 9].

A direct link between MSVLab and NARSS Weather and Climate Station (WCS) was established in 2004 to automatically process and utilize the station's remotely-sensed data. The obtained data sets were used to initialize and evaluate the numerical weather modeling systems, see for example [10, 11, 12, 13, 14]. In 2006, Four-Dimensional Data Assimilation (FDDA) was started in NARSS to

provide physically consistent estimates of the meteorological conditions [15]. In 2007, the integration of all efforts resulted in developing the first version of the fully automated framework for the numerical weather modeling system for Egypt, see [16] for more details. This model may later be extended to develop a weather-related early warning system.

1.1 NUMERICAL WEATHER MODELING

Numerical Weather Prediction (NWP) is a multiscale problem. There are several NWP models ranging from global scale to local scale. The model classification depends mainly on the resolution used and the physics involved in the model. Global models simulate the entire globe using atmospheric current state conditions using data assimilation and previous global simulations. Regional models are limited to certain geographical areas. They are higher in resolution to capture the local phenomena more accurately. Thus, they need more sophisticated physics, which would require higher computational cost if solved globally.

Numerical weather prediction models use different scales such as macro-, meso and microscales. Mesoscale meteorology studies the atmospheric phenomena at typical spatial scales from 1 to 100 km. This range resolves most weather phenomena that directly impacts human activity. Microscale meteorology is the study of the short-lived atmospheric phenomena smaller than mesoscale, about 1 km or less. Moving from mesoscale to microscale weather modeling requires higher resolution in both horizontal and vertical directions and may require different handling of the physics. Table 1-1 shows a comparison between different modeling scales, grid lengths and locales.

Table 1-1 Different modeling scales, grid lengths and Locales

Scale	Length	Locale
Micro	1 m - 1 km	Local
Meso	1 km - 100 km	Regional
Macro (Synoptic)	100 km - 10 000 km	Continental
Mega	> 10 000 km	Global

These global weather models provide the regional models with the required initial and boundary conditions. The regional weather models are typically used in weather forecasting and atmospheric applications in research studies. In general, these models solve the Euler equations for atmospheric dynamics using numerous physics schemes to resolve the planetary boundary layer, microphysics, cumulus, radiation in addition to land surface processes. Figure 1-1 shows the relation between global and regional weather models where the initial and boundary conditions for the regional models are obtained from the global models. This is usually done using some form of interpolation and map projection such as Mercator, Lambert, Polar Stereographic or Geographic Lat/Lon.

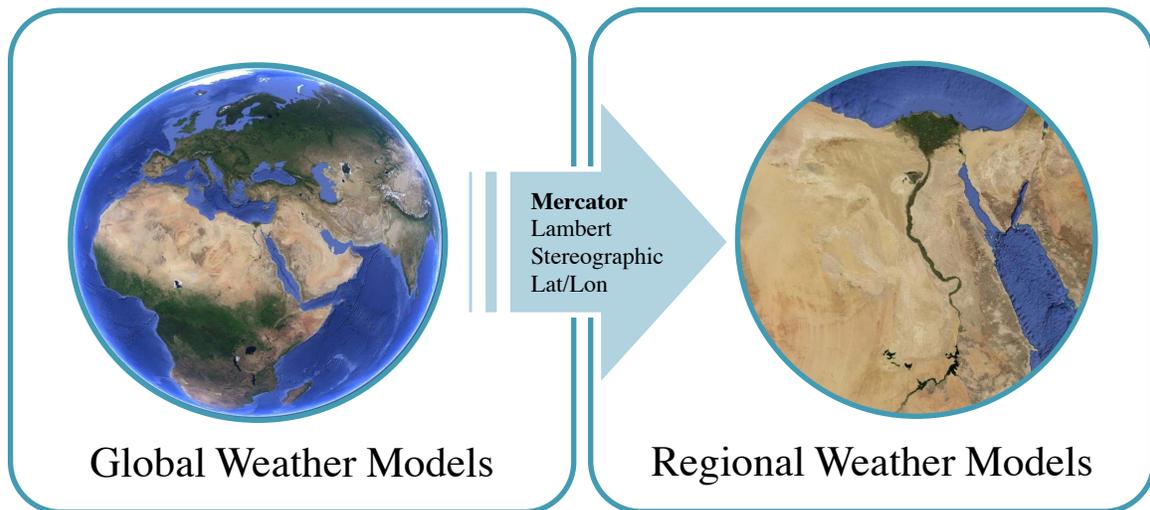


Figure 1-1 The relation between global and regional weather models

On the other hand, climate models use quantitative methods to simulate the interactions of the atmosphere, oceans, land surface, and ice. They are used for many applications ranging from the study of the dynamics of the climate system to the prediction of future climate. There are various types of climate models. Climate models that look at few variables of the climate system may be simple enough to run on a personal computer. Other climate models take into account many factors of the atmosphere, biosphere, geosphere, hydrosphere, and cryosphere such as to model the entire Earth system. They take into account the

interactions and feedbacks between these different parts of the planet. They also include so many mathematical calculations that they must be run on supercomputers. In general, the more complex a model, the more factors it takes into account and the fewer assumptions it makes. There are many number of different climate and global/regional weather models are available. Table 1-2 lists some of the available models.

Table 1-2 Different climate and global/regional weather models

Type	Models
Climate	IGCM · HadCM3 · GFDL CM2.X · CGCM · CCSM · ECHAM
Global	IFS · GEM · GFS · NOGAPS · UM · JMA · GME · ARPEGE
Regional	MM5 · NAM · RUC · RAMS · WRF · RAQMS · HIRLAM · LAPS

Selecting a NWP model is based on the scientific evaluation and the quality control. Selecting the inputs datasets and model configurations based on operational evaluation are among the most important factors that determine how accurate the meteorology is estimated. The Fifth-Generation Mesoscale Model (MM5) [17, 18] was selected in this study because of its broadening usage and as a continuation of the related local studies. MM5 is a limited-area, nonhydrostatic, terrain-following sigma-coordinate model designed to simulate or predict mesoscale and regional-scale atmospheric circulation.

The MM5 model has the following features:

- (i) a multiple-nest capability,
- (ii) nonhydrostatic dynamics which allows the model to be used at a few-kilometer scale,
- (iii) multitasking capability on shared- and distributed-memory machines,
- (iv) a four-dimensional data-assimilation (FDDA) capability, and
- (v) multiple physics options.

MM5 is ported for a variety of platforms such as Cray, Sun and SGI workstations, and Linux based PCs.

Numerical weather prediction is an initial boundary value problem. The weather is a chaotic system: small errors in the initial conditions of a forecast grow rapidly and affect predictability. Furthermore, predictability is limited by model errors linked to the approximate simulation of atmospheric processes used in the numerical models. These two sources of uncertainty limit the skill of single deterministic forecasts in an unpredictable way, with days of high/poor quality forecasts followed by days of poor/high quality forecasts [19]. Multiple, individual forecasts may be produced using different initial/boundary conditions, different physics formulations of a NWP model and/or different NWP models [20].

The problem of determination of the initial conditions, the available observational data, the assimilation of such data in the model, and the method of forecasting that takes into account the uncertainties associated with the initial conditions and/or the forecast model will be presented in the next subsections.

1.1.1 Initial Boundary Value Problem

The problem of determining the initial conditions for a NWP model is very important and complex, and it has become a science by itself. Interpolation methods were developed for fitting the observational data to the model grid. Usually, the available data are not enough to initialize the models. Moreover, data distribution is very non-uniform in space and time. The number of observations is less than the number of the model degrees of freedom (DOF) by at least two orders of magnitude. This is clear in local-scale modeling for areas of poor data (e.g., Egypt) rather than for data-rich areas (e.g., USA and Europe) [21]. Table 1-3 shows a typical model DOFs and the corresponding numbers of observations at global, regional and local scales.

Table 1-3 A typical number of the model DOFs and available observations

Scale	Global	Regional	Local
Model DOFs	$O(10^7: 10^8)$		
Observations	$O(10^5: 10^6)$	$O(10^3: 10^5)$	$O(10^1: 10^3)$

The lateral boundary conditions for a model run require that the gridded analyses data covers the entire time period of the run. The land surface model employed in these models requires high resolution landuse/vegetation as well as the soil temperature data [22].

1.1.2 Observational Data

The Weather Meteorological Organization (WMO) Operational Network [23], shown in Figure 1-2, comprises 188 Members' National Meteorological and Hydrological Services (NMHSs), 3 world meteorological centers, 40 regional specialized meteorological centers, 10 scientific and technical programs, and 30 regional meteorological training centers. The network enables the delivery of accurate, timely and reliable observations, forecasts and warnings, and analysis of meteorological, hydrological and climate -related hazards. The WMO Operational Network includes

- (i) WMO Global Observing System (GOS),
- (ii) WMO Global Telecommunication System (GTS) and
- (iii) WMO Global Data Processing and Forecasting System (GDPFS).

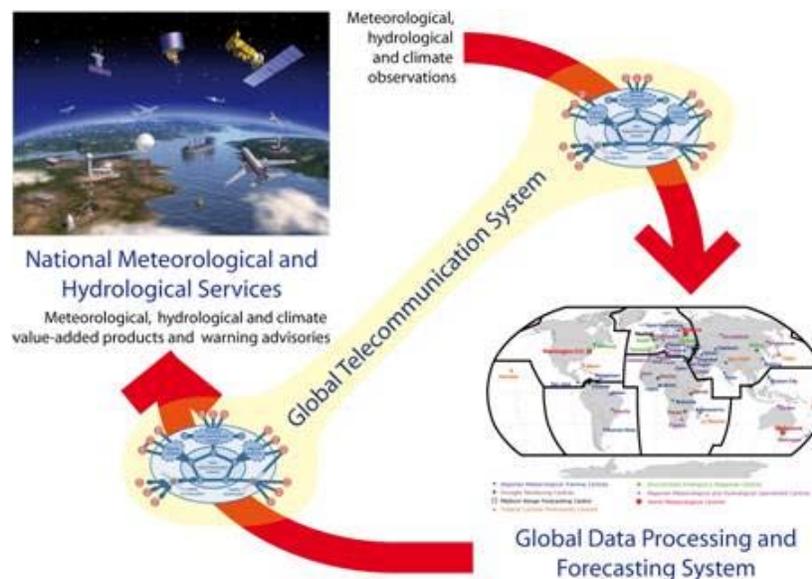


Figure 1-2 WMO operational network [23]

The WMO GOS involves two subsystems:

- (i) Surface-based subsystem that is operated mainly by Member NMHSs,
- (ii) Space-based subsystem that is operated by either national or international space agencies.

Since 1963, the WMO GOS has enabled coordination of the observation and collection of weather, water and climate information from around the globe. Through this system, data are collected from 17 satellites, hundreds of ocean buoys, aircraft, ships and nearly 10 000 land-based stations and are exchanged and archived in near-real time. Everyday more than 50,000 weather reports and several thousand charts and digital products are disseminated among countries to provide services for societies benefit.

WMO GTS is composed of a dedicated network of surface-based and satellite-based telecommunication links and centers operated by countries 24 hours a day, seven days a week all year round. It interconnects all NMHS for round-clock reliable and near-real-time collection and distribution of all meteorological and related data, forecasts and alerts. WMO GTS is the backbone system for global exchange of data and information in support of multi-hazard, multipurpose early warning systems, including all meteorological and related data; weather, water and climate analyses and forecasts. National Meteorological and Hydrological Services are linked among themselves through the GTS network. This secured communication network enables real-time exchange of information, critical for forecasting and warning of hydro-meteorological hazards.

The WMO GDPFS is organized as a network of 3 World Meteorological Centers (WMCs) and 40 Regional Specialized Meteorological Centers (RSMCs). Operated or supported by NMHSs, those centers carry out data archiving, processing and forecasting functions at the global and regional levels, respectively. This coordinated system provides analysis, modeling, forecasting and other products and services in support of forecasting and early warnings of weather- and climate-

related hazards to all countries. WMO specifically ensures that capacities are implemented within the NMHSs of developing and least developed countries for issuance of forecasts and warning of various hydro-meteorological hazards.

1.1.3 Data Assimilation

Data Assimilation (DA) is a technique in which the observed information is accumulated into the model state using advantage of consistency constraints with laws of time evolution and physical properties [24]. It is a promising tool to enhance the results of NWP models. There are two main types of DA, three-dimensional DA, 3DDA, at certain time and four-dimensional DA, 4DDA, where time dimension is added. The observations of the current, and possibly past, state of a system are combined with the forecast to produce an analysis. The analysis is considered as the best estimate of current state of the system.

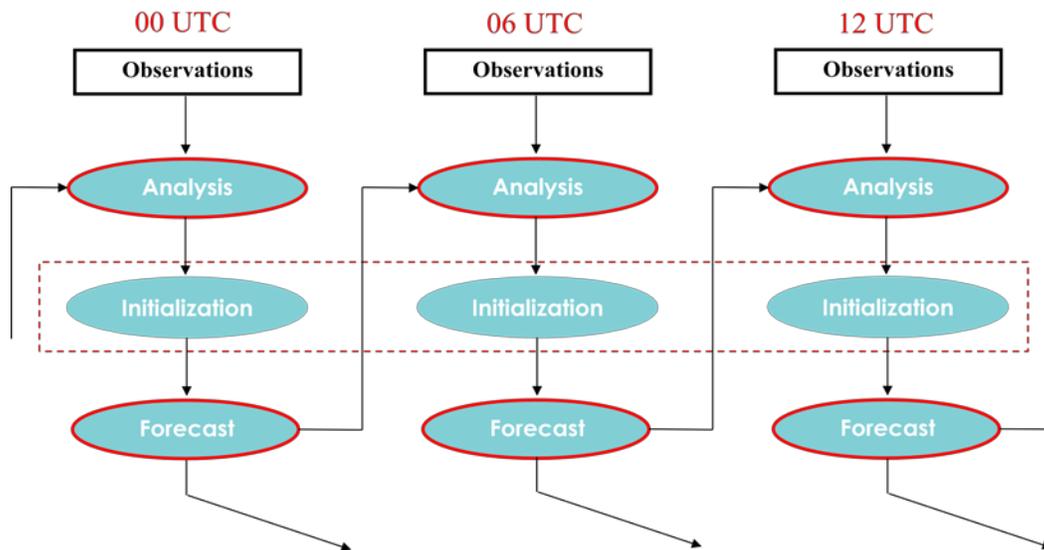


Figure 1-3 Typical 6-hour analysis cycle [21]

To overcome the lack of enough observations to prepare the required initial conditions for the forecasts, additional information (the background) is used. Initially climatology was used as a first guess. Then, short-range forecast was used where the forecasts became better. DA proceeds by analysis cycles including analysis and forecast steps. The analysis step is performed to balance the

uncertainty in the data and the forecast. The model is then advanced in time and its results become the forecast for the next analysis cycle. Figure 1-3 shows a schematic for a typical DA analysis cycle.

The forecast (the prior) is combined with the new observations to create the analysis (the posterior), and the initialization imposes a dynamic balance between mass and wind fields. DA methods are divided into empirical, statistical and advanced methods. The different methods perform differently and generally it is not clear which approach is better. So, the selection of a method depends on the application and the available computational resources. The empirical DA methods are used in this work because they are simple and computationally more efficient than the other computationally-expensive methods.

In this work, DA is used to implement the observed information into the PSU/NCAR NWP model (MM5). The MM5 model is initialized from the NCEP Final Reanalysis (FNL) or Global Forecasting System (GFS) data. One-way nesting type for 3 nested domains with resolutions of 81, 27 and 9 km is used. The available conventional and remotely-sensed observations are implemented using empirical DA methods. Objective Analysis (OA) and Four-Dimensional Data Assimilation (FDDA) are available in the public releases of NWP models. The processing and visualization of satellite data are also included. This is starting by the investigation of the benefits of using DA in NWP for Egypt in the four seasons. The effects of DA via different nudging options on the qualities of simulations/predictions are investigated at locations with different observation site density inside Egypt, stressing on the effects on accuracy at the observational-data-void areas and at the locations with complex terrain.

1.1.4 Ensemble Forecasting

Ensemble Forecasting (EF) is a method of forecasting that takes into account the uncertainties associated with the initial conditions and/or the forecast model.

Ensemble Forecasting was first introduced to represent initial error related uncertainties. Then, model related errors were introduced in the modeling to capture variations in the forecast uncertainties arising from model imperfections [25]. This means that an ensemble forecast is a collection of many forecasts that all verify at the same time. Ensemble members represent possibilities given the uncertainties associated with the forecasting. These possibilities can be used to estimate probabilities of various events as well as an average forecast (ensemble mean) [26]. Figure 1-4 shows the cyclic dependency of EF, DA and NWP.

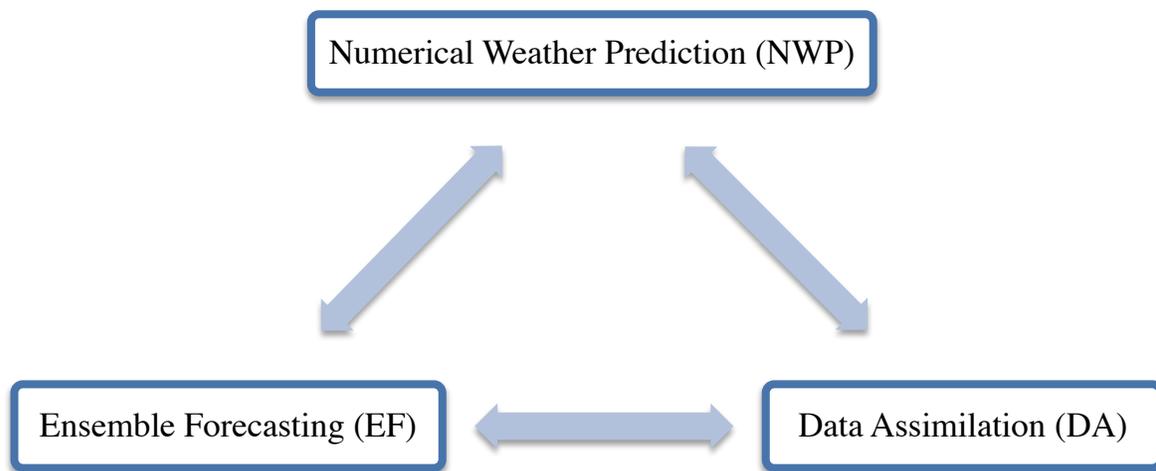


Figure 1-4 Cyclic dependency of EF, DA and NWP

In this work, the effects of changing different model configurations are studied. This represents a preliminary study for ensemble forecasting which takes into account the uncertainties associated with the initial conditions and/or the NWP model. Nine different physics options are used in the study based on 3 Planetary Boundary Layer (PBL) schemes and 3 radiation schemes. The total number of ensemble members is 36 forecast. Nine forecasts are used as reference forecasts and the other 27 forecasts utilize data assimilation which is used to provide 3 different sets of initial conditions. The first set is produced using objective analysis of the first-guess and the observations. The second and third sets are produced using objective analysis and FDDA dynamic initialization for 6 and 12 hours,

respectively. The ensemble mean is evaluated based on the percentage difference in the maximum, minimum and average temperatures.

1.2 MOTIVATION

Egypt as a developing country has special needs for integrating Numerical Weather Prediction, NWP, to several of its national projects. Examples include the ambitious renewable energy plan to 2020 and beyond, the treatment of air pollution in mega cities, the efforts of health sector to manage trans-boundary diseases and infections, the management of the agriculture and irrigation activities, including crop health, and the needed optimal utilization of natural resources. Accuracy of the computations is expected to increase with the density of the observations but may be countered by the distribution of the observation sites and the grid size used.

Use of techniques like Data Assimilation, DA, and Ensemble Forecasting, EF, in NWP practices in developed countries usually results in better weather simulations/predictions. The application of such techniques in Egypt may not always produce the expected better accuracy. The main reason for this problem is the scarcity of observational data and the non-uniformity of the sparse observation sites. Moreover, the simulations/predictions may have different qualities at different locations of the same region depending on whether the location is in the vicinity of an observation site, between observations sites, far from observations sites, or at a site with complex terrain.

The motivation of this work is to test the effects of using data assimilation in NWP regional model for Egypt. This includes the study of the qualities of weather simulations for different weather regimes. Special interest is given to the observational data-void areas and areas with difficult terrain properties. Moreover, the study of the effects of changing the model physics options subjected to the complexity of terrain properties in Egypt is essential. This is to find whether there

are optimum physics options or special treatment should be taken into consideration for each area in Egypt.

1.3 LITERATURE REVIEW

In the early 20th century, Vilhelm Bjerknes argued that atmospheric physics had advanced sufficiently to allow weather to be forecast using calculations. He thought that weather could be described by seven primary variables, namely: pressure, temperature, air density, air water content, and the three components of wind velocity, and he developed a set of seven equations. In 1922, Lewis Fry Richardson developed the first NWP system. His method was based on simplified versions of Bjerknes's "primitive equations" of motion and state with adding an eighth variable for atmospheric dust. Still, this task remained so large that Richardson did not imagine that it as a weather forecast technique. His own attempt to calculate the weather for a single eight-hour period took six weeks and ended in failure. Only in the 1940s, when digital computers made possible automatic calculation on an unprecedented scale, did Richardson's technique become practical [27].

Although the early methods continued to be used and developed until the 1950s, both the lack of faster calculating methods and the dearth of accurate observational data limited their success as forecasting techniques. In 1950, the first successful NWP was performed by a team composed of the American meteorologist Jule Charney, Norwegian meteorologist Ragnar Fjørtoft and applied mathematician John von Neumann, using the first large-scale, electronic, digital computer capable of being reprogrammed to solve a full range of problems called the Electronic Numerical Integrator and Computer (ENIAC). They used a simplified form of atmospheric dynamics based on the barotropic vorticity equation.

In 1954, the Royal Swedish Air Force Weather Service in Stockholm was the first in the world to begin routine real-time NWP. The Institute of Meteorology at the

University of Stockholm, associated with the eminent meteorologist Carl-Gustaf Rossby, developed the numerical model. Forecasts for the North Atlantic region were made three times a week on the Swedish BESK computer using a barotropic model. This was followed, 6 months later, by the first operational NWP in USA. A historical overview of NWP is presented in Appendix A.

1.3.1 Data Assimilation

In early experiments, Richardson (1922) and Charney et al. (1950) performed hand interpolations of the available observations to grid points, and the fields of initial conditions were manually digitized. This was a very time consuming procedure. The need for an automatic “Objective Analysis” quickly became apparent. The Objective Analysis (OA) is the first 3DDA method and it is based on simple interpolation approaches. Interpolation methods fitting data to grids were developed by many researchers in different countries (e.g., Panofsky (1949), Gilchrist and Cressman (1954), Barnes (1964, 1973, 1974, 1978 and 1994)).

The first 4DDA method was developed by Bergthorsson and Doos (1955) in Sweden and by Cressman (1959) of the US Weather Service known as “Successive Correction Method (SCM)”. This method is based on the estimated error statistical covariance of the forecast and observations and it was used in the operational NWP systems in 1950’s – 1960’s [28].

In 1958, Sasaki [29] introduced the variational approach to meteorological analysis. The 3DDA Variational method (3DVar) is based on the minimization of a cost function. The cost function is proportional to the square of the distance between the analysis and both the background and observations [30]

The breakthrough in the field of DA was achieved by Gandin (1963) who introduced the Statistical/Optimal Interpolation (OI) method. This method is a 3DDA method and it depends on the regression analyses, which utilizes

information about the spatial distributions of covariance functions of the errors of the "first guess" field (previous forecast) and "true field". These functions are not defined. However, different approximations were assumed. OI is similar to SCM but its weights are not computed empirically [31].

Attempts to introduce Kalman Filter (KF) algorithms as 4DDA tool came later. KF is an efficient recursive filter that estimates the state of a dynamic system from a series of incomplete and noisy measurements. This was is a very difficult task since the full version of KF algorithm requires solution of the enormous number of additional equations. The number of the model degrees of freedom is of order 10^7 while the number of observations is of order 10^5 - 10^6 . Some examples are Extended Kalman Filter (EKF) (Jazwinski (1970); Gelb (1974); Ghil and Malanotte-Rizzoli (1991)), Reduced-Rank Square-Root Kalman Filter (RRSQT KF) (Todling and Cohn (1994)), and Ensemble Transform Kalman Filter (ETKF) (Bishop et al. (2001)) [32, 33].

A 4DDA method similar to OA was developed by Kistler (1974) which is called Four-Dimensional Data Assimilation (FDDA). FDDA is based on the simple idea of nudging or Newtonian relaxation (the second axiom of Newton). The idea is to add in the right handside of dynamical equations of the model a term that is proportional to the difference between the calculated meteorological variable and the observed value. This term that has a negative sign keeps the calculated state vector closer to the observations [34].

In 1986, Lorenc [35] showed that all the above-mentioned Kalman filtering or variational methods are in some limit equivalent, i.e. under some assumptions they minimize the same cost function. However, in practical applications these assumptions are never fulfilled. Different methods perform differently and generally it is not clear what the approach is better. The fundamental questions also arise in application of the advanced data assimilation techniques such as

convergence of the computational method to the global minimum of the functional to be minimized.

In 1990, Stauffer and Seaman developed the original FDDA for the MM5 model and then it was modified and developed for the Weather Research and Forecast (WRF) model [36]. In 1997, the MM5 adjoint which is a tool effectively computes the gradient of any MM5 forecast aspect with respect to the model's control variables) and 4DDA Variational method (4DVar) were released. The first 4DVar was implemented at European Center for Medium-range Weather Forecasts (ECMWF) at the end of 1997 by Bouttier and Rabier [37]. A beta version of the Weather Research and Forecasting (WRF) model was released in 2000. In 2001, the MM5-3DVar was adopted as a starting point for WRF-3DVar. In 2002, the MM5-3DVar, developed as part of Taiwan's CAA project, was released [38].

The first public release of WRF Variational data assimilation (WRF-Var) was in 2003. Xiao assimilated Doppler radar radial observational data using WRF-Var (2005-2008) [39, 40, 41, 42]. In 2009, WRF 3.1 included 4DVar and radiance data assimilation capabilities. The WRF-Var is updated with the development of the WRF model till the current public release WRF 3.2 (April 2, 2010).

1.3.2 Ensemble Forecasting

By the early 1950s, some meteorologists considered applying statistical methods to weather prediction to cope with the uncertainties inherent in forecasting. In 1952, Sutcliffe stated that the scientific problem of forecasting shall be solved not by becoming over more accurate in our forecasts but by knowing what we should be able to predict in theory and what is required to make the prediction in practice. In 1957, Malone discussed in some detail the barriers to improved prediction such as enormous complexity of atmospheric processes, the growth of major disturbances from small perturbations, and the incompleteness of observations. He notes that the observational network is at best a rather crude sampling device and

agreed with Sutcliffe that a form of statistical physics applicable to meteorological prediction seems to be required [43].

In 1960, Gleeson developed a statistical theory that attempts to cope with the prediction difficulties directly. It is based on the assumption that weather-observing stations are distributed at random relative to features of hypothetically true synoptic patterns. In 1963, Lorenz investigated fundamental aspects of atmospheric predictability. He demonstrated that weather, even when viewed as a deterministic system, may have a finite prediction time. Further, predictability varies with different weather situations in a way not easily discernible by naked eye examination of weather maps [44].

The more realistic systems of equations which have subsequently been used in predictability studies, such as the models of Smagorinsky (1963), Mintz (1964), and Leith (1965), have been described by Charney et al. (1966). These models do not agree with each other, but Charney et al. concluded that a reasonable estimate of the time required for small errors to double, in the root mean square sense, is five days. In 1969, Lorenz also calculated that the average limit to atmospheric predictability at planetary scales is on the order of 10 days [45].

By 1970, Tatarskiy (1969), Epstein (1969), and Gleeson (1970) had proposed methods to forecast probabilities. These procedures do not involve an ensemble but rather forecast statistical quantities directly. The difficulty with these approaches is that they require an enormous amount of calculations, even for computer resources available in the foreseeable future. As an alternative, Leith (1974) demonstrated that an ensemble of roughly 10 forecasts seemed to be large enough to make real improvements in 6–10-day forecasts. With the advances in computing that were taking place at that time, ensemble forecasting, or “Monte Carlo” forecasting as Leith called it, became a distinct possibility.

During the 1980s, ensemble forecasts were computed in a research mode. This was to establish procedures and assess utility. One of the simplest approaches to ensemble forecasting is to consider a collection of forecasts issued at different times but they are verified at the same time. This technique, known as the Lagged Average Forecast (LAF) method, was discussed by Hoffman and Kalnay (1983). For example, a 24-h forecast made this morning could be considered along with a 36-h forecast made last night, a 48-h forecast from yesterday morning, etc. The advantage of the LAF method is that it uses forecasts that already exist. Its chief disadvantage is that the forecasts in an LAF ensemble are not even close to being equal contenders since the newest (shortest range) forecast will almost always be considerably more accurate than the oldest (longest range) forecast.

As computer power increased, it became possible to compute multiple forecasts that all start at the same time. Medium-range ensemble forecasts have been produced operationally in the United States National Centers for Environmental Prediction (NCEP, United States) and European Centre for Medium-Range Weather Forecasts (ECMWF, Europe) since late 1992 [47, 48]. In December 1992, the NCEP operational forecast ensembles consisted of 14 forecasts, 4 of which were computed at 0000 UTC and the remaining 10 were computed 12–48 h earlier (Tracton and Kalnay 1993), see [46].

The potential utility of short-range ensemble forecasting was discussed at a workshop in 1994. The principal recommendation of that workshop was to perform a pilot study in which an ensemble of regional 48-hour forecasts would be computed weekly. The ensemble required an adjustment because the range of values in the ensemble underestimated the range of values in the verifications. In 1996, the NCEP ensemble contained 17 forecasts of which 12 were computed at 0000 UTC and 5 were computed 12 hours earlier (Kalnay and Toth 1996).

Each of the worldwide meteorological centers produces the ensembles using different forecast models and different ensemble construction techniques. Ensemble forecasts are also produced operationally at several other centers around the world as well. As computer power continues to increase, the number of forecasts in an ensemble, the complexity of the model, the model resolution, and the length of the forecast are all expected to increase. At the Canadian Meteorological Centre (CMC), ensemble forecasts have been produced operationally since January 1996. For more details, see [49, 50, 51, 52].

As of 1997, both NCEP and ECMWF used ensemble forecasts primarily for synoptic and planetary scales in the multiday forecast range, however ensembles are potentially useful at all space and time scales. Unfortunately, even at forecast lead times of a few hours, mesoscale features in an ensemble of forecasts will differ. After a few days, synoptic-scale forecasts will exhibit noticeable differences, with planetary scale forecasts diverging after that. This is the cascade of uncertainty mentioned earlier.

During the 1990s, using simplified versions of Kalman filtering in ensemble forecasting were discussed. In 1995, the use of ensembles in a reduced-rank extended Kalman filter was discussed by Fisher and Courtier. In 1997, the use of bred mode information for improving analyses was discussed in Pu et al. In 1999, Buizza and Palmer described the use of ensemble forecast statistics for specifying improved stationary background error statistics to four-dimensional variational analysis (4DVar). At the end of 1990s, Burgers et al., van Leeuwen, Mitchell and Houtekamer, and Hamill and Snyder discussed the use of an ensemble of forecasts using a technique called the Ensemble Kalman Filter (EnKF) [53].

The Ensemble Transform Kalman Filter (ETKF) was initially applied to the adaptive sampling problem; for example, Majumdar et al. (2001 and 2002). ETKF is one variant of ensemble-based Kalman square root filters. In 2003, Lorenc

reviewed and compared different ensemble Kalman filters such as ETKF, EAKF, EnSR and PO method and 4DVar for data assimilation [54]. In 2004, Etherton and Bishop showed how ETKF ensemble perturbations enabled a highly efficient hybrid data assimilation scheme [55]. A local Ensemble Kalman Filter (LEKF) was proposed by Ott et al. (2004) and then by Szunyogh et al. (2005) [56, 57].

In the last five years (2005-2010), the Ensemble Forecasting (EF) and data assimilation (DA) were considered as two problems that have the same solution. Recently, Hybrid ETKF/Var ensemble data assimilation has been developed, but it is not available in the current public releases of the NWP models because it is computationally very expensive. Hybrid EF/DA system may be more robust for small ensemble size and/or large model errors [58].

1.4 THESIS OBJECTIVES

Remote sensing technology provides a vast amount of data about the daily changing state of the atmosphere, ocean, and land surface. This remotely-sensed data with the conventional observations can be implemented into NWP models via Data Assimilation (DA). Ensemble Forecasting (EF) is a NWP method that is used to attempt to generate a representative sample of the possible future states of the atmosphere. Probabilistic forecasts of the weather may then be generated from the relative frequencies of events in the ensemble. EF and DA are very useful and practical ways for estimating the uncertainty of a weather forecast and improving the accuracy of NWP results.

In this thesis, the benefits of using data assimilation in NWP for Egypt based on the MM5 model are investigated in the four seasons of the year. The available real-time conventional and satellite observations are implemented using empirical data assimilation methods, Objective Analysis (OA) and Four-Dimensional Data Assimilation (FDDA), which are available in the public releases of NWP models.

The processing and visualization of satellite data, mainly NOAA/ATOVS to retrieve temperature and water vapor profiles, is included.

The effects of DA via different nudging options on the results of simulations/predictions are investigated at locations with different observation site density inside Egypt. This study will stress on the accuracy at the observational-data-void areas and at the locations with complex terrain. Then, a preliminary short-range EF experiment for Egypt will be built, tested and evaluated in winter and summer.

In summary, the main objectives of this thesis are:

1. Using DA to enhance the results of NWP for Egypt by implementing the available remotely-sensed and conventional observations.
 - Collection and processing of remotely-sensed data and conventional observations to be used for DA and validation purposes.
2. Evaluating the effects of DA on NWP for Egypt.
 - Finding the locations of large errors and trying to enhance the results by choosing different DA or physics options.
3. Studying the sensitivity of the NWP model to different configurations.
 - Evaluating the effect of changing model initialization and/or physics options on the accuracy of the results.
4. Building a preliminary short-range ensemble forecasting system for Egypt which may be further developed for operational use.
 - Testing the benefits of EF by a preliminary experiment using different sets of physics options and initial conditions.

1.5 THESIS ORGANIZATION

This thesis consists of seven chapters and seven appendices. In chapter two, the numerical weather modeling system for Egypt will be discussed including the

governing equations, the initial and boundary conditions, the solution technique, and the model evaluation methodology.

Chapter three presents the sources, processing, analysis and quality of the observational conventional and satellite data available at NARSS Weather and Climate Station (WCS) and/or the Research Data Archive (RDA).

Chapter four introduces the analysis problem and shows different DA methods and types. Then, the empirical DA methods used in this thesis will be presented. This includes the 3DDA method which is called Objective Analysis (OA), the 4DDA method which is called Four-Dimensional Data Assimilation (FDDA), the different schemes, and the quality control for observations. Then, the work done to utilize DA in NWP for Egypt is presented.

Chapter five introduces the atmospheric predictability, the sources of NWP errors, and construction of an ensemble. It investigates the ensemble forecasting philosophy, different EF methods and a preliminary short-range EF system for Egypt based on the available observational data and computational resources.

The results are discussed in chapter six. It includes the enhancement and evaluation of using FDDA in NWP for Egypt, the MM5 model sensitivity to different physics options and/or initial conditions, and a preliminary SREF experiment for Egypt which may be further developed for operational use. Most of the results are carried out in the four seasons to evaluate the modeling system at different locations and weather regimes.

Lastly, Chapter seven presents the conclusions from this study and recommendations for future work.

Appendix A presents the early history of numerical weather prediction and the evolution of forecast skill.

Appendix B presents the different physics options in the MM5 model including cumulus parameterizations, planetary boundary layer and diffusion schemes, explicit moisture schemes, radiation schemes, and surface schemes.

Appendix C presents the components of the global observing system including surface, upper-air, marine, aircraft, and satellite observations.

Appendix D presents the Cressman analysis including standard, ellipse and banana schemes.

Appendix E presents the operational ensemble forecasting methods including breeding, singular vectors, multiple data assimilation systems, perturbed physical parameterization, multiple system ensembles, and other methods.

Appendix F presents some detailed results at selected locations inside Egypt.

Appendix G presents the automation scripts written in this study for the initialization and running of the numerical modeling system and the visualization of its outputs.

CHAPTER TWO

MODELING SYSTEM

In this chapter, the numerical weather modeling system for Egypt will be discussed including the governing equations, the initial and boundary conditions, the solution technique, and the model evaluation methodology.

2.1 INTRODUCTION

The numerical weather modeling system for Egypt in NARSS is based on the fifth-generation PSU/NCAR Mesoscale Model (MM5) and the Weather Research and Forecasting (WRF) model. The modeling started with using only the MM5 model because the WRF model was under development at the time of starting the modeling. Then, the WRF model was included in the system. The two models may be coupled later to optimize the system performance producing more accurate/physically-consistent results.

2.1.1 A Modeling System Framework

The framework of the numerical modeling system for Egypt involves five main components. The five components of the framework are linked automatically using a control script and many scripts for each component. The first component is the inputs which are divided into basic and advanced inputs. The basic inputs are the data required to initialize the numerical model. The advanced inputs are the remotely-sensed data used to update the surface boundary conditions. The second component is the core of the system solver or numerical model (MM5/WRF). The third component is the observational data used to enhance the model results via data assimilation and it also helps to evaluate the model. The fourth component is the evaluation unit which checks the quality of the observations and validates the

model. The fifth component is the outputs. Figure 2-1 shows the framework of the modeling system.

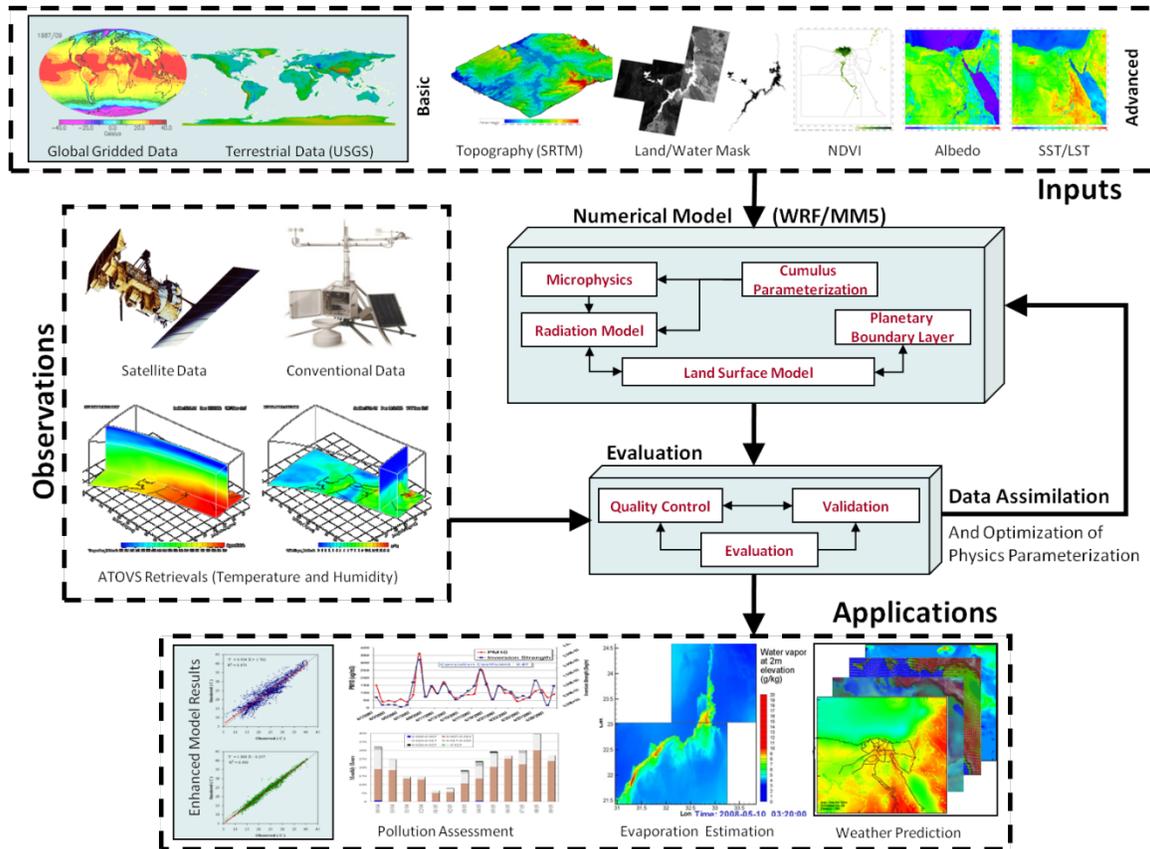


Figure 2-1 Framework of the numerical weather modeling system at NARSS

2.1.2 The MM5 Modeling System

The MM5 model has been developed as a community mesoscale model. It is a limited-area, nonhydrostatic, terrain-following sigma-coordinate model designed to simulate or predict mesoscale atmospheric circulations. Originally, MM5 was developed by Anthes [59] and then updated by Anthes and Warner [60]. The nonhydrostatic version was developed by Dudhia [61] which employs reference pressure as the basis for a terrain-following vertical coordinate and the fully compressible system of equations. For more details about MM5 development, see Grell et al. [62].

The last major MM5 release version 3.7 was in December 2004, with the last bug fix release in October 2006. At this point NCAR stopped developing the MM5 model and the code was frozen. NCAR stopped providing tutorials for the MM5 model in January 2005. Then, the WRF model, developed at NCAR and operated by the University Corporation for Atmospheric Research (UCAR), become fully supported by NCAR including online tutorials and documentation. The WRF modeling system now has all the capabilities that MM5 system had. Table 2-1 summarizes the history of MM5 and WRF models [63].

Table 2-1 History of the MM5 and WRF models updated from [63]

Date	Subject
Late 60's	R. Anthes developed a 3-layer hurricane model
1970s	Anthes and his students developed a general mesoscale model based on the hurricane model. The model evolved from MM0 to MM3.
1980s	Penn State and NCAR jointly developed the MM4 (as part of the Regional Acid Deposition Modeling Project).
Late 80's	NCAR/MMM started supporting MM4 as a community model (with annual workshops and tutorials).
1992	MM5 (with nonhydrostatic, multi-nesting, FDDA, improved physics and numerics) was released.
1996	R. Gall and B. Kuo visited NCEP, began the discussion on the joint development of a next-generation WRF model.
1997	MM5 adjoint and 4DVar were released.
2000	Beta release of the Weather Research and Forecasting (WRF) model.
2002	MM5 3DVar (developed as part of Taiwan's CAA project) was released. MM5 3DVar served as the basis for WRF 3DVar.
2004	WRF 2.0 released including nesting capability. Final version of MM5 3.7 frozen. No further development on MM5.
2006	WRF 2.2 released including 3D grid analysis nudging, observation nudging and more physics options.
2008	WRF 3.0 released including more physics options and bug fixes.
2009	WRF 3.1 released including Gridded surface analysis nudging, spectral nudging and more physics options. 4DVar and radiance data assimilation capabilities.
2010	WRF 3.2 released including new features, updates, bug fixes, and software improvement.

The MM5 solves the continuity (pressure tendency), Euler Equations, and the energy equations to predict the pressure, velocity field and temperature. Other parameterized submodels are used to simulate the atmospheric boundary layer, short wave and long-wave radiations, water vapor transport, cloud and precipitation processes and transfer of heat momentum and moisture between land surface and atmosphere. The schematic diagram in Figure 2-2 shows the flowchart of the complete MM5 modeling system. It is intended to show the order of the programs, flow of data, and to briefly outline their primary functions.

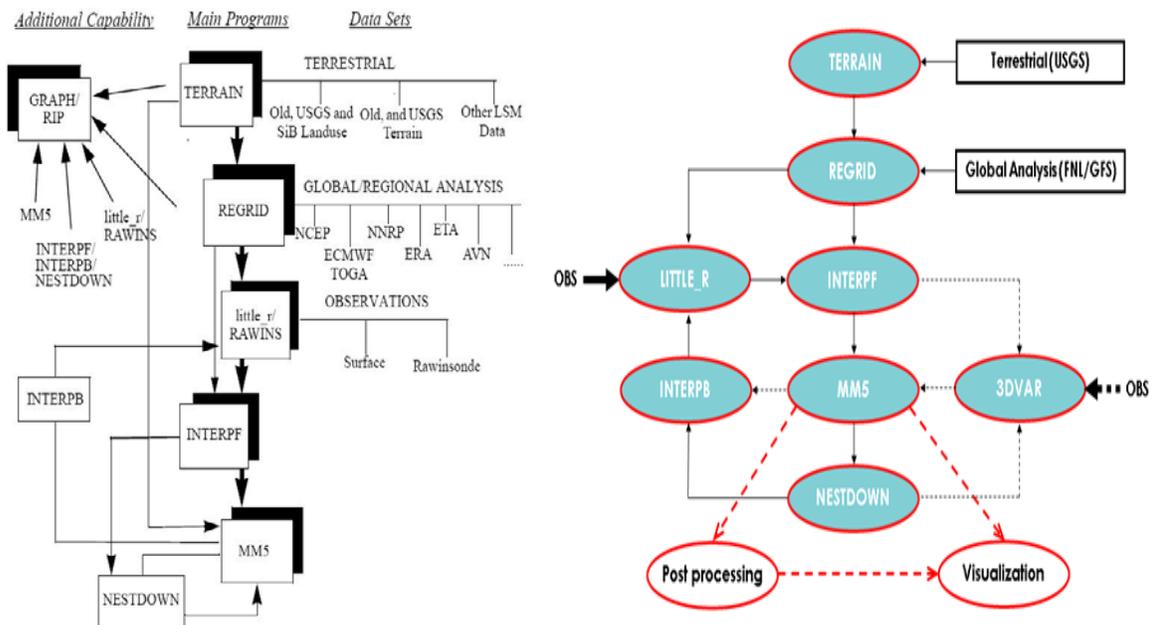


Figure 2-2 The MM5 flow chart [17; on left] and current implementation [on right]

The dashed lines, in the current implementation of the MM5 flow chart, in the Figure 2-2 represent alternative programs that were not used in the current study. The post processing and visualization in this work utilize many software packages like the NCAR Command Language (NCL), Vis5d and many Fortran codes developed at NARSS or downloaded from the open-source community. The function, inputs and outputs of each program of the MM5 modeling system are summarized in Table 2-2.

Table 2-2 Summary of the functions, input and output of MM5 programs [17]

Module	Function	Input from	Output to
TERRAIN	<ul style="list-style-type: none"> Define model domain and map projection. Generate terrain, and land use category data on model grids. Generate vegetation/soil category data for MM5 model's land-surface model option. Calculate map-scale factors and Coriolis parameter for the model. 	Global Terrestrial Datasets (USGS)	REGRID
REGRID	Generate first-guess pressure-level fields on model grids from another model dataset.	TERRAIN and Global Datasets (FNL/GFS)	INTERPF or LITTLE_R
INTERPF	Interpolate pressure-level data from either REGRID or LITTLE_R to model's sigma coordinate.	REGRID or LITTLE_R	MM5 or 3DVar
MM5	Performed the time integration.	TERRAIN and INTERPF or 3DVar	Binary output or NESTDOWN
NESTDOWN	<ul style="list-style-type: none"> Generate fine mesh model input from coarse mesh model output. Generate fine mesh model input from coarse mesh model input. 	MM5	INTERPB or 3DVar
INTERPB	Interpolate model's sigma coordinate to pressure-level data for LITTLE_R.	NESTDOWN or MM5	LITTLE_R
LITTLE_R	Perform objective analysis: blend first-guess fields with radiosonde and surface observations	INTERPB or REGRID, Observations	INTERPF
3DVar	Perform 3D variational data assimilation: performs a cost-function minimization in order to produce an "optimal" analysis.	INTERPF or NESTDOWN, Observations, Background error statistics	MM5

To automate the execution of the programs, several Linux scripts were developed. The scripts take into consideration different options in the programs to generate the required namelists. This made it easier to perform many simulations required in this study and to check the system sensitivity to many different options. Moreover, the developed scripts automatically apply the required changes based on the user inputs.

2.1.3 The MM5 Forecast Variables

During the integration process, the MM5 model writes its variables outputs in binary file format. Table 2-3 describes the MM5 forecast variables.

Table 2-3 The MM5 forecast variables [17]

Variable	Description	Units
U	Horizontal velocity component in West-East direction	m/s
V	Horizontal velocity component in South-North direction	m/s
T	Temperature	K
Q	Water vapor mixing ratio	kg/kg
CLW	Cloud water mixing ratio	kg/kg
RNW	Rain water mixing ratio	kg/kg
ICE	Ice cloud mixing ratio	kg/kg
SNOW	Snow mixing ratio	kg/kg
W	Vertical velocity	m/s
PP	Perturbation pressure	Pa
GROUND T	Ground temperature	K
PBL HGT	Planetary Boundary Layer height	m
SHFLUX	Surface sensible heat flux	W/m ²
LHFLUX	Surface latent heat flux	W/m ²
SWDOWN	Surface downward shortwave radiation	W/m ²
LWDOWN	Surface downward longwave radiation	W/m ²
T2	Temperature at 2 meters elevation	K
Q2	Water vapor mixing ratio at 2 meters elevation	kg/kg
U10	Horizontal velocity component in West-East direction at 10 meters elevation	m/s
V10	Horizontal velocity component in South-North direction at 10 meters elevation	m/s

2.2 GOVERNING EQUATIONS

The basic governing equations of the MM5 model are described by Anthes, updated by Dudhia and they are summarized later in this section. The fully compressible system of equations employs reference pressure as the basis for a terrain-following vertical coordinate system. The model uses pressure perturbation and temperature as prognostic variables. In combination with the existing initialization techniques and physics of the current non-hydrostatic model, this provides a model capable of real-data simulations which is limited only by data resolution, data quality and computer resources.

The old version of the MM5 model uses the following hydrostatic equation:

$\frac{\partial p}{\partial z} = -\rho g$, which assumes an exact equilibrium in the vertical between the pressure gradient force and the gravitational force. Hydrostatic approximation is based on the assumption that the horizontal scale is large compared to the vertical scale, such that the vertical pressure gradient may be determined as the product of density times the gravitational acceleration. For the non-hydrostatic version of the MM5 model, constant reference state and perturbations are defined as

$$p(x, y, z, t) = p_0(z) + p'(x, y, z)$$

$$T(x, y, z, t) = T_0(z) + T'(x, y, z)$$

$$\rho(x, y, z, t) = \rho_0(z) + \rho'(x, y, z)$$

The terms p , T and ρ are the pressure, temperature and density of the air, respectively. While $p_0(z)$, $T_0(z)$ and $\rho_0(z)$ are the reference properties which are calculated from the reference state standard hydrostatic atmosphere.

The model uses an Arakawa-Lamb B-staggering horizontal grid. The scalar variables (T , q etc.) are defined at the center of the grid cell, while the eastward and northward velocity components, u and v , are collocated at the nodes. Figure 2-3 shows a schematic representation of the horizontal grid structure. The center

points of the grid cells are the cross points and the grid nodes are the dot points. The smaller inner box is a representative mesh staggering for a 3:1 coarse-grid distance to fine-grid distance ratio.

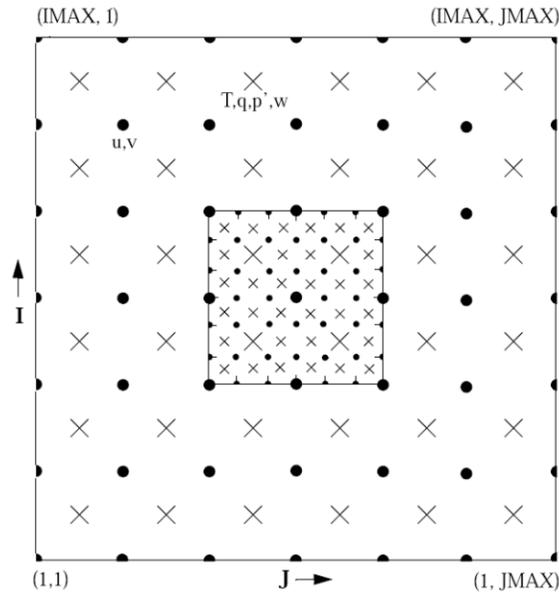


Figure 2-3 Schematic representation of the horizontal grid structure [17]

The vertical coordinate is terrain-following with the lower grid levels following the terrain while the upper surface is flat, as shown in Figure 2-4. Intermediate levels progressively flatten as the pressure decreases toward the chosen top pressure. A dimensionless quantity σ is used to define the model levels, where

$$\sigma = \frac{p - p_{\text{top}}}{p^*}, \text{ and } p^* = p_{\text{surf}} - p_{\text{top}} \quad (2.1)$$

Where p^* is the reference-state pressure, p_{top} is a specified constant top pressure, and p_{surf} is the reference-state surface pressure. In defining the σ levels it is the full levels that are listed, including levels at 0 and 1. The number of layers is therefore always one less than the number of full σ levels. The variables are defined in the middle of each vertical layer, referred to as half-levels and represented by the dashed lines in Figure 2-4. Vertical velocity is carried at the full levels (solid lines).

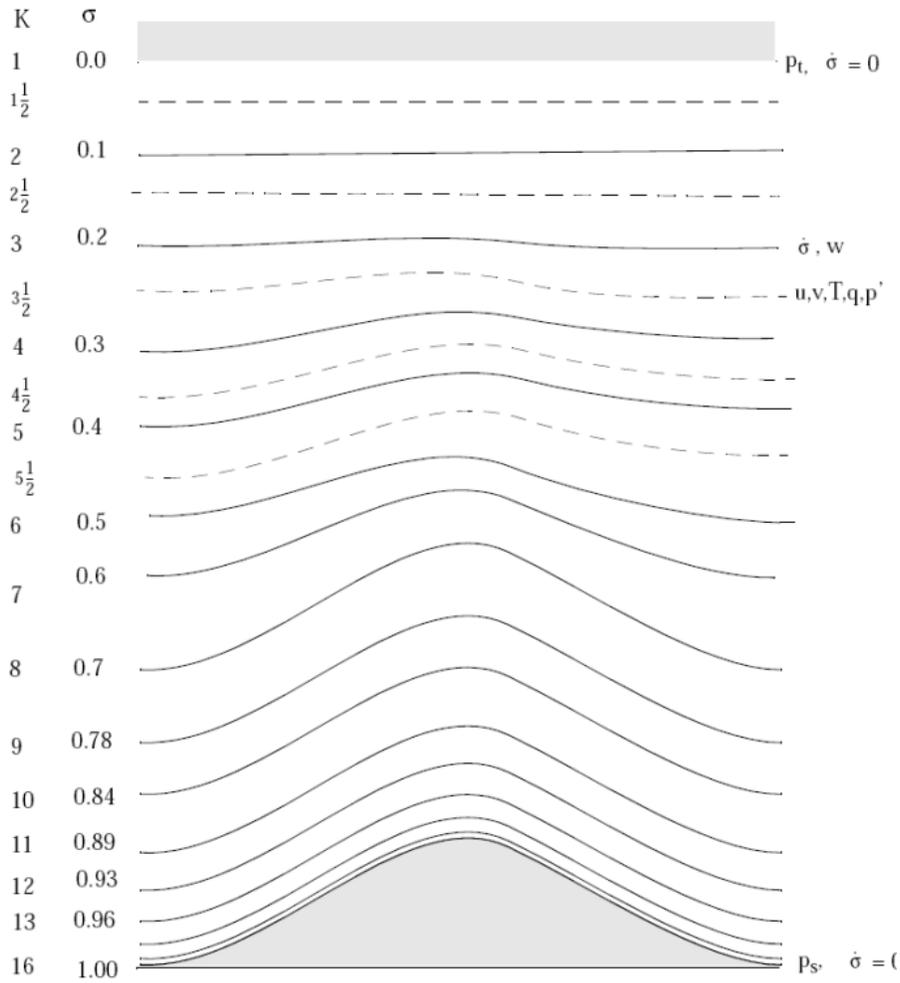


Figure 2-4 Schematic representation of the vertical grid structure [17]

2.2.1 Pressure Tendency Equation

The pressure field is calculated in MM5 by solving the pressure tendency equation expressed as

$$\frac{\partial p'}{\partial t} - \rho_0 g w + \gamma p \nabla \cdot v = -v \nabla p' + \frac{\gamma p}{T} \left(\frac{\dot{Q}}{c_p} + \frac{T_0}{\theta_0} D_\theta \right) \quad (2.2)$$

Where, the last term in equation (2.2) represents the increase in air pressure due to heating which forces the air to expand. This term is negligible in normal meteorological regimes and it will be neglected in the computations since it only forces a small divergence (expansion) in regions of heating.

2.2.2 Momentum Equations

The momentum equations for the compressible inviscid flow over complex terrains in the model coordinate system described by equation (2.1) and sketched in Figure 2-4 are;

Momentum (x-component)

$$\frac{\partial u}{\partial t} + \frac{m}{\rho} \left(\frac{\partial p'}{\partial x} - \frac{\sigma}{p^*} \frac{\partial p^*}{\partial x} \frac{\partial p'}{\partial \sigma} \right) = -V \cdot \nabla u + v \left(f + u \frac{\partial m}{\partial y} - v \frac{\partial m}{\partial x} \right) - ew \cos \alpha - \frac{uw}{Re} + D_u \quad (2.3)$$

Momentum (y-component)

$$\frac{\partial v}{\partial t} + \frac{m}{\rho} \left(\frac{\partial p'}{\partial y} - \frac{\sigma}{p^*} \frac{\partial p^*}{\partial y} \frac{\partial p'}{\partial \sigma} \right) = -V \cdot \nabla v - u \left(f + u \frac{\partial m}{\partial y} - v \frac{\partial m}{\partial x} \right) + ew \sin \alpha - \frac{vw}{Re} + D_v \quad (2.4)$$

Momentum (σ -component)

$$\frac{\partial w}{\partial t} - \frac{\rho_0 g}{\rho p^*} \frac{\partial p'}{\partial \sigma} + \frac{g p'}{\gamma p} = -V \cdot \nabla w + g \frac{p_0 T'}{p T_0} - \frac{g R_d p'}{c_p p} + e (u \cos \alpha - v \sin \alpha) + \frac{u^2 + v^2}{Re} + D_w \quad (2.5)$$

Where the terms $\frac{m}{\rho} \left(\frac{\partial p'}{\partial x} - \frac{\sigma}{p^*} \frac{\partial p^*}{\partial x} \frac{\partial p'}{\partial \sigma} \right)$ and $\frac{m}{\rho} \left(\frac{\partial p'}{\partial y} - \frac{\sigma}{p^*} \frac{\partial p^*}{\partial y} \frac{\partial p'}{\partial \sigma} \right)$ in equations (2.3) and (2.4) are the pressure gradient transformed into the sigma-pressure coordinates.

Advection terms are expressed as

$$V \cdot \nabla A = mu \frac{\partial A}{\partial x} + mv \frac{\partial A}{\partial y} + \dot{\sigma} \frac{\partial A}{\partial \sigma} \quad (2.6)$$

where

$$\dot{\sigma} = -\frac{\rho_0 g}{p^*} - \frac{m\sigma}{p^*} \frac{\partial p^*}{\partial x} u - \frac{m\sigma}{p^*} \frac{\partial p^*}{\partial y} v \quad (2.7)$$

Divergence term are expressed as

$$\nabla \cdot V = m^2 \frac{\partial}{\partial x} \left(\frac{u}{m} \right) - \frac{m\sigma}{p^*} \frac{\partial p^*}{\partial x} \frac{\partial u}{\partial \sigma} + m^2 \frac{\partial}{\partial x} \left(\frac{v}{m} \right) - \frac{m\sigma}{p^*} \frac{\partial p^*}{\partial y} \frac{\partial v}{\partial \sigma} - \frac{\rho_0 g}{p^*} \frac{\partial w}{\partial \sigma} \quad (2.8)$$

Sub grid eddies and other atmospheric physical processes such as turbulence fluxes, radiation, moisture transport, cloud and rain formations are represented by the divergence terms D_u , D_v and D_w in equations (2.3), (2.4) and (2.5) for x-component, y-component and z-component, respectively.

2.2.3 Temperature Tendency Equation

The temperature tendency equation, driven from the first law of thermodynamics, is expressed as

$$\frac{\partial T}{\partial t} = -v \cdot \nabla T + \frac{1}{\rho c_p} \left(\frac{\partial p'}{\partial t} + v \cdot \nabla p' - \rho_0 g w \right) + \frac{\dot{Q}}{c_p} + \frac{T_0}{\theta_0} D_\theta \quad (2.9)$$

The term $\frac{\dot{Q}}{c_p}$ represents the heating rate due to adiabatic processes represented in latent heat and radiation. The terms θ_0 and T_0 in equation (2.9) represent the reference state of potential temperature and temperature, respectively.

2.3 INITIAL AND BOUNDARY CONDITIONS

The initial and boundary conditions for the regional models are nested from global gridded datasets with horizontal resolution of 1° and time interval of 6 hours, as shown in Figure 1-1. In this study, the NCAR Final Reanalysis Data (FNL) [64] and the Global Forecast System (GFS) [65] are used. The modeling system usually gets and analyzes its input data on pressure surfaces. These have to be interpolated to the vertical coordinate before being input to the model.

The MM5 Model has the capability of multiple nesting with up to nine domains running at the same time and completely interacting. Figure 2-5 shows the three nested domains used in MM5 for Egypt. The domains have 81, 27 and 9 km horizontal resolutions with number of grid cells 70×70 , 96×96 , and 174×174 , respectively. The three domains were designed to have the same center. The inner domain of resolution 9 km encompasses entire Egypt.

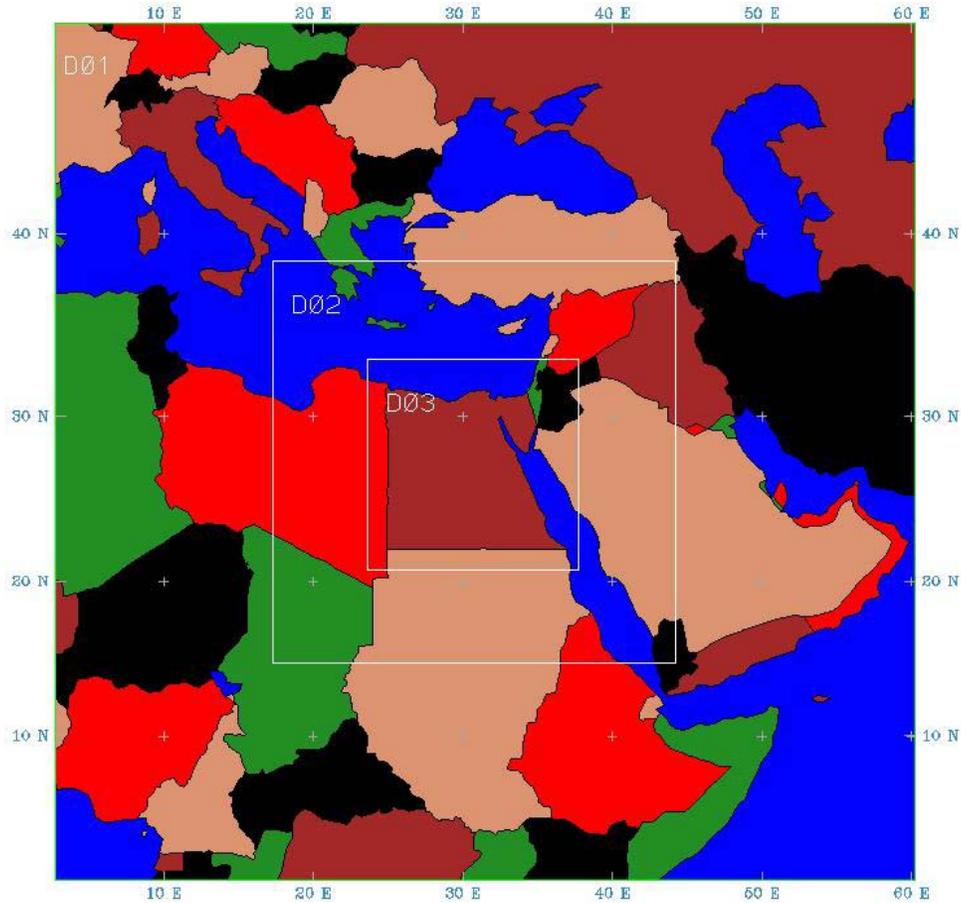


Figure 2-5 Nested domains used in MM5 for Egypt

The horizontal grid of the MM5 model is generated by the TERRAIN program which can perform three types of earth's projections, namely; Stereographic, Lambert and Mercator. The Mercator projection is a cylindrical map projection presented by the Flemish geographer and cartographer Gerardus Mercator, in 1569. It became the standard map projection for nautical purposes because of its ability to represent lines of constant course, known as rhumb lines or loxodromes, as straight segments. The Mercator projection is defined such that each rhumb line on a sphere is represented as a straight line. The rhumb line is a curve on the surface of a sphere that cuts all meridians at the same angle. The Mercator projection is suitable to equatorial latitudes and up to mid latitudes (45°) [66]. So, Mercator projection was used in this study for Egypt domains.

In Mercator projection, the grid point latitude is defined as

$$\varphi = 2 \cdot Conv \cdot \tan^{-1} \left[\exp \left(\frac{Y_c + (I - I_c) \cdot ds}{R_e} \right) \right] - 90 \quad (2.10)$$

And the grid point longitude is defined as

$$\lambda = \lambda_c + Conv \cdot \frac{J - J_c}{R_e} \cdot ds \quad (2.11)$$

Where I and J represent the model grid indices. The variable Y_c is the distance from the pole to the center of the coarse domain. λ_c is the center latitude and longitude of the domain respectively. I_c and J_c are the grid center indices. $Conv$ is the conversion factor from radians to degrees. Figure 2-6 shows a demonstration for the Mercator projection in the MM5 model while Figure 2-7 shows the generated horizontal grid for Egypt.

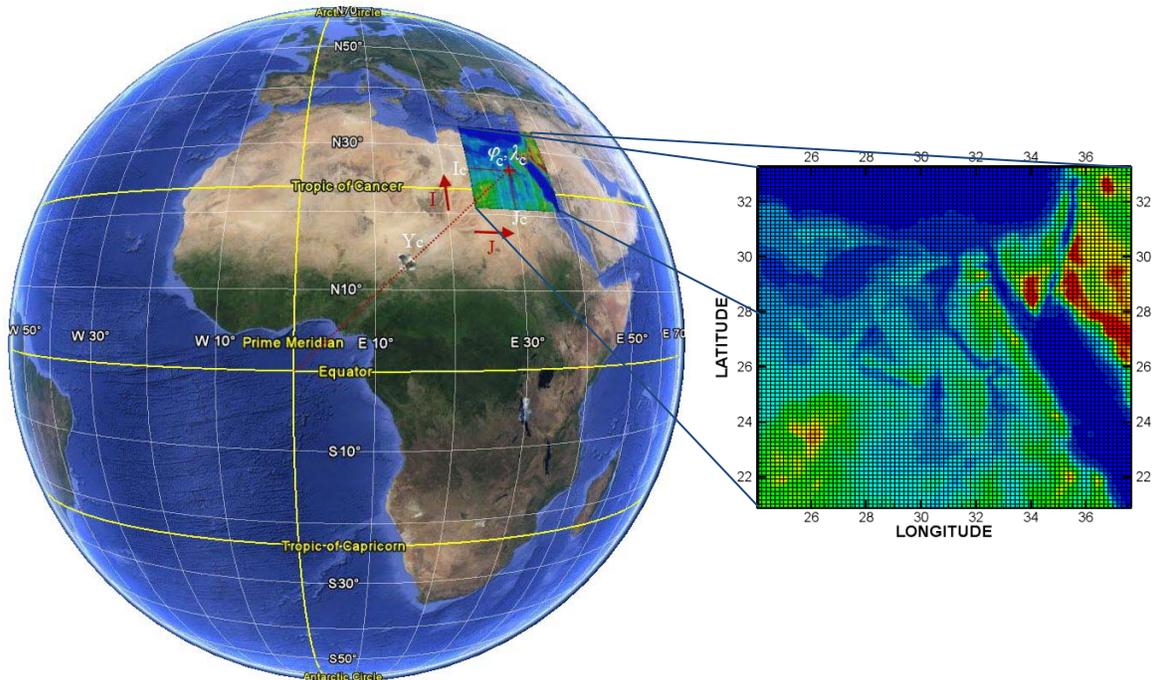


Figure 2-6 Demonstration for the Mercator projection in the MM5 model

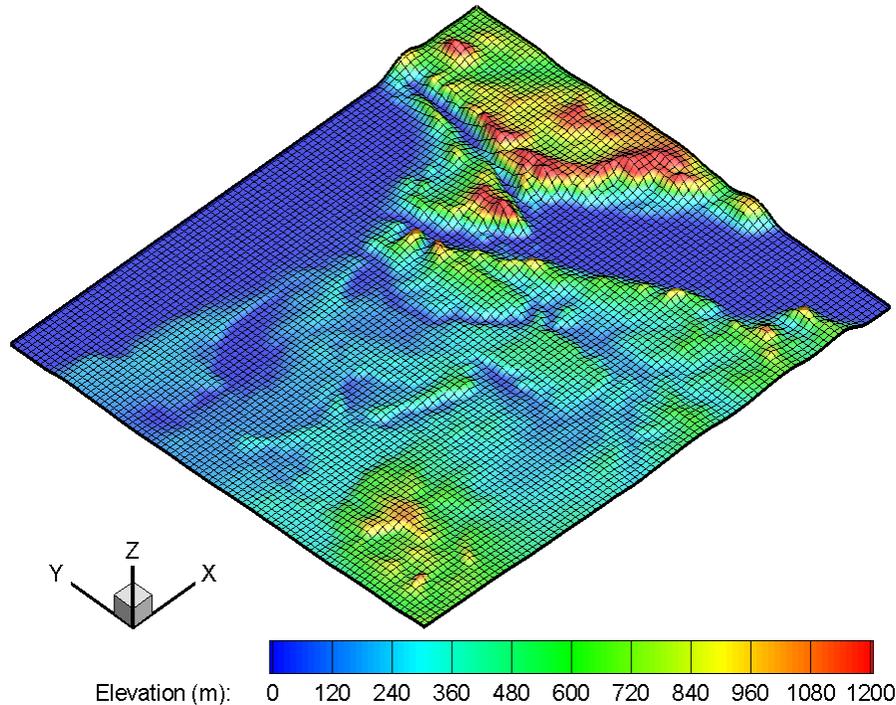


Figure 2-7 The MM5 horizontal grid for Egypt domain

The MM5 modeling system interpolates the pressure surfaces data to its vertical coordinates in two steps. The REGRID reads archived gridded meteorological analyses on pressure levels and interpolates them from some native grid to the horizontal grid as defined by the TERRAIN program. The INTERPF program handles the data transformation required to go from the analysis vertical grid to the MM5 vertical grid. The vertical grid of the MM5 model is generated using equation (2.1) and the base state.

The base state for the MM5 model is constructed from several values prescribing a surface level temperature and pressure, a temperature profile which may include an isothermal layer above the tropopause, and analytic expressions for a reference pressure and the height of the nonhydrostatic σ surfaces. Other than the terrain elevation, only these values are required by the modeling system as user input to completely define the base state. Figure 2-8 shows the generated vertical grid for Egypt domain.

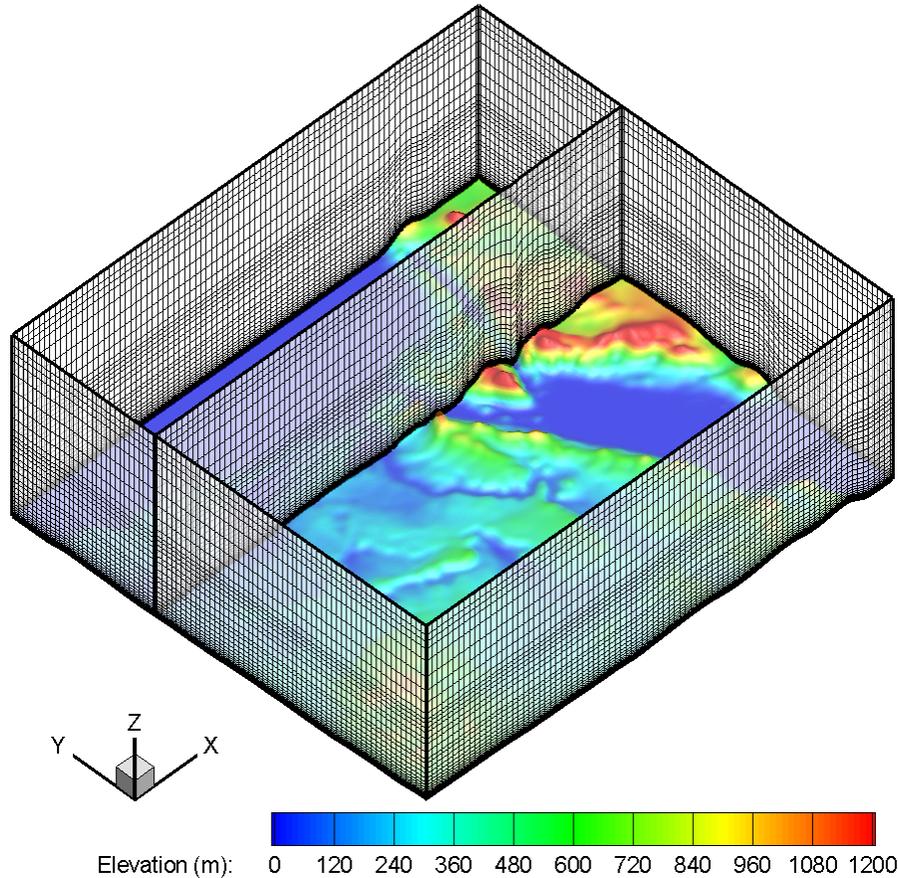


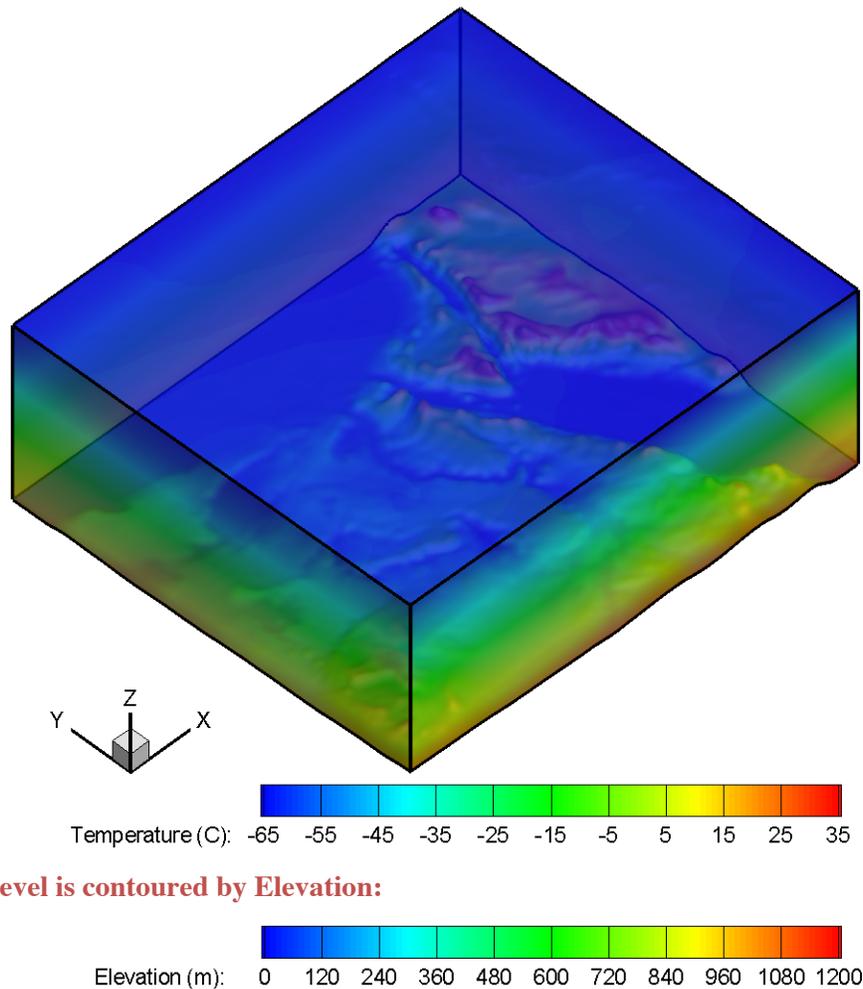
Figure 2-8 The vertical grid in the MM5 model for Egypt domain

2.3.1 Initial Conditions

The output data from the REGRID program including horizontal velocities, relative humidity and potential temperature is interpolated by the INTERPF program from pressure coordinates to sigma (σ) coordinates. The INTERPF program also calculates the base state properties including pressure, temperature and density, and the vertical velocity. The horizontal velocities, temperature and perturbation pressure are re-interpolated to satisfy the governing equations.

The vertical velocity equation of the model in finite difference form is used with the acceleration and advection terms set to zero. This ensures that the initial vertical acceleration is zero in each grid column. The initial conditions of finer-resolution domains can be interpolated directly from the output of coarser-

resolution domains using the NESTDOWN program. Figure 2-9 shows the initial conditions for Egypt domain.



The lower level is contoured by Elevation:

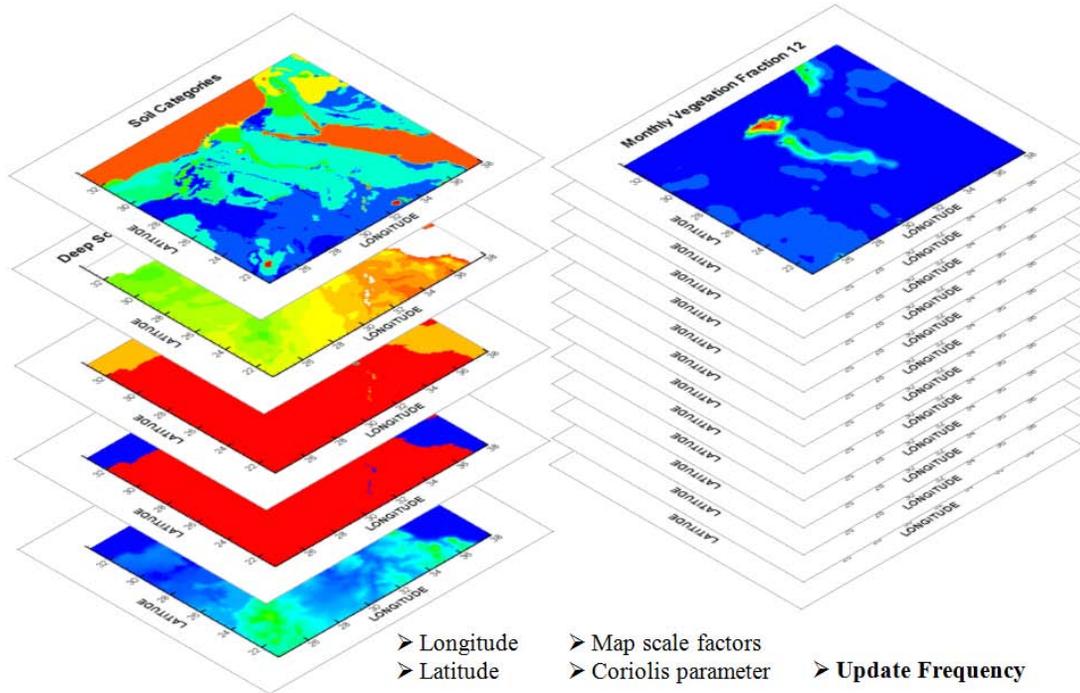
The initial conditions are nested from global gridded meteorological datasets.

Figure 2-9 The initial conditions of temperature for Egypt domain.

2.3.2 Surface Boundary Conditions

The surface boundary conditions in the MM5 model are produced by the TERRAIN program from USGS global datasets. These data has six different resolutions which are 1 degree, 30, 10, 5 and 2 minutes, and 30 seconds. For the Land Surface Model (LSM), 17-category soil, vegetation categories, and annual deep soil temperature data is needed, and the 12 monthly averages global

vegetation fraction data file is needed at 10 minutes resolution. Figure 2-10 summarizes the different levels of surface boundary conditions needed by the MM5 model for Egypt domain.



The TERRAIN program generates the soil categories, annual deep soil temperature, land use, land-water mask, terrain [to the left] and 12 monthly vegetation fraction [to the right].

Figure 2-10 Levels of surface boundary conditions in the MM5 model

The LSM option in the MM5 model was utilized using NOAH LSM. The required fields, such as soil temperature and soil moisture at various depths, were interpolated, using the REGRID program, from the input gridded datasets (FNL/GFS). Satellite observational data has been used, in a previous work [67], for updating the green vegetation fraction and sea/land surface temperature, identification of lakes in Toshka, and detecting the urban areas in the Nile Delta.

2.3.3 Lower Boundary Conditions

The lower boundary conditions in MM5 are computed from the LSM model for the temperature and horizontal velocity component, as shown in Figure 2-11. The LSM computes the energy budget between the atmosphere lower level and the

land surface. As a result, the temperature at 2m elevation and the horizontal velocity component at 10m elevation are produced. The lower temperature boundary condition and horizontal velocity components are then interpolated to the grid lowest vertical level. Other variables such as pressure and water vapor are directly interpolated from meteorological datasets for the first grid level.

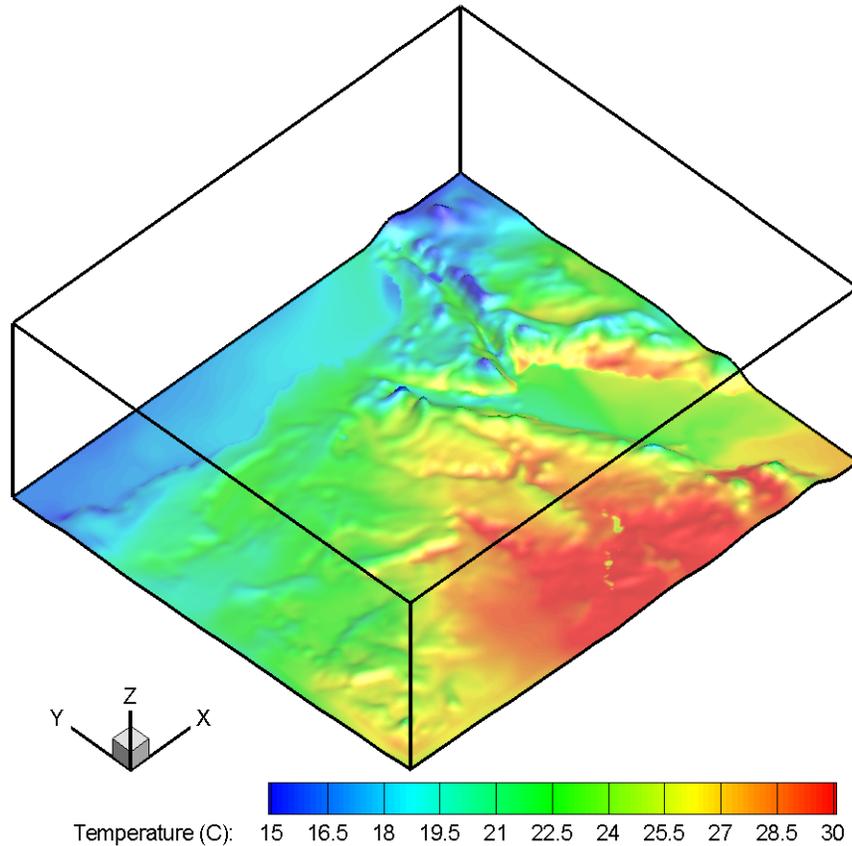


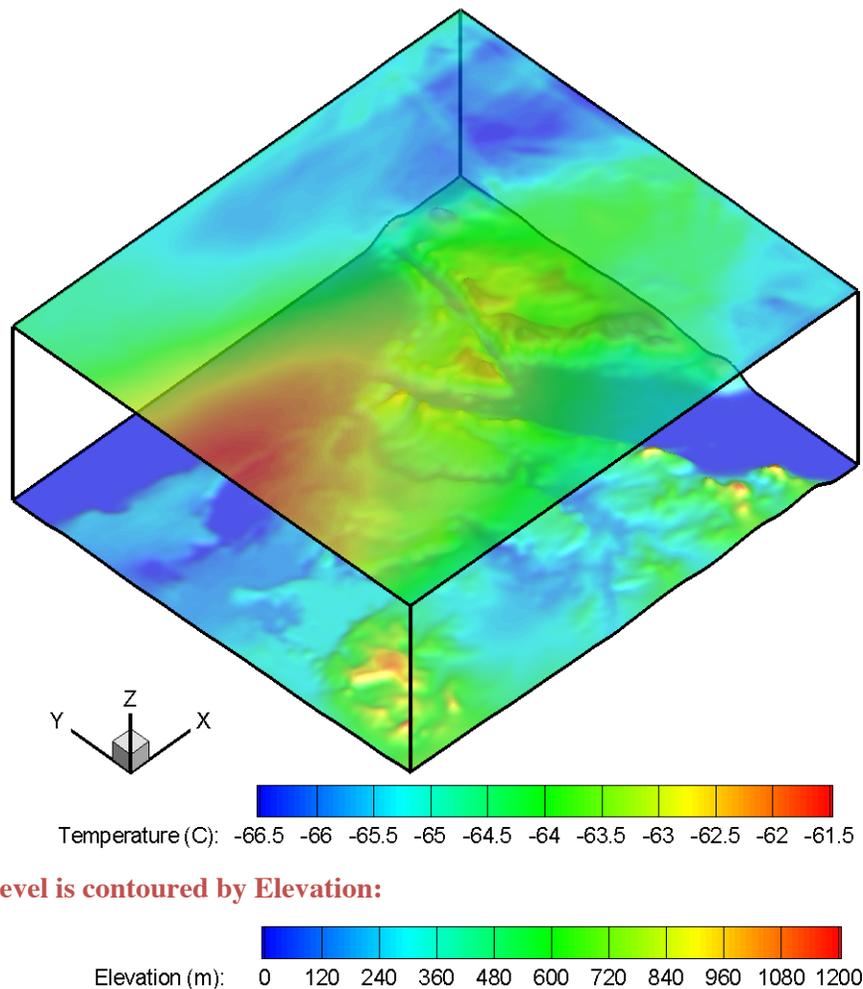
Figure 2-11 The lower temperature boundary conditions for Egypt domain

2.3.4 Upper Boundary Conditions

The upper boundary condition is a nonreflecting boundary that allows the wave energy to pass to outside the computational domain. It was described by Klemp [68] for hydrostatic wave as

$$\hat{P} = \frac{\rho N}{K} \hat{W} \quad (2.12)$$

Where \hat{P} and \hat{W} are the horizontal Fourier components of pressure and vertical velocity respectively. ρ and N are the density and buoyancy frequency near the grid top boundary. K is the total horizontal wave number of Fourier component. Equation (2.12) should be enforced for all components if the energy transport is to be purely upward with no reflection. Figure 2-12 shows the upper boundary conditions in the MM5 model at the model top vertical level.



The lower level is contoured by Elevation:

Figure 2-12 Temperature distribution at the upper boundary for Egypt domain

The upper boundary condition is combined with the implicit pressure / vertical momentum calculations. Before any value at time $(n + 1)$ is determined, the

values at the top grid level W_1 which are at staggered half a grid length above P_1 can be expressed as

$$P_1^{n+1} = b + a \cdot W_1^{n+1} \quad (2.13)$$

The coefficient $a(x, y, t)$ is dependent upon the thermodynamic structure and the bottom boundary conditions on W in the grid column. It varies within only 5 per cent of a constant value even with high terrain, and it is also not strongly time-dependent. The value of $b(x, y, t)$ depends on pressure and most of the pressure tendency terms.

2.3.5 Lateral Boundary Conditions

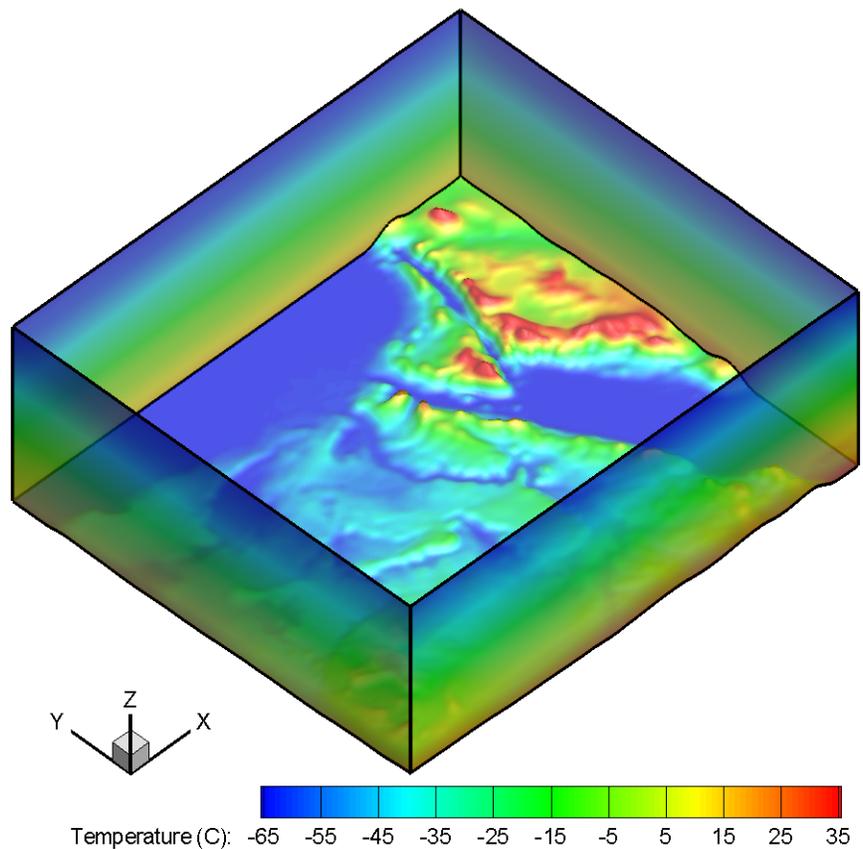
The lateral boundary conditions are used to initialize the atmospheric circulations from the regional scale and down to the local scale. The majority of atmospheric models, such as MM5 model, have one way boundary conditions, i.e. the model with coarse resolution provides information about the boundary values to the nested regional model.

In the MM5 model, the flow relaxation boundary condition type is used. In this method, the governing equations are modified by adding Newtonian relaxation term over the boundary conditions zone such that

$$\left(\frac{\partial \alpha}{\partial t}\right)_n = F(n)F_1(\alpha_{LS} - \alpha_{MC}) - F(n)F_2\Delta_2(\alpha_{LS} - \alpha_{MC}) \quad (2.14)$$

where $n = 1;2;3;4$ for cross-point variables, $n = 1;2;3;4;5$ for dot-point variables, α represents any variable, MC denotes the model calculated tendency, LS the large-scale tendency which is obtained either from observations or large-scale model simulations (oneway nesting), and n is the displacement in grid-points from the nearest boundary ($n = 1$ on the boundary). F decreases linearly from the lateral boundary, such that $F(n) = (5 - n)/3$, $n = 2,3,4$ and $F(n) = 0$, $n > 4$.

The flow relaxation scheme is used in this study for the most course mesh and all nested meshes. Figure 2-13 shows the boundaries where the flow relaxation scheme is applied. However, the vertical velocity is not relaxed. It can vary freely except for the outermost rows and columns, where zero gradient conditions are specified. For the velocity components, the values at the inflow points are specified in a manner similar to the specification of temperature and pressure. The values at the outflow boundaries are obtained by extrapolation from the interior points. These boundary values are required only in the computation of the nonlinear horizontal momentum flux divergence terms.



The lower level is contoured by Elevation:

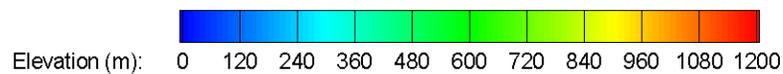


Figure 2-13 Lateral boundary conditions for temperature

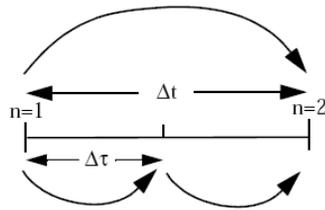
2.4 SOLUTION TECHNIQUE

The finite differencing in the MM5 model is, of course, crucially dependent upon the grid staggering wherever gradients or averaging are required to represent terms in the equations which are finite differenced on the B grid sketched in Figure 2-3.

2.4.1 Finite Differencing

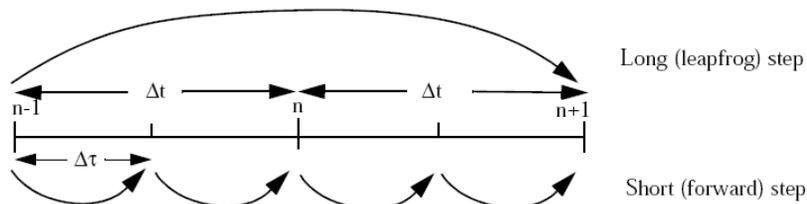
Second-order centered finite differences represent the gradient terms except for the precipitation fall term which uses a first-order upstream scheme for positive definiteness. For time integration, time-splitting scheme is used on fast terms and forward step is used for diffusion and microphysics. Figure 2-14 shows the temporal finite differencing in the MM5 model.

First time step:



Time step n:

T, qv, qc, etc., advection, physics, boundary, coriolis, diffusion terms



u, v, w, p' advanced (pressure gradients, divergence terms)

Time step n+1:

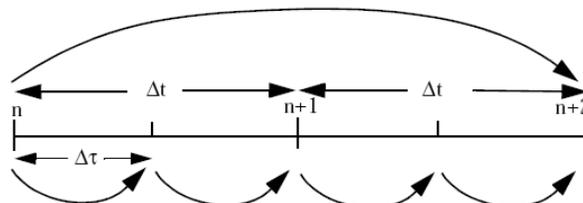


Figure 2-14 Temporal finite differencing in the MM5 model [18]

The nonhydrostatic equations are fully compressible and so they permit sound waves. These waves are fast and they require a small time step for numerical stability. It is possible to separate terms directly involved with acoustic waves from comparatively slowly varying terms, and to handle the former with smaller time steps while updating the slow terms less frequently. Using the reduced equation set for the small time step makes the model more efficient. Some radiation and cumulus options use a constant tendency over many time steps.

2.4.2 Physical Parameterizations

The MM5 model is available with different physics schemes to model different physical processes. They are applied for different grid resolutions and may vary according to the grid spacing. This involves cumulus parameterization, Planetary Boundary Layer (PBL), moisture, radiation and surface. Appendix B describes the MM5 physics options. While the model physics parameterizations are categorized in a modular way, it should be noted that there are many interactions between them via the model state variables such as potential temperature, moisture, wind, etc. and their tendencies, and via the surface fluxes. The direct interactions of parameterizations in the MM5 model are shown in Figure 2-15.

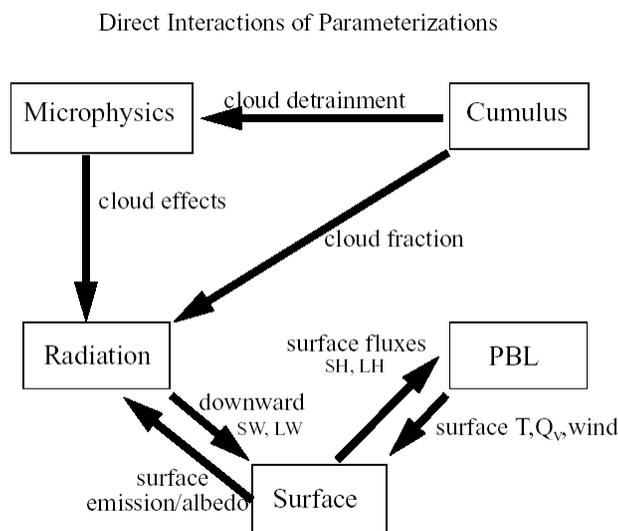


Figure 2-15 Interactions of parameterizations in the MM5 model [61]

It is not clear which scheme is the best scheme the performance of each scheme depends on the location and weather regime. There are different physics packages available with MM5 model. In this study, the cumulus parameterization utilizing the Anthes-Kuo scheme for the coarse domains of a resolution of 81 and 27 km, respectively, and the Grell scheme is used for Egypt domain of a resolution of 9 km. The Medium Range Forecast (MRF) PBL was used to model the PBL and diffusion with different formulations of the roughness length for heat/moisture. The simple-cloud (Dudhia) and mixed-phase (Reisner 1) schemes were used for explicit moisture. Cloud, CCM2 and RRTM longwave schemes were used for radiation. The unified NOAA/OSU Land Surface Model (LSM) was used for surface parameterization.

2.4.3 Automation Scripts

Three sets of scripts were written in this study for the automation of the execution of the different MM5 programs following the modeling system framework described in Figure 2-1. All scripts were written to be easily modified by the user. Moreover, the input arguments of them automatically controls the function of the MM5 programs by generating an updated namelist file for each program. These scripts are summarized in Appendix G. The run directory has a naming convention as shown in Figure 2-16.

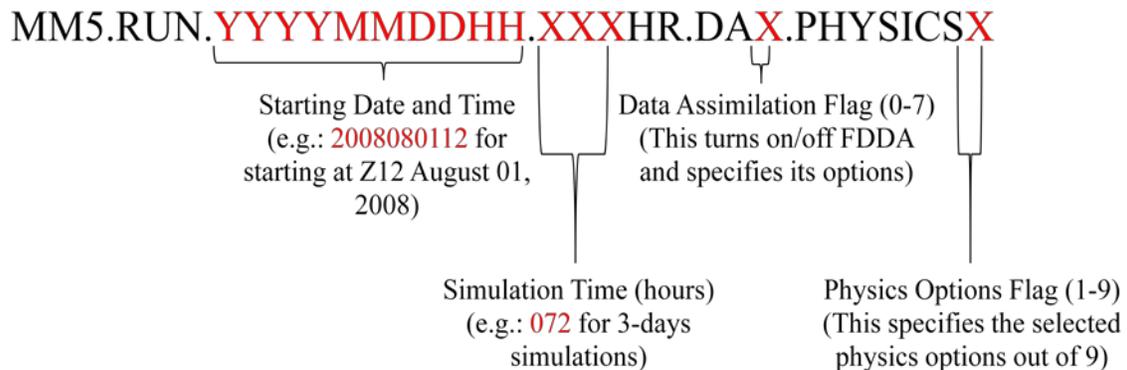


Figure 2-16 The naming convention for the MM5 run directory

The first set of scripts was combined in the initialization script, denoted by “*MM5.INIT*”. This script is responsible for the initialization of the MM5 main inputs and scripts (decks) by running the TERRAIN, REGRID and INTERPF programs. It is also responsible for the preparation of the observational data in the format required for FDDA run utilizing the NCAR Command Language (NCL). Figure 2-17 shows a schematic flowchart for this script.

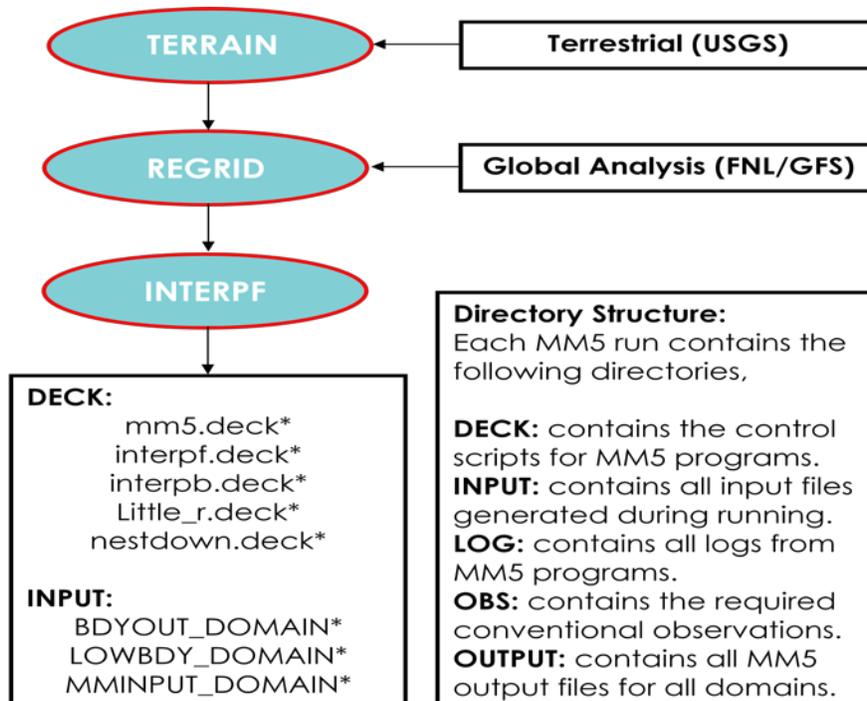


Figure 2-17 Schematic flowchart for the initialization script

The second set of scripts was combined in the running script, denoted by “*MM5.RUN*”. This script is responsible for the execution the MM5 modeling system shown in Figure 2-2 for all initialized run directories. It is also responsible for the running of the MM5 solver for all domains using the decks generated by the initialization script. The third set of scripts was combined in the visualization script, denoted by “*readmm5nc.m*”, written in MATLAB (MATrix LABoratory) numerical computing environment. This script is responsible for the visualization of the MM5 outputs in MATLAB and the generation of files suitable for the Tecplot numerical simulation and visualization software.

2.5 MODEL EVALUATION

Evaluation of a NWP model is usually based on the comparison of its outputs with the observations. This may be based on statistical analysis applied to either the basic model output fields (e.g., temperature, water vapor or wind components) or to the derived atmospheric fields (e.g., thermal advection and moisture convergence) [69]. Comparing observational data with the model results helps identifying model strengths and deficiencies. This also can be used to assess the quality of the different types of observations [70].

The atmospheric physics are parameterized in the model and some physical phenomena such as up-slope/down-slope winds, sea/land breezes, and mountainous terrain areas are not resolved. Therefore, the model evaluation may not be easy and it is dependent on the application of interest such as temperature, rainfall, hurricanes, and wind energy. In this study, the model evaluation was based mainly on the temperature field.

Quantitative evaluation of the model may be based on the root-mean-square error (RMSE) and maximum-absolute-bias error (MABE). The RMSE and MABE for a variable (α) were calculated as follows,

$$\text{RMSE} = \sqrt{\frac{1}{N} \sum_{i=1}^N (\alpha_i^s - \alpha_i^o)^2} \quad (2.15)$$

$$\text{MABE} = \max_{i=1,N} |\alpha_i^s - \alpha_i^o| \quad (2.16)$$

Where α_i^s is the i^{th} simulated variable, α_i^o is the i^{th} observed variable, and N is the number of observations. In this study, the model evaluation was monitored by comparing the results of the performed simulations with the observations at 24 surface stations and 5 upper-air stations inside Egypt. Any other evaluation criteria, such as comparing the model results with the average, minimum, and maximum observed values, were used to study the model sensitivity.

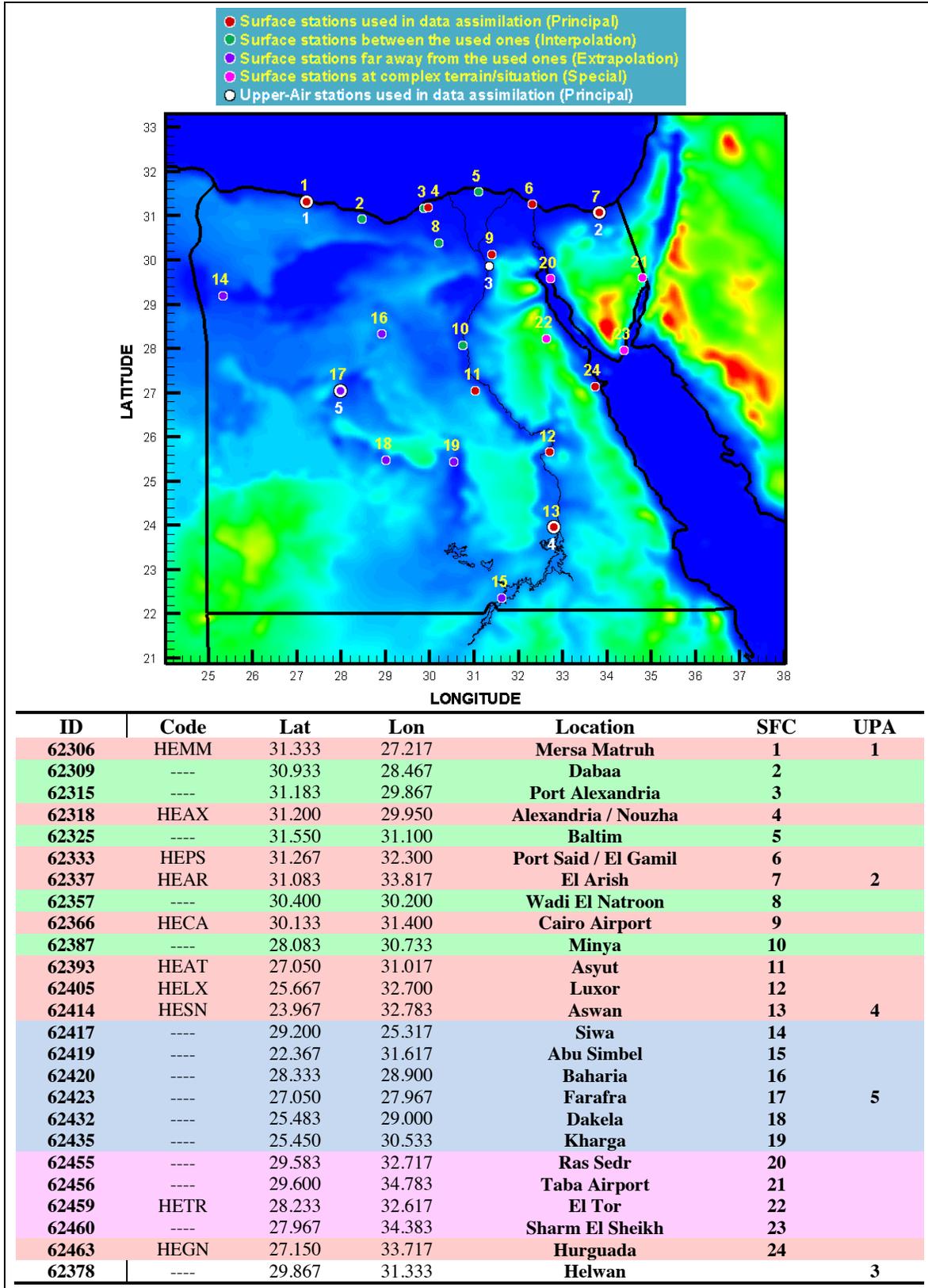


Figure 2-18 The locations of WMO stations used for the model evaluation

Figure 2-18 shows the locations of surface and upper-air World Meteorological Organization (WMO) stations inside Egypt used for the model evaluation. The stations were divided into four sets of surface observations and one set for the upper-air observations. The first set of surface observations, “principal set”, is the set of stations used in data assimilation. The second set of surface observations, “interpolation set”, is the set of stations located in-between the principal stations. The third set of surface observations, “extrapolation set”, is the set of stations located far from the principal stations. The fourth set of surface observations, “special set”, is the set of stations located at complex terrain. The stations near coast lines were included in different sets according to their locations with respect to the principal stations. Figure 2-18 shows the distribution and locations of the different observation sets.

CHAPTER THREE

OBSERVATIONAL DATA

This chapter investigates the sources, processing, analysis and quality of the observational conventional and satellite data available at NARSS Weather and Climate Station (WCS) and/or the Research Data Archive (RDA).

3.1 INTRODUCTION

Observational data is the data collected from field (conventional data) or remote sensing (satellite data). WMO GOS is the coordinated system of methods and facilities for making meteorological and other environmental observations on a global scale in support of all WMO Programs. The system is comprised of operationally reliable surface-based and space-based subsystems. The GOS comprises observing facilities on land, at sea, in the air and in outer space. The main components of the WMO GOS are surface, upper-air, marine, aircraft and satellite observations, as shown in Figure 3-1 and described in Appendix C [71].

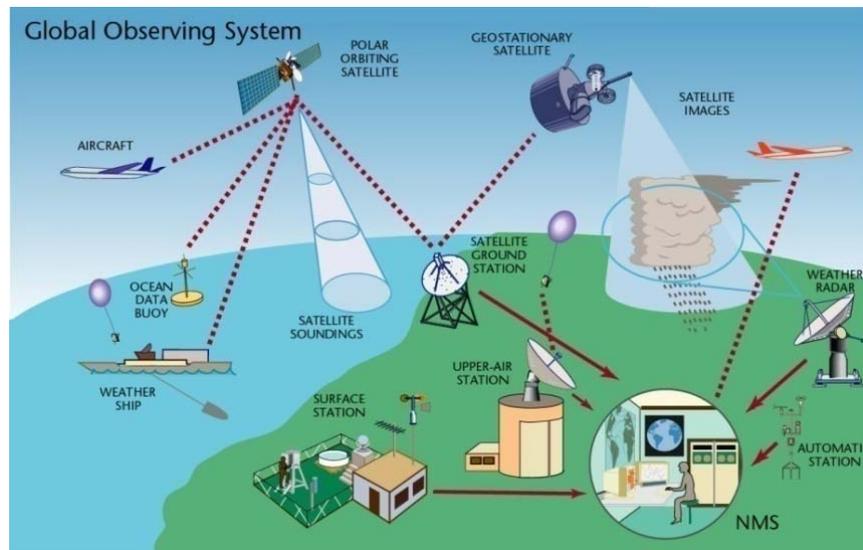


Figure 3-1 WMO Global Observing System (GOS) [71]

3.2 AVAILABLE DATA

The main sources of data used in this study are the NCAR's Computational and Information Systems Laboratory (CISL) Research Data Archive (RDA) and the NARSS High Resolution Picture Transmission (HRPT) station. The RDA datasets [72] are the main source of conventional observations. The NARSS HRPT station receives data from the weather satellites such as the American National Oceanic and Atmospheric Administration (NOAA) weather satellite series. The current system records multi-spectral images (several spectral channels) of the Earth surface, and allows the computation of temperature and humidity profiles of the atmosphere. NOAA data can be alternatively downloaded from The Comprehensive Large Array-data Stewardship System (CLASS) [73].

3.2.1 Research Data Archive

CISL integrates world-class high performance computing with research and development and applied mathematics to extend the reach of the atmospheric and related sciences community. CISL RDA contains a large and diverse collection of meteorological and oceanographic observations, operational and reanalysis model outputs, and remote sensing datasets to support atmospheric and geosciences research, along with ancillary datasets, such as topography/bathymetry, vegetation, and land use.

In this study, the RDA datasets DS336.0, DS353.4 and DS464.0 were used for research runs (analysis), while datasets DS337.0, DS351.0 and DS461.0 were used for operation runs (forecast). Table 3-1 shows the name, availability dates, format, software, scale, resolution, scope, sources and documents of the different RDA datasets used in this study. There are other RDA datasets which were used in this study but there are not observational data such as NCEP Final Reanalysis (FNL) that was used for model initialization.

Table 3-1 Research Data Archive (RDA) datasets

Dataset	Research			Operation		
	DS336.0	DS353.4	DS464.0	DS337.0	DS351.0	DS461.0
Name	Historical Unidata IDD Global observational Data	NCEP ADP Operational Global Upper Air observations	NCEP ADP Operational Global Surface observations	NCEP ADP Global Surface and Upper Air Weather observations	NCEP ADP Global Upper Air Observational Weather Data	NCEP ADP Global Surface Observational Weather Data
Start Date	May, 2003	Dec, 1972	Feb, 1975	Apr, 2008	Apr, 2000	Apr, 2000
End Date	Current	Feb, 2007	Feb, 2007	Current	Current	Current
Format	NetCDF	ON29	ON29	PREPBUFR	BUFR	BUFR
Software	NetCDF Decoders	FETCH Utility		PREPBUFR Decoders	BUFR Decoders	
Scale	Global					
Resolution	Synoptic/Stations					
Scope	UPA/SFC	UPA	SFC	UPA/SFC	UPA	SFC
Variables	Pressure, temperature, dew point temperature, wind speed and direction					
Sources	Archive: http://dss.ucar.edu/datasets/dsxxx.x					
	Real-time: ftp://ftpprd.ncep.noaa.gov/pub/data/nccf/com					
Documents	http://dss.ucar.edu/datasets/dsxxx.x/docs/					

3.2.2 NOAA Satellite Data

The American polar orbiting weather satellites are controlled by NOAA. The first launch occurred in 1978 with the TIROS N satellite, followed by NOAAxx series (from NOAA6 to NOAA14), and since 1998 by NOAA-KLM and NOAA-NN' series. Each satellite has a lifetime of about 4 years. Two satellites must be operational at the same time. The system is composed of two sun synchronous polar orbiting spacecraft, flying at an altitude close to 835 km. One is planned to fly in the afternoon orbit (PM), and the other in the morning orbit (AM). Due to the Earth rotation, the satellite track is translated westward with each pass. The nodal period is about 102 minutes, which means that each satellite describes 14.2 orbits per day, so the sub satellite track is not daily periodic (the same orbital configuration is achieved every 9 days). The satellites orbital parameters drift slowly, provoking observation conditions and pass time changes [74].

The main suite of instruments onboard the NOAA-KLM and NOAA-NN' satellites includes the Advanced Very High Resolution (AVHRR), the Advanced

Television Infrared Observation Satellites (TIROS) Operational Vertical Sounder (ATOVS) with the High resolution Infrared Radiation Sounder (HIRS) and the Advanced Microwave Sounding Unit (AMSU-A and AMSU-B). The Microwave Humidity Sounder (MHS) replaces AMSU-B in NOAA-NN'. The instrumentation onboard NOAA satellites are shown in Table 3-2.

Table 3-2 Instrumentation onboard NOAA satellites [74]

NOAA (6-14)		NOAA (15, 16 and 17)		NOAA (18 and 19)	
TOVS	HIRS/2 MSU SSU	ATOVS	HIRS/3 AMSU-A AMSU-B	ATOVS	HIRS/4 AMSU-A MHS
AVHRR/2		AVHRR/3		AVHRR/3	

Narrow band radiometers, wedged on specific wavelengths, are designed to measure spectral radiances. If the wavelength energy measured by the instrument is not absorbed by the atmosphere (i.e. atmospheric window), but only by clouds, the instrument is called an imaging radiometer such as AVHRR, whose purpose is to give information, reflectance or temperature, on different types of surfaces including sea, vegetation, ice, snow, or on cloud cover (top). If the wavelength energy of the instrument is absorbed by the atmosphere, the instrument is called a sounder (e.g. TOVS/ATOVS with HIRS, AMSU-A and AMSU-B/MHS), which is dedicated to measurements of the atmospheric vertical structure.

In this study, the NOAA/AVHRR data were used to provide the Normalized Difference Vegetation Index (NDVI), albedo, Sea Surface Temperature (SST) and Land Surface Temperature (LST). The NOAA/ATOVS data were used to provide vertical temperature and humidity profiles. Table 3-3 shows the name, availability dates, format, software, scale, resolution, scope, sources and documents of the NOAA satellite data used in this study.

Table 3-3 NOAA/AVHRR and NOAA/ATOVS satellite data

Data	AVHRR	ATOVS		
Name	Advanced Very high Resolution Radiometer	Advance TIROS (Television Infrared Observation Satellites) Operational Vertical Sounders		
Date	The first launch of NOAA-KLM and NOAA-NN' series since 1998			
Format	L0	Level 0: Raw Data (straight from the satellite)		
	L1	Level 1a: separated data for each instrument		
		Level 1b: geo-referenced and calibrated data (reversible: calibration coefficients are separated from raw data)		
		Level 1c: geo-referenced and calibrated temperatures and albedo (non-reversible: calibration coefficients are applied to numerical data)		
		Level 1d: mapped and filtered data (cloud mask)		
L2	Level 2: Final Products (converts the data into physical parameters)			
Software	L0	Local HRPT Receiving Station		
	L1	ATOVS and AVHRR Processing Package (AAPP)		
	L2	AVHRR Decoders	International ATOVS Processing Package (IAPP)	
Instrument	AVHRR	HIRS	AMSU-A	AMSU-B/MHS
Channels	5	20	15	5
Scale	Regional	Regional	Regional	Regional
Resolution	1.1 km	17 km	50 km	16 km
Scope	Clouds, clear-sky Surface	Clouds, clear-sky Temperature, Moisture, and Surface	All-sky Temperature, TPW, Cloud Water, and Surface	All-sky Moisture, and Surface
Products	NDVI, Albedo, SST	Vertical Temperature and Humidity Profiles		
Sources	Local Receiving System: NARSS HRPT Receiving Station			
	NOAA Satellite and Information Service: http://www.class.ngdc.noaa.gov			
Documents	ftp://ftp2.ncdc.noaa.gov/pub/doc/klmguide/pdf/			

Different sources of data have different formats, horizontal/vertical resolution, temporal resolution and variables. Therefore, it is essential to build up an automatic interface layer to convert these data to the format of the model. The next section will describe the processing of these data to build such layer.

3.3 DATA PROCESSING

The main purpose of the data processing is to retrieve different meteorological parameters from indirect observations and to prepare the observational data for the numerical model in an automated way. Data processing includes retrieval, calibration, quality control and validation of observational data for use in

preparing model inputs. Local calibration of satellite data was not taken into consideration in this study because of the lack of accurate local information inside the domain of interest. Moreover, the available WMO conventional observations do not cover the entire domain with a good density, see Figure 2-18.

3.3.1 Processing of RDA Data

The observations in the RDA data are encoded from the raw GTS format into many different standard formats for different datasets, see Table 3-1. The upper-air reports (ROAB/PIBAL/AIRCAR/AIRCFT) present a global synoptic set of 3 or 6 hourly data subsets. The surface reports (METAR/SYNOP/SHIP/BUOY) represent a global synoptic set of 1 or 3 hourly data subsets. The processing of each dataset was based on its format and the decoders of this format. Many modifications of the open-source decoders were carried out to make sure that they are working correctly. Then, a set of Linux scripts were written to automate the processing of the datasets and prepare them for the modeling system.

The LITTLE_R program specifies its own format for input which is better suited for users to adapt their own data. The observational data are written into data files at certain times. Each file contains the observations at this time within a time window. The time window is chosen to be 0.5 hour for hourly files. Each file is converted to LITTLE_R format which is one entry per observation (surface/upper-air): Header; Data; End. Each time period is stored in a separate file. Then, the observations are ready to be used in another component of the modeling system framework. Figure 3-2 shows the processing of RDA conventional data at the NARSS weather and climate station.

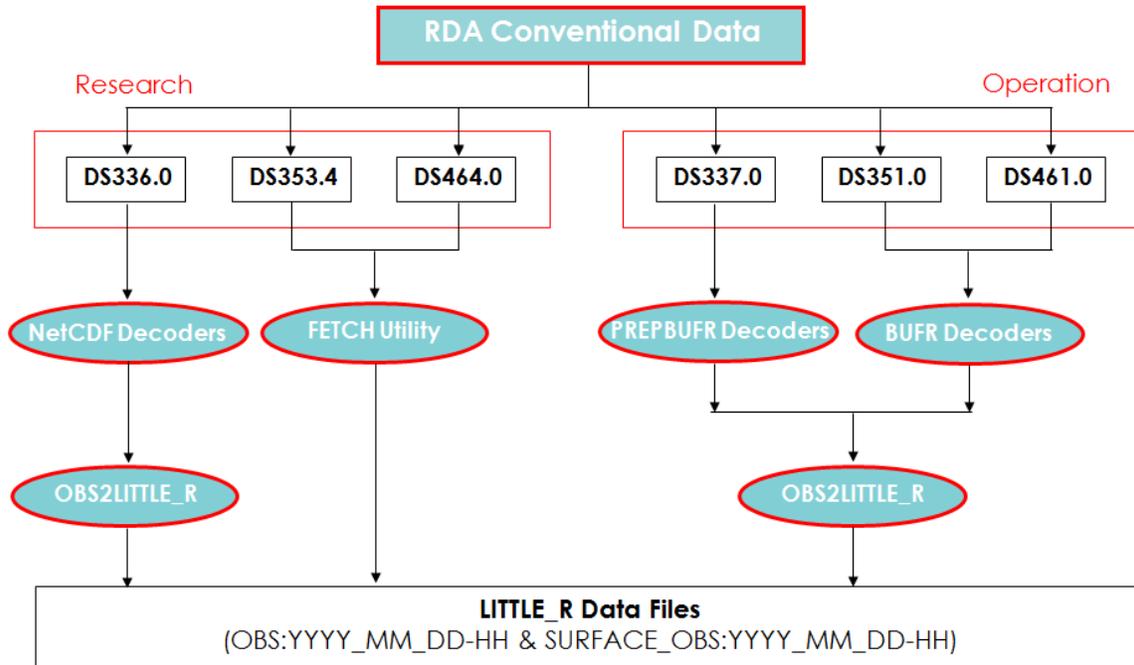


Figure 3-2 Processing of the RDA conventional data at NARSS

The DS353.4 and DS464.0 datasets are commonly used with the MM5/WRF models. These datasets were not updated since February, 2007. The rest of datasets are available in real time. In this study, most software, which was used to process different datasets, was downloaded from the RDA site. Then, many modifications were applied to adapt them for automatic processing as shown in Figure 3-2. A control Linux script was written to automatically process the data according to the user choices such as region of interest, time period and frequency. The output files have a standard naming convention.

3.3.2 Processing of NOAA Data

The NOAA data undergo three levels of data processing; Level-0, 1 and 2. In Level-0 data processing, data products are unprocessed telemetry data as received from the observing platform excluding communication artifacts introduced by ground system. In Level-1 data processing, calibration, geometric correction and cloud removal are included. In Level-2 data processing, data products are derived in a form of geophysical variables at the same resolution and locations as that of

level 1 source data. The satellite data formats (Level-0, 1 and 2) are described in Table 3-3. The processing of NOAA HRPT data at NARSS uses the software of the receiving system for level-0 data processing. The ATOVS and AVHRR Processing Package (AAPP) is used for level-1 data processing. The International ATOVS Processing Package (IAPP) and different AVHRR decoders are used for level-2 data processing. Figure 3-3 shows the processing of NOAA HRPT data at NARSS to retrieve temperature and water vapor profiles from NOAA/ATOVS data and Normalized Difference Vegetation Index (NDVI), Sea Surface Temperature (SST), Land Surface Temperature (LST) and Albedo from NOAA/AVHRR data.

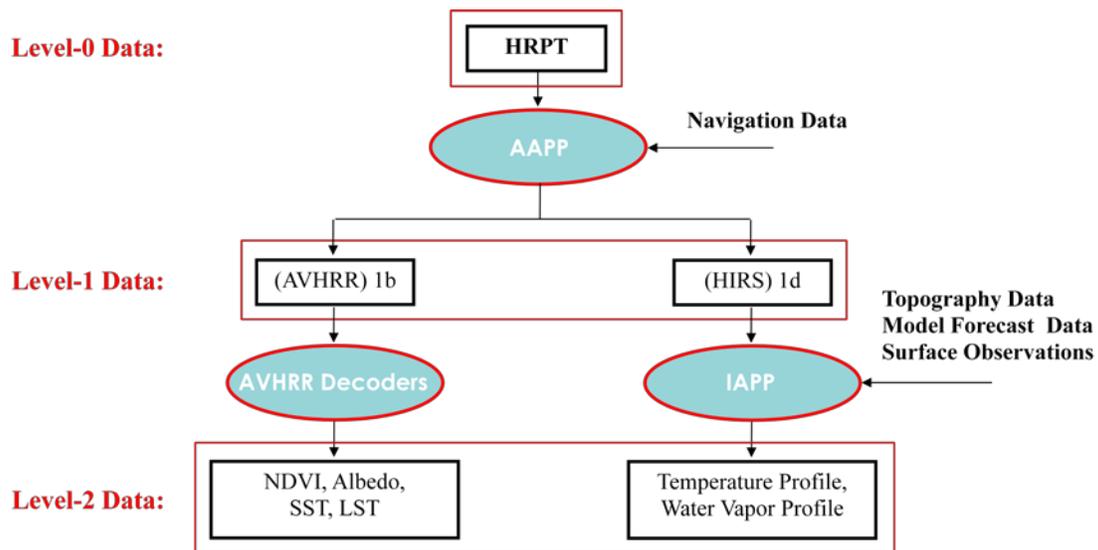


Figure 3-3 Processing of NOAA data at NARSS

AAPP is a software package for the processing of ATOVS and AVHRR data from the NOAA series of weather satellites. The package is intended for the ingestion and the pre-processing tasks of received HRPT data up to the retrieval of the satellite brightness temperatures. AAPP contains software for the processing of raw HRPT data (**level-0**) successively to

- **Level-1a:** separated data for each instrument,
- **Level-1b:** geo-referenced and calibrated data (reversible) where calibration coefficients are separated from raw data,

- **Level-1c:** geo-referenced and calibrated temperatures and albedo (non-reversible) where calibration coefficients are applied to numerical data, and
- **Level-1d:** mapped and filtered data (cloud mask).

Levels 1a, 1b, 1c and 1d are internationally recognized data formats [75]. They contain modules for de-commutation, navigation, calibration, pre-processing, mapping to a common instrument grid, and AVHRR cloud-masking. They were developed by a number of European meteorological organizations including Météo France, UK Met Office, and ECMWF, and they are now maintained as a deliverable of the EUMETSAT NWP SAF.

On the other hand, IAPP has been developed by The Cooperative Institute for Meteorological Satellite Studies (CIMSS) to retrieve atmospheric temperature and moisture profiles, total ozone and other parameters in both clear and cloudy atmospheres from the TOVS/ATOVS radiance measurements [76]. It is written to use the ATOVS and AVHRR Processing Package (AAPP) for signal de-commutation, navigation, calibration and data collocation to level-1d. The HIRS, AMSU-A, and AMSU-B AAPP level-1d formatted files are all used in the IAPP. All other input and output files are in NetCDF format [77].

The main AVHRR decoder used in this study is the Envi software package [78]. ENVI is one of the products of the ITT Company for Visual Information Solutions. It is the premier software solution for processing and analyzing geospatial imagery used by GIS professionals, scientists, researchers, and image analysts around the world. ENVI software combines the latest spectral image processing and image analysis technology with an intuitive, user-friendly interface to help the user get meaningful information from imagery.

3.3.2.1 Transmission of NOAA Data

The radiometric and sounding data are stored onboard NOAA satellites until the satellite transmits them to a ground station. Two kinds of AVHRR data are transmitted: full resolution data LAC and sampled data (mean of 4/5 pixels 1/3 line) GAC. Global TIROS Information Processor (TIP) information is also transmitted. The main acquisition stations are located near the North Pole, which provide facilities for the acquisition of a great number of recorded orbits. The stations transfer these data onto another satellite that can transmit the information to the main NOAA/NESDIS centre, where the data are processed and archived before distribution.

The satellite also transmits multiplexed AVHRR, TIP and AMSU data, i.e. HRPT real time data, toward a line of sight acquisition stations located anywhere on the Earth surface. The frames are continuously transmitted, line by line, during approximately 15 minutes for an orbit segment. The generated data file contains about 5000 AVHRR lines (2048 pixels per line for 5 channels) representing 100 Mbytes (10 bit words) of data. The data transmission is illustrated on Figure 3-4.

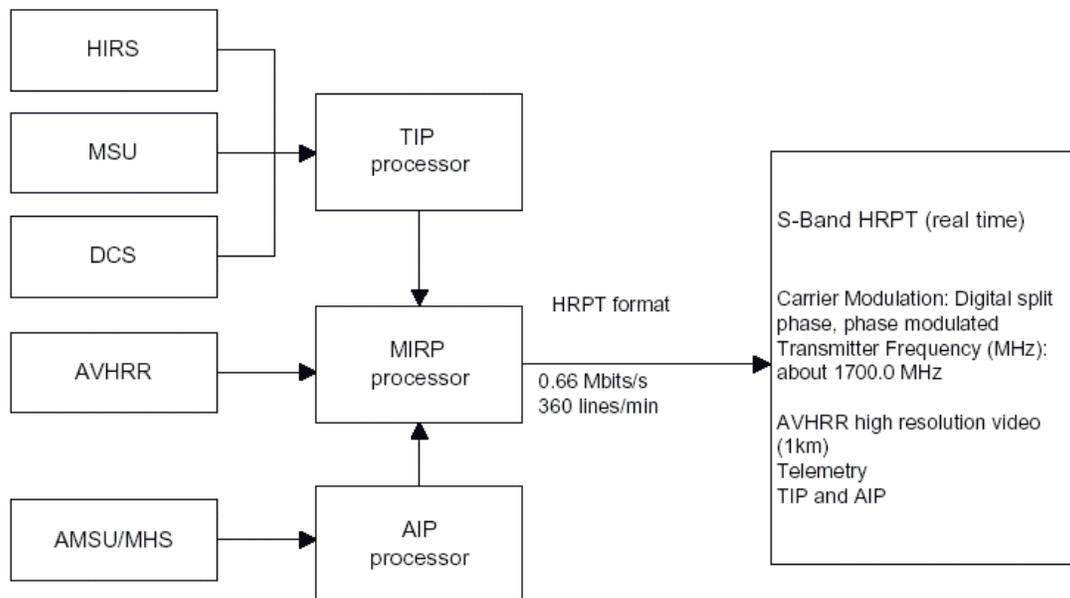


Figure 3-4 Data transmission in HRPT format [74]

The data of the MHS instrument are re-formatted by the onboard MHS Interface Unit (MIU). The MIU then passes the data to the AIP, such that the MHS data occupies the same HRPT slots that are used for AMSU-B – albeit with a time delay of approximately 1 scan line. Thus, from the user’s point of view the processing of MHS appears similar to the processing of AMSU-B.

3.3.2.2 Geometric Correction

Satellite images have different distortions due to the earth's curvature and rotation, satellite parameters (e.g., speed, attitude and altitude), scan skew and projection of spherical surface on a flat image. To use this remotely sensed image, geometric correction must be first applied. Image data needs to be rectified to a standard map projection and datum to create a faithful representation of the digital image in terms of position. Rectification is the process of transforming the data from one grid system (image pixels grid) into another grid system using a geometric transformation. Geometric transformation was performed through polynomial transformation or triangle-based methods. Figure 3-5 shows the application of geometric correction on an AVHRR true-color image.

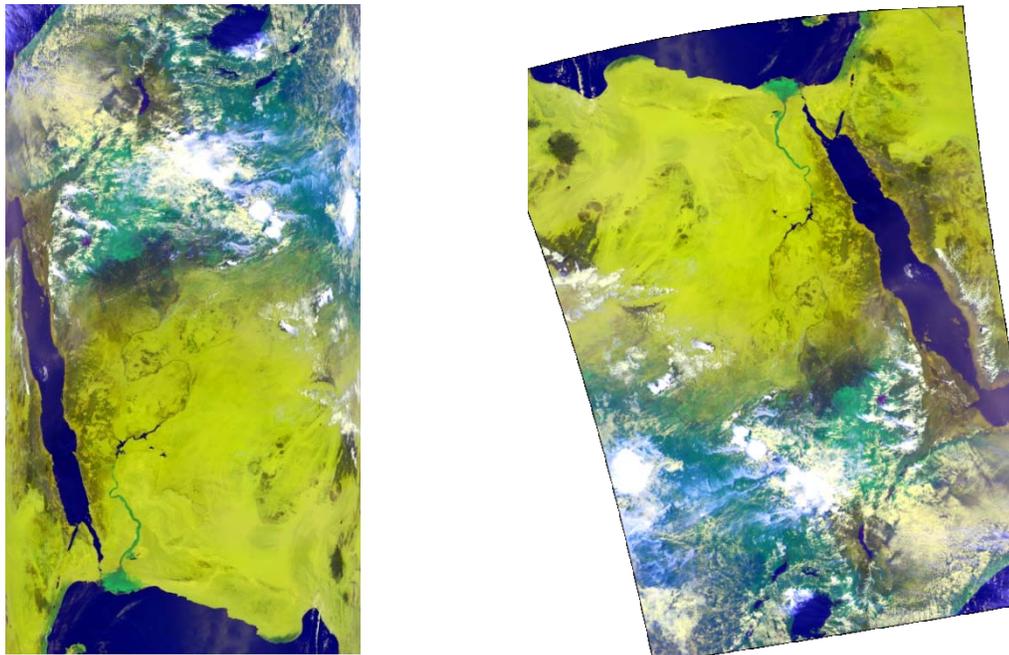


Figure 3-5 Geometrically uncorrected [on left] and corrected [on right] images

3.3.2.3 Cloud Masking

The earth-surface images observed by satellites may be masked by clouds. The cloudy pixels represent the reflected energy from the clouds and not those from the earth. To obtain accurate estimates of surface parameters from satellite data, a scheme has to be devised which identifies cloud-free and cloud-filled pixels. The cloud identifying scheme consists of many tests applied to each individual pixel to determine whether the pixel is cloud-free, partially-cloudy or cloud-filled.

Surface parameters such as vegetation indexes, albedo and surface skin temperature can be then inferred from the cloud-free pixels and cloud parameters such as cloud top temperature, optical depth and liquid water contents are obtained from the cloud-filled pixels. Figure 3-6 shows the AVHRR true-color image and the generated cloud mask. The cloud-filled pixels are represented by red color. The partially-cloudy pixels are represented by the green color.

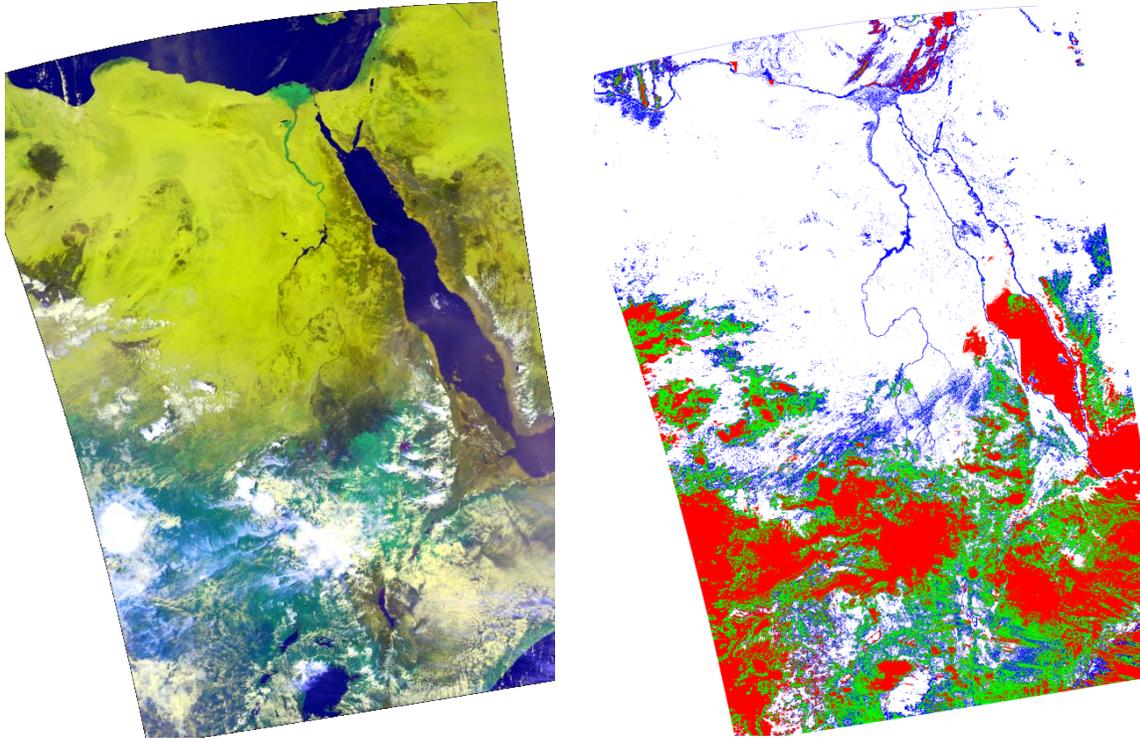


Figure 3-6 AVHRR true-color image [on left] and the cloud mask [on right]

3.4 DATA ANALYSIS

The horizontal and vertical structures of the meteorological parameters form the basic input to define the initial conditions for NWP models. The initialization task requires the use of satellite-derived products that are global data by nature and that are denser, more uniform, and more homogeneous than other data sources. However, the measurements of remote sensing based satellite radiometer do not yield the meteorological parameters directly. These parameters are obtained through a process of mathematical inversion which is called observational/forward model. The accuracy of the retrieved products is dependent upon the accuracy of the satellite measurements, the prescribed atmospheric transmittance functions, and the inversion algorithm. A precise knowledge of the instrument performance and of the accuracy of the atmospheric transmittance functions for the various spectral channels used are crucial to the retrieval accuracy.

3.4.1 ATOVS Data Retrievals

The variables which are currently calculated by the IAPP retrieval code and which are written to the output NetCDF file are: the temperature profile, water vapor profile, total atmospheric ozone, surface skin temperature, the total precipitable water, the microwave emissivity, total cloud liquid water, surface ice index, and the surface snow index. The pressure levels of the retrievals, scan line day of year, scan line UTC time of day, land-ocean index and the latitude and longitude of the center pixel are also written to the output file.

The ATOVS data retrievals have a spatial resolution of about 47-147 km and 42 pressure levels for vertical profiles of temperature and water vapor. The vertical profiles have 42 points from near the ground up to about 45 km. Figure 3-7 shows the three-dimensional retrievals from NOAA/ATOVS data at NARSS while Figure 3-8 shows the two-dimensional retrievals. The retrievals were processed from images received by the NARSS HRPT station.

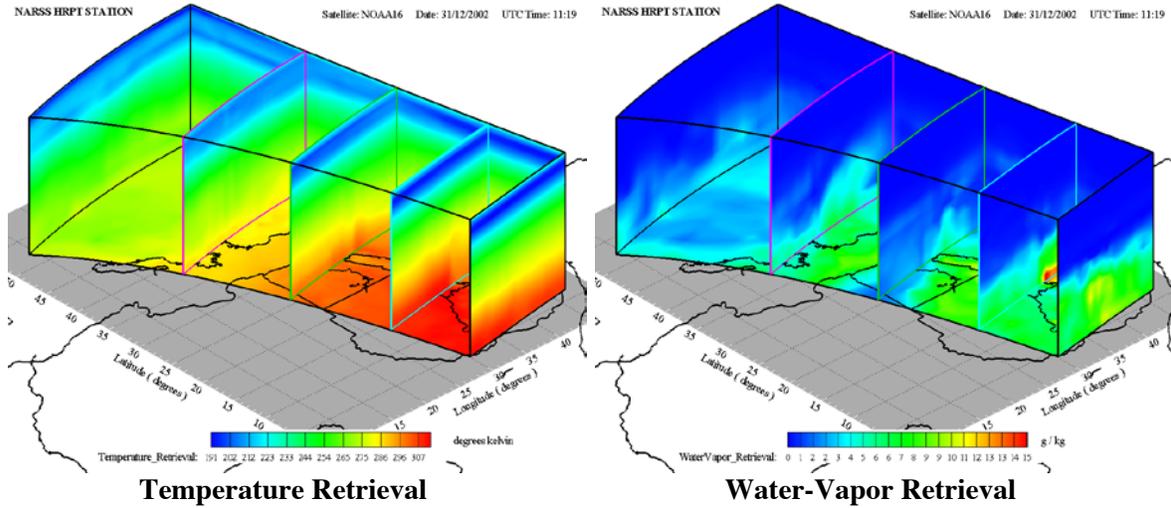


Figure 3-7 Three-dimensional retrievals from NOAA/ATOVS data at NARSS

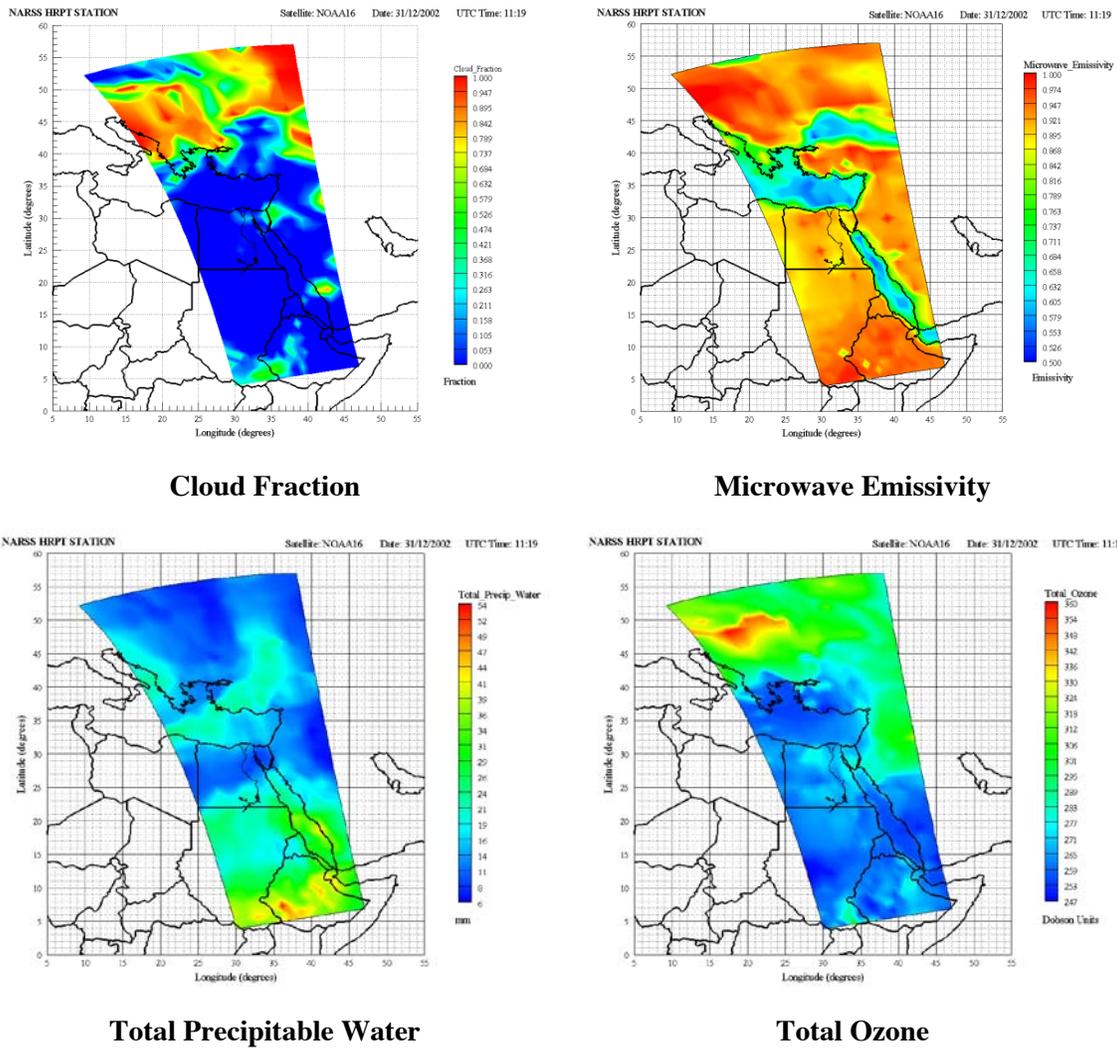


Figure 3-8 Two-dimensional retrievals from NOAA/ATOVS data at NARSS

3.4.2 AVHRR Data Retrievals

The AVHRR image files consist of five bands within the visible, near-, middle-, and thermal-infrared spectral regions. Image spatial resolution is 1.1 km along swath centerlines, but it degrades to several kilometers at the outer edges of the swaths. Combination of the five bands can be used to generate the different parameters, such as the Normalized Difference Vegetation index (NDVI), albedo, Sea Surface Temperature (SST) and Land Surface Temperature (LST). The data is spatially resized for the region as shown in Figure 3-9.

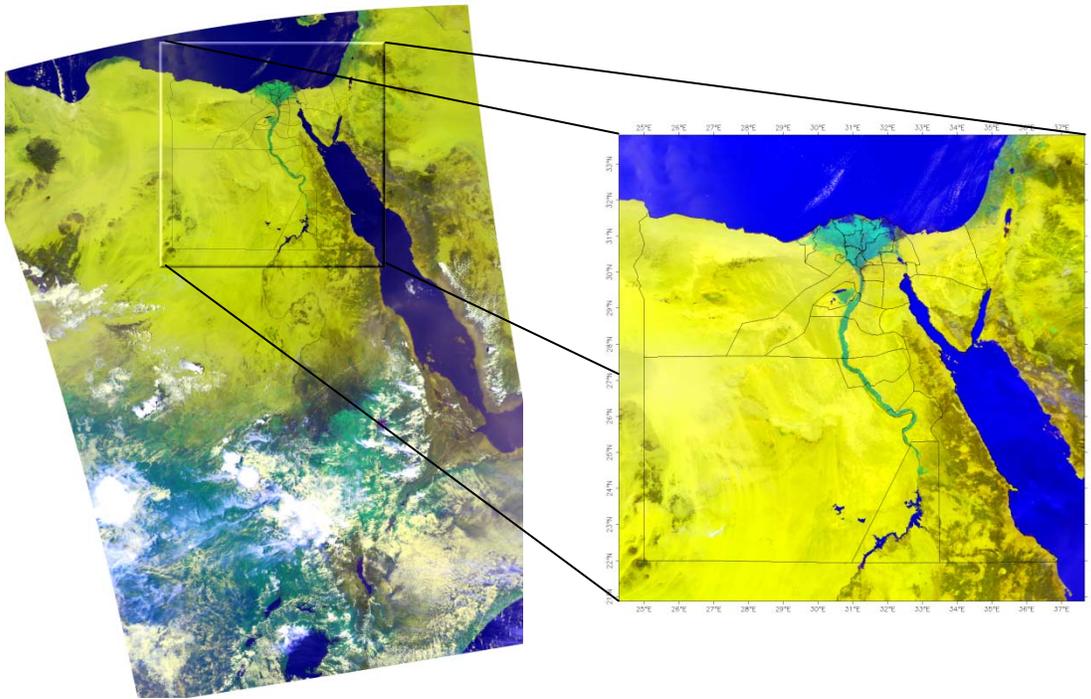


Figure 3-9 AVHRR true-color image and a spatial subset for the region of interest

The AVHRR five channels for the region of interest are then calibrated by finding the transfer function between the digital counts from the instrument and the scene input spectral radiance for all the instrument's spectral channels. Channel 1 and 2 are calibrated to reflectance (percentage) while channels 3, 4 and 5 are calibrated to brightness temperatures (K). Table 3-4 shows the calibration of visible and near-infrared channels which are used to calculate the NDVI and albedo while

Table 3-5 shows the calibration of middle- and thermal-infrared channels which are used to calculate SST and LST.

The produced retrievals are then filtered by the cloud mask. The images are analyzed and displayed using the Envi software. Many overlays may be added to the digital images such as grid lines, color ramps, political/governorate limits and/or coast lines. Moreover, the images are written in the standard Envi format and they can be exported in many other formats for a consequent application.

Table 3-4 Calibration of AVHRR visible and near-infrared channels

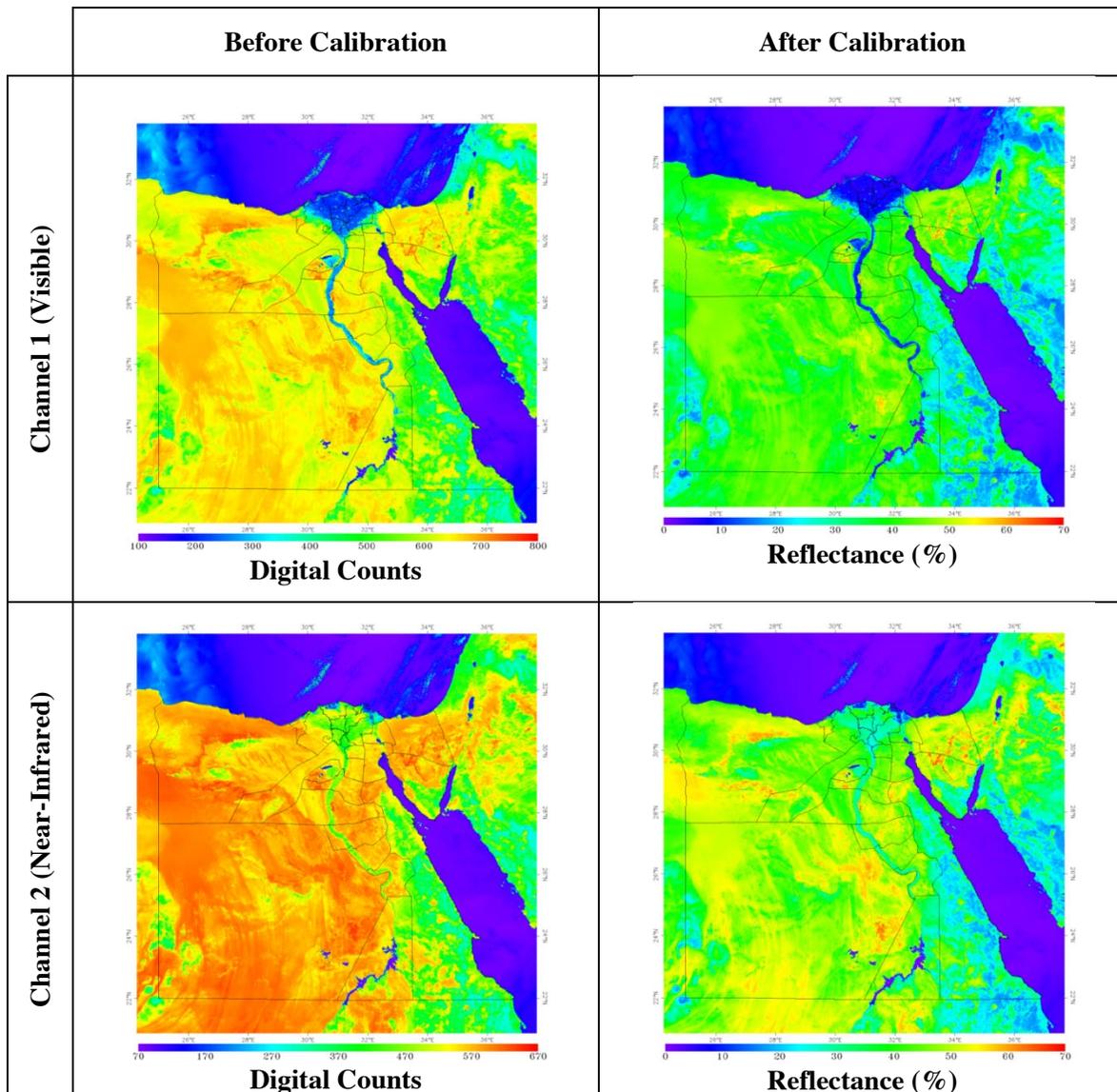
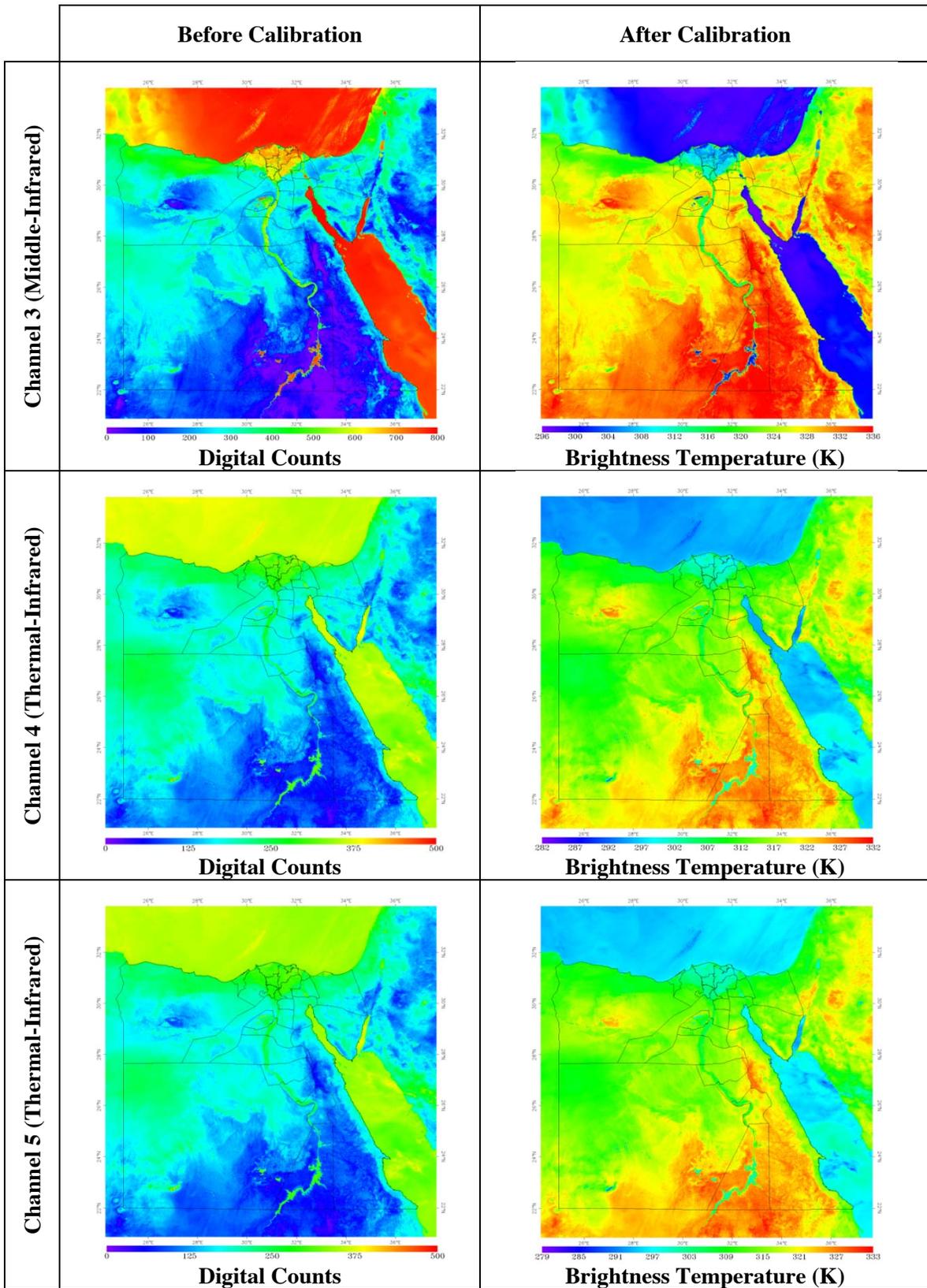


Table 3-5 Calibration of AVHRR middle- and thermal- channels



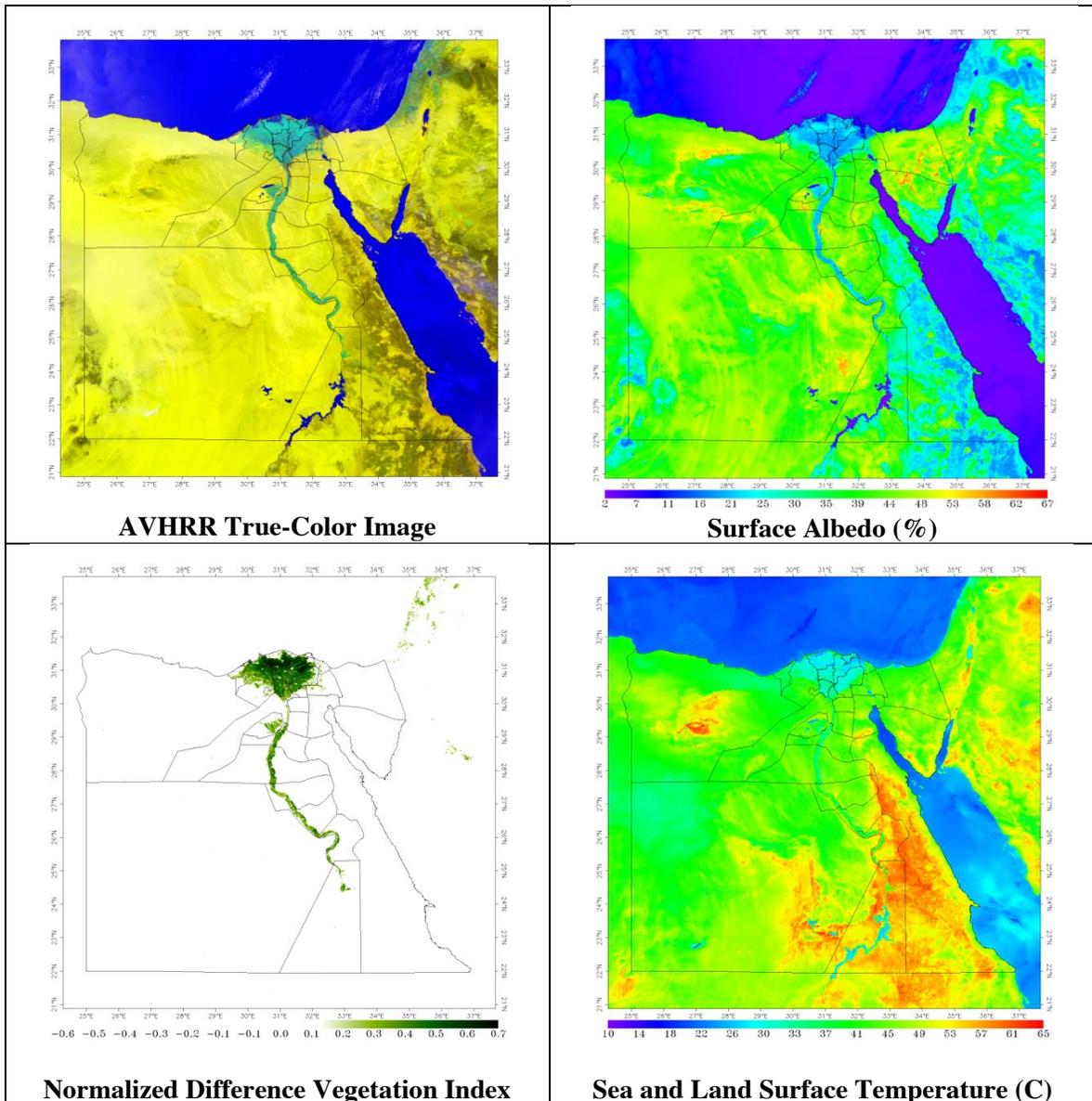
The NDVI and surface albedo are calculated using the following equations:

$$NDVI = \frac{Ch_2 - Ch_1}{Ch_2 + Ch_1} \tag{3.1}$$

$$Albedo = 0.5Ch_1 + 0.5Ch_2 \tag{3.2}$$

Where Ch_1 and Ch_2 are the spectral reflectance measurements acquired from channel 1 and 2, respectively. The SST and LST are estimated using the multi channel sea surface temperature (MCSST) [79] and split window [80] algorithms, respectively. Table 3-6 shows sample retrievals of the AVHRR data at NARSS.

Table 3-6 Sample retrievals of the NOAA/AVHRR data at NARSS



3.5 DATA QUALITY

The quality of NOAA/ATOVS retrievals, which are the main source of satellite observations in this study, was based on the comparison with the radiosonde observations. Figure 3-10 shows the comparison of NOAA/ATOVS temperature retrievals from NOAA-15 and NOAA-16 satellites with the observations at Helwan station at 00Z 01 October, 2005. The correlation with the observations was 0.998 and 0.999 for NOAA-15 and NOAA-16, respectively. The root-mean-square bias was 0.454 °C for NOAA-15 and 0.422 °C for NOAA-16.

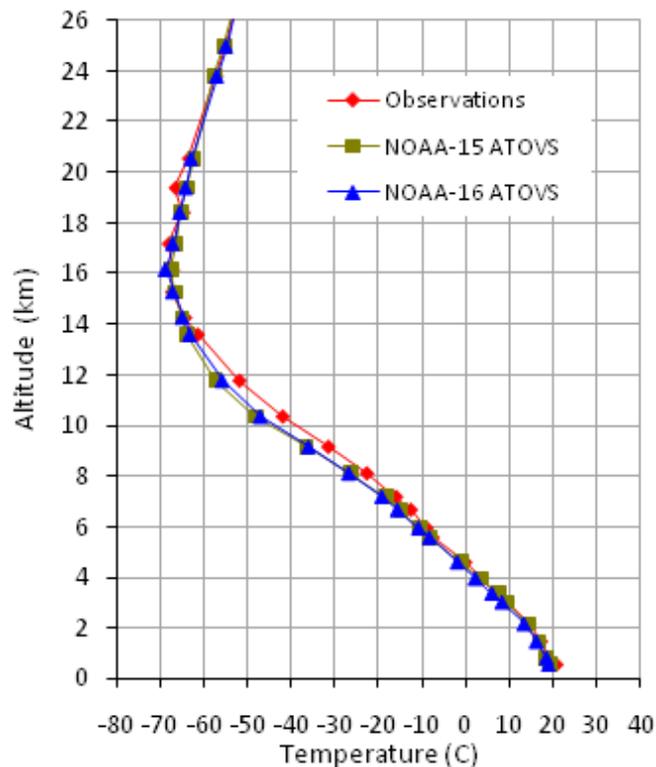


Figure 3-10 Validation of ATOVS Retrievals at Helwan station at 00Z 01 Oct 2005

In the work of Schroeder et al. [81], global IAPP retrievals were compared to radiosondes and CHAMP radio occultation data. They found that Water vapor and temperature products exhibit a very high quality, Total Precipitable Water (TPW) bias fluctuates around 0 mm with a mean value of 0.2 mm, Layer Precipitable Water (LPW) bias generally <0.5 mm (max. of 0.8 mm at 850-700 hPa) and Temperature (T) bias usually <0.5 K (max. of -1 K at 300-200 hPa).

CHAPTER FOUR

DATA ASSIMILATION

This chapter introduces the analysis methods and shows the different methods and types of data assimilation (DA). Then, it shows the empirical DA methods used in this study. This includes the 3DDA method which is called “Objective Analysis (OA)”, the 4DDA method which is called “Four-Dimensional Data Assimilation (FDDA)”, different schemes used in the analysis, and the quality control for observations. Then, the work done to utilize DA in NWP for Egypt is presented.

4.1 INTRODUCTION

The mathematical formalization of the analysis method starts with the definition of the work space. The values of the model DOFs needed to represent the state of the system are collected as a column matrix called *state vector* x . The components of this vector are related to the real state based on the choice of discretization method. The best possible representation of the real state as a state vector may be denoted by the *true state* x^t . The first guess of this state, before the analysis, is denoted by the *priori* or the *background* x^b . The analysis problem tries to find the corrections δx (or *analysis increments*) to the background which make the *analysis* ($x^a = x^b + \delta x$) as close as possible to the true state x^t .

4.1.1 Observational Increments

The observations are gathered into an *observation vector* y^o . These observations are compared with the state vector data using a scheme called the *observation operator* or *forward operator* H . This scheme is a collection of interpolation operators from the model discretization to the convertors points, and conversions from model variables to the observed parameters. The key of data analysis is the use of the discrepancies between observations and state vector. This is determined

from the vector of *departures* at the observation points ($y^o - H(x)$). When calculated with the background x^b , it is called *innovations* or *observational increments*, and with the analysis x^a , *analysis residuals*. This provides important information about the quality of DA.

4.1.2 Observational Errors

The observational errors may occur from:

- Having a less than adequate number of observations,
- Inability to resolve meteorological features of interest,
- Poor receiving frequency for observations to monitor phenomena evolution,
- Inaccurate instrument readings or data-transmission problems, and
- Observations although correct, may represent local phenomena too small to be resolved by the analysis, i.e. the resulting representation for a larger area is biased by extremely localized phenomena.

Thus, the observational errors (ε^o) may be modeled as two main types:

- instrumental errors (ε_{inst}^o) and
- representativeness errors (ε_{repr}^o); and it may be written as

$$\varepsilon^o = y^o - H(x^t) = \varepsilon_{inst}^o + \varepsilon_{repr}^o + \varepsilon_H^o \quad (4.1)$$

where ε_H^o is the error in the forward operator H .

These errors with any other NWP errors may be propagated by the model and it may be present within the subsequent forecast which is used as a background for the next forecast. This results in the cyclic dependency; i.e. DA depends on a reasonable background field and the background field depends on the result of DA and/or NWP. Observations are often incomplete and/or imperfect. Therefore, more/better observations have to be implemented.

4.2 ASSIMILATION METHODS

In DA cycle, the background field is a short-range forecast, x^b . This field is interpolated to the corresponding locations of observations and then converted from model variables to observed variables, y^o . The first guess at the observations locations is therefore $H(x^b)$. The difference between the observations and the model first guess is the innovations (observational increments), $y^o - H(x^b)$. The analysis, x^a , is obtained by adding the innovations to the model forecast (background) with a weighting function, W , based on the following equation:

$$x^a = x^b + W[y^o - H(x^b)] \quad (4.2)$$

Different analysis schemes are based on equation (4.2) but they differ by changing the approach taken to combine the background and the observations to produce the analysis. The matrix of weights, W , may be determined empirically, statistically or using Kalman Filtering (KF). KF is a development of the least-squares methods which are based on the statistical estimation theory.

4.2.1 Types of Data Assimilation

Data Assimilation (DA) in NWP is usually classified into two main types which are 3-dimensional (3DDA) and 4-dimensional (4DDA) types. In 3DDA, only those observations available at the time of analysis are used based on equation (4.2). The weight, W , in 3DDA is a function of the three independent spatial variables (three-dimensional weighting function). In 4DDA, the temporal (past) observations are included in two steps as follows:

- Forecast Step: from time t_n to time t_{n+1} which is governed by

$$x_{t_{n+1}}^b = M[x_{t_n}^a] \quad (4.3)$$

- Analysis Step: at time t_{n+1} which is governed by

$$x_{t_{n+1}}^a = x_{t_{n+1}}^b + W[y_{t_{n+1}}^o - H(x_{t_{n+1}}^b)] \quad (4.4)$$

The weight, W , in 4DDA is a function of the three independent spatial variables and the time (four-dimensional weighting function). In the analysis cycle, the use of model forecast is essential in achieving 4DDA. This means that DA cycle is working as a long model integration, in which the model is “nudged” by observational increments in such a way that it remains close to the real atmosphere. The importance of the model cannot be overemphasized: it transports information from data-rich to data-poor regions, and provide a complete estimation of the four-dimensional state of the atmosphere.

4.2.2 Methods of Data Assimilation

DA methods may be further divided into empirical (theoretical), statistical (variational/sequential) and advanced methods according to the approach taken to determine the weighing functions to produce the analysis. The empirical DA methods are simple and useful starting tools. The other two methods are computationally very expensive because they try to find the optimal analysis. These methods are needed due to the following problems in empirical methods:

- Good-quality preliminary estimations of the analysis may be replaced by values provided from poor-quality observations and hence
 - Quality control checks should be applied to help in using only the observations with high quality level.
- It is not clear how to relax the analysis toward the arbitrary state away from the observations especially in data-poor areas and hence
 - The weighting function, W , should be selected carefully.
- The physical constraints are not guaranteed by the analysis and hence
 - Smoothing should be applied on the analysis.

The different methods of DA perform differently and generally it is not clear which approach is better than the others. Table 4-1 summarizes different empirical, statistical and advanced DA methods.

Table 4-1 Empirical, statistical and advanced DA methods

DA	Method	Approach	History
Empirical	Objective Analysis (OA)	Simple interpolation techniques (fits observational data to grids with a weighting average)	Charney, 1951
	Successive Correction Method (SCM)	The estimated error statistical covariance of the first guess and observations (adds innovations (observational increments) to the first guess with empirical weights)	Bergthorsson and Doos, 1955
	Four-Dimensional Data Assimilation (FDDA)	Simple idea of nudging or Newtonian relaxation (nudges the first guess to the observations with empirical weights)	Kistler, 1974
Statistical	Optimal Interpolation (OI)	Similar to SCM but the weights are determined using the minimization of analysis error at grid points	L Gandin, 1963
	3D Variational DA (3DVar)	The minimization of a cost function in three spatial dimensions	Sasaki, 1958
	4D Variational DA (4DVar)	The minimization of a cost function in three spatial dimensions and time	Bottier and Rabier, 1997
Advanced	Ensemble Kalman Filter (EnKF)	Application of an ensemble of forecasts to represent the model error statistics (solves the KF error covariance update equation)	Evensen, 1994
	Ensemble Transform Kalman Filter (ETKF)	A modified scheme of the EnKF to solve the KF error covariance update equation (useful for EF but not for DA)	Bishop, 2001
	Hybrid DA	Hybrid ETKF and Variational DA (ETKF updates ensemble deviations but not ensemble mean)	Wang, 2008

Table 4-2 shows a comparison between the available empirical and variational methods in the current public releases of NWP models. The advanced methods are not available and they are computationally very expensive.

Table 4-2 Comparison between empirical and variational assimilation methods

Empirical Methods	Variational Methods
Aim to nudge the first guess toward the observations at analysis time.	Aim to produce <i>optimal</i> estimate of the true atmospheric state.
3DDA methods (OA) 4DDA methods (FDDA)	3DDA methods (3DVar) 4DDA methods (4DVar)
Used to assimilate direct data (conventional data / satellite retrievals)	Used to assimilate direct data and indirect data (radiances/radar)
Some decoders were developed to utilize the PREPBUFR observations.	Direct utilization of the PREPBUFR observations.
Require dense distribution of observations for better results.	Require more observations for the huge number of DOFs.
Empirical / Time efficient.	Approximate / Time consuming.

In this study, the empirical DA methods were used. The main reason for this choice is that they are simple and computationally not expensive. Therefore, it is convenient to start with them and then the sensitivity of using different DA methods and/or different types of observations in NWP may be studied later. Moreover, the available observations inside Egypt are sparse and they are not uniformly distributed in the domain.

4.3 OBJECTIVE ANALYSIS

The goal of Objective Analysis (OA) in meteorological modeling is to improve the global meteorological analyses (the first guess) on the mesoscale grid by incorporating information from observations. OA is the process of interpolating observed values onto the grid points used by the model/analysis in order to define the initial conditions of the state of the atmosphere. The basic problem of OA is that the observations are irregularly spaced while the values for points on a regularly-spaced grid must be provided. Figure 4-1 shows a schematic of grid points (circles) and irregularly-distributed observations (squares).

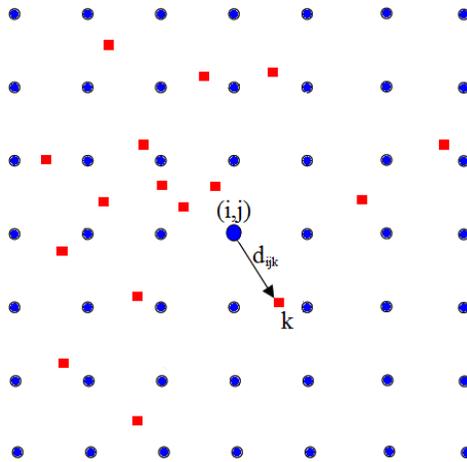


Figure 4-1 Schematic of grid points and irregularly-distributed observations [21]

Traditionally, these observations were in the form of “direct” observations of temperature, humidity, and wind from surface and radiosonde reports. As remote sensing techniques improved, more and more “indirect” observations became available. Effective use of these indirect observations for objective analysis is not an easy task. The MM5 modeling system includes the LITTLE_R program for OA of the direct observations of temperature, water vapor and winds. Figure 4-2 shows a demonstration for the OA and cyclic dependency for a cross-section of the temperature field near the surface.

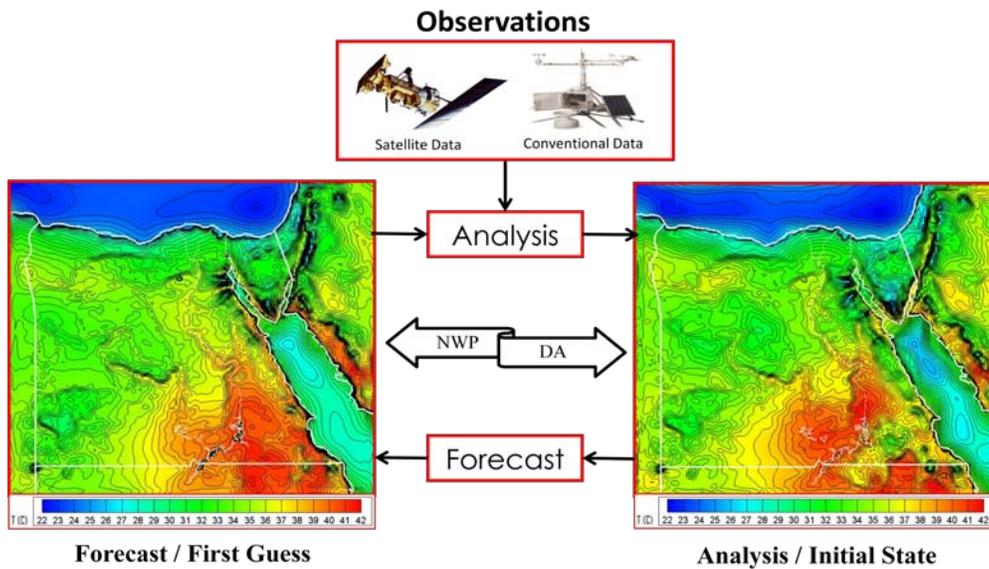


Figure 4-2 Objective analysis and cyclic dependency

4.3.1 Objective Analysis Techniques

The weight function in equation (4.2) may be determined empirically using four different OA techniques using the LITTLE_R program in the MM5 modeling system. Three techniques are based on the “Cressman scheme” [82], in which several successive scans nudge the first-guess field toward the neighboring observed values. The fourth technique is based on “Multi-quadric scheme” [83] which uses hyperboloid radial basis functions to perform the OA.

4.3.1.1 Cressman Schemes

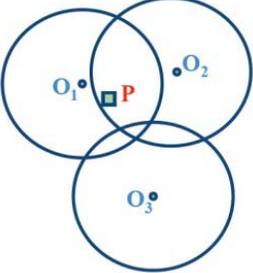
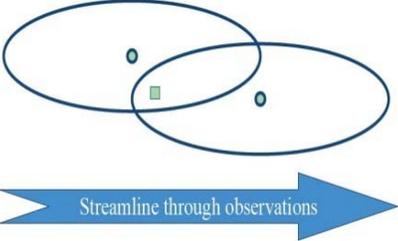
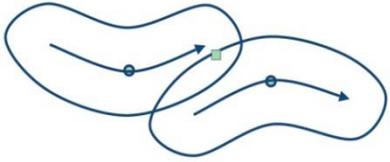
The OA of direct observations for the variable α can be obtained using a reduced form of equation (4.2) which may be expressed, in an update form, as

$$\alpha_{ij}^a = \alpha_{ij}^b + \frac{\sum_{k=1}^N w_{ijk} \Delta \alpha_k}{\sum_{k=1}^N w_{ijk}} \quad (4.5)$$

where α_{ij}^a is the analysis for the variable α at the grid point (i, j) , α_{ij}^b is the first guess for the variable α at the grid point (i, j) , N is the total number of observations used by the OA, $\Delta \alpha_k$ is the difference between the k^{th} observed value and the first guess value for the variable α at the station location, and w_{ijk} is weight function of the k^{th} observation at the grid point (i, j) .

The standard Cressman scheme assigns a circular radius of influence to each observation. In analyses of wind and relative humidity which are the fields strongly deformed by the wind, at pressure levels, the circles from the standard scheme are elongated into ellipses oriented along the flow (ellipse scheme) or curved along the streamlines (banana scheme). Table 4-3 shows schematics for the different schemes and the corresponding weighting functions. The Banana scheme converts to the ellipse scheme under straight-flow conditions. The ellipse scheme converts to the standard Cressman scheme under low-wind conditions. Appendix D describes the mathematical formulations for different Cressman schemes.

Table 4-3 Different Cressman schemes and the corresponding weighting functions [58]

Standard Scheme		<p>The circular weighting function:</p> $w_{ijk} = \begin{cases} \frac{R^2 - d_{ijk}^2}{R^2 - d_{ijk}^2} & \text{for } d_{ijk} < R \\ 0 & \text{for } d_{ijk} \geq R \end{cases}$ $d_{ijk}^2 = (x_{ij} - x_k)^2 + (y_{ij} - y_k)^2$
Ellipse Scheme		<p>The elliptical weighting function:</p> $w_{ijk} = \begin{cases} \frac{R^2 - d_m^2}{R^2 - d_m^2} & \text{for } d_m^2 < R^2 \\ 0 & \text{for } d_m^2 \geq R^2 \end{cases}$ $d_m^2 = \frac{x^2}{E_k^2 V } + y^2$
Banana Scheme	 <p>• Observations ■ Grid Point</p>	<p>The banana weighting function:</p> $w_{ijk} = \begin{cases} \frac{R^2 - d_m^2}{R^2 - d_m^2} & \text{for } d_m^2 < R^2 \\ 0 & \text{for } d_m^2 \geq R^2 \end{cases}$ $d_m^2 = \frac{r_k^2 (\theta_k - \theta_{ij})^2}{E_k^2 V } + (r_k - r_{ij})^2$

The first-guess field at each grid point P is adjusted by taking into account all the observations which influence P . The differences between the first-guess field and the observations are calculated, and a distance weighted average of these difference values is added to the value of the first-guess at P . Once all grid points have been adjusted, the adjusted field is used as the first guess for another adjustment cycle. Subsequent passes use smaller radius of influence.

4.3.1.2 Multi-quadric Scheme

The Multi-quadric scheme works on the whole domain, adjusting every grid point according to the following equation:

$$\alpha_{ij}^a = \alpha_{ij}^b + Q_{ijk} [Q_{km} + (N\lambda\sigma_k^2\delta_{km})]^{-1} \Delta\alpha_m \quad (4.6)$$

where λ is the smoothing parameter, δ_{km} is the Kronecker delta, and σ_k^2 is the statistical error associated with the meteorological variables. Q_{km} and Q_{ijk} are the matrices of the radial functions depending on the distance between the k^{th} observation and, respectively, the other station (m) and the grid point (i, j). The normalized radial functions depend on the multi-quadric parameter c ($0 < c < 1$) that is an index of their sharpness, as shown in the following equation:

$$Q_k(x_m, y_m) = - \left(\frac{|x_m - x_k|^2 + |y_m - y_k|^2}{c^2} + 1.0 \right)^{1/2} \quad (4.7)$$

Cressman and Multi-quadric techniques use different approaches. Cressman technique corrects the values at each grid point accounting only for the observations enclosed in the circular/elliptical/banana area and it does not account for the errors coming from the measurements. On the other hand, Multi-quadric technique corrects the values at each grid point using the whole dataset and it accounts for the statistical error in the measurements. Multi-quadric analysis does not do well in domains that have large no-data areas [84]. However, reasonably good sets of initial and/or boundary conditions were possible to obtain, using Multi-quadric technique, over the western desert in Egypt which is a relatively large data-poor area extending to domain boundaries.

4.3.2 Quality Control for Observations

A critical component of the OA is the screening for bad observations. Even a single bad observation can ruin initial conditions and produce a series of failed forecasts due to cyclic dependency (bad observations = bad analysis which results in forecast failure). LITTLE_R program, in the MM5 modeling systems, are responsible for the OA. A number of checks in LITTLE_R are optionally performed including removal of observations outside the domain, excluding location/time duplicates and incomplete observations such as no location, and ensuring vertical consistency of upper-air profiles.

Three different categories of tests may be performed to check the quality of observations. The first is the quality control on individual reports which does not compare the observations with the first-guess field or the other observations. The second is the ERRMAX test which compares the observations to the first-guess field. The third is the Buddy test which compares the observations to both the first guess and the neighboring observations influence a grid point.

4.3.2.1 Quality Control on Individual Reports

The quality control on individual reports includes the following checks and features:

- Gross Error Checks.
 - Sane values, pressure decreases with height, and so on.
- Filter and smooth temperature and wind profiles.
 - Optional but not recommended.
- Adjust temperature profiles to remove super-adiabatic layers.
 - Super-adiabatic layers are unstable layers:
 - The temperature decreases with height at a rate of greater than 10 degrees Celsius per kilometer.
 - Optional but not recommended.
- No comparisons to other reports or to the first-guess field.

4.3.2.2 The ERRMAX Test

The quality control for observations using the ERRMAX test includes the following checks and features:

- Limited user control over data removal.
 - The thresholds may be set varying the tolerance of error check.
- Observations are compared to the first-guess field.
- If the difference value between the observation and the first-guess exceeds a certain threshold, the observation is discarded.

- Discard if $\Delta\alpha = \alpha^o - \alpha^b > \Delta\alpha|_{max}$, $\Delta\alpha|_{max}$ is a user-defined threshold for the variable α .
- Threshold varies depending on the field (α), level, and time of day.
- Works well with a good first-guess field.

4.3.2.3 The Buddy Test

The quality control for observations using the Buddy test includes the following checks and features:

- Limited user control over data removal.
 - Weighting factors may be set varying the tolerance of error check.
- Observations are compared to both the first guess and the neighboring observations influence a grid point.
- If the difference value between the observation and the first-guess ($\Delta\alpha$) varies significantly from the distance-weighted average of the difference values of neighboring observations, the observation is discarded.
- Works well in regions with good data density.

4.4 FOUR-DIMENSIONAL DATA ASSIMILATION

Four-Dimensional Data Assimilation (FDDA) is a concept of combining current and past observational data in an explicit dynamical model such that the model's prognostic equations provide time continuity and dynamic coupling among the various fields. It is used for either model initialization (dynamic initialization) or for using the model as an analysis/research tool (dynamic analysis).

4.4.1 Applications of FDDA

FDDA may be used for dynamic initialization during a pre-forecast time period, dynamic analysis during the whole simulation time period or boundary conditions by nudging the outer domain towards analysis.

4.4.1.1 Dynamic Initialization

FDDA is applied during a pre-forecast time period, as shown in Figure 4-3, for which additional observations and/or analyses exist. Then it will be turned off as the forecast begins. This has two advantages over the static initialization, namely;

- (i) It can make use of asynoptic data during the pre-forecast period and it generally contains more observational information at the forecast start time, and
- (ii) There is a reduced spin-up or shock effect at the forecast starting owing to the better balance of the initial conditions of the model.

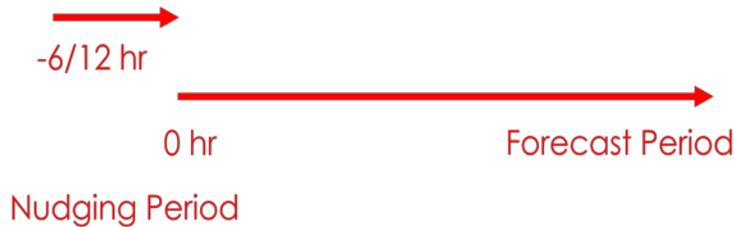


Figure 4-3 Using FDDA for dynamic initialization

4.4.1.2 Dynamic Analysis

Dynamic analysis is the same as dynamic initialization except that the intent is to produce a four-dimensional consistent analysis taking into account dynamical balances that are provided by the model and observations that are introduced by nudging, as shown in Figure 4-4. It may be used to initialize higher-resolution simulations or for kinematic studies such as chemical and tracer transports.

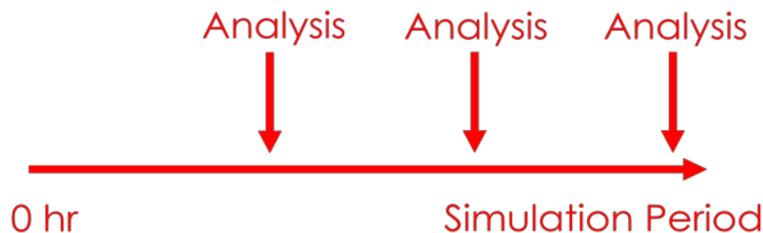


Figure 4-4 Using FDDA for dynamic analysis

4.4.1.3 Boundary Conditions

By using DA on the coarse mesh and nesting the results with a finer mesh, as shown in Figure 4-5, the fine mesh is provided with superior boundary conditions compared to the standard linear interpolation of analyses. The boundaries have a much-higher-time-resolution features passing through them to the fine mesh.

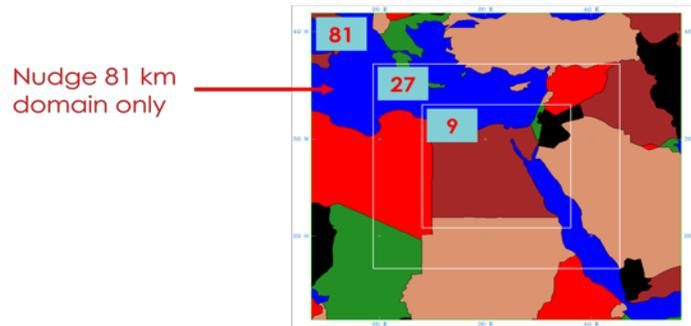


Figure 4-5 Using FDDA for boundary conditions

4.4.2 FDDA Methods

FDDA is based on the basic idea of nudging or Newtonian relaxation. Nudging is a relatively simple but very flexible technique which relaxes the model state toward the observed state by adding, to one or more of the prognostic equations, artificial tendency terms based on the difference between the two states. There are two major types of FDDA; analysis/grid and station/observational nudging. The model can use these types individually or combined.

4.4.2.1 Analysis or Grid Nudging

The grid nudging is an intermittent process of initializing the model, using the subsequent forecast (typically 3-12 h) as a first guess in a static three-dimensional OA step. Newtonian relaxation terms are added to the prognostic equations for wind, temperature, and water vapor. These terms relax the model value towards a given analysis. The analysis-nudging term for a given variable is proportional to the difference between the model simulation and an analysis of observations,

calculated at each grid point. The non-hydrostatic predictive equation of a variable $\alpha(\mathbf{x}, t)$ may be written as

$$\frac{\partial p^* \alpha}{\partial t} = F(\alpha, \mathbf{x}, t) + G_\alpha \cdot W_\alpha(\mathbf{x}, t) \cdot \epsilon_\alpha(\mathbf{x}) \cdot p^*(\hat{\alpha}_o - \alpha) \quad (4.8)$$

where,

t	The model-relative time variable
\mathbf{x}	The three independent spatial variables
p^*	The reference pressure computed from the hydrostatic reference state which is constant with time
α	The model's dependent variable
$\hat{\alpha}_o$	The estimate of the observation for α analyzed to the model grid
F	The model's physical forcing terms (advection, Coriolis effects, etc.)
G_α	The nudging factor for the variable α which determines the magnitude of the nudging term relative to all the other model processes in F . <ul style="list-style-type: none"> • It is usually defined to be similar in magnitude to the Coriolis parameter, • It must also satisfy the numerical stability criterion $G_\alpha \leq \frac{1}{\Delta t}$. • Typical values of G_α are from $10^{-4} s^{-1}$ to $10^{-3} s^{-1}$, where values of $3 \times 10^{-4} s^{-1}$ to $6 \times 10^{-4} s^{-1}$ are usually "large enough". • A value of G_α which is too large will force the model state too strongly toward the observations. • A value of G_α which is too small will minimize the effect of the observations on the evolution of the model state.
W_α	The four-dimensional weighting function: $W_\alpha(\mathbf{x}, t) = w_{xy} \cdot w_\sigma \cdot w_t$
w_{xy}	The horizontal weighting function which is usually set to unity at each equivalent grid point and zero everywhere
w_σ	The vertical weighting function: $w_\sigma = w_\sigma^R + w_\sigma^S \leq 1$
w_σ^R	The vertical weighting function for assimilation of 3-D radiosonde data <ul style="list-style-type: none"> • $w_\sigma^R = 0.0$ in the PBL, 0.1 in the transition layer and 1.0 aloft.
w_σ^S	The vertical weighting function for assimilation of 2-D surface data <ul style="list-style-type: none"> • The surface data are assimilated with full strength ($w_\sigma^S = 1.0$) within the PBL layers and with reduced strength ($w_\sigma^S = 0.9$) one layer above.
w_t	The temporal weighting function which is usually set to unity, except when decreasing the nudging at the end of a dynamic-initialization period.
ϵ_α	The analysis quality factor which is based on the quality and distribution of the data used to produce the gridded analysis: $0 \leq \epsilon_\alpha \leq 1$

The last term in equation (4.8) is the analysis-nudging relaxation term which has an artificial contribution. It must not be a dominant term and should be scaled by the slowest physical adjustment process in the model (inertial effects).

4.4.2.2 Station or Observational Nudging

The observational nudging is a continuous dynamical assimilation. It is useful in situations where analysis nudging is not practical such as at high resolution or with asynoptic data. It uses relaxation terms based on the model error at observational stations. Relaxation is designed to reduce this error using the following equation

$$\frac{\partial p^* \alpha}{\partial t} = F(\alpha, \mathbf{x}, t) + G_\alpha \cdot p^* \frac{\sum_{i=1}^N W_{\alpha i}^2(\mathbf{x}, t) \cdot \gamma_i \cdot (\alpha_o - \hat{\alpha})_i}{\sum_{i=1}^N W_i(\mathbf{x}, t)} \quad (4.9)$$

where,

α_o	The locally observed value of the variable α
$\hat{\alpha}$	The variable α interpolated to the observation location in three dimensions
i	The subscript denotes the i^{th} observation of a total of N
N	The total number of observations within a preset radius of a given grid point
W_α	The four-dimensional weighting function: $W_\alpha(\mathbf{x}, t) = w_{xy} \cdot w_\sigma \cdot w_t$
w_{xy}	The horizontal weighting function: $w_{xy} = \begin{cases} \frac{R^2 - D^2}{R^2 + D^2}, & 0 \leq D \leq R \\ 0, & D > R \end{cases}$
R	The radius of influence for the i^{th} observation
D	The distance from observation modified by elevation difference
w_σ	The vertical weighting function: $w_\sigma = \begin{cases} 1 - \frac{ \sigma_o - \sigma }{R_\sigma}, & \sigma_o - \sigma \leq R_\sigma \\ 0, & \sigma_o - \sigma > R_\sigma \end{cases}$
R_σ	The vertical radius of influence of the i^{th} observation
σ	The vertical position of the model grid point
σ_o	The vertical position of the i^{th} observation
w_t	The temporal weighting function: $w_t = \begin{cases} 1, & t - t_o < \tau/2 \\ \frac{\tau - t - t_o }{\tau/2}, & \tau/2 \leq t - t_o \leq \tau \\ 0, & t - t_o > \tau \end{cases}$
t_o	The model-relative time of the i^{th} observation
τ	The half-period of a predetermined time window over which an observation will influence the model simulation
γ	The observational quality factor which accounts for characteristic errors in measurement systems and representativeness: $0 \leq \gamma \leq 1$

The last term in equation (4.9) is the observational-nudging relaxation term.

4.5 APPLICATION OF DATA ASSIMILATION

In this study, the performance of the MM5 model, when using different FDDA techniques, was tested and evaluated during the period from July 1, 2005 to July 1, 2006 using three-day simulations per season to provide physically-consistent meteorological simulations with 9 km resolution for Egypt. The model evaluation, when using FDDA in NWP for Egypt at different locations, was performed in summer 2008 and winter 2006. This includes different nudging options for the third domain and in the PBL as well as Cressman and Multi-quadric nudging schemes. Then, the performance of the model is tested for different initialization times as well as in medium-range simulations. Table 4-4 presents the selected time periods for which the MM5 simulations were performed.

Table 4-4 Selected time periods for testing FDDA enhancement of the MM5 simulations

#	Starting Date and Time				Period (days)	Description
	Year	Month	Day	Time		
1	2005	7	1	Z00	3	FDDA enhancement of the NWP MM5 model for the four seasons and using different FDDA techniques
2	2005	10	1	Z00	3	
3	2006	1	1	Z00	3	
4	2006	4	1	Z00	3	
5	2006	1	1	Z12	3	Model Evaluation when using different FDDA options
6	2008	8	1	Z12	3	
7	2006	1	1	Z12	6	Medium-range simulation

Four sets of simulations were performed. The first simulation is the reference case, FDDA0, where no FDDA is utilized. The second simulation, FDDA1, utilized grid FDDA from OA of NCEP Final Analysis (FNL) or the coarser-domain output with conventional data. The third simulation, FDDA2, utilized observational FDDA from remotely-sensed observations (ATOVS data). Finally, FDDA3 utilized combined grid and observational FDDA. From the analysis nudging options, both the 3D and surface nudging was applied for winds, temperature and moisture. Table 4-5 summarizes the description of different sets of DA simulations.

Table 4-5 Description of different sets of DA simulations

Simulation	Description
FDDA0	Reference Simulation, no FDDA is utilized
FDDA1	Grid FDDA was utilized from OA of FNL with conventional data
FDDA2	Observational FDDA was utilized from ATOVS data
FDDA3	Combined grid and observational FDDA were implemented

The MM5 model was initialized with the NCEP Final Analysis (FNL) datasets that has a $1^\circ \sim 111$ km horizontal resolution, and with the NOAH land surface model (LSM). The physics of all MM5 simulations utilized the MRF scheme for planetary boundary layer (PBL), the Grell and Anthes-Kuo schemes for cumulus parameterization, simple ice (Dudhia) scheme for explicit moisture, a cloud radiation scheme, and the NOAH LSM for ground temperature. Table 4-6 shows the grid and physics options for the three domains used in DA simulations.

Table 4-6 Grid and physics options for the three domains used in DA simulations

Domain ID	1	2	3
Resolution	81 km	27 km	9 km
Grid Size	$71 \times 71 \times 38$	$97 \times 97 \times 38$	$175 \times 175 \times 38$
Physics	Cumulus	Anthes-Kuo	Grell
	PBL & Diffusion	MRF PBL	MRF PBL
	Microphysics	Dudhia	Dudhia
	Radiation	Cloud Radiation	Cloud Radiation
	Surface	NOAH LSM	NOAH LSM

The grid nudging technique was used to nudge the simulated aloft fields every six hours and the surface fields every three hours. FDDA options through grid nudging were utilized in the free atmosphere to nudge horizontal winds, temperature, and moisture for all domains. In the PBL, the winds, temperature, and moisture were nudged using surface data through grid nudging.

The observational nudging technique was used to nudge the simulated aloft fields at the time of ATOVS data within one minute time window ($\tau = 0.5$ minute). FDDA options through observational nudging were utilized in the free atmosphere

to nudge temperature and humidity profiles retrieved from ATOVS. Recommended nudging coefficients were used where $G_\alpha = 3 \times 10^{-4} s^{-1}$ for all variables. The radii of influence were set to one grid distance of each domain ($R = \Delta x = \Delta y$). The output frequency of the model was set to one hour.

The ATOVS data frequency over Egypt is around 2-5 images daily, at day and at night, from different NOAA satellites. This data is received during 15 minutes for each image and was mapped, then, into 30-minutes intervals. Two experiments were tested using different observational-nudging frequency. In the first experiment, the nudging weights were computed for each time step (20 seconds for Egypt domain). Then, in the second experiment, they were computed every 30 minutes. This saved around 10-15% of computational time without affecting the accuracy of the simulations.

Five sets of simulations were performed to investigate the effects of data assimilation via different nudging options on the results of simulations at locations with different observation site density inside Egypt, stressing on the effects of using data assimilation on accuracy at the observational-data-void areas and at the locations with complex terrain.

These five sets of simulations are grid-nudging experiments that were used to test the effects of using different nudging options. The first case is the reference case, “Ref”, where no FDDA was utilized for all domains. The second case, “No D3”, utilized FDDA with Multi-quadric nudging scheme, MQ, for the coarser domains only (no nudging for the third domain). The third case, “No PBL”, utilized FDDA with MQ for all domains but not in the PBL. The fourth case, “DA MQ”, utilized FDDA with MQ for all domains. Finally, “DA CM” utilized FDDA with Cressman nudging scheme, CM, for all domains. Table 4-7 summarizes the descriptions of these simulations.

Table 4-7 Description of the simulations performed for the evaluation of FDDA

Simulation	Description
Ref	No FDDA was utilized for all domains
No D3	FDDA with MQ was utilized for the coarser domains only
No PBL	FDDA with MQ was utilized for all domains but not in the PBL
DA MQ	FDDA with MQ was utilized for all domains
DA CM	FDDA with CM was utilized for all domains

Four sets were selected for data sampling locations. The first set, a “principal set”, is the set of stations used in data assimilation. The second set, “interpolation set”, is the set of stations located in-between the principal stations. The third set, “extrapolation set”, is the set of stations located far from the principal stations. The fourth set, “special set”, is the set of stations located at complex terrain. The distribution and locations of the different observation sets are shown Figure 2-18.

The results of using FDDA in numerical weather modeling for Egypt including the comparison of different FDDA methods in the four seasons and the evaluation study are presented in section 6.1 and 6.2.

CHAPTER FIVE

ENSEMBLE FORECASTING

This chapter is considered with introducing the atmospheric predictability, the sources of errors in NWP models, and the construction of an ensemble. It also covers the investigation of the ensemble forecasting philosophy, different Ensemble Forecasting (EF) methods and a preliminary short-range EF system for Egypt based on the available observational data and computational resources.

5.1 INTRODUCTION

Until 1991, operational NWP centers were using a single control forecast, Deterministic Forecasting (DF), starting by the analysis. In December 1992, both NCEP and ECMWF started using EF from slightly perturbed initial conditions. EF provides a range of possible solutions (ensemble members) to forecast the future atmospheric state. The average of these solutions (ensemble mean) is generally more accurate than the single deterministic forecast and the spread of them gives information about the forecast errors. It may also provide a quantitative basis for probabilistic forecasting. Figure 5-1 shows the evolution of the forecasting skills.

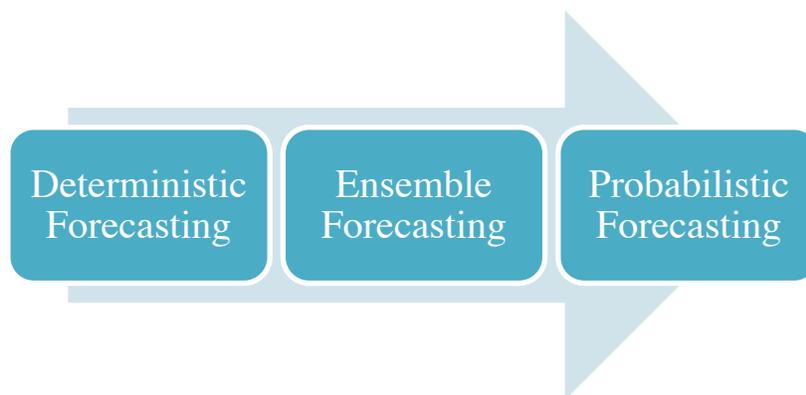


Figure 5-1 Evolution of forecasting skills to probabilistic forecasting

5.1.1 Atmospheric Predictability

In 1951, Charney indicated that he expected that even as models improved there would still be a limited range to skillful atmospheric predictions. He attributed this to inevitable model deficiencies and finite errors in the initial conditions. In 1963, Lorenz discovered the fact that the atmosphere, like any dynamical system with instabilities, has a finite limit of predictability, which he estimated to be about two weeks, even if the model is perfect, and even if the initial conditions are known almost perfectly. He introduced a three-variable model governed by

$$\frac{dx}{dt} = \sigma(y - x) \quad (5.1)$$

$$\frac{dy}{dt} = rx - y - xz \quad (5.2)$$

$$\frac{dz}{dt} = xy - bz \quad (5.3)$$

Where the constants were chosen to be: $\sigma = 10$, $b = 8/3$, and $r = 28$.

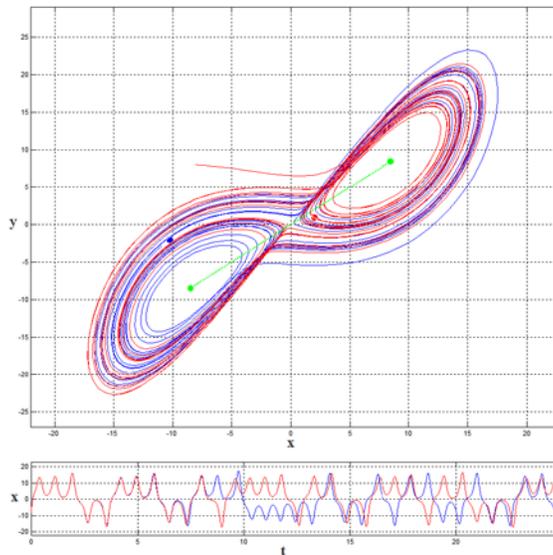


Figure 5-2 “Identical twins” experiment with Lorenz three-variable model

Lorenz performed what is now denoted as “identical twin” experiment using two runs with very slightly different initial conditions (order of round-off errors). He found that after a few weeks the two solutions were as different from each other as two random trajectories of the model as shown in Figure 5-2.

5.1.2 Sources of Errors in NWP Model

Scientific, computational and practical limitations prevent the construction of a perfect NWP model. The main two sources of NWP errors are the discrepancy between the model and nature (continuous with time) and the discrepancy between the estimated and actual state of nature (initial time). It is also important to realize that there is a finite limit to predictability from NWP models. The loss of predictability is not uniform in space and time. The different Sources of NWP errors are summarized in Table 5-1 [85].

Table 5-1 Different sources of NWP errors [85]

Errors in the initial conditions	Lack of observations: <ul style="list-style-type: none"> • Data voids: oceans, Polar Regions, deserts, mountains, less- developed countries • Insufficient data density to resolve important mesoscale features
	Observational errors: <ul style="list-style-type: none"> • Instrumental errors: instrument errors, errors in coding, transmission • Representativeness errors: observations of sub-grid-scale variability not represented in the grid-average values of the model and analysis
	Analysis and initialization errors: <ul style="list-style-type: none"> • Errors in interpolation to the analysis grid • Errors in interpolation to the model grid • Errors caused by balancing assumption in the initialization
Errors in the numerical model	Limitation of the equations: <ul style="list-style-type: none"> • Complete governing laws not fully known or used (e.g., turbulence modeling)
	Errors in numerical approximation: <ul style="list-style-type: none"> • Horizontal resolution: truncation error • Vertical resolution: truncation error • Boundary conditions: surface boundary conditions and lateral boundary conditions in Limited-Area Models (LAM)
	Incomplete or erroneous physics: <ul style="list-style-type: none"> • Mountains: not steep or rugged enough • Latent heat release, especially convection • Boundary layer processes • Radiation; surface energy balance

5.1.3 Construction of an Ensemble

Ensemble Forecasting (EF) may take into account all known sources of uncertainty in NWP as illustrated in Figure 5-3. The ensemble of forecasts should have a spread around its mean that is approximately equal to the error in the ensemble mean forecast. This is a necessary condition for having a perfect ensemble where the verifying analysis, in a statistical sense, is indistinguishable from the ensemble members [86].

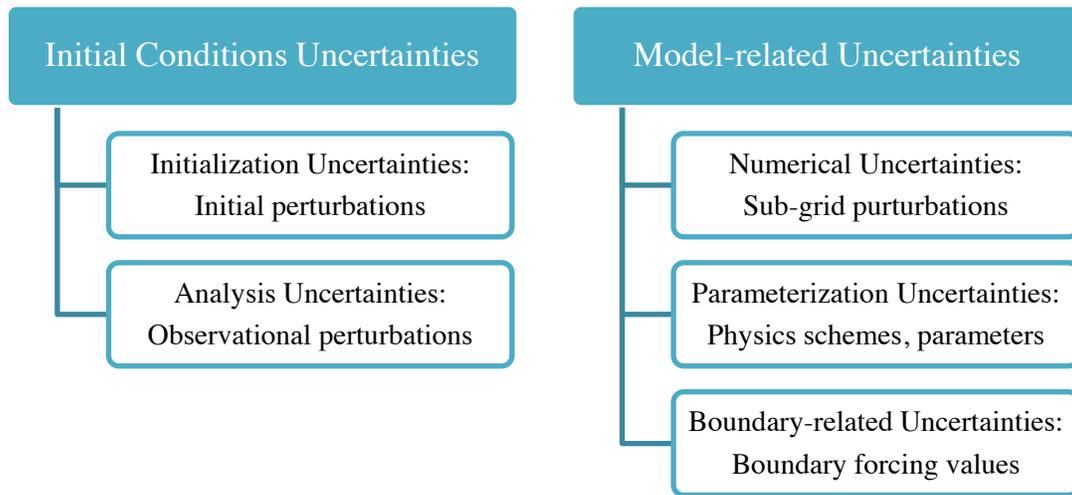


Figure 5-3 Initial conditions and model-related uncertainties in EF

Figure 5-4 shows the essential components of an ensemble which are a control forecast started from the analysis, forecasts started from two perturbations to the analysis (in this case equal and opposite), the ensemble average, and the “truth”, or forecast verification, which may become available later.

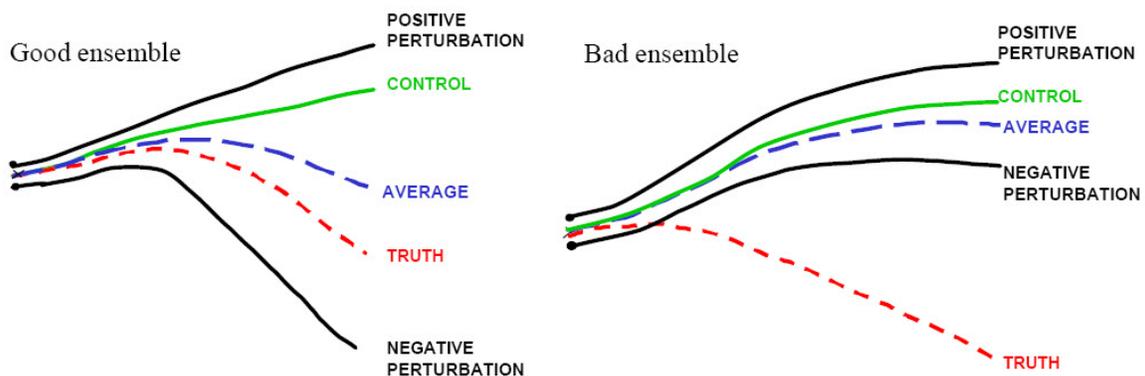


Figure 5-4 Schematic of the essential components of an ensemble of forecasts [87]

The first schematic (on left) shows an example of a “good ensemble” in which “truth” looks like a member of the ensemble. In this case, the ensemble average is closer to the truth, due to nonlinear filtering of errors, and the ensemble spread is related to the forecast error. The second schematic (on right) is an example of a “bad ensemble”. Due to poor initial perturbations and/or model deficiencies, the forecasts are not able to track the verifying truth, and remain relatively close to each other. In this case the ensemble is not helpful, since the lack of ensemble spread would give unjustified confidence in the erroneous forecast [87].

Nevertheless, for NWP development, the “bad” ensemble is still very useful. After the verification time arrives, it clearly indicates the presence of a deficiency in the forecasting system. A single “deterministic” forecast, in contrast, would not be able to distinguish between a deficiency in the system and errors in the initial conditions as the reason of failure. Ideally, the initial perturbations should sample well the analysis “errors of the day” and the spread among the ensemble members should be similar to that of the forecast errors. The two essential problems in the design of an ensemble forecasting system are how to create effective initial perturbations, and how to handle model deficiencies, which make the ensemble forecast spread smaller than the forecast error.

5.2 ENSEMBLE FORECASTING METHODS

Ensemble Forecasting (EF) methods differ mostly in the way the initial perturbation are generated, and may be essentially classified into two classes:

1. The initial perturbations are random: chosen to be realistic, i.e. they have
 - a. Horizontal and vertical structures similar to the forecast error, and
 - b. Amplitudes, compatible with the estimated analysis uncertainty.
2. The initial perturbations depend on the dynamics of the flow, and they include “errors of the day”, i.e. they are characterized by including the growing errors that depend on the evolving underlying atmospheric flow.

5.2.1 Early Methods

In 1969, Epstein [88] introduced the idea of Stochastic-Dynamic forecasting (SDF) to explicitly acknowledge the uncertainty of atmospheric model predictions. He pointed out that it could be also used in the analysis cycle to provide the forecast error covariance. In 1974, Leith [89] proposed the idea of performing EF with a limited number m of ensemble members instead of the conventional DF. He suggested using directly a Monte Carlo Forecasting (MCF), where random perturbations sampling the estimated analysis error covariance are added to the initial conditions. In 1983, Hoffman and Kalnay suggested, as an alternative to MCF, the Lagged Averaged Forecasting (LAF) method, in which forecasts from earlier analyses were included in the ensemble.

5.2.1.1 Stochastic-Dynamic Forecasting

Epstein designed SDF as a shortcut to estimate the true probability distribution of the forecast uncertainty. He derived a continuity equation for the probability density, $\varphi(X, t)$, of a model solution, X , of a dynamical model, $\dot{X} = G(X(t))$, where the model has dimension, D :

$$\frac{\partial \varphi}{\partial t} + \nabla_D \cdot (\dot{X}\varphi) = 0 \quad (5.4)$$

An ensemble starting from an infinite number of perturbed integrations spanning the analysis uncertainty gives the “true” probability distribution. Epstein introduced and tested an approximation of equation (5.4) for the Lorenz three-variable model. The “true” probability distribution was computed from a Monte Carlo ensemble of 500 members. However, since SDF involves the integration of forecast equations for each element of the covariance matrix. This method is still not computationally feasible for models with large number of DOFs. In a model with N DOFs, it requires $N(N + 1)/2 + N$ forecast equations, equivalent to making $(N + 3)/2$ model forecasts.

5.2.1.2 Monte Carlo Forecasting

Leith suggested using directly a Monte Carlo Forecasting (MCF), where random perturbations sampling the estimated analysis error covariance are added to the initial conditions as shown in Figure 5-5. He cast his analysis using, the deviation u of the model variables with respect to climatology (also known as forecast *anomalies*). He noted that in an infinitely large ensemble, the average forecast error variance at long time leads converges to the climatological error variance U , whereas the error variance of individual forecasts converges to twice the climatological error variance.

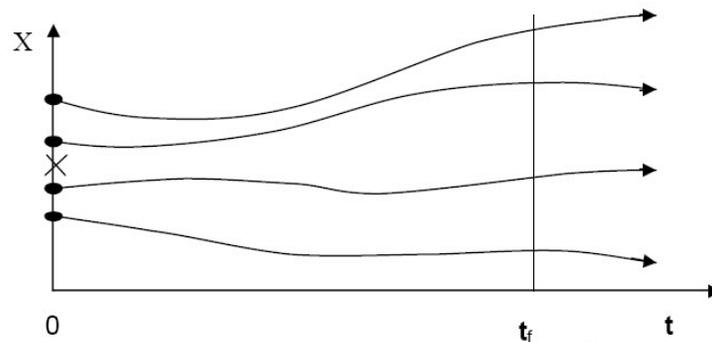


Figure 5-5 Schematic time evolution of MCF [87]

The true state of the atmosphere is denoted u_0 , and \hat{u} then denotes an unbiased estimate of u_0 , whose expected value is equal to zero: $\langle \hat{u} \rangle = 0$. If $\bar{u} = (1/m) \sum_{i=1}^m u_i$ is the average of an ensemble of m forecasts, then its error covariance evolves following this equation:

$$\langle (\bar{u} - u_0) \langle \bar{u} - u_0 \rangle^Y \rangle = \langle \bar{u} \bar{u}^T + u_0 u_0^T + \bar{u} u_0^T + u_0 \bar{u}^T \rangle \xrightarrow{t \rightarrow \infty} \left(1 + \frac{1}{m} \right) U \quad (5.5)$$

Leith suggested that adequate accuracy would be obtained for the best estimate of the forecast, i.e. the *ensemble mean*, with sample sizes as small as 8, but that the estimation of *forecast errors* may require a large number of ensemble members. MCF is thus a feasible approach for EF, requiring only a definition of the initial perturbations and m forecasts.

5.2.1.3 Lagged Average Forecasting

The forecasts initialized at the current initial time, $t = 0$, as well as at previous times, $t = -\tau, -2\tau, \dots, -(N - 1)\tau$, are combined to form an ensemble as shown in Figure 5-6. The time window τ is typically 6, 12, or 24 hours, so that the forecasts are already available and the perturbations are generated automatically from the forecast errors. So, they include “errors of the day”. Since the ensemble members are forecasts of different “ages” they should be averaged with weights estimated from their average forecast errors. Hoffman and Kalnay found that compared to MCF, LAF resulted in a better prediction skill, i.e. a stronger relationship between ensemble spread and error.

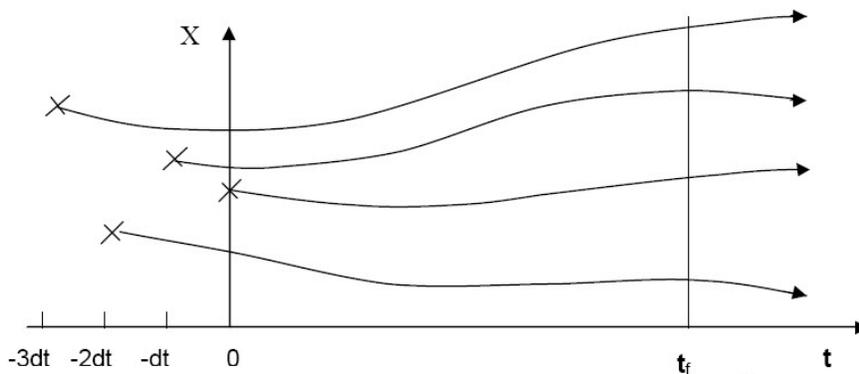


Figure 5-6 Schematic time evolution of LAF [87]

The main disadvantage of LAF is that the “older” forecasts are less accurate, was corrected by the Scaled LAF (SLAF) approach of Ebisuzaki and Kalnay [90], in which the LAF perturbations (difference between the forecast and the current analysis) are scaled by their “age”, so that all the SLAF perturbations have errors of similar magnitude. They also suggested that the scaled perturbations should be both added and subtracted from the analysis, thus increasing the ensemble size and the probability of “encompassing” the true solution within the ensemble. SLAF can be easily implemented in both global and regional models, including the impact of perturbed boundary conditions [91].

5.2.2 Operational Methods

The first two optimal perturbations methods implemented operationally are known as breeding and singular vectors which include errors of the day like LAF. Other methods that are also very promising are based on ensembles of data assimilations, and ensembles based on operational systems from different centers, combining different models and data assimilation systems. These and other methods that have become operational or are under consideration in operational centers are briefly discussed in Appendix E.

5.2.3 Ensemble Kalman Filtering

In the case of linear dynamics, the mathematically consistent technique to define a background error covariance matrix is the Kalman filter [92, 93], which utilizes the dynamical equations to evolve the most probable state and the error covariance matrix in time. In the case of linear systems with unbiased normally distributed errors, the Kalman filter provides estimates of the system state that are optimal in the mean square sense. The method has also been adapted to nonlinear systems, but, in this case, optimality no longer applies [94].

There are two basic approaches to Ensemble Kalman Filtering (EnKF). In the first one, known as “perturbed observations”, an ensemble of data assimilations is carried out using the same observations to which random have been added. The ensemble is used to estimate the forecast error covariance needed in the Kalman Filter [95, 96, 97]. This approach has been shown to be very competitive with the operational 3DVar, an important milestone, given that 3DVar has the benefit of years of improved developments [98].

The second approach is the class of square root filters [99], and does not require perturbing the observations. Several groups have recently independently developed square root filters [99, 101].

5.3 ENSEMBLE FORECASTING EXPERIMENT

Ensemble Forecasting (EF) entails the production of many forecasts to reflect the uncertainty in the initial conditions. The number of initial conditions in an ensemble will always be small compared to the infinite number of possible atmospheric states. Moreover, the model deficiencies, due to the assumptions in the governing equations and/or the physical parameterizations have to be taken into consideration. Constructing an ensemble of forecasts is an important issue where the accuracy of the produced forecast (ensemble mean) is proportional to the number and the quality of the ensemble members, subject to computer limitations and time constraints.

In this study, the sensitivity of the MM5 model to different physics options was investigated in the four seasons. This may help in choosing the most appropriate physics package to be used for a single control forecast. Nine different physics options were used in this study based on three PBL schemes and three radiation schemes. The MRF PBL scheme with the three different formulations for the thermal roughness length was used. The first formulation, IZOTOPT=0, is the original Carlson-Boland formulation. The second formulation, IZOTOPT=1, is the Garratt formulation. The third formulation, IZOTOPT=2, is the Zilitinkevich formulation. The Cloud-Radiation, CCM2-Radiation and RRTM-Longwave radiation schemes were used. The other physical parameterizations utilized the Grell scheme, mixed-phase scheme and NOAH LSM. The different physics options were tested in the four seasons, as shown in Table 5-2.

Table 5-2 Selected time periods for testing model sensitivity to different physics options

Period	Starting Date and Time				Ending Date and Time			
	Year	Month	Day	Time	Year	Month	Day	Time
1	2006	1	1	Z00	2006	1	4	Z00
2	2006	4	1	Z00	2006	4	4	Z00
3	2006	7	1	Z00	2006	7	4	Z00
4	2006	10	1	Z00	2006	10	4	Z00

Table 5-3 Ensemble members used in the EF experiment for Egypt
(The member number 1 represents the configurations used as a single control forecast)

Member Number	Physics Options		Data Assimilation	
	MRF PBL	Radiation	OA	FDDA
1	IZ0TOPT = 0	Cloud-Radiation	Off*	Off
2	IZ0TOPT = 1	Cloud-Radiation	Off*	Off
3	IZ0TOPT = 2	Cloud-Radiation	Off*	Off
4	IZ0TOPT = 0	CCM2-Radiation	Off*	Off
5	IZ0TOPT = 1	CCM2-Radiation	Off*	Off
6	IZ0TOPT = 2	CCM2-Radiation	Off*	Off
7	IZ0TOPT = 0	RRTM-Longwave	Off*	Off
8	IZ0TOPT = 1	RRTM-Longwave	Off*	Off
9	IZ0TOPT = 2	RRTM-Longwave	Off*	Off
10	IZ0TOPT = 0	Cloud-Radiation	On	Off
11	IZ0TOPT = 1	Cloud-Radiation	On	Off
12	IZ0TOPT = 2	Cloud-Radiation	On	Off
13	IZ0TOPT = 0	CCM2-Radiation	On	Off
14	IZ0TOPT = 1	CCM2-Radiation	On	Off
15	IZ0TOPT = 2	CCM2-Radiation	On	Off
16	IZ0TOPT = 0	RRTM-Longwave	On	Off
17	IZ0TOPT = 1	RRTM-Longwave	On	Off
18	IZ0TOPT = 2	RRTM-Longwave	On	Off
19	IZ0TOPT = 0	Cloud-Radiation	On	On, -06 hr
20	IZ0TOPT = 1	Cloud-Radiation	On	On, -06 hr
21	IZ0TOPT = 2	Cloud-Radiation	On	On, -06 hr
22	IZ0TOPT = 0	CCM2-Radiation	On	On, -06 hr
23	IZ0TOPT = 1	CCM2-Radiation	On	On, -06 hr
24	IZ0TOPT = 2	CCM2-Radiation	On	On, -06 hr
25	IZ0TOPT = 0	RRTM-Longwave	On	On, -06 hr
26	IZ0TOPT = 1	RRTM-Longwave	On	On, -06 hr
27	IZ0TOPT = 2	RRTM-Longwave	On	On, -06 hr
28	IZ0TOPT = 0	Cloud-Radiation	On	On, -12 hr
29	IZ0TOPT = 1	Cloud-Radiation	On	On, -12 hr
30	IZ0TOPT = 2	Cloud-Radiation	On	On, -12 hr
31	IZ0TOPT = 0	CCM2-Radiation	On	On, -12 hr
32	IZ0TOPT = 1	CCM2-Radiation	On	On, -12 hr
33	IZ0TOPT = 2	CCM2-Radiation	On	On, -12 hr
34	IZ0TOPT = 0	RRTM-Longwave	On	On, -12 hr
35	IZ0TOPT = 1	RRTM-Longwave	On	On, -12 hr
36	IZ0TOPT = 2	RRTM-Longwave	On	On, -12 hr

* When OA is off, FDDA is also off. This represents the default MM5 runs without nudging.

Then, a preliminary Short-Range EF (SREF) experiment for Egypt based on the MM5 model was built, tested and evaluated in two seasons: winter, 2006 and summer, 2008. Nine forecasts were used as reference forecasts and the other 27 forecasts utilized data assimilation. The available real-time conventional and satellite observations were implemented using OA and FDDA DA methods. They were used to provide 3 different sets of initial conditions. The first set was produced using OA of the first-guess (interpolated from FNL datasets) and the available observations. The second and third sets were produced using OA and FDDA dynamic initialization for 6 and 12 hours, respectively.

The total number of ensemble members was 36 forecast. Table 5-3 shows the configurations used in constructing the ensemble members. The ensemble size may be further increased including different configurations using the WRF model. Moreover, the initial perturbations may be produced using Bred Vectors (BVs) and/or Singular Vectors (SVs) as described in Appendix E. The ensemble mean was evaluated based on the RMSE and total/daily percentage difference in the maximum, minimum and average values of the near-surface temperature at WMO stations shown in Figure 2-18. The comparisons with reference analyses were utilized using OA from FNL datasets (or MM5 outputs from coarser domains) and the available observations.

The results of sensitivity of the model to different physics options and the ensemble forecasting experiment are presented in sections 6.3 and 6.4.

CHAPTER SIX

RESULTS AND DISCUSSIONS

This chapter shows the results of the current study. It includes the enhancement and evaluation of using FDDA in NWP for Egypt, the MM5 model sensitivity to different physics options and/or initial conditions, and a preliminary SREF experiment for Egypt which may be further developed for operational use. Most of the results are carried out in the four seasons to evaluate the modeling system at different locations and weather regimes. The amount of results from the performed simulations is huge and it is not practical to present all of them. The detailed results are presented in Appendix F to focus on the main conclusions.

6.1 USING FDDA IN NWP FOR EGYPT

The sensitivity of the MM5 model to different FDDA techniques was tested and evaluated in the four seasons during the period from July 1, 2005 to July 1, 2006 to provide accurate meteorological simulations with 9 km resolution for Egypt. Four sets of simulations were performed, as shown in Table 6-1. The first case is the reference case, FDDA0, where no FDDA is implemented. The second case, FDDA1, utilized grid FDDA from OA of NCEP Final Analysis (FNL). The third case, FDDA2, utilized observational FDDA from ATOVS observations. The fourth case, FDDA3, utilized combined grid and observational FDDA. These simulations represent dynamic analysis only.

Table 6-1 Description of the performed simulations for evaluating FDDA benefits

#	Run Name	Description
1	FDDA0	Reference Run, No FDDA was implemented
2	FDDA1	Grid FDDA was utilized from OA of FNL and observational data
3	FDDA2	Observational FDDA was utilized from ATOVS observations
4	FDDA3	Combined grid and observational FDDA was utilized

6.1.1 Comparison of Different FDDA Methods

The performance of the NWP model (MM5) is sensitive to the technique of data assimilation used and the quality, quantity and distribution of the implemented observations. The four sets of data assimilation techniques used in this study were chosen to test the sensitivity and to find the most appropriate and time efficient configuration. Computational time represents the cost of using certain FDDA technique while the reduction in RMSE/MABE errors represents the gain. In this study, the cost versus gain of different FDDA techniques, as described in Table 6-1, the errors at different atmospheric layers and the scatter plots of all performed simulations are investigated.

6.1.1.1 Cost versus Gain Statistics

The Computational time is inversely proportional to the computational power. This study was done using a cluster of two PCs and the percentage increase in the computational time is important.

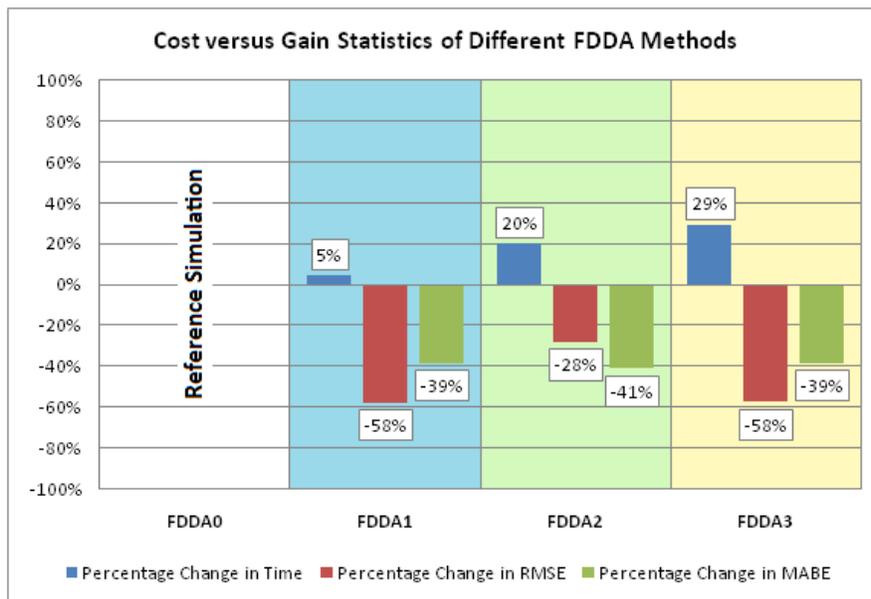


Figure 6-1 Cost versus gain statistics of different FDDA techniques

To quantify the gains obtained from FDDA, the RMSE and MABE were calculated for each location. Figure 6-1 shows the computational cost in time

versus computational gains in RMSE and MABE for each simulation. Using FDDA1 in the model increases the computational time by about 5% while it reduces the RMSE and MABE by about 58% and 39% respectively. Using FDDA2 in the model increases the computational time by about 20% while it reduces the RMSE and MABE by about 28% and 41% respectively. Using FDDA3 in the model increases the computational time by about 29% while it reduces the RMSE and MABE by about 58% and 39% respectively.

It is clear that the combined grid and observational FDDA, FDDA3, as used in this study is not time efficient and reproduces temperatures similar to those of grid FDDA, FDDA1. This may be explained by noting that the first guess is very dense compared to the ATOVS data. The large cost of observational FDDA, FDDA2, may be due to the high frequency at which the model checks for the validity of the implemented observations. This effect may be reduced by tuning the observational FDDA options and preparing the observations at certain specified time periods.

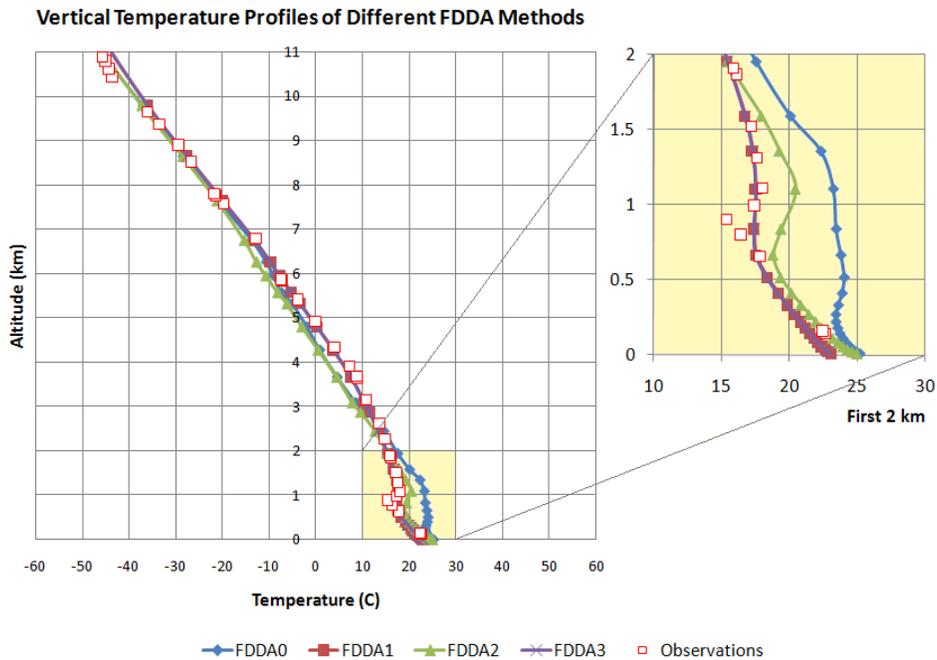


Figure 6-2 Vertical temperature profiles for different FDDA techniques

Figure 6-2 compares the simulated and observed vertical temperature profiles for different FDDA techniques at Helwan station in Z24 on October 3, 2005. From the result it is apparent that using FDDA has improved the model accuracy and the computed temperatures correlate better with the observed temperatures, especially near the surface. Such an enhancement is necessary to accurately compute the thermal inversion layers during air pollution episodes and to compute the particle transport near the surface.

6.1.1.2 Errors at Different Atmospheric Layers

Table 6-2 lists the RMSE and MABE and the associated gains at the total, troposphere, first 2 km, and first 1 km atmospheric layers. Figure 6-3 shows the RMSE and MABE of temperature at different atmospheric layers. It is clear that there are great enhancements near the surface.

Table 6-2 RMSE and MABE and associated gains at different atmospheric layers

Simulation	Total		Troposphere		First 2 km		First 1 km		
	RMSE C	MABE C	RMSE C	MABE C	RMSE C	MABE C	RMSE C	MABE C	
Z24 3/10/2005	FDDA0	2.963	7.435	3.172	7.435	4.054	7.435	4.083	7.435
	FDDA1	1.242	4.561	0.772	1.740	0.812	1.376	0.948	1.376
	Gain	58.10 %	38.65 %	75.65 %	76.60 %	79.98 %	81.49 %	76.77 %	81.49 %
	FDDA2	2.119	4.360	2.208	4.360	1.325	3.372	1.191	3.372
	Gain	28.10 %	41.36 %	30.38 %	41.36 %	67.32 %	54.65 %	67.32 %	54.65 %
	FDDA3	1.251	4.561	0.791	1.734	0.848	1.379	0.995	1.379
	Gain	57.78 %	38.65 %	75.06 %	76.68 %	79.08 %	81.45 %	79.08 %	81.45 %

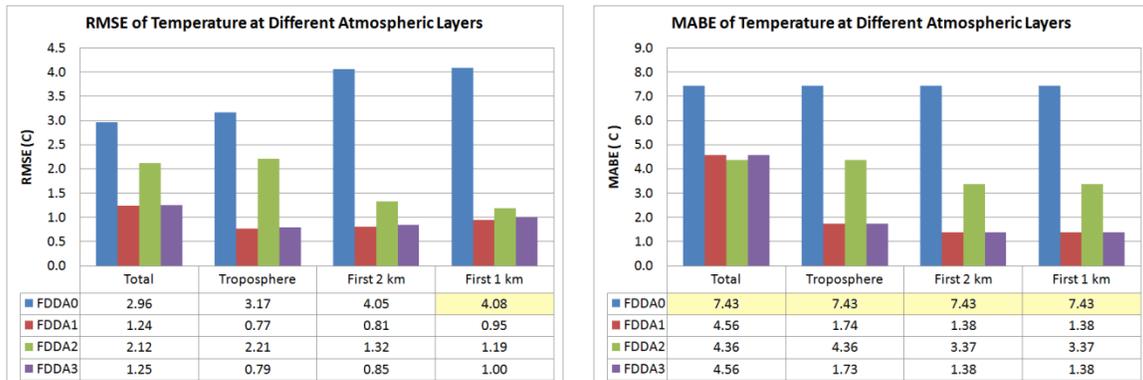


Figure 6-3 RMSE and MABE of Temperature at different atmospheric layers

6.1.1.3 Scatter Plots for Different FDDA methods

The scatter plots of the simulated and observed near-surface temperature at all stations in the four seasons are shown in Figure 6-4. It shows the great enhancements of the performance of the model when FDDA is used.

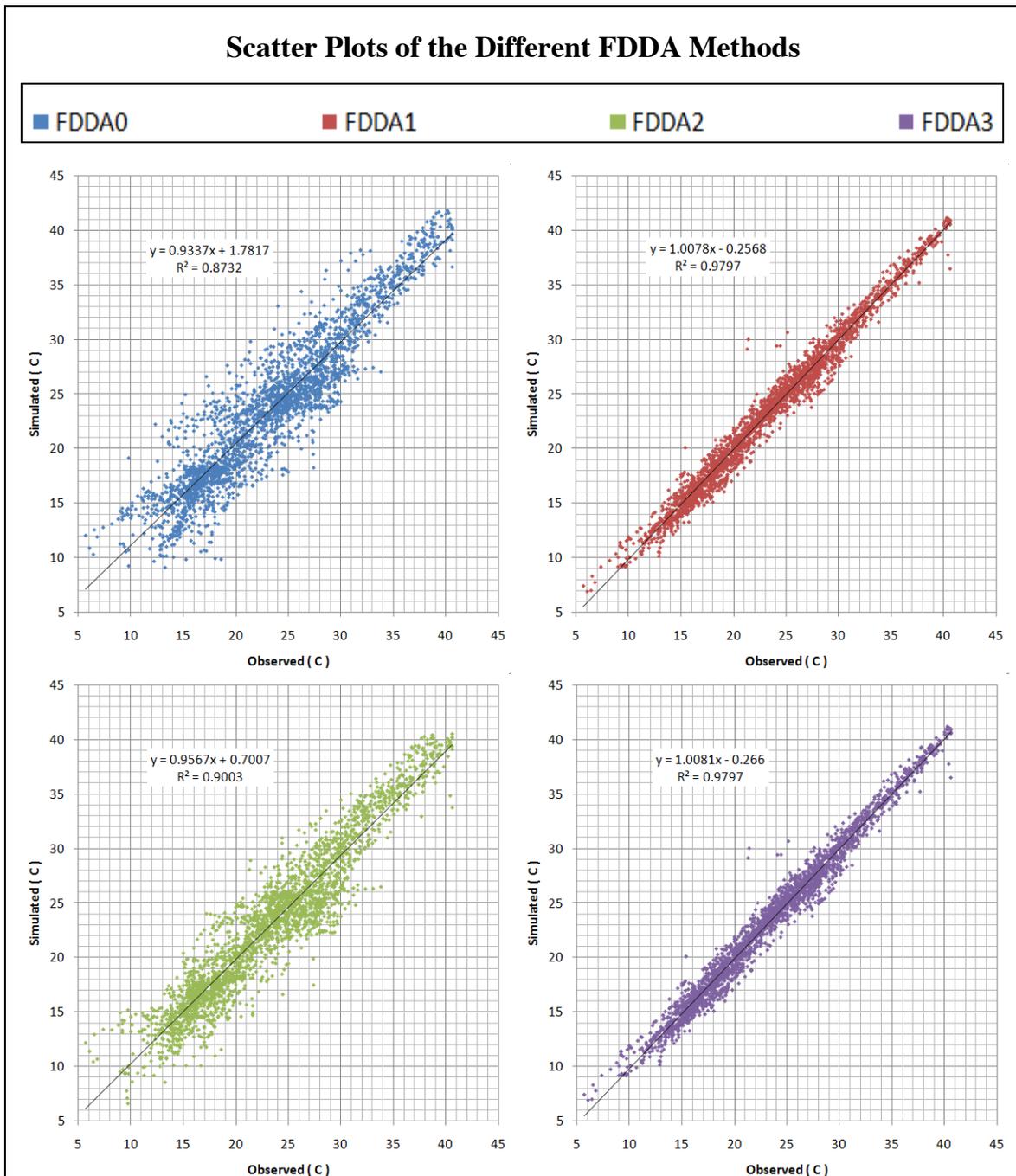


Figure 6-4 Scatter plots of near-surface temperature for different FDDA techniques

Since the results of FDDA1 and FDDA3 are almost identical, FDDA1 was used for the evaluation of the model for near-surface temperatures. Comparisons were performed using either the 2-m temperature or the temperature at the first pressure level near the surface which is approximately at 8 m above the ground.

6.1.2 FDDA Results in the Four Seasons

Since the results of FDDA1 and FDDA3 are identical, FDDA1 was used for the evaluation of the model performance for near surface temperatures. The comparisons were performed using either the 2-m temperature or the temperature at the first pressure level near the surface which is approximately at 8 m above the ground since it was less computationally demanding. Figure 6-5 shows RMSE and MABE of near-surface temperature for all seasons at Alexandria/Nouzha station. The results show that the RMSE has reduced to the order of 1.5 degrees Celsius, during the 72 hours.

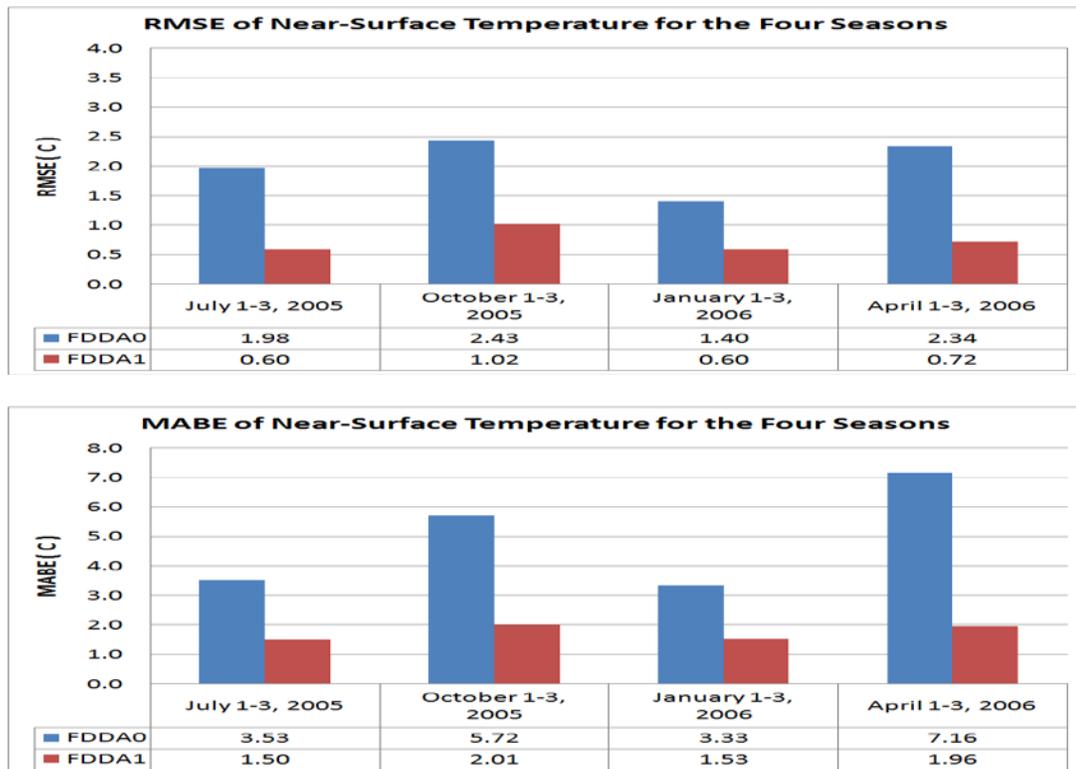


Figure 6-5 RMSE and MABE of near-surface temperature for all seasons in Alexandria

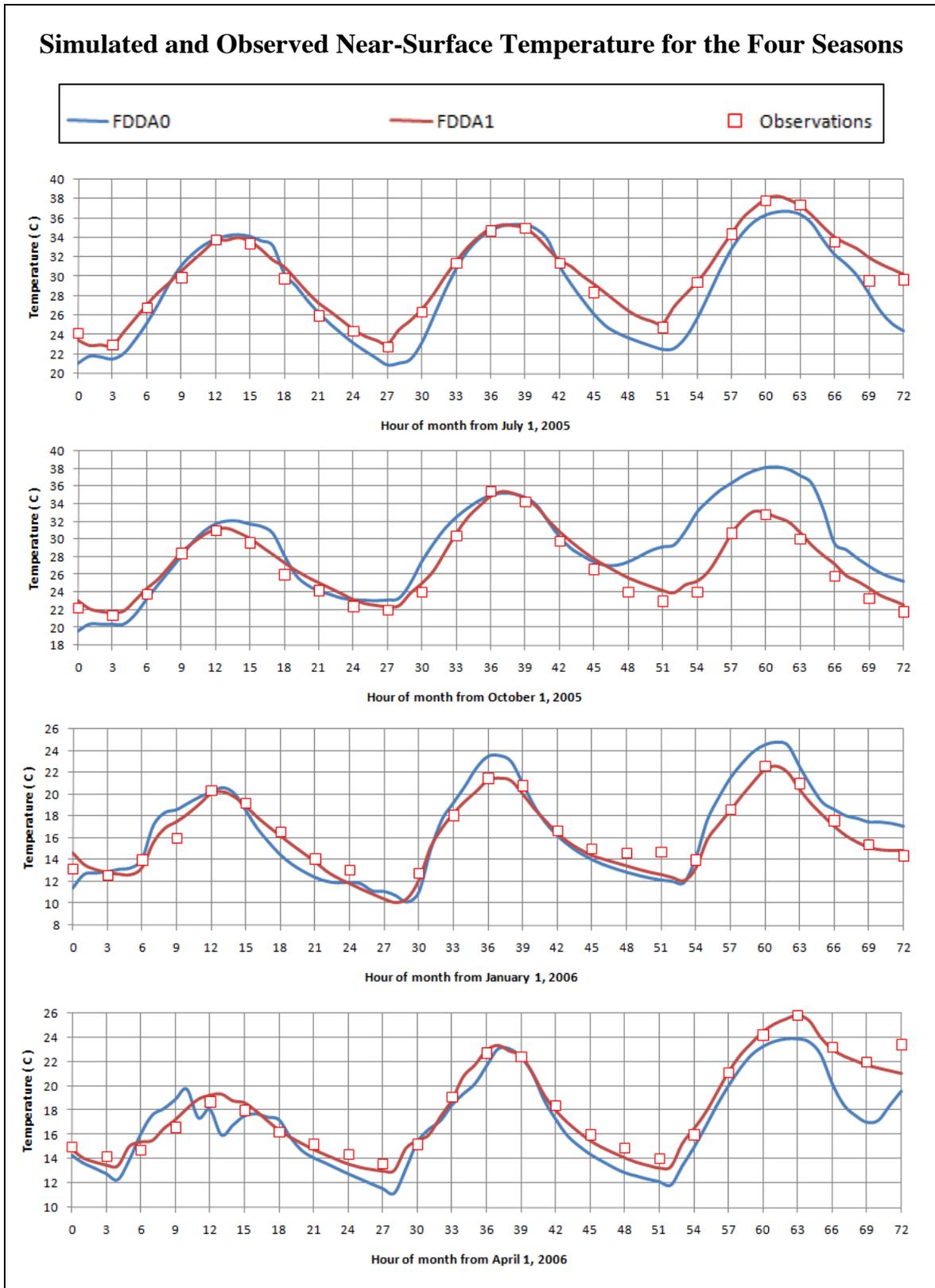


Figure 6-6 Near-surface temperatures for the four seasons in Alexandria

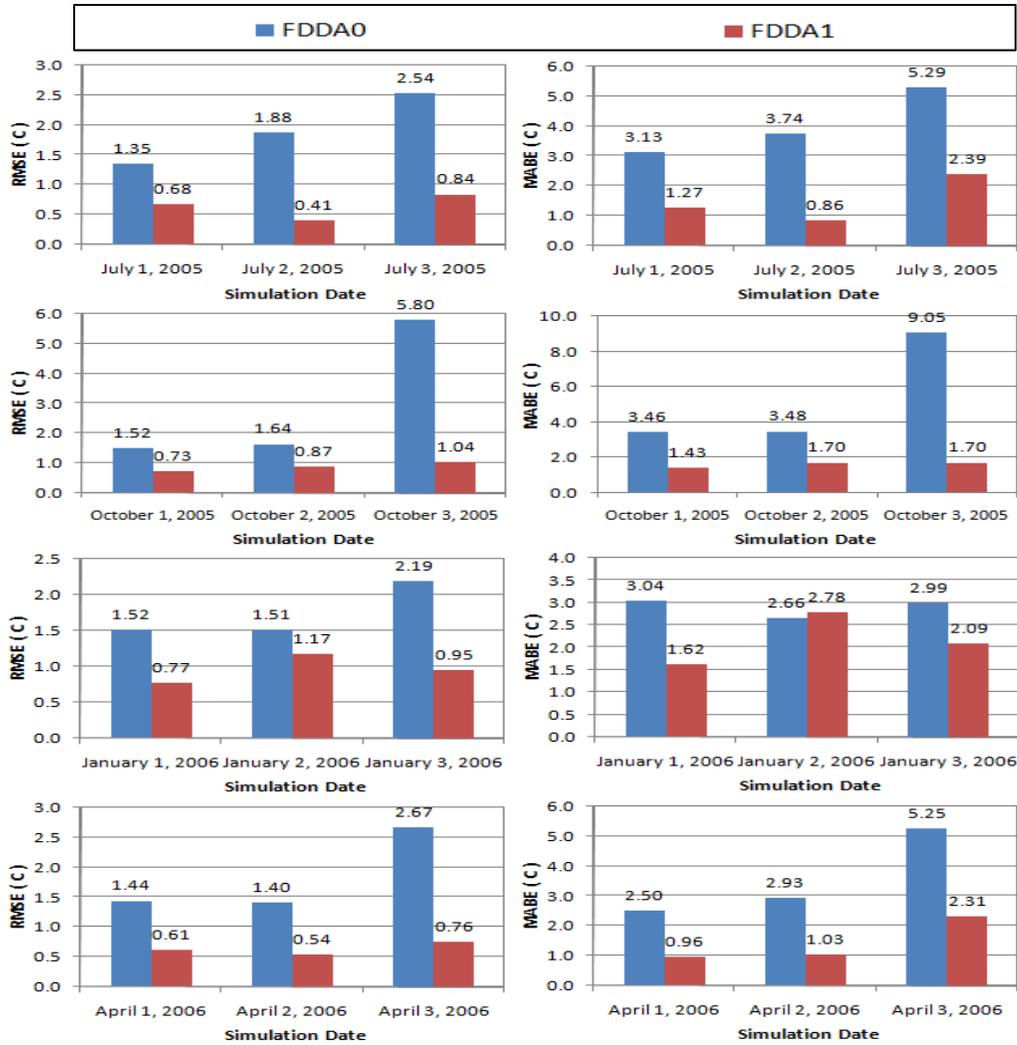


Figure 6-7 Daily RMSE and MABE of near-surface temperature in Alexandria

Figure 6-6 shows the simulated and observed near-surface temperature at Alexandria/Nouzha station for twelve days, three days from each season, while Figure 6-7 shows the daily errors. The large differences in error that occur without data assimilation have greatly reduced by an order of 4-6 degrees Celsius). This reduction in error is quite large especially for the last 24 hours in the simulation, where the error increases.

6.1.3 Summary of the FDDA Benefits

A statistical summary of the MM5 performance in all seasons at all stations is listed in Table 6-3. Large errors in Atmospheric Boundary Layer (ABL) simulations can arise due to inaccurate surface parameters as well as simplification

in the boundary layer formulations and other model deficiencies [102, 103]. These errors are generated at stations near coast lines and locations of complex topography especially at night hours with the current resolution, 9 km. As shown in Table 6-3, most gains are greater than 50% and the maximum gains are located at south part of Egypt such as in Asyut, Luxor and Aswan stations in winter.

Many questions may arise from these results: why the maximum gains, or maximum errors in the reference simulations, are located at south of Egypt and what is the error distribution inside Egypt and how these errors can be eliminated? Many attempts were carried out to reduce the errors by using different configurations of the NWP model or by using different initial/boundary conditions and then moving towards the ensemble forecasting.

Figure 6-8 shows the distribution of temperature difference between FDDA1 and FDDA0 simulations after 52 hours from starting in winter. It is clear that FDDA1 produces higher values of temperature near surface around the Red Sea and produces lower values at the south east of Egypt. This may be due to the selection of physics options that are not suitable for these areas or due to the surface boundary conditions (e.g., land-use).

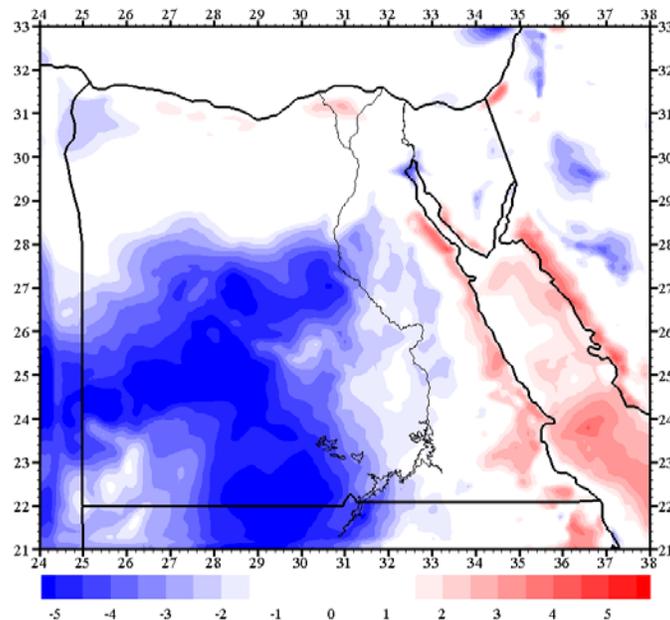


Figure 6-8 Temperature difference between FDDA1 and FDDA0 simulations

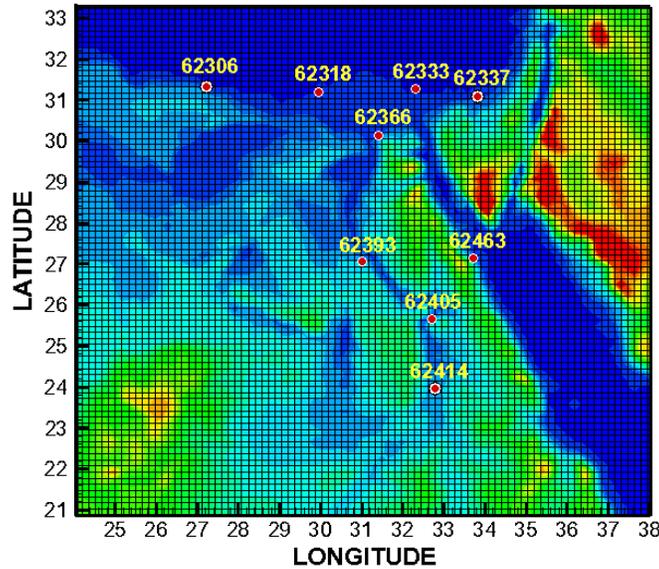


Figure 6-9 Stations map for the statistical summary of the FDDA benefits

Table 6-3 Statistical summary of the FDDA benefits in the four seasons at all stations

Simulation	July 1-3, 2005		October 1-3, 2005		January 1-3, 2006		April 1-3, 2006		
	RMSE C	MABE C	RMSE C	MABE C	RMSE C	MABE C	RMSE C	MABE C	
62306	FDDA0	1.442	3.459	1.823	5.074	1.948	4.561	2.399	9.160
	FDDA1	1.129	2.789	0.594	1.373	0.822	1.644	0.910	3.626
	Gain	21.72 %	19.37 %	67.44 %	72.94 %	57.81 %	63.96 %	62.09 %	60.41 %
62318	FDDA0	1.976	3.529	2.434	5.716	1.403	3.330	2.340	7.156
	FDDA1	0.596	1.503	1.016	2.008	0.596	1.534	0.723	1.964
	Gain	69.85 %	57.41 %	58.26 %	64.88 %	57.52 %	53.93 %	69.11 %	72.55 %
62325	FDDA0	3.035	5.702	1.840	4.034	1.476	2.871	2.169	5.020
	FDDA1	1.664	3.184	1.020	2.752	0.808	2.055	1.611	3.067
	Gain	45.16 %	44.16 %	44.57 %	31.78 %	45.21 %	28.42 %	25.73 %	38.90 %
62337	FDDA0	2.963	6.767	2.991	6.734	2.979	6.332	1.153	2.638
	FDDA1	1.171	2.374	1.084	1.906	0.992	2.181	0.770	1.742
	Gain	60.48 %	64.92 %	63.77 %	71.70 %	66.68 %	65.56 %	33.21 %	33.97 %
62366	FDDA0	1.977	5.293	3.612	9.053	1.777	3.040	1.931	5.249
	FDDA1	0.675	2.393	0.875	1.703	0.972	2.777	0.639	2.313
	Gain	65.87 %	54.79 %	75.77 %	81.19 %	45.28 %	8.64 %	66.92 %	55.93 %
62393	FDDA0	1.487	4.023	2.202	4.477	5.178	9.885	1.619	4.172
	FDDA1	0.616	1.505	0.567	1.531	1.041	2.139	0.778	1.722
	Gain	58.56 %	62.58 %	74.23 %	65.80 %	79.90 %	78.36 %	51.94 %	58.73 %
62405	FDDA0	1.772	5.615	1.583	4.268	5.243	8.767	1.727	5.617
	FDDA1	0.702	1.881	0.695	2.466	0.767	2.569	0.916	2.476
	Gain	60.41 %	66.50 %	56.10 %	42.23 %	85.38 %	70.70 %	46.94 %	55.92 %
62414	FDDA0	2.180	5.864	2.040	6.245	3.366	5.830	3.299	8.754
	FDDA1	1.018	4.101	0.905	1.883	0.665	1.888	1.448	2.965
	Gain	53.30 %	30.06 %	55.64 %	69.84 %	80.24 %	67.62 %	56.09 %	66.13 %
62463	FDDA0	1.500	4.122	1.532	3.500	2.058	4.648	1.985	4.356
	FDDA1	1.057	2.489	0.817	2.117	1.078	4.683	1.192	2.430
	Gain	29.54 %	39.62 %	46.67 %	39.51 %	47.59 %	-0.75 %	39.97 %	44.21 %

Legend < 25 % 25-50 % 50-75 % > 75 %

Table 6-3 summarizes the FDDA benefits and different errors at different locations inside Egypt while Figure 6-9 shows the stations map. It is clear that the maximum gains (>75%) are located at the south part of Egypt in winter. Therefore, the model has to be evaluated at locations different than the sites used for observations. The next section will show the evaluation of FDDA on NWP for Egypt using different observation sets.

6.2 EVALUATION OF FDDA ON NWP FOR EGYPT

To test the sensitivity of the model to different nudging options and to find the best simulation set for different weather regimes, five sets of simulations were performed in winter and summer as described in Table 6-4.

Table 6-4 description of the performed simulations

Simulation	Description
Ref	Reference simulation, no nudging was utilized for all domains
No D3	Nudging was utilized for the coarser domains only with Multiquadric analysis
No PBL	Nudging was utilized for all domains not in the boundary layer with Multiquadric analysis
DA MQ	Nudging was utilized for all domains with Multiquadric objective analysis
DA CM	Nudging was utilized for all domains with Cressman analysis

Four sets of observational stations (totaling 24) were selected to test the qualities of simulations [Figure 2-18];

1. Principal set: stations used in data assimilation (9).
2. Interpolation set: stations located in-between the principal stations (5).
3. Extrapolation set: stations located far from the principal stations, out of interpolation zone (6).
4. Special set: stations located at complex terrain (4).

Five sets of simulations were performed to evaluate FDDA via different nudging options on the results of simulations at locations with different observation site

density inside Egypt. The effects on accuracy at the observational-data-void areas and at the locations with complex terrain were also examined.

Sample results for the different simulations and observation sets at selected stations are shown in Figure 6-10 for summer. Cairo International Airport station represents the principal set, Minya Station represents the interpolation set, Farafra station represents the extrapolation set, and Sharm Elsheikh station represents the Special set. The results at Cairo International Airport station were the best while the results at Sharm Elsheikh station were the worst. Minya and Farafra stations had almost similar results.

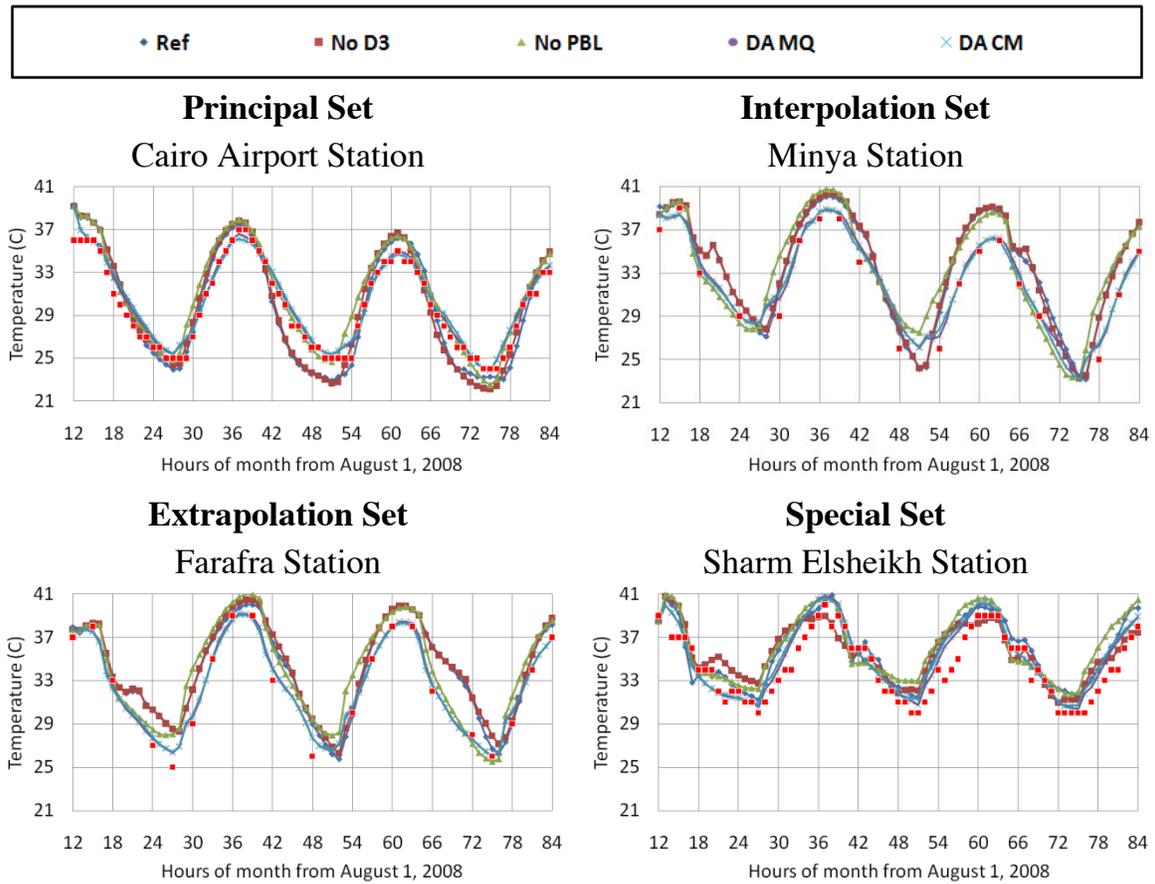


Figure 6-10 Sample results for the different simulations and observation sets in summer

Figure 6-11 shows the scatter plots of the simulated and observed near-surface temperature for the four observation sets while Figure 6-12 shows the correlations of all simulations with the observation for all observation sets.

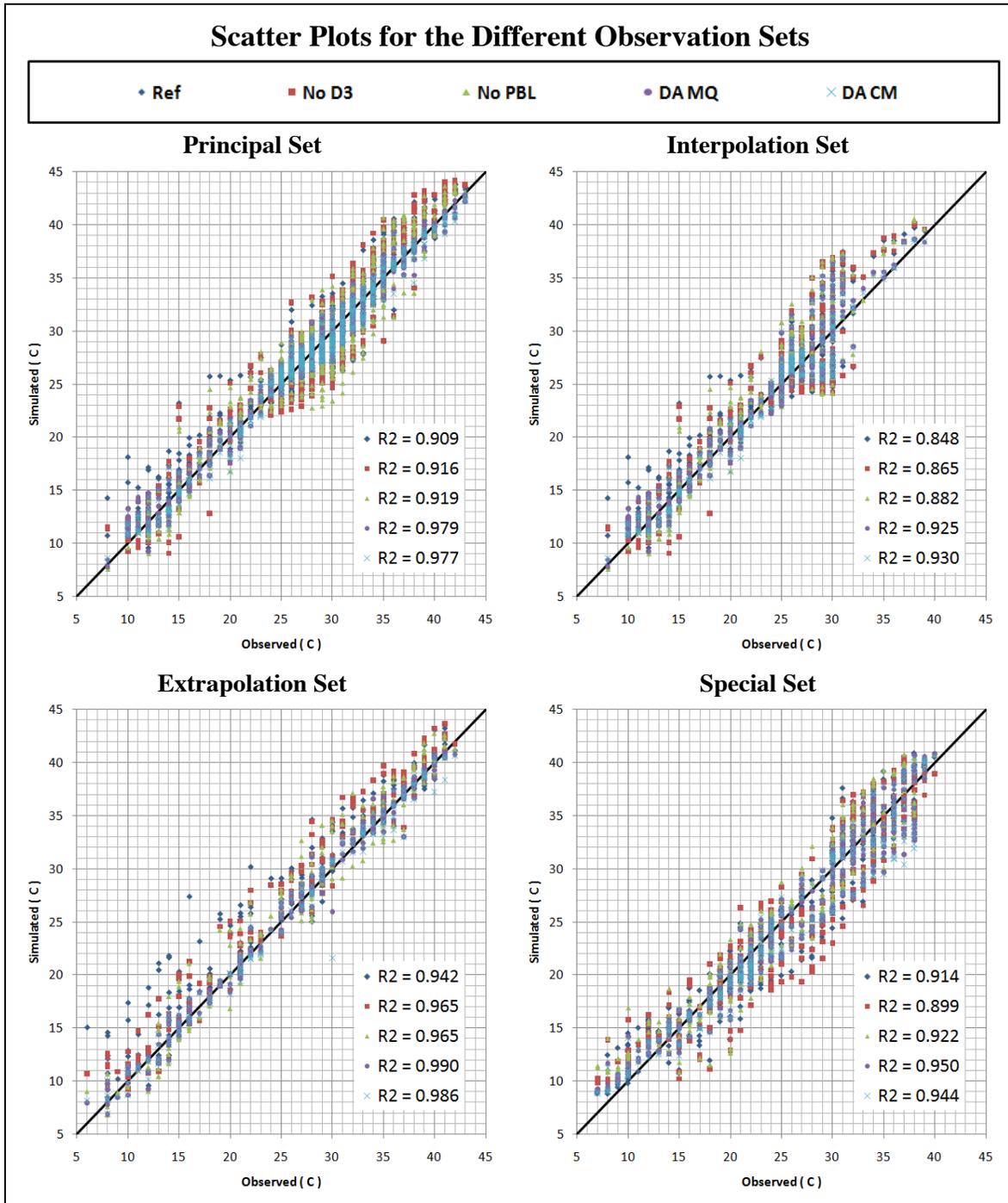


Figure 6-11 Scatter plots of near-surface temperature for the different observation sets

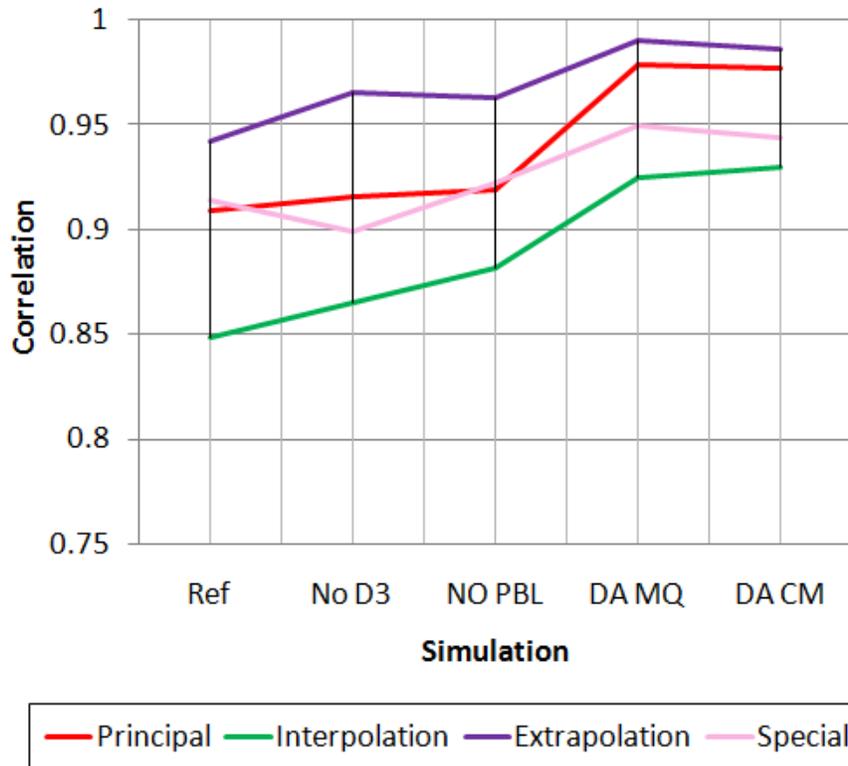


Figure 6-12 Correlations of all simulations for the four observation sets

The results for the principal set were the best. The accuracies of simulations when using nudging with Multiquadric or Cressman were better than the reference simulations for all observation sets. The worst results were associated with the special set, the winter simulations of the extrapolation set, and near the coast lines. The summer simulations of the extrapolation set were found to have relatively good results. The results of the interpolation set were good when the station is located between observation sites with almost similar surface boundary conditions (terrain and land-use). This set includes the stations near coast lines which were found to have large errors.

Figure 6-13 shows the scatter plots for the interpolation set with and without the coastal stations. It is clear that the correlations were better when the coastal stations were excluded from the interpolation set areas.

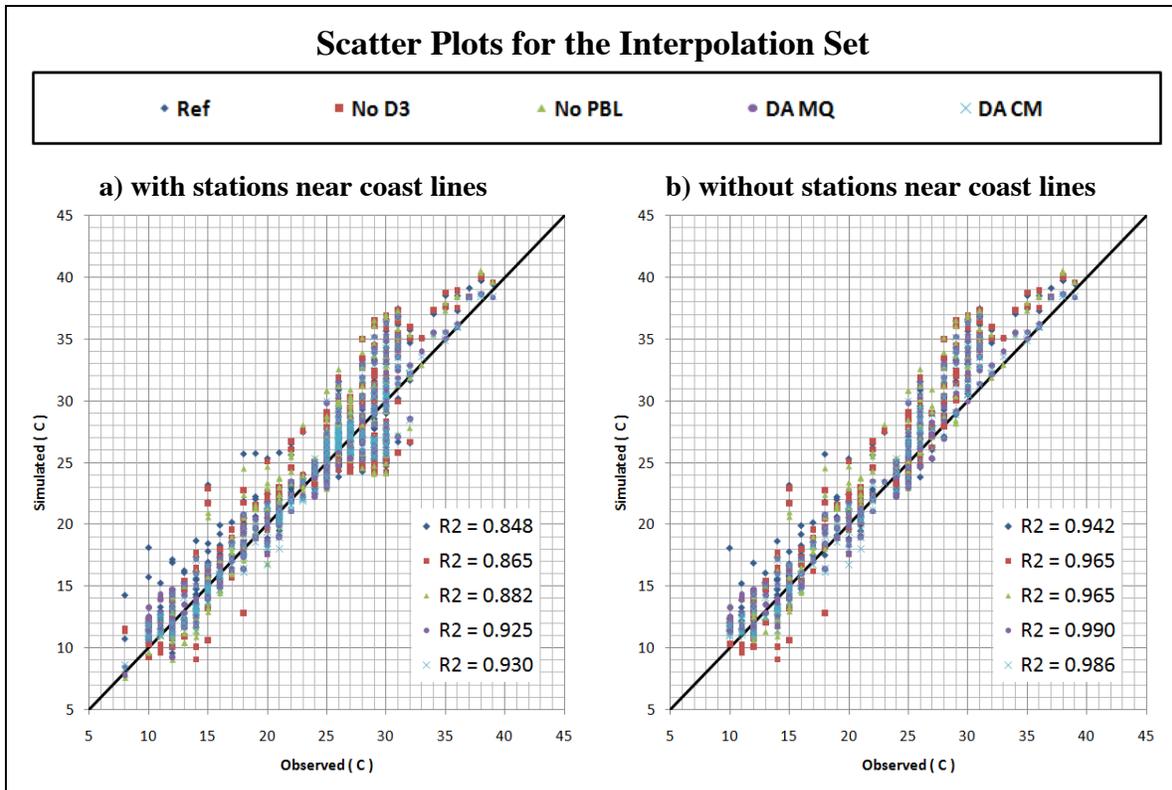


Figure 6-13 Scatter plots for the interpolation set with and without coastal stations

Figure 6-14 shows the average RMSE for all simulations at the four observational sets in summer and winter. The reference simulations at extrapolation set had the largest errors (in winter). These errors were reduced by about 56% in summer and 73% in winter when nudging with Multiquadric analysis was used. In general, the simulations had the best accuracies and correlations with the observations among all other simulations when nudging was used with multiquadric analysis.

The accuracy of simulations when using FDDA with MQ or CM is better than the reference simulations for all observation sets. When FDDA is turned off for the finest grid domain, the results are good in some locations and are bad in other locations. This may be due to the uncertainties in initial and boundary conditions produced from the coarser domains. The simulations with no nudging in the PBL are also sensitive to the location and weather regime and it may probably produce

better results when finer resolution, such as 1 km resolution, with better surface boundary conditions were used.

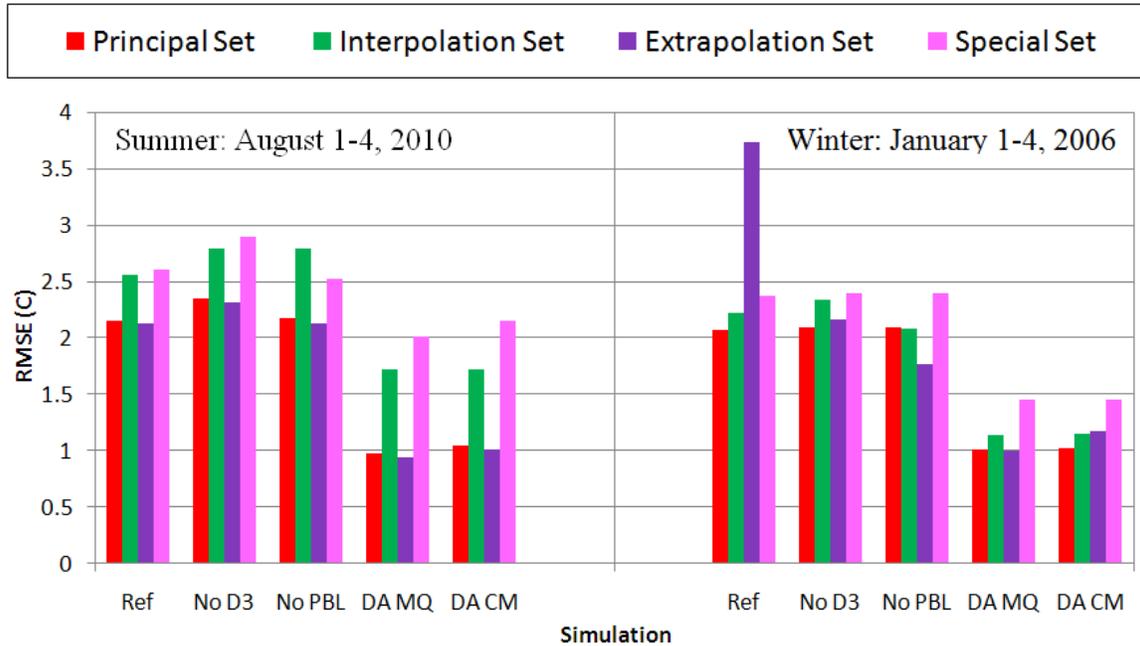


Figure 6-14 Average RMSE in summer and winter

Table 6-5 summarizes the average, minimum and maximum RMSE (°C) for all observation sets and all simulation in the two seasons, namely winter 2006 and summer 2008. It is clear that the simulations “DA MQ” have average errors less than 2 degrees Celsius for all observation sets. The maximum errors appear for the interpolation set in summer or for the special set in winter. The simulations “DA CM” have almost similar values of errors when compared with the simulations “DA MQ”. FDDA can reduce the errors at the locations of assimilated observations and the locations near them provided that they have almost similar surface boundary conditions. Again, the large errors occur at the locations near the coast lines or the locations that have complex terrain/land-use. These errors may be reduced by increasing the density and quality of observations in these areas.

Table 6-5 Summary of RMSE for all observation sets and all simulations

Set	Season	RMSE	Ref	No D3	No PBL	DA MQ	DA CM
Principal	Winter	Average	2.074	2.089	2.094	1.008	1.023
		Minimum	1.088	1.394	1.453	0.787	0.767
		Maximum	3.674	2.691	2.641	1.293	1.343
	Summer	Average	2.147	2.344	2.174	0.973	1.052
		Minimum	1.617	1.756	1.571	0.782	0.867
		Maximum	2.951	2.820	3.075	1.164	1.266
Interpolation	Winter	Average	2.221	2.336	2.081	1.140	1.157
		Minimum	1.349	1.943	1.217	0.748	0.905
		Maximum	4.322	3.119	2.990	1.491	1.336
	Summer	Average	2.554	2.791	2.791	1.722	1.725
		Minimum	1.228	1.752	1.613	0.906	0.852
		Maximum	3.443	3.598	3.915	2.981	2.929
Extrapolation	Winter	Average	3.735	2.161	1.764	0.997	1.169
		Minimum	2.724	1.013	1.360	0.698	0.776
		Maximum	5.227	3.685	2.533	1.271	1.930
	Summer	Average	2.127	2.309	2.123	0.938	1.015
		Minimum	1.633	1.587	1.745	0.697	0.738
		Maximum	2.631	3.053	2.840	1.235	1.218
Special	Winter	Average	2.369	2.396	2.394	1.449	1.450
		Minimum	1.646	1.655	1.598	1.028	1.058
		Maximum	3.489	4.010	3.877	2.250	2.187
	Summer	Average	2.607	2.899	2.525	2.014	2.148
		Minimum	1.252	1.920	2.276	1.328	1.528
		Maximum	3.922	3.853	3.021	3.388	3.538

Figure 6-15 shows the spatial distribution of the near-surface temperature difference between the reference simulations and the simulations when nudging was used with multiquadric analysis. Summer simulations are in the left of the figure while winter simulations are in the right. Both are after 24 hours from the initial time. The maximum errors were associated with the special set areas and near coast lines. Also, large errors were associated with the extrapolation set in winter.

The accuracy of simulations when using nudging with Multiquadric or Cressman analysis was found to be almost the best simulations for all observation sets. The

accuracy of the simulations when nudging was turned off for the finest domain and the simulations with no nudging in the planetary boundary layer were found to be sensitive to the location and weather regime but probably will produce better results when finer resolutions and better surface boundary conditions are used.

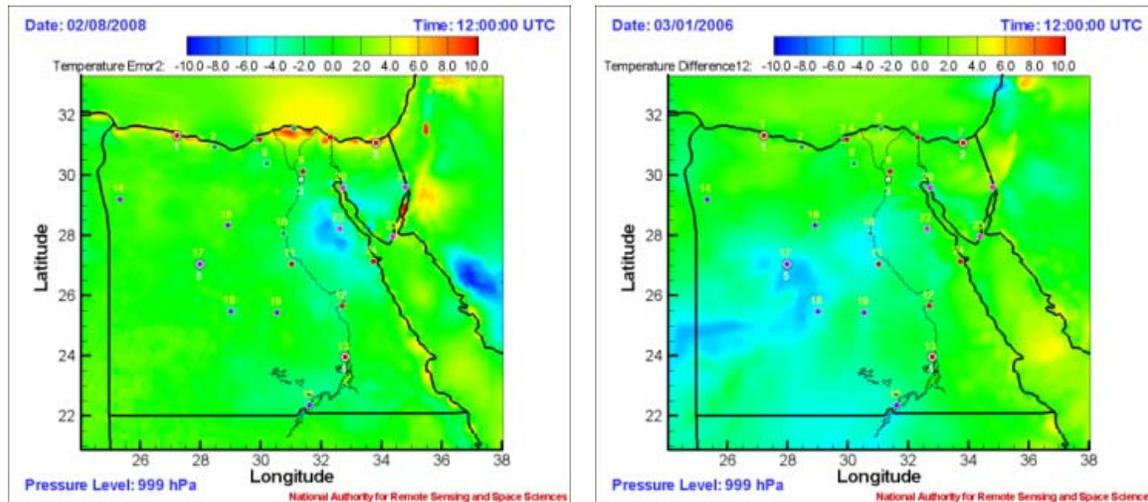


Figure 6-15 Temperature difference between reference and “DA MQ” simulations

It is found that the principal set provided the best accuracies and correlations with observations. The main reason for this is that it is used to nudge the model state during the integration time. The results of the interpolation set were found to be good provided that the stations have similar surface boundary conditions as the neighboring principal stations. This is also apparent for the extrapolation set. The worst results were found to be at the locations with complex terrain/land-use (special set) and near the coast lines for all simulations. This is expected to be improved by increasing the density and quality of observations.

Determination of the optimal number, locations, and types of new additional observational stations is a major recommendation of these results. Extension to adaptive (targeted) observations is planned.

6.3 EFFECT OF CHANGING MODEL CONFIGURATIONS

The FDDA enhancement of the model for Egypt requires the availability of observational data and FNL datasets and thus it is suitable for the research applications and not for operational predictions. So, it is essential to enhance the model results when the observational data is not available using different configurations of the NWP model. In this section, the model sensitivity to initialization time, dynamic initialization and different physics options is presented. During the sensitivity studies, the best configurations of the model were selected and then different scenarios were used to construct an ensemble of forecasts. The results of the ensemble forecasting experiment is presented in the consequent section.

6.3.1 Model Sensitivity to Initialization Time

Initializing the model from Z12 compared to Z00 slightly improves the accuracy of simulations when FDDA is turned off and has no effect when FDDA is implemented. Figure 6-16 shows the simulated and observed near-surface temperature and absolute-bias-error (ABE) for different initialization times in January, 2006 at Alexandria/Nouzha station while Figure 6-17 shows the associated RMSE and MABE.

Table 6-6 lists a statistical summary of the effect of initialization time on the MM5 performance. RMSE, for reference simulations, is reduced by about 13% when initializing the model at Z12 rather than Z00 while MABE is reduced by about 6%. The simulations, when FDDA is turned on, are almost insensitive to the initialization time. It is clear that FDDA eliminates the sensitivity of the model to initialization time where the model is nudged towards the observation and the effect of initializations is eliminated with time.

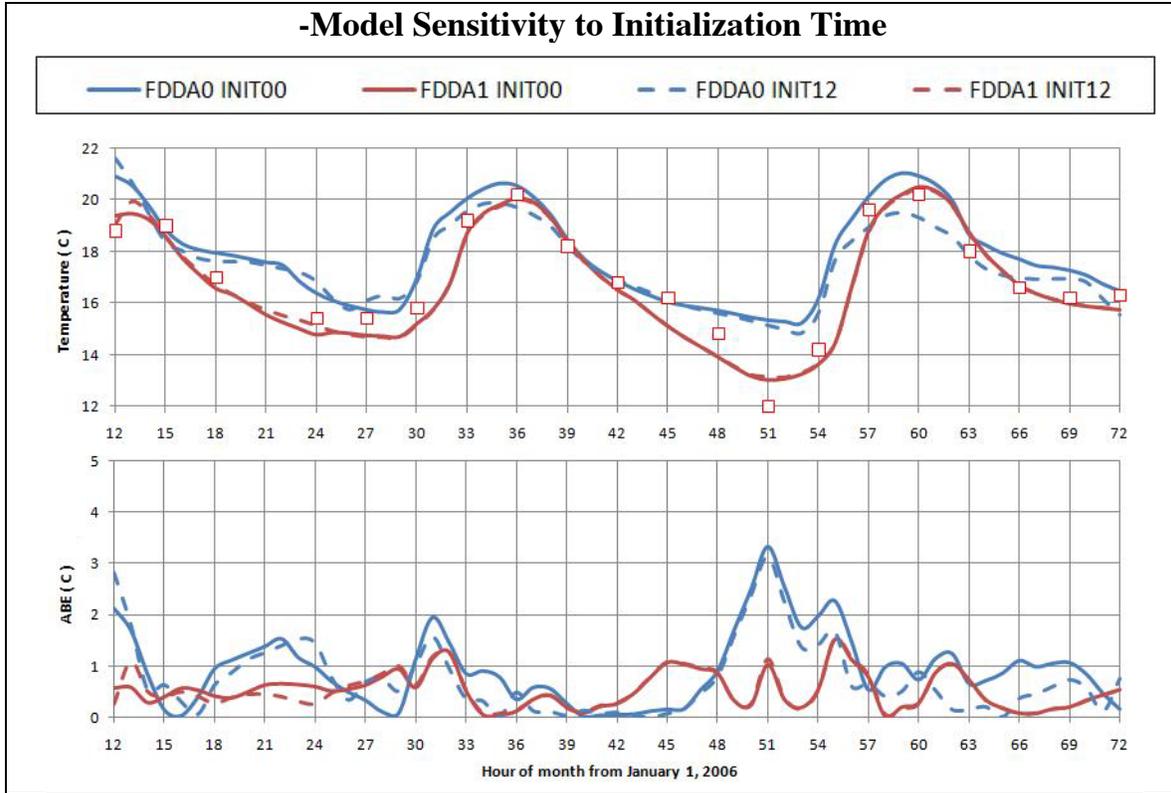


Figure 6-16 Near-surface temperature and ABE for different initialization times

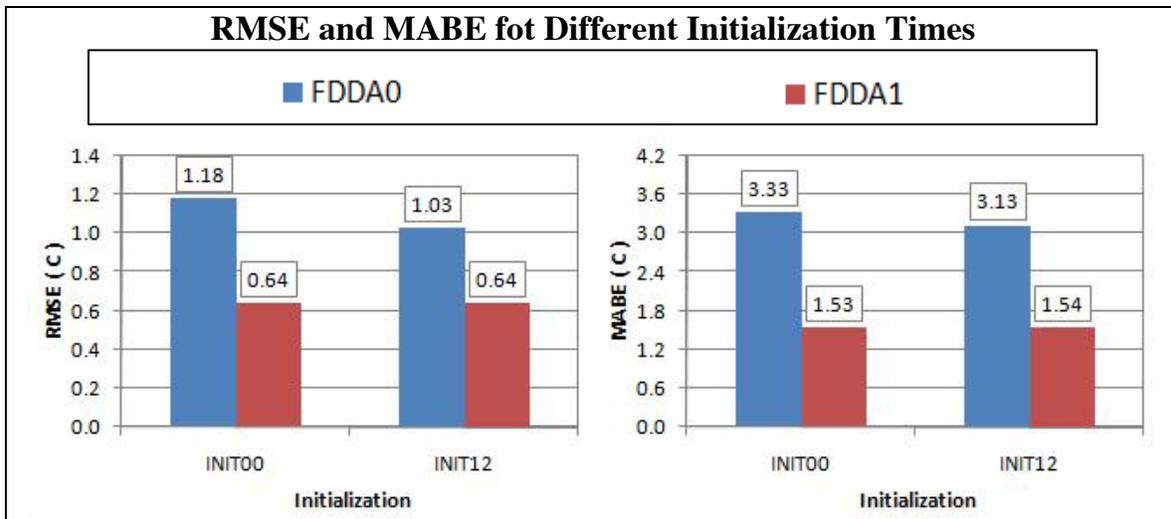


Figure 6-17 RMSE and MABE for different initialization times

Table 6-6 Summary of the model sensitivity to initialization time

Initialization	RSME C			MABE C		
	FDDA0	FDDA1	Gain	FDDA0	FDDA1	Gain
INIT00	1.179	0.638	45.92 %	3.330	1.534	53.93 %
INIT12	1.027	0.639	37.80 %	3.129	1.541	50.75 %
Gain	12.88 %	-0.21 %		6.04 %	-0.46 %	

6.3.2 Model Sensitivity to Dynamic Initialization

Figure 6-18 shows the total and daily RMSE, Figure 6-20 shows the simulated and observed near-surface temperature and ABE for 6-days simulations from Z12 January 1, 2006 at Alexandria/Nouzha Station, while Figure 6-19 shows the total and daily MABE when using dynamic initialization. Figure 6-20 shows the simulated and observed near-surface temperature and ABE. When turning off FDDA after one day, the errors is reduced for the other days which indicates that dynamic initialization improves the accuracy of forecasts where it provides accurate and smooth initial conditions for the model run. This improvement is caused by reducing the RMSE & MABE of the near-surface temperature by about 38% & 35% and limiting the bias error within 4 °C for about 30 more hours.

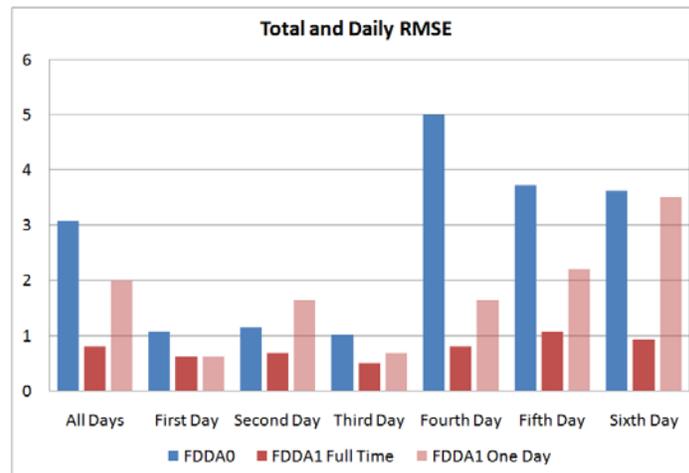


Figure 6-18 Total and daily RMSE when using dynamic initialization

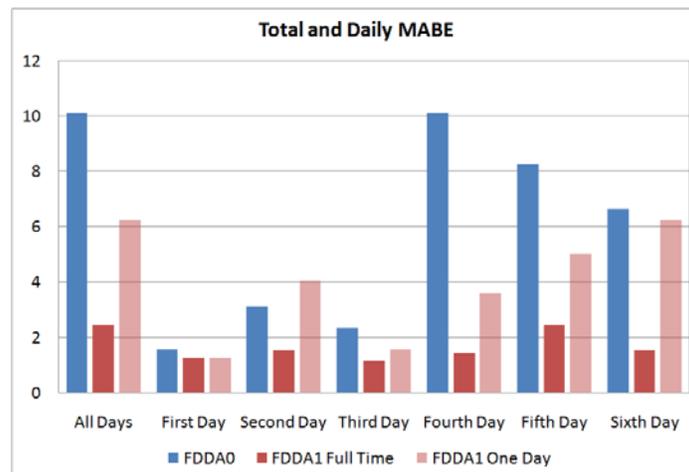


Figure 6-19 Total and daily MABE when using dynamic initialization

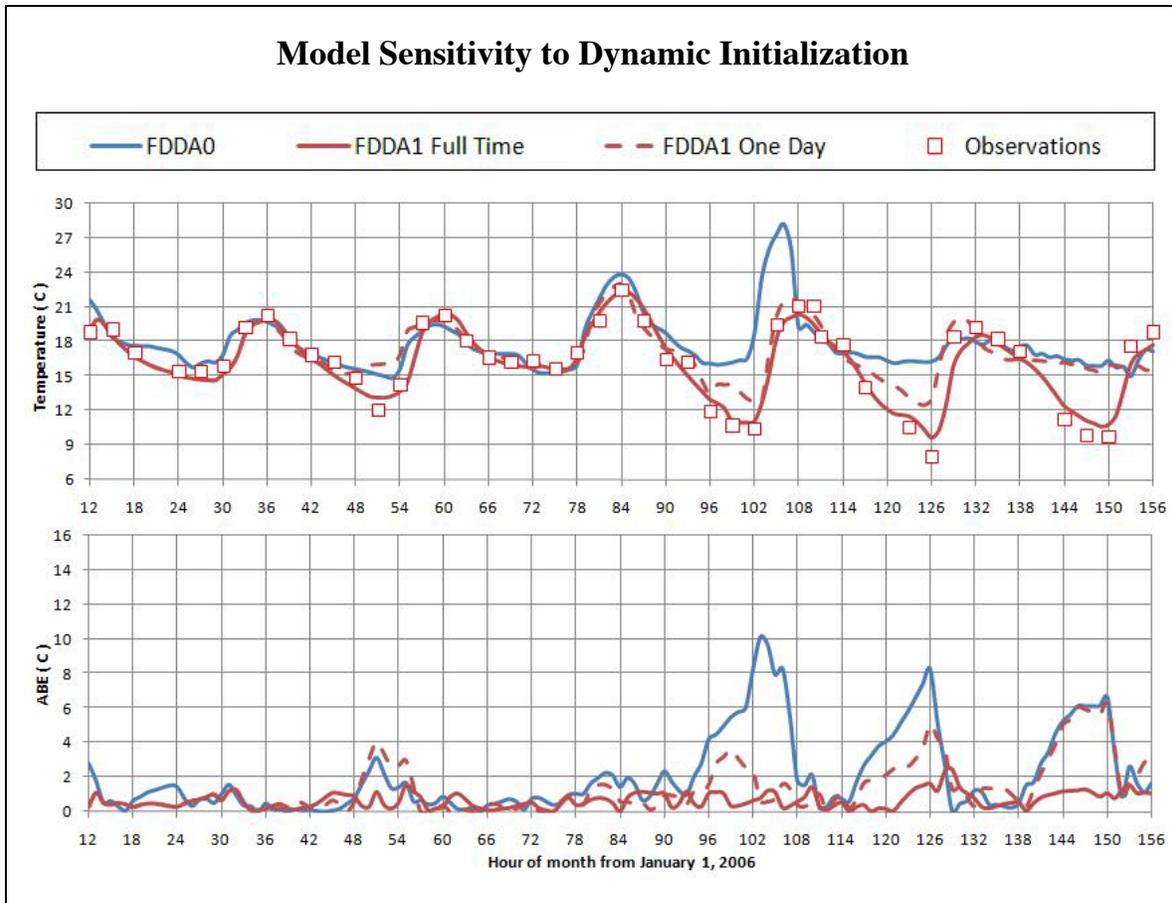


Figure 6-20 Near-surface temperature and ABE when using dynamic initialization

Table 6-7 Summary of the results of testing the model sensitivity to dynamic initialization

Period	FDDA1 Full Time			FDDA1 One Day			
	FDDA0	FDDA1	Gain	FDDA0	FDDA1	Gain	
RMSE	All Days	3.082	0.807	73.83 %	3.082	2.010	34.79 %
	First Day	1.081	0.626	42.09 %	1.081	0.626	42.09 %
	Second Day	1.146	0.690	39.81 %	1.146	1.646	- 43.5 %
	Third Day	1.017	0.507	50.18 %	1.017	0.691	32.10 %
	Fourth Day	4.997	0.820	83.61 %	4.997	1.645	67.09 %
	Fifth Day	3.716	1.073	71.13 %	3.716	2.207	40.60 %
MABE	Sixth Day	3.623	0.933	74.26 %	3.623	3.509	03.14 %
	All Days	10.100	2.461	75.63 %	10.100	6.248	38.14 %
	First Day	1.568	1.263	19.46 %	1.568	1.263	19.46 %
	Second Day	3.129	1.541	50.75 %	3.129	4.028	- 28.7 %
	Third Day	2.346	1.150	50.98 %	2.346	1.567	33.19 %
	Fourth Day	10.100	1.429	85.85 %	10.100	3.595	64.41 %
Fifth Day	8.253	2.461	70.18 %	8.253	5.025	39.11 %	
Sixth Day	6.644	1.549	76.69 %	6.644	6.248	05.96 %	

Table 6-7 lists the RMSE and MABE and associated gains for a 6-days simulation at Alexandria/Nouzha Station. It is clear that the meteorological conditions when

FDDA is implemented are greatly improved compared to observations in long-term simulations. The initialization time of forecasts, which is the ending time of dynamic initialization, must be taken into consideration because it affects the accuracy of the results. While the errors of the NWP model increase nonlinearly with time in the reference simulations, FDDA nudges the model solution toward the observation even for long ranges. However, this may be inaccurate where the effects of nudging terms are more than the model dynamics.

6.3.3 Model Sensitivity to Physics Options

The atmosphere is a chaotic system and the numerical modeling of the atmospheric physics has many sources of errors. The main sources of errors are the model deficiencies and the uncertainty in the initial conditions. The model deficiencies can be tested by studying the model sensitivities to different physics parameterizations and options. The sensitivity of the model to different physics options was tested in the four seasons and the results are compared to observations at nine monitoring locations inside Egypt, as shown in Figure 2-18. Nine different physics packages were used based on three different PBL schemes and three different radiation schemes. The reference initial conditions described in section 5.3 were used without data assimilation.

This section displays only a sample of the results. The rest of the results for other stations are shown in Appendix F. Conclusions were drawn based on the full set. Figure 6-21 shows near-surface temperature while Figure 6-22 shows percentage change in the average temperature for the nine different physics packages, which are the first nine members of the ensemble forecasting experiment described in Table 5-3, at Luxor station in autumn. It is clear that different radiation schemes produce different temperatures especially at day time while different PBL thermal roughness length formulations slightly change the simulated temperature for the same initial and boundary conditions.

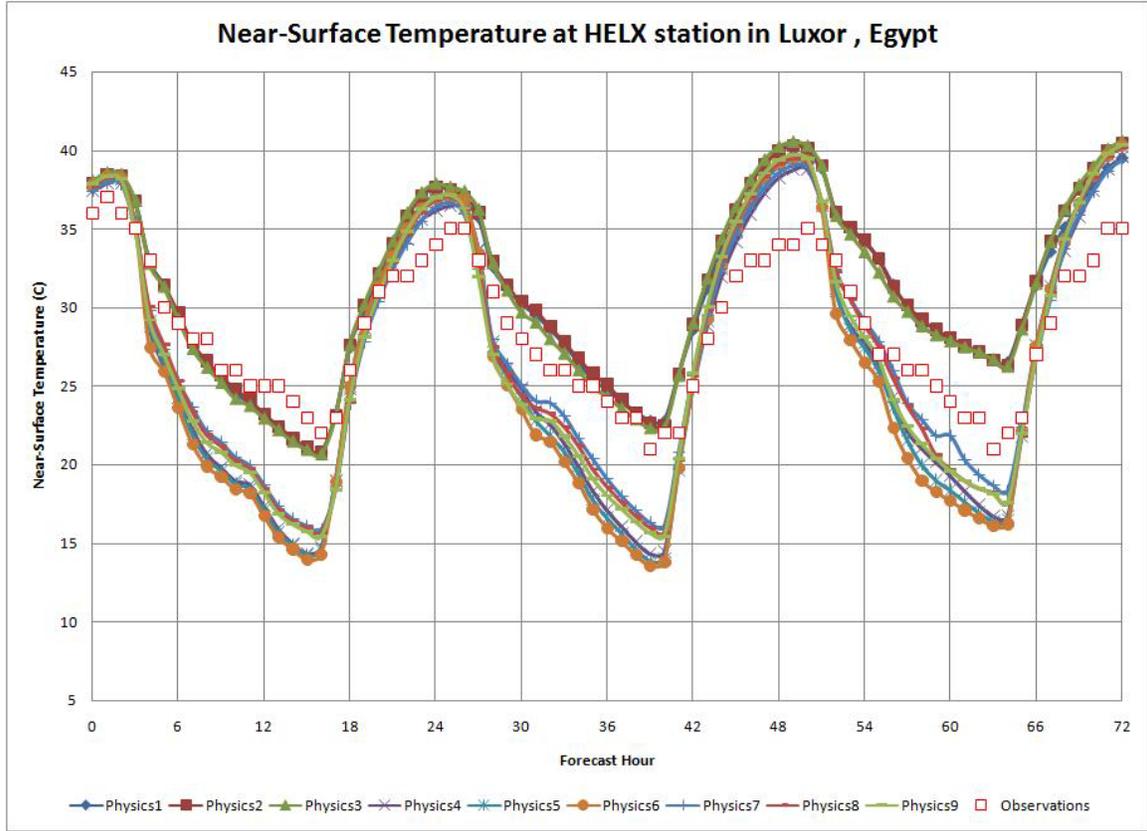


Figure 6-21 Near-surface temperature at Luxor station in autumn

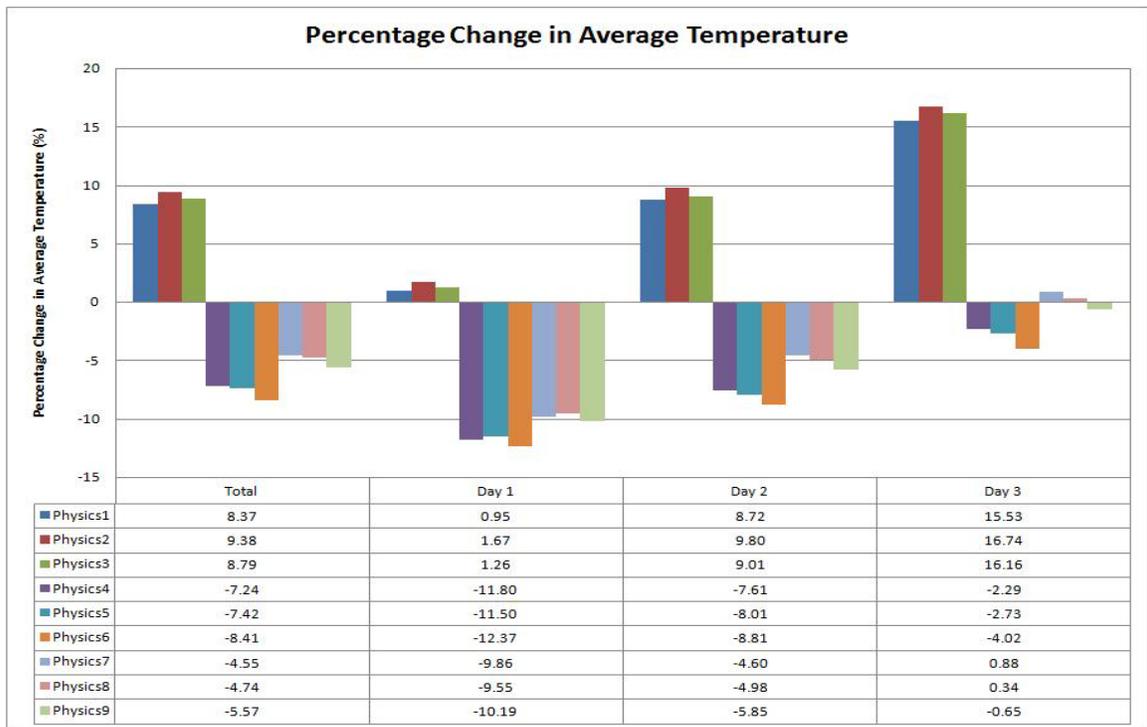


Figure 6-22 Change in average temperature at Luxor station in autumn

6.4 ENSEMBLE FORECASTING EXPERIMENT

The present section presents sample results of the preliminary SREF experiment for Egypt and its effects on the quality of the forecasts. The accuracy of the produced forecast is proportional to the number and quality of the ensemble members. The ensemble mean was evaluated based on the Root-Mean-Square Error, percentage change in average, minimum and maximum temperatures at 9 monitoring locations.

This section displays only a sample of the results. The rest of the results for other stations are shown in Appendix F. Conclusions were drawn based on the full set. Figure 6-23 shows near-surface temperature for all ensemble members and ensemble mean compared to the observations at Luxor station in summer. Figure 6-24 shows the minimum and maximum envelopes and the ensemble mean while Figure 6-25 shows the percentage change in the average temperature.

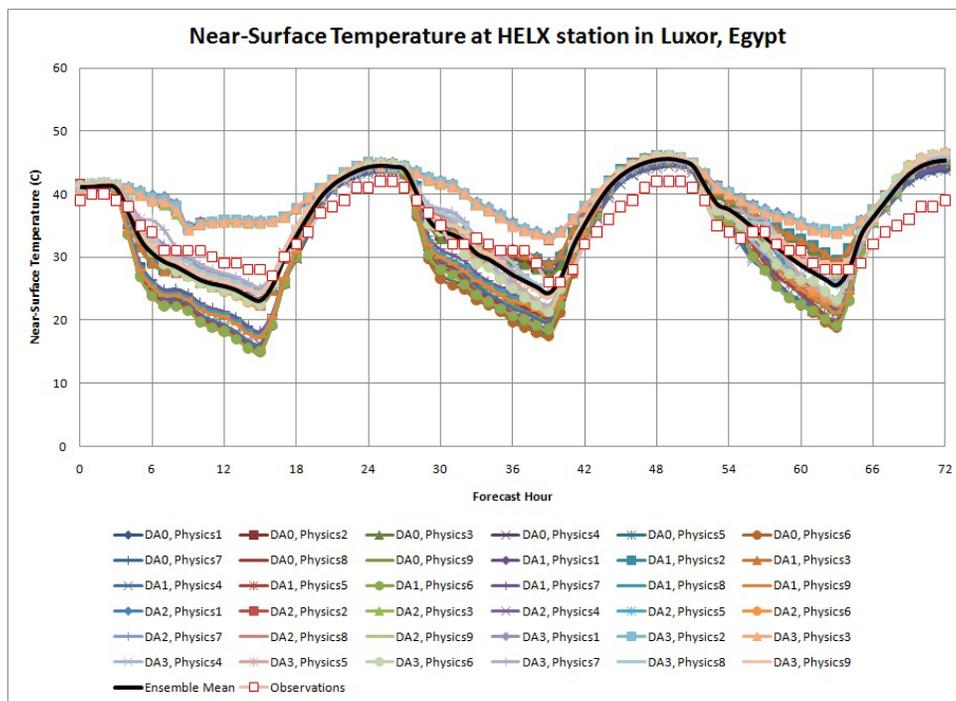


Figure 6-23 Near-surface temperature at Luxor station in summer

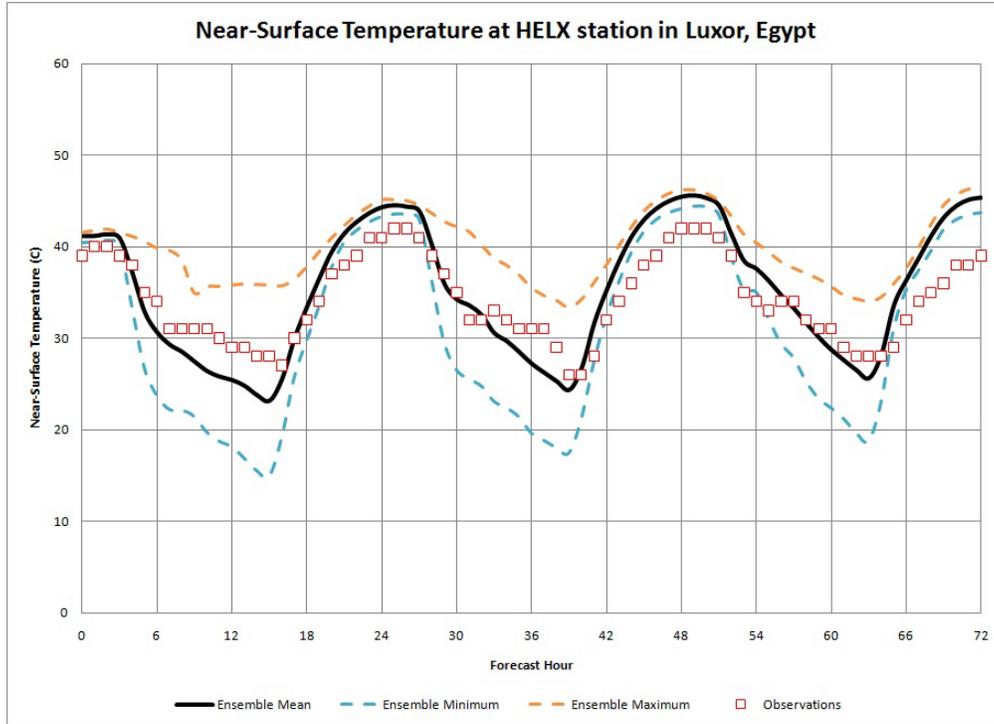


Figure 6-24 Min, Max and Mean temperature at Luxor station in summer

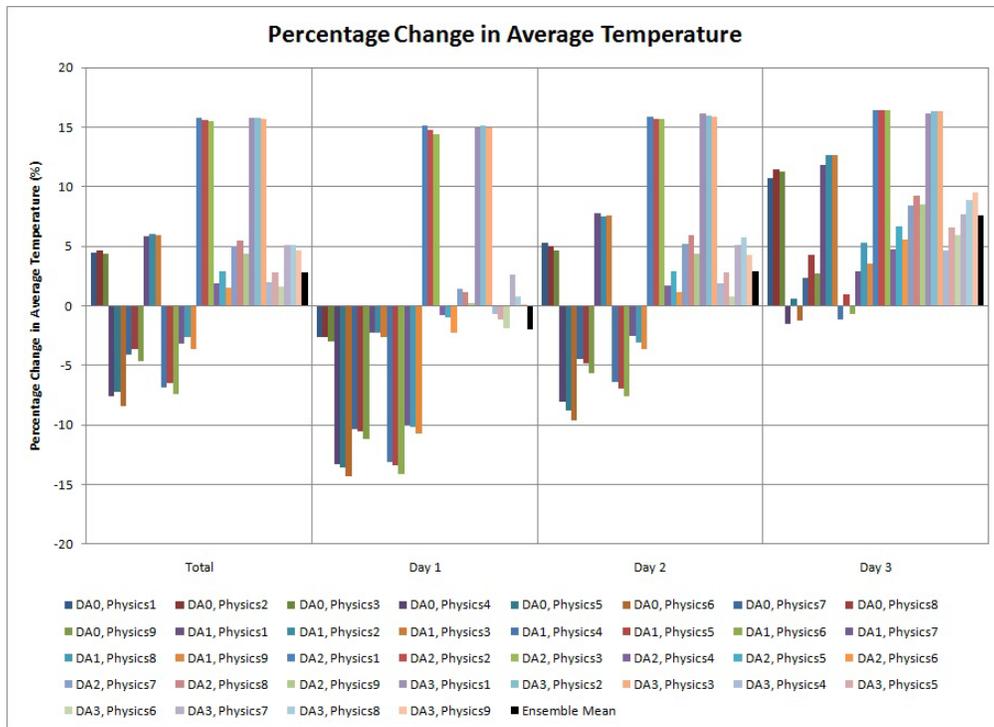


Figure 6-25 Change in average temperature at Luxor station in summer

It is clear that the ensemble mean is almost the best forecast with small errors and provides better estimation for the average, minimum and maximum temperatures at all locations for the different weather regimes. Moreover, using FDDA to produce different sets of initial conditions for the model run affects the behavior of the model. Different physics options produce different temperatures when the different initial conditions were used.

CHAPTER SEVEN

CONCLUSIONS AND FUTURE WORK

Four main topics were studied in this study. The first is the FDDA enhancement of the mesoscale numerical weather modeling for Egypt. The second is the evaluation of the work done in the first section. The third is the sensitivity of the MM5 numerical weather model to different configurations. And finally the work was extended to a preliminary Short-Range Ensemble Forecasting (SREF) experiment. More work is required to refine this study and to extend it for operational use. This chapter presents different conclusions drawn from this study and suggestions for future work.

7.1 USING FDDA IN NWP FOR EGYPT

- The performance of the NWP model (MM5) is sensitive to the technique of data assimilation used and the quality, quantity and distribution of the implemented observations.
- Grid FDDA increases the computational time by about 5% while it reduces the RMSE and MABE by about 58% and 39%, respectively.
- Observational FDDA increases the computational time by about 20% while it reduces the RMSE and MABE by about 28% and 41%, respectively.
- Combined grid and observational FDDA as used in this study is not time efficient and reproduces temperature results, similar to those of grid FDDA. This may be changed by tuning the observational FDDA options and prepare the observations at certain times (say every half hour).
- FDDA improves the model accuracy and the computed temperatures correlate better with the observed temperatures, especially near the surface.

- The large errors are generated at stations near coast lines and locations of complex topography especially at night hours.
- FDDA increases the near-surface temperature around the Red Sea and reduces it at the south east of Egypt. This may be due to the selection of physics options that are not suitable for these areas or due to the surface boundary conditions, such as land-use.

7.2 EVALUATION OF FDDA ON NWP FOR EGYPT

- The accuracy of the model was affected at the observational-data-void areas, near coast lines, and at the locations with complex terrain. This may be fixed by increasing the number and quality of uniformly-distributed observational stations at these locations.
- The best accuracies are obtained at the location where enough observational data density is available.
- The results for of the locations between observational sites, suspected to interpolation, are good provided that the station is located between observation sites that have similar surface boundary conditions. This is also apparent for the location far away from the observational sites, suspected to extrapolation.
- The worst results are, as expected, in the locations with complex terrain/land-use. Moreover, large errors are produced near the coast lines for all simulations.
- The accuracy of simulations when using FDDA with MQ or CM is better than the reference simulations anywhere.
- When FDDA is turned off for the finest domain, the results are good in some locations and bad in other locations. This may be due to the

uncertainties in initial and boundary conditions produced from the coarser domains.

- The simulations with no nudging in the PBL is also sensitive to the location and weather regime and probably they may produce better results when finer resolution such as 3 km or 1 km resolution with better surface boundary conditions are used.

7.3 EFFECT OF CHANGING MODEL CONFIGURATIONS

- Initializing the MM5 numerical weather model from Z12 compared to initializing Z00 slightly improves the accuracy of simulations when FDDA is turned off and has no effect when FDDA is implemented.
- Dynamic initialization for one day improves the accuracy of forecasts by reducing the RMSE & MABE of the near-surface temperature by about 38% & 35% and limiting the bias error within 4 °C for about 30 more hours. This may be because it provides accurate and smooth initial conditions for the model run.
- The meteorological conditions when FDDA is implemented are greatly improved compared to observations in long-term simulations.
- The initialization time of forecasts, which is the ending time of dynamic initialization, must be taken into consideration where it affects the accuracy.
- The MM5 model is sensitive to the selected physics options.
- Different radiation schemes produce different temperatures especially at day time.

- Different PBL thermal roughness length formulations slightly change the simulated temperature for the same initial and boundary conditions.

7.4 ENSEMBLE FORECASTING EXPERIMENT

- The deterministic forecasting is not helpful in the applications where better accuracies are essential like the pollution assessment and wind energy potential.
- The accuracy of the produced forecast, ensemble mean, is proportional to the number and quality of the ensemble members.
- The construction of an ensemble is the main problem in EF and computational cost must be taken into consideration.
- The ensemble mean is almost the best forecast with small errors and provides better estimation for the average, minimum and maximum temperatures at all locations for the different weather regimes.
- Using FDDA to produce different sets of initial conditions for the model run affects the behavior of the model. The different physics options produce different temperatures when different initial conditions were used.

7.5 SUGGESTIONS FOR FUTURE WORK

- Extending the work to cover a complete year or more rather than taking samples of time to evaluate the modeling system during a continuous coverage of weather regimes.
- Integrating this work to study the effect of data assimilation on the investigation of the temperature inversion phenomena and distribution of pollutants under the inversion layer.

- Repeating this work with different points of view such as changing the application from the temperature simulation/prediction to wind energy potential and/or precipitation prediction and so on.
- Implementing observations from different data sources especially to utilize the rapid increase of remote sensing technology providing a huge amount of satellite data which can be used for updating the surface boundary conditions or assimilated into model simulations.
- Increasing the spatial resolution for Egypt domain with two-way nesting technique which requires more computational resources.
- Comparing the results obtained in this study, using the MM5 model, with similar simulations performed with the WRF model.
- Evaluation of modeling using advanced data assimilation methods (e.g., 3/4DVar and/or KF) in the numerical weather modeling for Egypt.
- Constructing an ensemble using the advanced ensemble forecasting methods (e.g., BVs, SVs and KF).
- Extending to targeted/adaptive observations to find the locations of new observational data in specific areas at specific times with the aim of improving the quality of pre-selected NWP forecast features.
- Expanding the comparison of different physics options to say more about individual schemes.
- Trying to further develop this work for operational use.

REFERENCES

1. Mesinger, F., “Eta Model at NCEP: Challenges overcome and lessons learned”, Lecture notes, Workshop on Design and Use of Regional Weather Prediction Models, The Abdus Salam International Centre for Theoretical Physics, Miramare, Trieste, Italy, April 11-19, 2005.
2. Kandil, H. A., Abdelkader M. M., Abdelmoaty A., Elhadidi, B. M., and Sherif, A. O., “Building an Accurate Weather Prediction Model for Egypt”, The Egyptian Journal of Remote Sensing and Space Sciences, ISSN 1110-9823, Volume IX, 2006.
3. Elhadidi, B. M., Kandil, H. A., Abdelmoaty, A., Abdelkader, M. M., and Sherif, A. O., ” Modeling the Possible Meteorological Changes Induced by the Formation of Toshka Lakes in Egypt”, Cairo 10th International Conference on Energy and Environment, Luxor, Egypt, March 13-15, 2007.
4. Elhadidi, B. M., Abdelmoaty, A., Abdelkader, M. M., Kandil, H. A., and Sherif, A. O., ” Numerical Experiments to Investigate Meteorological Changes Induced by Local Land Modification”, Proceeding of 41st U.S-Egypt Workshop on The Predictive Methodologies for Global Weather-Climate and Related Disasters, March 13–15, 2006.
5. Abdelkader, M. M., Elhadidi, B. M., and Sherif, A. O., “Seasonal Evaluation of Temperature Inversion over Cairo”, Proceedings of 2008 IEEE International Geoscience and Remote Sensing Symposium, Boston, USA, July 6-11, 2008.
6. Kandil, H. A., Abdelkader, M. M., Abdelmoaty, A., Elhadidi, B. M., and Sherif, A. O., “Simulation of Atmospheric Temperature Inversions over Greater Cairo Using the MM5 Mesoscale Atmospheric Model”, Proceeding of 41st U.S-Egypt Workshop on The Predictive Methodologies for Global Weather-Climate and Related Disasters, March 13–15, 2006.

7. Kandil, H. A., Abdelkader, M. M., Abdelmoaty, A., Elhadidi, B., and Sherif, A. O., "Seasonal Evaluation of Temperature Inversions" , Cairo 10th International Conference on Energy and Environment, Luxor, Egypt, March 13-15, 2007.
8. Kandil, H. A., Abdelkader, M. M., Abdelmoaty, A., Elhadidi, B. M., and Sherif, A. O., "Simulation of Atmospheric Temperature Inversions over Greater Cairo Using the MM5 Mesoscale Atmospheric Model", The Egyptian Journal of Remote Sensing and Space Sciences, ISSN 1110-9823, Volume IX, 2006.
9. Sherif, A. O., Kandil H. A., Elhadidi B. M., Abdelmoaty, A., and Abdelkader, M. M., "Simulation of Surface Temperature Inversions In Egypt Using The MM5 MesoScale Numerical Modeling System", Cairo 9th International Conference on Energy and Environment, Sharm ElSheikh, Egypt, March 13-15, 2005.
10. Badr, H. S., "ATOVS Data Products and Applications", General Organization for Remote Sensing (GORS), Damascus, Syria, February 10-15, 2008.
11. Badr, H. S., Elhadidi, B. M., and Sherif, A. O., "FDDA Enhancement of the Mesoscale Meteorological Modeling System for Egypt", Cairo 10th International Conference on Energy and Environment, Luxor, Egypt, March 13-15, 2007.
12. Sherif, A. O., Kandil, H. A., Elhadidi B. M., Abdelmoaty, A., and Abdelkader, M. M., "Using Remote Sensing Observations to Improve the Predictions of a High-Resolution Mesoscale Weather Modeling System for Egypt", Cairo 9th International Conference on Energy & Environment, Sharm ElSheikh, Egypt, March 13-15, 2005.
13. Sherif, A. O., Kandil, H. A., Elhadidi, B. M., Abdelmoaty A., and Abdelkader, M. M., "Regional Weather Prediction Models with Remotely Sensed Data Assimilation", Cairo University 2nd International Conference on Applied Research, Cairo, Egypt, December, 2005.
14. Sherif, A. O., Kandil, H. A., Elhadidi, B. M., Abdelmoaty, A., and Abdelkader, M. M., "Improving the Weather Prediction Capabilities Using the Remote-

- Sensing Technology on a Cluster of 64-bit Machines”, 2nd International Conference on Advances on Engineering Science and Technology, National Research Center (NRC), Egypt, November 12-14, 2005.
15. Badr, H. S., Elhadidi, B. M., and Sherif, A. O., “FDDA Enhancement of the Mesoscale Meteorological Modeling System for Egypt”, Cairo 10th International Conference on Energy & Environment, Luxor, Egypt, March 13-15, 2007.
 16. Badr, H. S., Elhadidi, B. M., and Sherif, A. O., “Evaluation of Data Assimilation on Numerical Weather Prediction for Egypt”, Proceedings of 2010 IEEE International Geoscience and Remote Sensing Symposium, Honolulu, Hawaii, USA, July 25-30, 2010.
 17. Dudhia, J., Grill, D., Manning, K., Wang, W., Bruyere, C., Kelly, S., and Lackey, K., “PSU/NCAR Mesoscale Modeling System, Tutorial Class Notes and User's Guide”, NCAR Technical Notes, January 2005.
 18. Grell, A. G., Dudhia, J., and Stauffer, R. D., "A Description of the Fifth Generation Penn State/NACR Mesoscale Model (MM5)", NCAR Technical Notes, June 1995.
 19. Buizza, R., Houtekamer, P. L., Toth, Z., Pellerin, G., Wei, M., and Zhu, Y., “A Comparison of the ECMWF, MSC, and NCEP Global Ensemble Prediction Systems”, *Monthly Weather Review*, Volume 133, pp. 1076-1097, 2005.
 20. Sivillo, J. K., and Ahlquist, J. E., “An Ensemble Forecasting Primer”, *Weather and Forecasting*, Volume 12, pp. 809-818, 1997.
 21. Kalnay, E., “Atmospheric Modeling, Data Assimilation and Predictability”, Cambridge University Press, ISBN 052179179 0, 2003.
 22. Wang, W., Barker, D., Bruyere, C., Duda, M., Dudhia, J., Gill, D., Michalakes, J., and Rizvi, S., “WRF Version 3.2 Modeling System User’s Guide”, http://www.mmm.ucar.edu/wrf/users/docs/user_guide_V3/, 2010.

23. World Meteorological Organization Operational Network:
http://www.wmo.ch/pages/prog/drr/wmoOppNetwork_en.html
24. Bouttier, F., and Courtier, P., “Data Assimilation Concepts and Methods”, Meteorological Training Course Lecture Series, ECMWF, 1999.
25. Model Errors and Ensemble Forecasting:
<http://www.emc.ncep.noaa.gov/gmb/ens/papers/ecmwf-finalwordweb.htm>
26. Sivillo, J. K., Ahlquist, J. E., “An Ensemble Forecasting Primer”, Weather and Forecasting, Volume 12, pp. 809-818, 1997.
27. Before 1955: Numerical Models and Prehistory of AGCMs:
<http://www.aip.org/history/sloan/gcm/prehistory.html#fn0>
28. Kalnay, E., “Atmospheric Modeling, Data Assimilation and Predictability”, Cambridge University Press, ISBN 052179179 0, 2003.
29. Sasaki, Y., “An Objective Analysis Based on the Variational Method”, Journal of the Meteorological Society of Japan, Volume 36, pp. 77-88, 1958.
30. Sasaki, Y., “Influences of Thermal Boundary Layer on Atmospheric Cellular Convection”, Journal of the Meteorological Society of Japan, Volume 48, pp. 492-502, 1970.
31. Daley, R., “Atmospheric Data Analysis”, Cambridge Atmospheric and Space Science Series, Cambridge University, ISBN 0521382157, 1991.
32. Bishop, C. H., Etherton, B. J., and Majumdar, S. J., “Adaptive Sampling with The Ensemble Transform Kalman Filter. Part I: Theoretical aspects”, Monthly Weather Review, Volume 129, pp. 420-436, 2001.
33. Todling, R., and Cohn, S. E., “Suboptimal Schemes for Atmospheric Data Assimilation based on the Kalman Filter”, Monthly Weather Review, Volume 122, pp. 2530-57, 1994.
34. Hoke, J. E., and Anthes, R. A., “The Initialization of Numerical Models by a Dynamical Initialization Technique”, Monthly Weather Review, Volume 104, pp. 1551–1556, 1976.

35. Lorenc, A. C., “The Potential of the Ensemble Kalman Filter for NWP – A Comparison with 4DVar”, *Journal of the Royal Meteorological Society*, Volume 129, pp. 3183–3203, 2003.
36. Stauffer, D. R., and Seaman, N. L., “Multi-scale Four-Dimensional Data Assimilation”, *Journal of Applied Meteorology*, Volume 33, pp. 416-434, 1994.
37. Bouttier, F., and Rabier, F., “The Operational Implementation of 4DVar”, *ECMWF Newsletter*, Volume 78, pp. 2–5, 1997.
38. Barker, D. M., Huang, W., Guo, Y.-R., Bourgeois, A. J., and Xiao, Q. N., “A Three-Dimensional Variational Data Assimilation System for MM5: Implementation and Initial Results”, *Monthly Weather Review*, Volume 132, pp. 897-914, 2004.
39. Xiao, Q. N., Kuo, Y. H. Sun, J., Lee, W. C. Lim, E., Guo, Y. R, and Barker, D. M., “Assimilation of Doppler Radar Observations with a Regional 3DVar System: Impact of Doppler Velocities on Forecasts of a Heavy Rainfall Case”, *Journal of Applied Meteorology*”, Volume 44(6), pp. 768–788, 2005.
40. Xiao, Q. N., Kuo, Y. H., Sun, J., Lee, W. C. Barker, D. M., and Lim, E., “An Approach of Radar Reflectivity Data Assimilation and its Assessment with the Inland QPF of Typhoon Rusa (2002) at Landfall”, *Journal of Applied Meteorology and Climatology*, Volume 46, pp. 14–22, 2007.
41. Xiao, Q., Lim, E., Won, D. J., Sun, J., Lee, W. C, Lee, M. S., Lee, W. J., Cjo, J. Y., Kuo, Y. H., Barker, D. M., Lee, D. K., and Lee, H. S., “Doppler Radar Data Assimilation in KMAs Operational Forecasting”, *Bulletin of the American Meteorological Society*, Volume 89, pp. 39–43, 2008.
42. Xiao, Q., and Sun, J., “Multiple Radar Data Assimilation and Short-Range Quantitative Precipitation Forecasting of a Squall Line Observed During IHOP 2002”, *Monthly Weather Review*, Volume 135, pp. 3381–3404, 2007.
43. Gleeson, T. A., “A Statistical Theory of Meteorological Measurements and Predictions”, *Journal of Meteorology*, Volume 18, pp. 192–198, 1961.

-
44. Lorenz, E. N., “A Study of the Predictability of a 28-Variable Atmospheric Model”, *Tellus*, Volume 17, pp. 321-333, 1965.
 45. Lorenz, E. N., “The Predictability of A Flow Which Possesses Many Scales of Motion”, *Tellus*, 21, pp. 289–307, 1969.
 46. Sivillo, J. K., and Ahlquist, J. E., “An Ensemble Forecasting Primer”, *Weather and Forecasting*, Volume 12, pp. 809-818, 1997.
 47. Toth, Z., and Kalnay, E., “Ensemble Forecasting at NCEP and the Breeding Method”, *Monthly Weather Review*, Volume 125, pp. 3297-3318, 1997.
 48. Toth, Z., and Kalnay, E., “Ensemble Forecasting at NMC: the Generation of Perturbations”, *Bulletin of the American Meteorological Society*, Volume 74, pp. 2317-2330, 1993.
 49. Houtekamer, P. L., and Lefaiivre, L., “Using Ensemble Forecasts for Model Validation”, *Monthly Weather Review*, Volume 125, pp. 2416–2426, 1997.
 50. Houtekamer, P. L., Lefaiivre, L., and Derome, J., “A System Simulation Approach to Ensemble Prediction”, *Monthly Weather Review*, Volume 124, pp. 1225-1242, 1996b.
 51. Houtekamer, P. L., Lefaiivre, L., and Derome, J., “The RPN Ensemble Prediction System”, *Proceeding of ECMWF Seminar on Predictability*, Reading, United Kingdom, , Volume II, pp. 121–146, 1996a.
 52. Sherif, A. O., Kandil, H. A., Elhadidi, B. M., Abdelmoaty, A., and Abdelkader, M. M., “Improving the Weather Prediction Capabilities Using the Remote-sensing Technology on a Cluster of 64-bit Machines”, 2nd International Conference on Advances on Engineering Science & Technology, National Research Center (NRC), Egypt, 12th -14th November, 2005.
 53. Mitchell, H. L., and Houtekamer, P. L., “An Adaptive Ensemble Kalman Filter”, *Monthly Weather Review*, Volume 128, pp. 416–433, 2000.
 54. Lorenc, A. C., “The Potential of the Ensemble Kalman Filter for NWP – A Comparison with 4DVar”, *Journal of the Royal Meteorological Society*, Volume 129, pp. 3183–3203, 2003.

55. Etherton, B. J., and Bishop, C. H., “Resilience of Hybrid Ensemble/3DVar Analysis Schemes to Model Error and Ensemble Covariance Error”, *Monthly Weather Review*, Volume 130, pp. 1065–1080, 2004.
56. Ott, E., Hunt, B. R., Szunyogh, I., Zimin, A. V., Kostelich, E. J., Corazza, M., Kalnay, E., Patil, D. J., and Yorke, J. A., “A Local Ensemble Kalman Filter for Atmospheric Data Assimilation”, *Tellus*, Volume 56A, pp. 415–428, 2004.
57. Szunyogh, I., Kostelich, E. J., Gyarmati, G., Patil, D. J., and Hunt, B. R., “Assessing a Local Ensemble Kalman Filter. Perfect Model Experiments with the NCEP Global Model”, *Tellus*, Volume 57A, pp. 528– 545, 2005.
58. Wang, W., Barker, D., Bruy`ere, C., Duda, M., Dudhia, J., Gill, D., Michalakes, J., and Rizvi, S., “WRF Version 3.2 Modeling System User’s Guide”, http://www.mmm.ucar.edu/wrf/users/docs/user_guide_V3/, 2010.
59. Anthes, A. R., “Development of Asymmetries in a Three-Dimensional Numerical Model of the Tropical Cyclone”, *Monthly Weather Review*, Volume 100, pp. 461–476, 1972.
60. Andorews, G. D., "An Introduction to Atmospheric Physics", Cambridge University Press, ISBN 0521629586, 2004.
61. Dudhia, J., "A Nonhydrostatic Version of the Penn State/NCAR Mesoscale Model: Validation Tests and Simulation of an Atlantic Cyclone and Cold Front", *Monthly Weather Review*, Volume 121, pp. 1493–1513, 1993.
62. Grell, A. G., Dudhia, J., and Stauffer, R. D., "A Description of the Fifth Generation Penn State/NACR Mesoscale Model (MM5)", NCAR Technical Notes, June 1995.
63. Kuo, B., “The Penn State/NCAR Mesoscale Model: MM5”, Board on Atmospheric Sciences and Climate (BASC) Panel, 14 October 2004.
64. Data set NCAR Final Reanalysis (FNL):
<http://dss.ucar.edu/datasets/ds083.2/data>
65. Data set Global Forecast System (GFS):

- <ftp://ftpprd.ncep.noaa.gov/pub/data/nccf/com/gfs/prod/>
66. Guo, Y., and Chen S., "Terrain and Land Use for the Fifth Generation Penn Stata/NCAR Mesoscale Modeling System (MM5) ", NCAR Technical Notes, January, 1994.
67. Abdelmoaty, A., "Modeling Land Cover Induced Climate Changes Using Satellite Data", M.Sc. Thesis, Department of Aerospace Engineering, Cairo University, Giza, Egypt, 2006.
68. Klemp, B. J., and Durran, R. D., "An Upper Boundary Condition Permitting Internal Gravity Wave Radiation in Numerical Mesoscale Models", Monthly Weather Review, Volume 1111, pp. 430-444, 1983.
69. Satyamurty, P., and Bittencourt, D., P., "Performance Evaluation Statistics Applied to Derived Fields of NWP Model Forecasts", Weather and Forecasting, Volume 14, pp. 726-740, 1999.
70. Trentmann, J., Seifert, A., Groenemeijer, P., Corsmeier, U., and Wernli, H., "Evaluation of highresolution NWP modelsimulations for midlatitude deep convection", 4th PAN-GCSS meeting, Météo-France Conference Centre, Toulouse, 2-6 June 2008.
71. World Meteorological Organization Global Observing System (GOS):
<http://www.wmo.ch/pages/prog/www/OSY/GOS.html>
72. The CISL Research Data Archive (RDA):
<http://dss.ucar.edu/>
73. The Comprehensive Large Array-data Stewardship System (CLASS):
<http://www.class.ngdc.noaa.gov>
74. Satellite and data description of NOAA's Polar-Orbiting satellites:
<http://www2.ncdc.noaa.gov/docs/podug>
75. ATOVS and AVHRR Processing Package (AAPP) Overview:
http://www2.met-office.gov.uk/research/interproj/nwpsaf/aapp/NWPSAF-MO-UD-004_Overview.pdf

-
76. Li, J., Wolf, W., Menzel, P., Zhang, W., Huang, H., and Achtor, T. H., "Global Soundings of the Atmosphere from ATOVS Measurements: The Algorithm and Validation", *Journal of Applied Meteorology*, Volume 39, pp.1248-1268, 2000.
77. Rew, R. K., "NetCDF User's Guide: An Interface for Data Access, Unidata Program Center", anonymous-ftp: unidata.ucar.edu, 1990.
78. ITT Company for Visual Information Solutions, Envi Software Package:
<http://www.ittvis.com/ProductServices/ENVI.aspx>
79. McClain, E. P., Pichel, W., and Walton, C., "Comparative Performance of AVHRR-based Multichannel Sea Surface Temperatures", *Journal of Geophysical Research*, Volume 90, pp. 11587–11601, 1985.
80. Price, J. C., 'Land Surface Temperature Measurements from The Split Window Channels of The NOAA 7 AVHRR", *Journal of Geophysical Research*, Volume 89, pp. 7231-7237, 1984.
81. Schroeder, M., Schulz, J., Jonas, M., and Lindau, R., "Long-term Application and Evaluation of IAPP using Global Radiosonde and CHAMP Measurements", EGU General Assembly, Volume 10, pp. 1607-7962, 2008.
82. Benjamin, S. G., "A Simple Scheme for Objective Analysis in Curved Flow", *Monthly Weather Review*, Volume 113, pp. 1184-1198, 1985.
83. Faccani, C., "High-Resolution Weather Forecasting over Complex Orography: Sensitivity to the Assimilation of Conventional Data", *Monthly Weather Review*, Volume 131, pp. 136-154, 2002.
84. Personal Communication with Dr. J. Dudhia.
85. Causes of Forecast *Errors* in *NWP*:
http://weather.ou.edu/~metr4424/Files/Forecast_Errors.pdf
86. Model Errors and Ensemble Forecasting:
<http://www.emc.ncep.noaa.gov/gmb/ens/papers/ecmwf-finalwordweb.htm>

87. Kalnay, E., “Ensemble Forecasting and Data Assimilation: Two Problems with The Same Solution?”, proceeding of the Symposium on the 50th Anniversary of NWP, University of Maryland, College Park, MD, 14-17 June 2004.
88. Epstein, E. S., “Stochastic Dynamic Prediction”, *Tellus*, Volume 21, pp. 739-759, 1969.
89. Leith, C. E., “Theoretical Skill of Monte Carlo Forecasts”, *Monthly Weather Review*, Volume 102, pp. 409–418, 1974
90. Ebisuzaki, W., and Kalnay, E., “Ensemble Experiments with a New Lagged Average Forecasting Scheme”, WMO, Research activities in atmospheric and oceanic modeling. Report #15, pp. 6.31-6.32, 1991.
91. Hou, D., Kalnay, E., and Drogemeier, K., “Objective Verification of the SAMEX '98 Ensemble Forecasts”, *Monthly Weather Review*, Volume 129, pp. 73-91, 2001.
92. Kalman, R., “A New Approach to Linear Filtering and Prediction Problems”, *Transactions of the ASME, Ser. D, Journal of Basic Engineering*, Volume 82, pp. 35-45, 1960.
93. Kalman, R., and Bucy, R., “New Results in Linear Filtering and Prediction Theory”, *Transactions of the ASME, Ser. D, Journal of Basic Engineering*, Volume 83, pp. 95-108, 1961.
94. Ott, E., Hunt, B. R., Szunyogh, I., Zimin, A. V., Kostelich, E. J., Corazza, M., Kalnay, E., Patil, D. J., and Yorke, J. A., “A Local Ensemble Kalman Filter for Atmospheric Data Assimilation”, *Tellus*, Volume 56A, pp. 415-428, 2004.
95. Evensen, G., “Sequential Data Assimilation with A Nonlinear Quasigeostrophic Model using Monte Carlo Methods to Forecast Error Statistics”, *Journal of Geophysical Research*, Volume 99 (C5), pp. 10143-10162, 1994.
96. Houtekamer, P. L., and Mitchell, H. L., “Data Assimilation Using An Ensemble Kalman Filter Technique”, *Monthly Weather Review*, Volume 126, pp. 796-811, 1998.

97. Hamill, T. M., Snyder, C., and Morss, R. E., “A Comparison of Probabilistic Forecasts from Bred, Singular-Vector, and Perturbed Observation Ensembles”, *Monthly Weather Review*, Volume 128, pp. 1835-1851, 2000.
98. Houtekamer, P. L., Mitchell, H. L., Pellerin, G., Buehner, M., Charron, M., Spacek, L., and Hansen, B., “Atmospheric Data Assimilation with The Ensemble Kalman Filter: Results with Real Observations”, *Monthly Weather Review*, under review, Volume 133, pp. 604–620, 2005.
99. Tippett, M. K., Anderson, J. L., Bishop, C. H., Hamill, T. M., and Whitaker, J. S., “Ensemble Square Root Filters”, *Monthly Weather Review*, Volume 131, pp. 1485-1490, 2002.
100. Bishop, C. H., Etherton, B. J., and Majumdar, S. J., “Adaptive Sampling with The Ensemble Transform Kalman Filter. Part I: Theoretical aspects”, *Monthly Weather Review*, Volume 129, pp. 420-436, 2001.
101. Anderson, J. L., “An Ensemble Adjustment Filter for Data Assimilation”, *Monthly Weather Review*, Volume 129, pp. 2884-2903, 2001.
102. Alapaty, K., Pleim, J., Raman, S., Niyogi, D., and Byun, D., “Simulation of Atmospheric Boundary Layer Processes using Local and Nonlocal Closure Schemes”, *Journal of Applied Meteorology*, Volume 36, pp. 214-233, 1997.
103. Niyogi, D., and Raman, S., “Comparison of Four Different Stomatal Resistance Schemes using FIFE Observations”, *Journal of Applied Meteorology*, Volume 36, pp. 903 – 917, 1997.

APPENDICES

A HISTORICAL OVERVIEW OF NWP

A.1. THE EARLY HISTORY OF NWP

In late 1945, Vladimir Zworykin, the “Father of Television”, who worked at RCA, joined with John von Neumann, the “Father of the Computer”, to suggest the use of the computer in meteorology. Zworykin’s interest was in weather modification, and von Neumann’s was in fluid dynamics. They also had the dream of connecting the TV and the computer into something we today know as a PC or Workstation. Their dream came partially true in Sweden in around 1955 when for the first time a forecast map that was made directly and automatically without any human intervention was produced on a screen (oscilloscope).

In early 1946, von Neumann contacted Rossby’s group. They told von Neumann why a zonally averaged dynamical model would not work, and instead suggested a barotropic model which had been manually tested by Vector Starr in his 1941 book on weather forecasting for a 72-h forecast at 700hpa. Von Neumann and Zworykin also appeared at the annual meeting of the AMS.

In the summer of 1946, the Princeton meeting took place. Few if any had idea of what should be done. Not even the normally optimistic Rossby could see a solution to the problem. A working group was set up with Albert Cahn and Phil Thompson, with Hans Panofsky and Bernhard Haurwitz acting as advisors. By the autumn of 1946, there was still no clear idea of what to do. Cahn left meteorology to become a successful real estate agent in California, leaving Phil Thompson in despair. It was at this crucial state that Jule Charney moved to Chicago (on his way from Norway). Charney had attended the Princeton meeting, where he had offered some obscure ideas about having the whole atmosphere represented by a few singular levels.

In early 1947, Chaney, now in Oslo, wrote to Phil Thompson that he indeed saw light at the end of the tunnel taking a completely new approach. It is important to realize the practical (political/psychological) impact of L. F. Richardson's 1922 book was essentially to convince the meteorological community that NWP was impossible. This was further supported by the experience Phil Thompson and others had while trying to make use of Jack Bjerknes "tendency equation" (which was as much in vogue then as potential vorticity is today!).

Why Sweden?

The first real-time, operational NWP was run in Sweden in September 1954 (to 72 h at 500 hpa), half a year before the USA.

Two reasons:

- 1) For a short period in 1954, the Swedes were in possession of the world's most powerful computer, BESK. In 1950, they had already constructed a more basic one, BARK. One must again realize the thinking at that time: even among the most radical, it was felt that having just one computer in Sweden for the coming 20-30 years was sufficient. Even in USA the thought that four or five computers would be more than enough for foreseeable future. The "explosion" only came in 1955 when IBM launched their first machine.
- 2) Rossby moved to Sweden and wanted to repeat the ENIAC success of 1950 in his homeland. In this endeavor he was supported by: (a) the Swedish Air force and other national institutions (but not Meteorological Service!); (b) young enthusiastic scientists who worked at or visited his institutions, both Swedish and foreign; (c) the USA Air Force and Woods Hole.

The Swedish project was hampered or complicated by an internal political conflict. In 1954 a new Director of SMHI (the Swedish Meteorological Office) was to be elected by the government. Rossby would have been the obvious choice

but he was seen as a troublemaker. The “official” candidate was Alf Nyberg, who had taken a very skeptical attitude towards Rossby’s project. Against him, Rossby lobbied Herrlin, head of the Military Meteorological Service. Unfortunately the run-up to the selection of new Director coincided with the launch of the first real-time operational NWP, 29 Sep-2 Oct 1954. Those who supported Nyberg took a negative attitude; those who supported Rossby took a positive one. In the end, the government chose Nyberg. SMHI began slowly to support NWP 5 day/week barotropic forecasts to 72 h at 500 hpa started in early December 1954. The US operational NWP started in May 1955, but it was not until 1958 that they reached the same quality standard as the Swedish. Japan started in 1959 along the same lines as Sweden.

A.2. THE EVOLUTION OF FORECAST SKILL

Major milestones of operational numerical weather forecasting include the paper by Charney et al. (1950) with the first successful forecast based on the primitive equations, and the first operational forecasts performed in Sweden in September 1954, followed 6 months later by the first operational (real time) forecasts in the USA. We describe in what follows the evolution of NWP at NCEP, but as mentioned before, similar developments took place at several major operational NWP centers: in the UK, France, Germany, Japan, Australia and Canada.

The history of operational NWP at the NMC (now NCEP) has been reviewed by Shuman (1989) and Kalnay et al. (1998). It started with the organization of the Joint Numerical Weather Prediction Unit (JNWPU) on 1 July, 1954, staffed by members of the US Weather Bureau (later, the National Weather Service, NWS), the Air Weather Service of the US Air Force, and the Naval Weather Service (In 1960 the JNWPU reverted to three separate organizations: the National Meteorological Center (National Weather Service), the Global Weather Central (US Air Force) and the Fleet Numerical Oceanography Center (US Navy)).

Shuman pointed out that in the first few years, numerical predictions could not complete with those produced manually. They had several serious flaw, among them overprediction of cyclone development. Far too many cyclones were predicted to deepen into storms. With time, and with the joint work of modelers and practicing synopticians, major sources of model errors were identified, and operational NWP became the central guidance for operational weather forecasts.

Shuman (1989) included a chart with the evolution of the S1 score (Teweles and Wobus, 1954), the first measure of error in a forecast weather chart which, according to Shuman (1989), was designed, tested, and modified to correlate well with expert forecasters' opinions on the quality of a forecast. The S1 score measures the average relative error in the pressure gradient (compared to a verifying analysis chart). Experiments comparing two independent subjective analyses of the same data-rich North American region made by two experienced analysts suggested that a "perfect" forecast would have an S1 score of about 20%. It was also found empirically that forecasts with an S1 score of 70% more were useless as synoptic guidance.

Shuman pointed out some of the major system improvements that enabled NWP forecasts to overtake and surpass subjective forecasts. The first major improvement took place in 1958 with the implementation of barotropic (one-level) model, which was actually a reduction from the three-level model tried, but which included better finite differences and initial conditions derived from an objective analysis scheme (Bergthorsson and Döös, 1955, Cressman, 1959). It also extended the domain of the model to an octagonal grid covering the Northern Hemisphere down to 9-15° N. These changes resulted in numerical forecasts that for the first time were competitive with subjective forecasts, but in order to implement them JNWPU had to wait for acquisition of a more powerful supercomputers, an IBM 704, to replace the previous IBM 701.

This pattern of forecast improvements which depend on a combination of the better use of the data and better models, and would require more powerful supercomputers in order to be executed in a timely manner has been repeated throughout the history of operational NWP. Table A-1 summarizes the major improvements in the first 30 years of operational numerical forecasts at the NWS. The first primitive equations model was implemented in 1966.

The first regional system (Limited Fine Mesh or LFM model, howcroft, 1971) was implemented in 1971. It was remarkable because it remained in use for over 20 years, and it was the basis of model output Statistics (MOS). Its development was frozen in 1986. A more advanced model and data assimilation system, the Regional Analysis and Forecasting System (RAFS) was implemented as the main guidance for North America in 1982. The RAFS was based on the multiple Nested Grid Model (NGM, Phillips, 1979) and on a regional OI scheme (DiMego, 1988). The global spectral model (Sela, 1980) was implemented in 1980.

Table A-1 Major operational implementation & computer acquisitions at NMC (1955-1985)

Year	Operational Model	Computer
1955	Princeton three-level quasi-geostrophic model (Charney, 1954). Not used by the forecasters.	IBM 701
1958	Barotropic model with improved numerics, objective analysis initial conditions, and octagonal domain.	IBM 704
1962	Three-level quasi-geostrophic model with improved numerics	IBM 7090 (1960) IBM 7094 (1963)
1966	Six-layer primitive equations model (Shuman and Hovermale, 1968)	CDC 6600
1971	LFM model (Howcroft, 1971) (first regional model at NMC)	
1974	Hough functions analysis (Flattery, 1971)	IBM 360/195
1978	Seven-layer primitive equation model (hemispheric)	
1978	OI (Bergman, 1979)	Cyber 205
Aug 1980	Global spectral model, R30/12 layers (Sela, 1980)	
March 1985	Regional Analysis and Forecast System based on the NGM (Philips, 1979) and OI (DiMego, 1988)	

Table A-2 (from Kalnay et. Al, 1998 and P. Caplan, personal communications, 2000) summarizes the major improvements implemented in the global system starting in 1985 with the implementation of the first comprehensive package of

physical parameterization from GFDL (Geophysical Fluid Dynamics Laboratory). Other major improvements in the physical parameterizations were made in 1991, 1993, and 1995. The most important changes in the data assimilation were an improved OI formulation in 1986, the first operational 3DVar in 1991, the replacement of satellite retrievals of temperature with the direct assimilation of cloud-cleared radiances in 1995, and the use of “raw” (not cloud-cleared) radiances in 1998. The first operational ensemble system was implemented in 1992 and enlarged in 1994. The resolution of ensembles was increased in 2000.

Table A-3 contains a summary of the regional systems used for short-range forecasts (up to 48 h). The RAFS (triple nested NGM and OI) were implemented in 1985. The Eta model, designed with advanced finite differences, step-mountain coordinates, and physical parameterizations, was implemented in 1993, with the same 80-km horizontal resolution as the NGM. It was denoted “early” because of a short date cut-off. The resolution was increased to 48 km, and a first “mesoscale” version with 29 km and reduced coverage was implemented in 1995.

A cloud prognostic scheme was implemented in 1995, and a new land-surface parameterization in 1996. The OI data assimilation was replaced by a 3DVar in 1998, and at this time the early mesoEta models were unified into a 32-km/45-level version. Many other less significant changes were also introduced into the global and regional operational systems and are not listed here for the sake of brevity.

Table A-2 Major changes in the NMC/NCEP global model and DA system since 1985

Year	Operational Model	Computer
April 1985	GFDL physics implemented on the global spectral model with silhouette orography, R40/18 layers	
Dec 1986	New OI code with new statistics	
1987		2 nd Cyber 205
Aug 1987	Increased resolution to T80/18 layers, Penman-Montieth evapotranspiration and other improved physics (Caplan and White, 1989, Pan, 1990)	
Dec 1988	Implementation of hydrostatic complex quality control (CQC) (Gadian, 1988)	
1990		Cray YMP/8cpu/32 megawords
Mar 1991	Increased resolution to T126 L18 and improved physics, mean orography. (Kanamitsu <i>et al.</i> , 1991)	
June 1991	New 3DVar(Parrish and Derber, 1992, Derber <i>et al.</i> , 1991)	
Nov 1991	Addition of increments, horizontal and vertical OI checks to the CQC (Collins and Gandian, 1990)	
7 Dec 1992	First ensemble system: one pair of bred forecasts at 00Z to 10 days, extension of AVN to 10 days (Toth and Kalnay, 1993, Tracton and Kalnay, 1993)	
Aug 1993	Simplified Arakawa-Schubert cumulus convection (Pan and Wu, 1995). Resolution T 126/28 layers	
Jan 1994		Cray C90/16cpu/128 megawords
March 1994	Second ensemble system: five pairs of bred forecasts at 00Z, two pairs at 12Z, extension of AVN, a total of 17 global forecasts every day to 16 days	
10 Jan 1995	New soil hydrology (Pan and Mahrt, 1987), radiation, clouds, improved data assimilation. Reanalysis model	
22 Oct 1995	Direct assimilation of TOVS cloud-cleared radiances (Derber and Wu, 1998). New planetary boundary layer (PBL) based on nonlocal diffusion (Hong and Pan, 1996). Improved CQC	Cray C90/16cpu/256 megawords
5 Nov 1997	New observational error statistics. Changes to assimilation of TOVS radiances and addition of other data sources	
13 Jan 1998	Assimilation of noncloud-cleared radiances (Derber <i>et al.</i> , pers.comm.). Improved physics.	
June 1998	Resolution increased to T170/40 layers (to 3.5 days). Improved physics. 3D ozone data assimilation and forecast. Nonlinear increments in 3DVar. Resolution reduced to T 62/28 levels on Oct. 1998 and upgraded back in Jan.2000	IBM S V2 256 Processors
June 2000	Ensemble resolution increased to T 126 for the first 60 h	
July 2000	Tropical cyclones relocated to observed position every 6 h	

Table A-3 Major changes in the NMC/NCEP regional modeling and DA since 1985

Year	Operational Model	Computer
March 1985	RAFS based on triply NGM (Phillips, 1979) and OI (DiMego, 1988). Resolution: 80 km/16 layers.	Cyber 205
August 1991	RAFS upgraded for the last time: NGM run with only two two grids with inner grid domain doubled in size. Implemented Regional Data Assimilation System (RDAS) with three-hourly updates using an improved OI analysis using all off-time data including Profiler and Aircraft Communication Addressing and Reporting System (ACARS) wind reports (DiMego <i>et al.</i> , 1992) and CQC procedures (Gandian <i>et al.</i> , 1993).	Cray YMP 8 processors 32 megawords
June 1993	First operational implementation of the Eta model in the 00Z & 12Z early run for North America at 80-km and 38-layer resolution (Mesinger <i>et al.</i> , 1988, Janjic, 1994, Black <i>et al.</i> , 1993)	
September 1994	The RUC (Benjamin <i>et al.</i> , 1996) was implemented for CONUS domain with three-hourly OI updates at 60-km resolution on 25 hybrid (sigma-theta) vertical levels.	Cray C-90 16 processors 128 megawords
September 1994	Early Eta analysis upgrades (Rogers <i>et al.</i> , 1995)	
August 1995	A mesoscale version of the Eta model (Black, 1994) was implemented at 03Z and 15Z for an extended CONUS domain, with 29-km and 50-layer resolution and with NMC's first predictive cloud scheme (Zhao and Black, 1994) and new coupled land-surface-atmosphere package (two-layer soil).	Cray C-90 16 processors 256 megawords
October 1995	Major upgrade of early Eta runs: 48-km resolution, cloud scheme and Eta Data Assimilation System (EDAS) using three-hourly OI updates (Rogers <i>et al.</i> , 1996)	
January 1996	New coupled land-surface-atmosphere scheme put into early Eta runs (Chen <i>et al.</i> , 1997, Mesinger, 1997)	
July-August 1996	Nested capability demonstrated with twice-daily support runs for Atlanta Olympic Games with 10-km 60-layer version of Meso Eta.	
February 1997	Upgrade package implemented in the early and Meso Eta runs.	
February 1998	Early Eta runs upgraded to 32 km and 45 levels with four soil layers. OI analysis replaced by 3DVar with new data sources. EDAS now partially cycled (soil moisture, soil temperature, cloud water/ice & turbulent kinetic energy)	
April 1998	RUC (three-hourly) replaced by hourly RUC II system with extended CONUS domain, 40-km and 40-level resolution, additional data sources and extensive physics upgrades.	
June 1998	Meso runs connected to early runs as a single 4/day system for North America domain at 32-km and 45-level resolution, 15Z run moved to 18Z, added new snow analysis. All runs connected with EDAS, which is fully cycled for all variables.	IBM S V 2 256 processors

B THE MM5 PHYSICS OPTIONS

B.1. CUMULUS PARAMETERIZATION

A significant part of precipitation, especially during the warm season, is convective in nature. Since the model does not explicitly resolve and predict convection on the space and time scales where they occur, the effects of convection on the resolvable scales of the model must be parameterized. It must try to predict convection and account for the collective influence of small-scale convective processes on large-scale model variables in each grid box. Procedures to do this are called cumulus or convective parameterization schemes.

When deep convection occurs in the atmosphere, it strongly affects both meso and larger-scale dynamics. The effects of the atmospheric convection on vertical stability by redistributing heat, moisture, and momentum will affect the formation of temperature inversion. Additionally, the cloud cover associated with convection strongly affects both surface heating and other radiative processes.

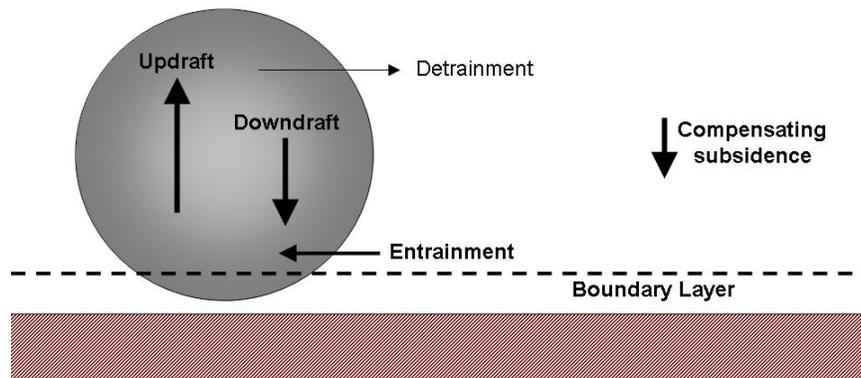


Figure B-1 Schematic for cumulus processes in atmosphere used in MM5 model

Figure B-1 shows a schematic for the cumulus processes modeled in the MM5 model. The different schemes and options for cumulus parameterization are summarized in Table B-1.

Table B-1 Cumulus parameterization schemes and options

Cumulus Parameterization	Description	Options	Schemes	Applicability
	<ul style="list-style-type: none"> • Represent sub-grid scale vertical fluxes and rainfall due to convective clouds • Generally produce column moisture and temperature tendencies and surface convective rainfall • May also produce column cloud tendencies (KF schemes) • Require trigger to determine where convection activates, and closure to determine strength 	<p>ICUPA represents the scheme</p>	<p>1. None</p>	<ul style="list-style-type: none"> • No cumulus scheme required if grid size is sufficient to resolve updrafts and downdrafts • May apply to grid lengths less than 5 km
		<p>Shallow Convection (ISHALLOW=1):</p> <ul style="list-style-type: none"> • Handles non-precipitating clouds • May help PBL-top clouds to mix • Not clear cost of this scheme is justified by its small effect on results • Adapted from Grell scheme • Updrafts with high entrainment rate • Driven by PBL tendencies only (not total rate of destabilization) 	<p>2. Anthes-Kuo</p>	<ul style="list-style-type: none"> • Oldest scheme in model • Moisture convergence closure • Tends to produce much convective rainfall • Less resolved-scale precip • Specified heating profile • Moistening depends on environment RH • Applicable to larger grid sizes (> 30 km)
		<p>3. Grell</p>	<ul style="list-style-type: none"> • Rate of destabilization closure (quasi-equilibrium) • simple single-cloud scheme with updraft and downdraft fluxes • Mass-flux type scheme with compensating subsidence • Tends to allow a balance between resolved scale rainfall and convective rainfall • Useful for smaller grid sizes 10-30 km 	
		<p>4. Arakawa-Schubert</p>	<ul style="list-style-type: none"> • Quasi-equilibrium closure • Multi-cloud scheme with updrafts and downdrafts (added by Grell to original scheme) • Suitable for larger grid sizes (>30 km) • Possibly expensive compared to other schemes. • Requires a library (not portable from Cray very easily) 	
		<p>5. Fritsch-Chappell</p>	<ul style="list-style-type: none"> • Old scheme: forerunner to Kain-Fritsch • Based on releasing instability (CAPE) over a given time scale • Updrafts and downdrafts represented • Mass-flux type scheme with compensating subsidence • Suitable for 20-30 km grids • Not used much since KF scheme became available 	
		<p>6. Kain-Fritsch</p>	<ul style="list-style-type: none"> • Uses sophisticated cloud-mixing scheme to determine updraft/downdraft properties • Releases CAPE in a given time scale • Mass-flux scheme • Also can detrain cloud and precipitation in addition to vapor 	
		<p>7. Betts-Miller</p>	<ul style="list-style-type: none"> • Relaxation adjustment to a post-convective mixed sounding in a given time scale • More suited to tropics but can be used anywhere. (Comes from Eta model BMJ scheme) • No explicit downdrafts (some surface cooling due to adjustment) • Suitable for > 30 km, but no explicit downdraft, so may not be suitable for severe convection 	
		<p>8. Kain-Fritsch 2</p>	<ul style="list-style-type: none"> • New scheme as of MM5 v3.5 • Adds shallow convection and other improvements to KF scheme 	

B.2. PBL AND DIFFUSION SCHEMES

The lowest layer of the atmosphere is called the troposphere. The troposphere can be divided into two parts. The first part is the Planetary Boundary Layer, PBL, extending upward from the surface to a height that ranges anywhere from 100 to 3000 m. The second part is the free atmosphere. The PBL is directly influenced by the presence of the Earth's surface, responding to such forcings as frictional drag, solar heating, and evapotranspiration. Each of these forcings generates turbulence of various-sized eddies, which can be as deep as the PBL itself, lying on top of each other. Therefore, the PBL must be parameterized in the model as a mechanism for turbulence.

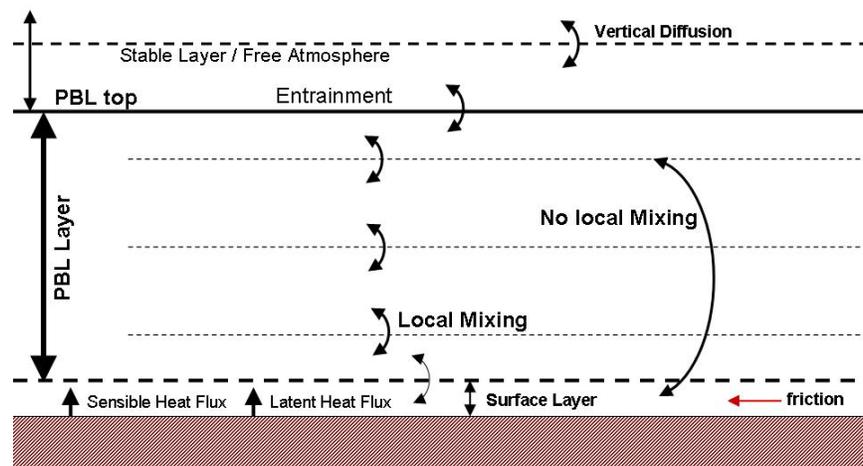


Figure B-2 Schematic for PBL used in the MM5 model

Figure B-2 shows a schematic for PBL processes in the MM5 model. The different schemes and options for PBL and diffusion are summarized in Table B-2. The Medium Range Forecast (MRF) and Blackadar PBL schemes have the option of using a different roughness length for heat/moisture than that used for momentum. Changing the thermal roughness length affects the partitioning of sensible and latent heat fluxes, and affects the total flux over water.

Table B-2 PBL and diffusion schemes and options

Planetary Boundary Layer (PBL) and Diffusion	Description	Options	Schemes	Applicability
	<ul style="list-style-type: none"> • Represent sub-grid vertical fluxes due to turbulence. Mostly distinguished by treatment of the unstable boundary layer. • Generally provide column tendencies of heat, moisture and momentum • May provide cloud tendencies • Surface layer, boundary layer, and free atmosphere • Interacts with fluxes from surface scheme • Provides frictional effects on momentum 	<p>IBLTYP represents the scheme</p> <p>Moist Vertical Diffusion (IMVDIF):</p> <ul style="list-style-type: none"> • Available only in Blackadar/MRF PBL • Default IMVDIF=1 accounts for vertical mixing in saturated layers • Produces moist-adiabatic mixed profile <p>Thermal Roughness Length (IZOTOPT):</p> <ul style="list-style-type: none"> • Only available with MRF and Blackadar schemes • Different treatments of thermal roughness length due to Garrett and Zilitinkevich • Affects sensible and latent heat flux, especially over water <p>IZOTOPT=0 is the original Carlson-Boland formulation IZOTOPT=1 is the Garratt formulation IZOTOPT=2 is the Zilitinkevich formulation</p> <p>Horizontal diffusion - (ITPDIF=0,1,2): ITPDIF=0 diffuses the full temperature (like all other fields) ITPDIF=1 (default) is to only horizontally diffuse the perturbation from the base-state temperature. ITPDIF=2 applies to temperature, moisture and cloud water, and is a purely horizontal diffusion accounting more accurately for coordinate slope and valley walls</p>	0.. None	<ul style="list-style-type: none"> • No surface layer • unrealistic in real-data simulations
			1. Bulk PBL	<ul style="list-style-type: none"> • Designed for coarse vertical resolution (dz > 250 m) • Stable and unstable regimes • Bulk aerodynamic drag and exchange coefficients
			2. High-Resolution (Blackadar) PBL	<ul style="list-style-type: none"> • Suitable for multi-layer PBL (e.g. 5 layers in lowest km) • Four stability regimes • Unstable regime has nonlocal mixing between surface layer and all other layers in PBL • PBL depth determined from temperature profile • Entrainment at PBL top due to overshooting thermals • Monin-Obukhov similarity theory for surface exchange coefficients
			3. Burk-Thompson PBL	<ul style="list-style-type: none"> • Also known as Navy PBL • Mellor-Yamada scheme • Predicts turbulent kinetic energy • Local vertical mixing • Has its own force-restore ground temperature routine (does not call SLAB) • Louis scheme for surface exchange coefficients
			4. Eta PBL	<ul style="list-style-type: none"> • Also known as Mellor-Yamada-Janjic PBL • Uses Mellor-Yamada • Predicts TKE • Local vertical mixing • Monin-Obukhov similarity theory • Can be used with Noah-LSM (v3.6)
			5. MRF PBL	<ul style="list-style-type: none"> • Also known as Hong and Pan PBL • Suitable for high-resolution in PBL • Based on Troen-Mahrt concept of nonlocal mixing (countergradient term and K profile in the well mixed PBL) • PBL depth determined from critical bulk Richardson number (shear and temperature profile) • Can be used with Noah LSM • Vertical diffusion uses an implicit scheme to allow longer time steps.
			6. Gayno-Seaman PBL	<ul style="list-style-type: none"> • Predicts TKE • Allows for cloud-topped PBL processes by using liquid water potential temperature and total water as its mixing variables
			7. Pleim-Chang PBL	<ul style="list-style-type: none"> • Currently can only be used with Pleim-Xiu LSM • Based on Blackadar scheme • Asymmetric Convective Model

B.3. EXPLICIT MOISTURE SCHEMES

Cloud and rain water fields predicted explicitly with microphysical processes. The different explicit moisture schemes are summarized in Table B-3. In this study, simple ice and mixed-phase schemes were used.

Table B-3 Explicit Moisture schemes

Explicit Moisture	Description	Options	Schemes	Applicability
	<ul style="list-style-type: none"> • Treatment of cloud and precipitation processes on the resolved scale • Process rates assume uniform grid-box • May or may not include ice phase and graupel/hail particles • Provides tendencies of temperature, and all moist variables, and surface non-convective rainfall • Provides information on clouds to radiation schemes 	<p>IMPHYS represents the scheme</p>	<p>1. Dry</p> <p>2. Stable Precipitation</p> <p>3. Warm Rain (Hsie)</p> <p>4. Simple Ice (Dudhia)</p> <p>5. Mixed-Phase (Reisner 1)</p> <p>6. Goddard microphysics</p> <p>7. Reisner 2 (graupel)</p> <p>8. Schultz</p>	<ul style="list-style-type: none"> • No vapor or clouds • If you want vapor as a passive advected variable, better to use IFDRY=1 (Fake dry) which turns off only latent heating, and is better for sensitivity studies. • Also known as the Nonconvective Rainfall scheme • Grid-scale saturation removed and immediately put into surface rainfall • No explicit clouds or rain evaporation • Namelist parameter CONF can be used to control maximum RH allowed • Predicts TKE • Allows for cloud-topped PBL processes by using liquid water potential temperature and total water as its mixing variables • Also known as Dudhia scheme • Adaptation of Hsie scheme to allow ice processes • Cloud and ice share one array, rain and snow share another. No additional memory. • Ice sedimentation • No supercooled water • Immediate snow melt at melting layer • Also known as Reisner 1 • Adds arrays for cloud ice and snow • Has same processes as Simple Ice • Treats supercooled water • Has gradual snow melt as it falls • Sophisticated scheme with graupel/hail as an additional variable • Suitable for cloud-resolving models • Also known as Reisner 2 • Additional variables for graupel and ice number concentration • Many differences in detail from Reisner 1 • Used in FSL's RUC runs • Still being developed by R. Rasmussen, J. Brown and G. Thompson • 3.4, 3.5 and 3.6 versions contain significant differences from each other • Also contains graupel field • Simple scheme designed for efficiency and tunability with a minimum number of parameters • Not well suited to vector machines

B.4. RADIATION SCHEMES

The radiation scheme is used to model the solar and atmospheric radiation in the MM5 model. Figure B-3 shows the radiation processes such as longwave radiation, shortwave radiation, atmospheric absorption and surface emissivity.

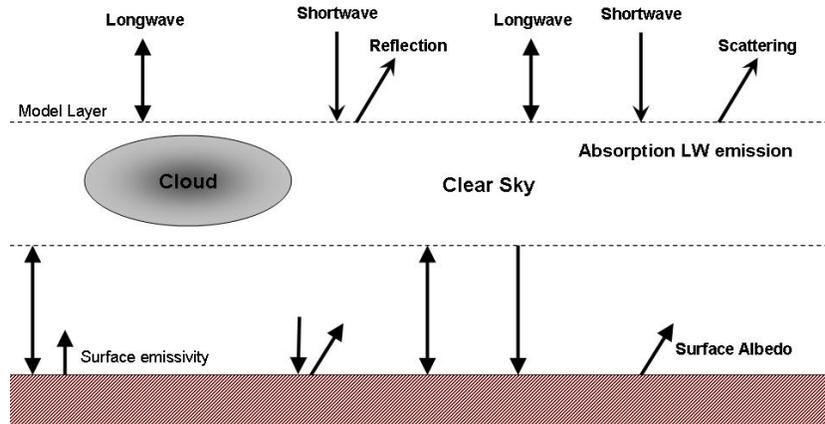


Figure B-3 Schematic for radiation processes in the MM5 model

Table B-4 shows the radiation schemes and options in the MM5 model. In this study, Cloud, CCM2 and RRTM longwave radiation schemes were used.

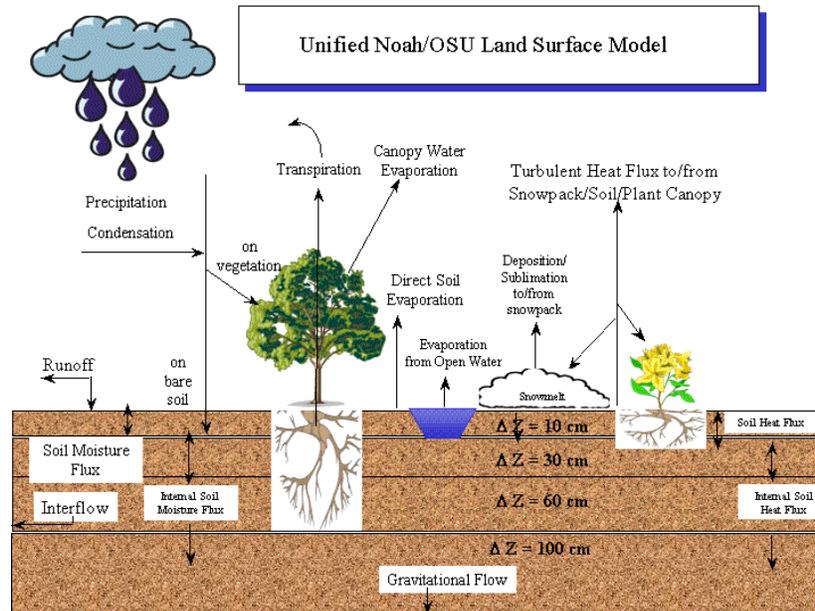
Table B-4 Radiation schemes and options

	Description	Options	Schemes	Applicability
Radiation	<ul style="list-style-type: none"> • Represent radiative effects in atmosphere and at surface • Provides surface downwelling longwave and shortwave fluxes • Provides column temperature tendencies due to vertical radiative flux divergence • May interact with model clouds or RH 	IFRAD represents the scheme Surface radiation: <ul style="list-style-type: none"> • Used if IFRAD=0/1 • Surface shortwave and longwave flux provided for ground temperature prediction • Uses column integrated water vapor • Uses RH to determine low/mid/high cloud fractions • Suitable for very coarse grids (> 50 km), or if microphysics is not being used 	0. None	<ul style="list-style-type: none"> • No radiation effects in the atmosphere • Surface radiation still active
			1. Simple Cooling	<ul style="list-style-type: none"> • Climatological mean cooling profile in the atmosphere • No diurnal dependence • Only a function of temperature • Surface radiation is active
			2. Cloud radiation	<ul style="list-style-type: none"> • Also known as Dudhia scheme • Provides atmospheric radiative effects due to modeled clouds • Provides surface longwave and shortwave fluxes itself (does not call Surface Radiation scheme)
			3. CCM2 radiation	<ul style="list-style-type: none"> • From CCM2 climate model (old scheme) • Better suited to coarse grid sizes and long time integrations • Interacts either with RH or with model clouds
			4. RRTM longwave radiation	<ul style="list-style-type: none"> • Rapid Radiative Transfer Model (AER, Inc.) • Sophisticated look-up table scheme for longwave radiation • Interacts with model clouds • Used with Dudhia shortwave scheme

B.5. SURFACE SCHEMES

The Land Surface Processes (LSP) plays an important role in atmospheric modeling not only on micro and regional scales but also on global scale. They are critical in influencing the PBL structure and thickness. The LSP are controlled by the energy exchange between the atmosphere and the land surface. In general, the LSP are parameterized in terms of turbulent fluxes of momentum, energy and moisture. The turbulent fluxes have two types which are; the mechanical turbulence (due to wind shear) and thermal turbulence (due to buoyancy).

The available surface schemes in MM5 are summarized in Table B-5. NOAH LSM has different processes for the biophysical, biogeochemical, hydrologic, and echo systems that interact with the atmosphere through boundary layer. These processes are used to predict the ground surface temperature and evaporation from open water areas. Figure B-4 shows a schematic for LSP in NOAH LSM.



NOAH LSM prognostic equations for soil moisture and temperature:

$$\frac{\partial \theta}{\partial t} = \frac{\partial}{\partial z} \left(D \frac{\partial \theta}{\partial z} \right) + \frac{\partial K}{\partial z} + F_{\theta} \quad \text{and} \quad C(\theta) \frac{\partial T}{\partial t} = \frac{\partial}{\partial z} \left(K_t(\theta) \frac{\partial T}{\partial z} \right)$$

D, K : functions (soil texture), F_{θ} : sources (infiltration) and sinks (evaporation) and C, K_t : functions (soil texture, soil moisture)

Figure B-4 Schematic for NOAH Land Surface Model

Table B-5 Surface schemes and options

	Description	Options	Schemes	Applicability
Surface	<ul style="list-style-type: none"> • Represent effects of land and water surfaces • Ground temperature based on heat budget using radiative fluxes and surface-layer atmospheric properties • Provides sensible and latent heat flux • May also represent sub-soil temperature and moisture profiles • May provide snow-cover tendencies and surface moisture availability variation 	<p>ISOIL represents the scheme</p> <p>Bucket Soil Moisture Model (IMOIAV=1,2): This can be run with ISOIL=0 or 1. It keeps a budget of soil moisture allowing moisture availability to vary with time, particularly in response to rainfall and evaporation rates. The soil moisture can be initialized from land-use type and season (LANDUSE.TBL) as before (IMOIAV=1), or a 10-cm soil moisture input as with the Noah LSM (IMOIAV=2).</p> <p>Snow Cover Model (IFSNOV=0,1,2): When the LSM is not used this switch determines how snow cover is handled. This can be used with IMOIAV=1 or 2, the bucket soil moisture. IFSNOV=0 means snow cover is ignored. IFSNOV=1 uses the input snow-cover (IFSNOV=2) to predict snow cover using an input water-equivalent snow depth.</p> <p>Polar Physics (IPOLAR=1):</p> <ul style="list-style-type: none"> • Uses 7-layers with ISOIL=1 soil model • Takes into account snow/ice ground properties • Accounts for sea-ice fraction (IEXSI switch) • Modifies simple-ice and Reisner 1 microphysics to use Meyers ice number conc formula • Should be used with Eta PBL 	<p>0.Force-restore (Blackadar)</p> <p>1.Five-layer Soil Temperature</p> <p>2.Noah LSM</p> <p>3.Pleim-Xiu LSM</p>	<ul style="list-style-type: none"> • Ground temperature prediction • 2-layer model with a constant-temperature substrate • Substrate (reservoir) temperature is specified in INTERPF as a diurnal average surface temperature • Tuned to represent diurnal cycle best • Predicts soil temperature in five layers • 1, 2, 4, 8, 16 cm thick • Can represent higher frequency changes than force-restore • Formerly Oregon State University (OSU) LSM • Same as that used by NCEP and AFWA in operational models • Four layers (10, 30, 60 and 100 cm thick) • Predicts soil temperature, soil water/ice, canopy water, and snow cover • Needs additional inputs of soil texture, annual mean surface temperature, and seasonal vegetation fraction, as well as initial soil temperature and moisture • Simple 2-layer model • Predicts soil temperature and soil moisture • Can use data assimilation to initialize soil moisture • Used at EPA

C GLOBAL OBSERVING SYSTEM

C.1. SURFACE OBSERVATIONS

The backbone of the surface-based sub-system continues to be about 11,000 stations on land making observations at or near the Earth's surface, at least every three hours and often hourly, of meteorological parameters such as atmospheric pressure, wind speed and direction, air temperature and relative humidity. Some 4000 of these stations comprise the Regional Basic Synoptic Networks (RBSNs) and over 3000 stations comprise the Regional Basic Climatological Networks (RBCNs) both drawn up by the six WMO Regional Associations. Data from these stations are exchanged globally in real time. A subset of these surface stations are used in Global Climate Observing System (GCOS) Surface Network (GSN).

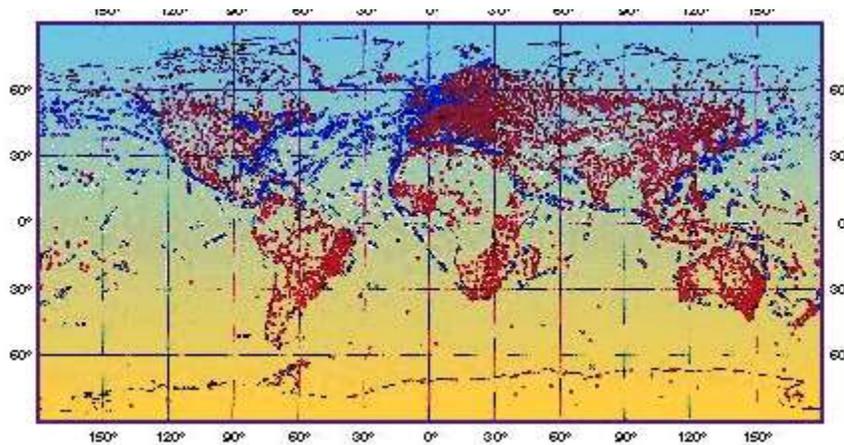


Figure C-1 WMO global surface observations.

C.2. UPPER-AIR OBSERVATIONS

From a global network of about 1,300 upper-air stations, radiosondes, attached to free-rising balloons, make measurements of pressure, wind velocity, temperature and humidity from just above ground to heights of up to 30km. Over two thirds of the stations make observations at 0000UTC and 1200UTC. Between 100 and 200

stations make observations once per day. In ocean areas, radiosonde observations are taken by about 15 ships, which mainly ply the North Atlantic, fitted with Automated Shipboard Aerological Programme (ASAP). A subset of upper-air stations comprises the GCOS Upper-air Network (GUAN).

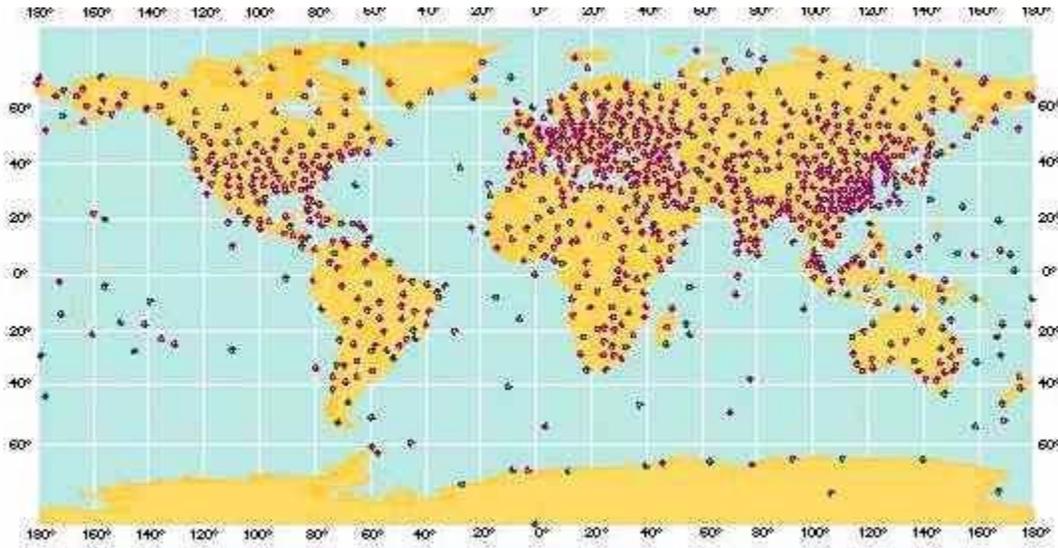


Figure C-2 WMO global upper-air observations

C.3. MARINE OBSERVATIONS

Over the oceans the Global Observing System (GOS) relies - in addition to satellites - on ships, moored and drifting buoys and stationary platforms. Observations made by ships recruited under the WMO Voluntary Observing Ship (VOS) Program, comprise much the same variables as at surface land stations with the important additions of sea surface temperature, wave height and period. The number of observing ships is around 4,000. About 40% are at sea at any given time. The operational drifting buoy program comprised of about 1,200 drifting buoys provides over 12,000 sea surface temperature and surface air pressure reports per day. These ships and buoys are part of the WMO Marine Program, which maintains lists of ships and observing standards.

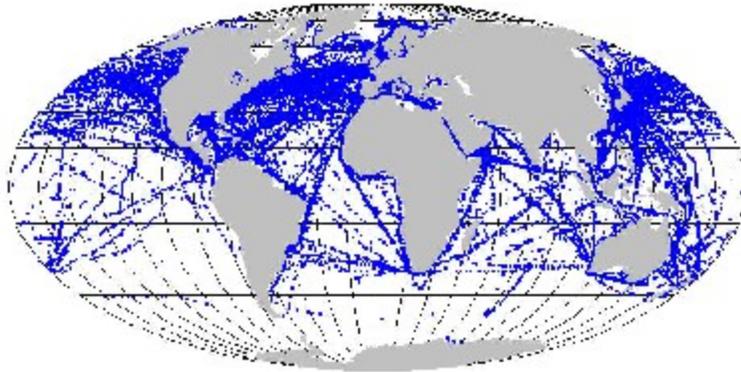


Figure C-3 WMO global marine observations

C.4. AIRCRAFT OBSERVATIONS

Over 3000 aircraft provide reports of pressure, winds and temperature during flight. The Aircraft Meteorological Data Relay (AMDAR) system makes high quality observations of winds and temperatures at cruising level as well as at selected levels in ascent and descent. The amount of data from aircraft has increased dramatically during recent years - from 78,000 reports in 2000 to almost 300,000 reports in 2005. Providing great potential for measurements in places where there is little or no radiosonde data, these systems are making a major contribution to the upper-air component of the GOS. See the WMO Aeronautical Program for more on aviation meteorology.

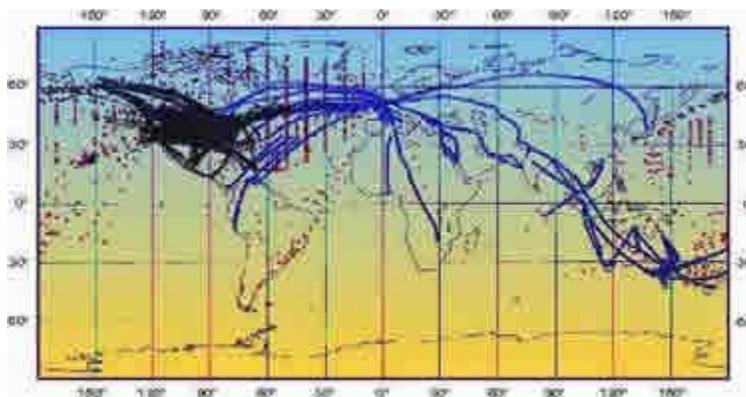


Figure C-4 WMO global aircraft observations

C.5. SATELLITE OBSERVATIONS

The Environmental Observation Satellite network included three operational near-polar-orbiting satellites and six operational geostationary environmental observation satellites as well as several Research and Development satellites. Polar orbiting and geostationary satellites are normally equipped with visible and infra-red imagers and sounders, from which one can derive many meteorological parameters. Several of the polar-orbiting satellites are equipped with sounders instruments that can provide vertical profiles of temperature and humidity in cloud free areas. Geostationary satellites can be used to measure wind velocity in the tropics by tracking clouds and water vapor. Research and Development (R&D) satellites comprise the newest constellation in the space-based component of the GOS. R&D missions provide valuable data for operational use as well as for many WMO supported programs. Instruments on R&D missions either provide data not normally observed from operational meteorological satellites or improvements to current operational systems. Figure C-5 shows the current and planned space-based component of the GOS.



Figure C-5 Current and planned space-based component of the GOS

C.6. OTHER OBSERVATION PLATFORMS

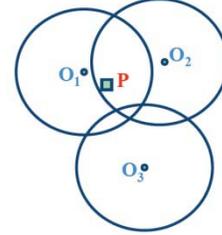
GOS also includes other observation platforms like solar radiation observations, lightning detection, and tide-gauge measurements. In addition, wind-profiling and Doppler radars are proving to be extremely valuable in providing data of high resolution in both space and time, especially in the lower layers of the atmosphere. Wind profilers are especially useful in making observations at times between balloon-borne soundings, and have great potential as a part of integrated networks. Doppler radars are used extensively as part of national, and increasingly of regional networks, mainly for short range forecasting of severe weather phenomena. Particularly useful is the Doppler radar capability of making wind measurements and estimates of rainfall amounts.

D CRESSMAN ANALYSIS

D.1. STANDARD CRESSMAN SCHEME

The circular weighting function:

$$w_{ijk} = \begin{cases} \frac{R^2 - d_{ijk}^2}{R^2 - d_{ijk}^2} & \text{for } d_{ijk} < R \\ 0 & \text{for } d_{ijk} \geq R \end{cases}$$



• Observations ■ Grid Point

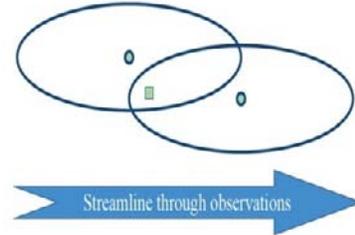
R is an arbitrary but constant radius of influence,

d_{ijk} is the distance between the (i, j) grid point to the k^{th} observation point,

D.2. ELLIPSE CRESSMAN SCHEME

The elliptical weighting function:

$$w_{ijk} = \begin{cases} \frac{R^2 - d_m^2}{R^2 - d_m^2} & \text{for } d_m^2 < R^2 \\ 0 & \text{for } d_m^2 \geq R^2 \end{cases}$$



• Observations ■ Grid Point

$$d_m^2 = \left(\frac{\hat{x}^2}{E_k^2 |V|} + \hat{y}^2 \right)$$

$$x' = (D_{ijk} \cdot V_k) / |V_k| = d_{ijk} \cos \theta = [(x_{ij} - x_k)u_k + (y_{ij} - y_k)v_k] / |V_k|$$

$$y' = \mathbf{k} \cdot (D_{ijk} \times V_k) / |V_k| = d_{ijk} \sin \theta$$

$$= [(x_{ij} - x_k)v_k - (y_{ij} - y_k)u_k] / |V_k|$$

$$E_k^2 |V| = (1 + \beta |V_k|)^{1/2}$$

β is a constant of elongation (0.02-0.20 s/m),

R is radius of influence in the direction normal to the wind at the observation point,

d_{ijk} is the distance between the (i, j) grid point to the k^{th} observation point,

x, y is the Cartesian coordinate position of grid point (i, j) or observation point (k) ,

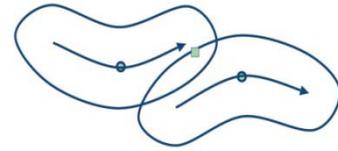
$V_k = \mathbf{i}u_k + \mathbf{j}v_k$ is the horizontal wind vector in Cartesian coordinates for the k^{th} observation point,

$\mathbf{i}, \mathbf{j}, \mathbf{k}$ are unit vectors for two horizontal and one vertical dimensions.

D.3. BANANA CRESSMAN SCHEME

The banana weighting function:

$$w_{ijk} = \begin{cases} \frac{R^2 - d_m^2}{R^2} & \text{for } d_m^2 < R^2 \\ 0 & \text{for } d_m^2 \geq R^2 \end{cases}$$



• Observations ■ Grid Point

$$d_m^2 = \left(\frac{r_k^2 (\theta_k - \theta_{ij})^2}{E_k^2 |V|} + (|r_k| - r_{ij})^2 \right) = \left(\frac{\dot{x}_b^2}{E_k^2 |V|} + \dot{y}_b^2 \right)$$

$$E_k^2 |V| = (1 + \beta |V_k|)^{1/2}$$

β is a constant of elongation (0.02-0.20 s/m),

R is an arbitrary radius of influence in the cross-flow direction at observation point (k) ,

r_k is the streamline radius of curvature at observation point (k) ,

r_{ij} is the distance from C_k to grid point (i, j) ,

C_k is the streamline center of curvature for observation point (k) ,

θ_k is the azimuthal angle at center of curvature (C_k) counterclockwise from the x-coordinate direction to the observation point (k) .

θ_{ij} is the azimuthal angle at center of curvature (C_k) counterclockwise from the x-coordinate direction to the grid point (i, j)

E OPERATIONAL ENSEMBLE FORECASTING

E.1. BREEDING

Breeding is a nonlinear generalization of the method to obtain leading Lyapunov Vectors (LVs), which are the sustained fastest growing perturbations. The Bred Vectors (BVs) are the differences between two nonlinear integrations while LVs are linear approximations of these differences. Figure E-1 shows how all perturbations will converge towards the leading local LV (LLV). BVs are related to LVs localized in both space and time and are independent of the norm of the initial perturbations. They find the most unstable normal modes in an evolving flow and may be used for “targeted observations”. In areas where the evolving flow is very unstable (and where forecast errors grow fast), the BVs tend to align themselves along very low dimensional subspaces (the locally most unstable perturbations).

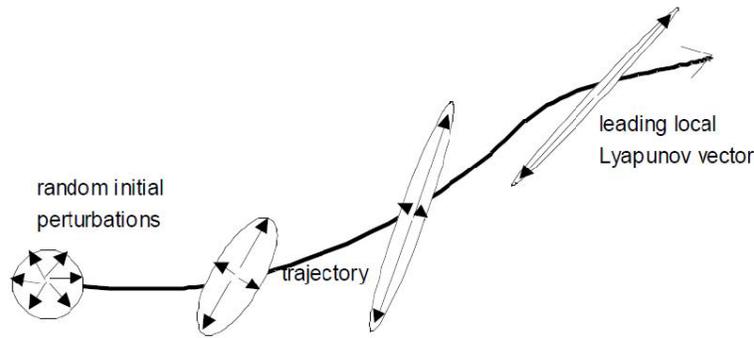


Figure E-1 Schematic for how all perturbations will converge towards the LLV

Consider an evolving basic solution $x(x, t)$, which is a function of space x and time t , that satisfies the equations of a nonlinear model. The discretization in space and time integration scheme is given by $x(x, t + \Delta t) = M(x(x, t))$. If the initial condition is perturbed, the linear evolution of the perturbation is given by

$$\delta x(x, t + \Delta t) = L(x, t + \Delta t)\delta x(x, t), \quad L = \frac{\partial M}{\partial x} \quad (\text{E.1})$$

Equation (E.1) is the Tangent Linear Model (TLM) and the matrix L is the propagator of the TLM. The leading LV is computed as follows:

- 1) Start with an arbitrary perturbation $\delta x(x, t)$ of arbitrary size
- 2) Evolve it from t to $t + \Delta t$ using the TLM (equation (E.1))
- 3) Repeat 2) for the succeeding time intervals

After a sufficiently long time $t_n = t + n\Delta t$, $n \rightarrow \infty$, the evolved perturbation $\delta x(x, t + n\Delta t)$ converges to the leading LV. If during the repeated application of the TLM the LV becomes too large it may be scaled down. Additional LVs can be obtained by the same procedure, except that after each time step the perturbation has to be orthogonalized with respect to the subspace of the previous LVs.

The BVs are computed as follows:

- 1) Start with an arbitrary initial perturbation $\delta x(x, t)$ of size A defined with an arbitrary norm. A is essentially the only tunable parameter of breeding.
- 2) Add the perturbation to the basic solution, integrate the perturbed initial condition with the nonlinear model, and subtract the original unperturbed solution from the perturbed nonlinear integration

$$\overline{\delta x(x, t + \Delta t)} = M[x(x, t) + \delta x(x, t)] - M[x(x, t)] \quad (\text{E.2})$$

- 3) Measure the size $A + \delta A$ of the evolved perturbation $\overline{\delta x(x, t + \Delta t)}$, and divide the perturbation by the measured amplification factor so that its size remains equal to A :

$$\delta x(x, t + \Delta t) = \overline{\delta x(x, t + \Delta t)} A / (A + \delta A) \quad (\text{E.3})$$

Steps 2) and 3) are repeated for the next time interval and so on. Additional BVs can be obtained by choosing different arbitrary initial perturbations and following the same procedure. Therefore all BVs are related to the leading LV, since the additional BVs are never orthogonalized. For global atmospheric models, the BVs remain distinct, rather than converging to a single leading BV.

A breeding cycle is started by introducing a random initial perturbation (which is denoted by "random seed"). The random seed is introduced only once. The local breeding growth rate is given by

$$g(t) = \frac{1}{n\Delta t} \ln \left(\frac{|\delta x(x, t + n\Delta t)|}{|\delta x(x, t)|} \right) \quad (\text{E.4})$$

Beyond an initial transient period of 3-4 days, the BVs acquire a large growth rate, faster than the growth rate for MCF, SLAF and forecast differences. Figure E-2 shows a schematic for a breeding cycle run upon an unperturbed integration.

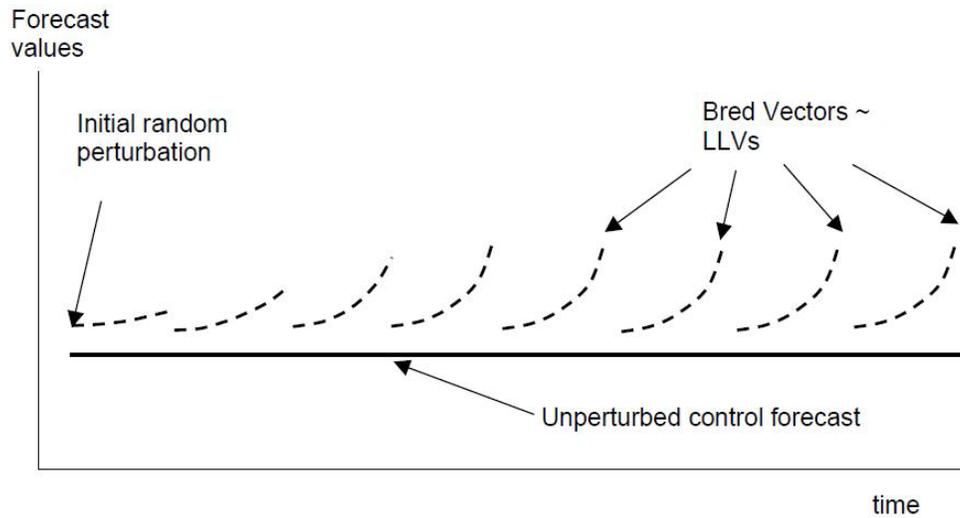


Figure E-2 Schematic for a breeding cycle run upon an unperturbed model integration

An alternative breeding method is “self-breeding” which is cost-free when performing ensemble forecasting. This approach uses pairs of ensemble forecasts to generate the perturbation at the new time:

$$\overline{\delta x(x, t + \Delta t)} = 1/2 \{M[x(x, t) + \delta x(x, t)] - M[x(x, t) - \delta x(x, t)]\} \quad (\text{E.5})$$

This difference is scaled down as before, and added and subtracted to the analysis valid at $t + \Delta t$. The two-sided self-breeding has the advantage that it maintains the linearity of the perturbation to second order compared to the one-sided generation of the bred vector which is linear to first order, but otherwise the procedures

produce similar results. Figure E-3 shows a schematic of a self-contained breeding pair of ensemble forecasts

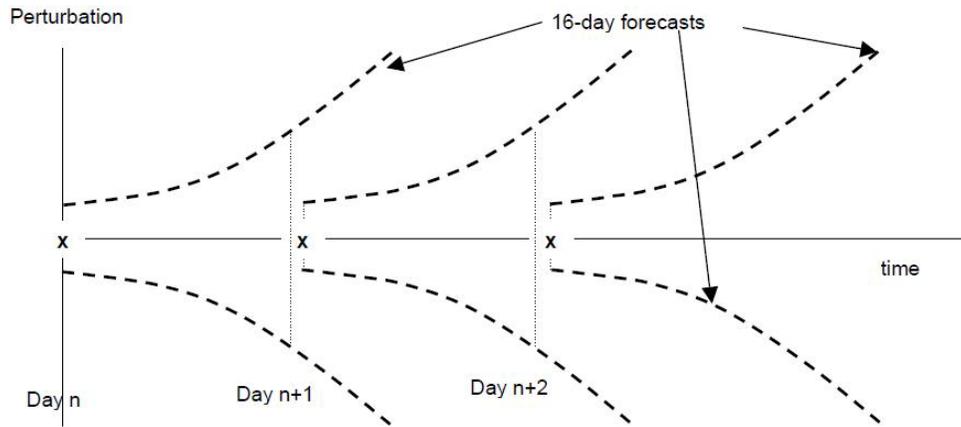


Figure E-3 Schematic of a self-contained breeding pair of ensemble forecasts

NCEP implemented an ensemble system based on breeding in 1992. Later, the US Navy, the National Centre for Medium Range Weather Forecasting in India, and the South African Meteorological Weather Service implemented similar EF systems. The Japanese Meteorological Agency EF system is also based on breeding, but imposing a partial global orthogonalization among the BVs, thus reducing the tendency of the BVs to converge towards a low dimensional space of the most unstable directions.

E.2. SINGULAR VECTORS

Singular Vectors (SVs) are the linear perturbations of a control forecast that grow fastest within a certain time interval, known as “optimization period”, using a specific norm to measure their size. SVs are strongly sensitive to the length of the interval and to the choice of norm. If the initial norm used to derive the SVs is the analysis error covariance norm, the initial SVs evolve into the eigenvectors of the forecast error covariance at the end of the optimization period. This indicates that if the analysis error covariance is known, then SVs based on this specific norm are the ideal perturbations.

Singular Value Decomposition (SVD) theory indicates that for any matrix such as the propagator matrix L of the TLM in equation (E.1) there exist two orthogonal matrices U, V such that

$$U^T L V = S, \quad S = \begin{bmatrix} \sigma_1 & 0 & \cdot & 0 \\ 0 & \sigma_2 & \cdot & 0 \\ \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & \cdot & \sigma_n \end{bmatrix} \quad \text{and} \quad U U^T = V V^T = I \quad (\text{E.6})$$

The elements σ_i of the diagonal matrix S are the singular values of L . The columns v_i of the matrix V are the right (or initial) SVs of L , they are indeed valid at the beginning of the optimization interval over which L is defined. The columns u_i of the matrix U are the left (final/evolved) SVs of L , they correspond to the end of the interval of optimization $(t, t + \Delta t)$. The initial SVs can be obtained as the eigenvectors of $L^T L$, a normal matrix whose eigenvalues are the squares of the singular values of L .

The SVs used to create the initial perturbations $\delta x(x, t)$ are obtained as the leading eigenvectors of

$$(W^{-1})^T L^T P^T P L W^{-1} \delta \hat{x}(x, t) = \sigma^2 \delta \hat{x}(x, t) \quad (\text{E.7})$$

subject to

$$\delta \hat{x}^T(x, t) \delta \hat{x}(x, t) = 1, \text{ and } \delta x(x, t) = W^{-1} \delta \hat{x}(x, t) \quad (\text{E.8})$$

The final norm could be a projection operator P at the end of the interval of optimization. The norm of the initial perturbations is defined using a weight matrix W applied to initial perturbations $\delta x(x, t)$ to be of equal size:

$$\|\delta x(x, t)\|^2 = (W \delta x(x, t))^T W \delta x(x, t) = \delta x^T(x, t) W^T W \delta x(x, t) = 1$$

ECMWF implemented an ensemble system with initial perturbations based on SVs using a total energy norm.

E.3. MULTIPLE DATA ASSIMILATION SYSTEMS

Houtekamer et al had developed a very promising EF system based on running an ensemble of DA systems to create the initial conditions. In their different DA systems they add (additional) random errors to the observations and include different parameters in the physical parameterizations of the model in different ensembles. This is a promising approach which is related to but more general than breeding. They introduced the use of perturbations in the physical parameterizations in the models used in different analysis cycles. Through a careful combination of changes in major parameterizations, it is possible to use the ensemble forecasts to isolate the impact of particular parameterizations. With respect to the computational cost, the multiple data assimilation system and the SV approach are comparable, whereas the breeding is essentially cost-free.

E.4. PERTURBED PHYSICAL PARAMETERIZATION

The methods discussed above only include perturbations in the initial conditions, assuming that the error growth due to model deficiencies is small compared to that due to unstable growth of initial errors. In addition, several groups have introduced changes in the physical parameterizations to allow for the inclusion of model uncertainty. Buizza et al developed a perturbation approach that introduces a stochastic perturbation of the impact of subgrid-scale physical parameterizations by simply multiplying the time derivative of the “physics” by a random number normally distributed with mean 1 and standard deviation 0.2.

E.5. MULTIPLE SYSTEM ENSEMBLES

Both the perturbations of the initial conditions and of the subgrid-scale physical parameterizations have been shown to be successful towards achieving the goals of ensemble forecasting. However, since they both introduce perturbations in the best estimate of the initial conditions and the model, which are in the control

forecast, it can be expected that the individual perturbed forecasts should be worse than the control. A typical ensemble average for a season (Figure E-4) shows that, indeed, the individual perturbed forecasts have less skill than the unperturbed control. Nevertheless, the ensemble average is an improvement over the control, especially after the perturbations grow into a nonlinear regime that tends to filter out some of the errors.

Anomaly Correlation for the winter of 1997/98 (from 00Z)
Controls (T126 and T62) and 10 perturbed ensemble forecasts

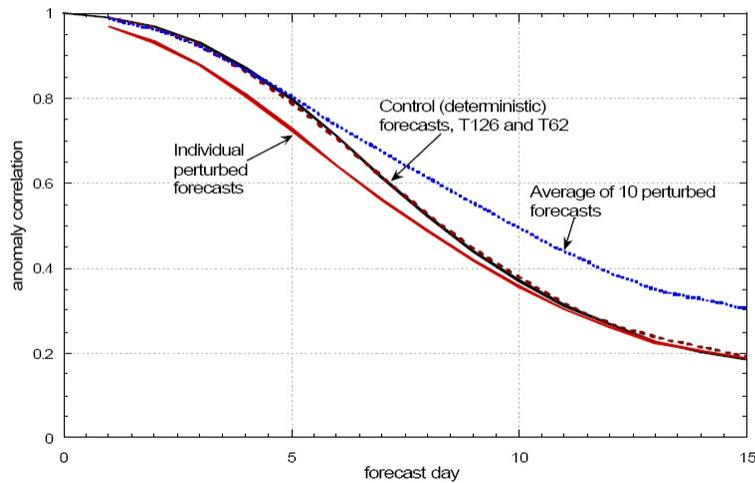


Figure E-4 Anomaly correlation of the ensembles during the winter of 1997/98

An alternative to the introduction of perturbations is the use of multiple systems developed independently at different centers. In principle, an ensemble of forecasts from different operational or research centers, each aiming to be the best and choosing different competitive approaches, should sample well the uncertainty in our knowledge of both the models and the initial conditions. It has long been known that the ensemble average of multiple center forecasts is significantly better than even the best individual forecasting system. This has also been shown to be true for regional models. Krishnamurti introduced the concept of “super-ensemble”, using linear regression and past forecasts of different systems as predictors to minimize the ensemble average prediction errors. This method is also

called “poor person’s” method to reflect that it does not require running a forecasting system.

E.6. OTHER METHODS

This field is changing quickly, and improvements and changes to the operational systems are under development. For example, ECMWF has implemented changes in the length of the optimization period for the SVs, a combination of initial and final or evolved SVs (which are more similar to BVs), and the introduction of a stochastic element in the physical parameterizations, all of which contributed to improvements in the ensemble performance. NCEP is considering the implementation of the Ensemble Transform Kalman Filtering (ETKF) to replace breeding. A recent comparison of the ensemble performance of the Canadian, US and ECMWF systems suggests that the ECMWF ensembles based on SVs behave well beyond the optimization period, at which time the model advantages of the ECMWF system are also paramount. The NCEP BVs are better at shorter ranges, and the multiple analyses Canadian method also seems to perform well.

F DETAILED RESULTS

F.1. USING FDDA IN NWP FOR EGYPT

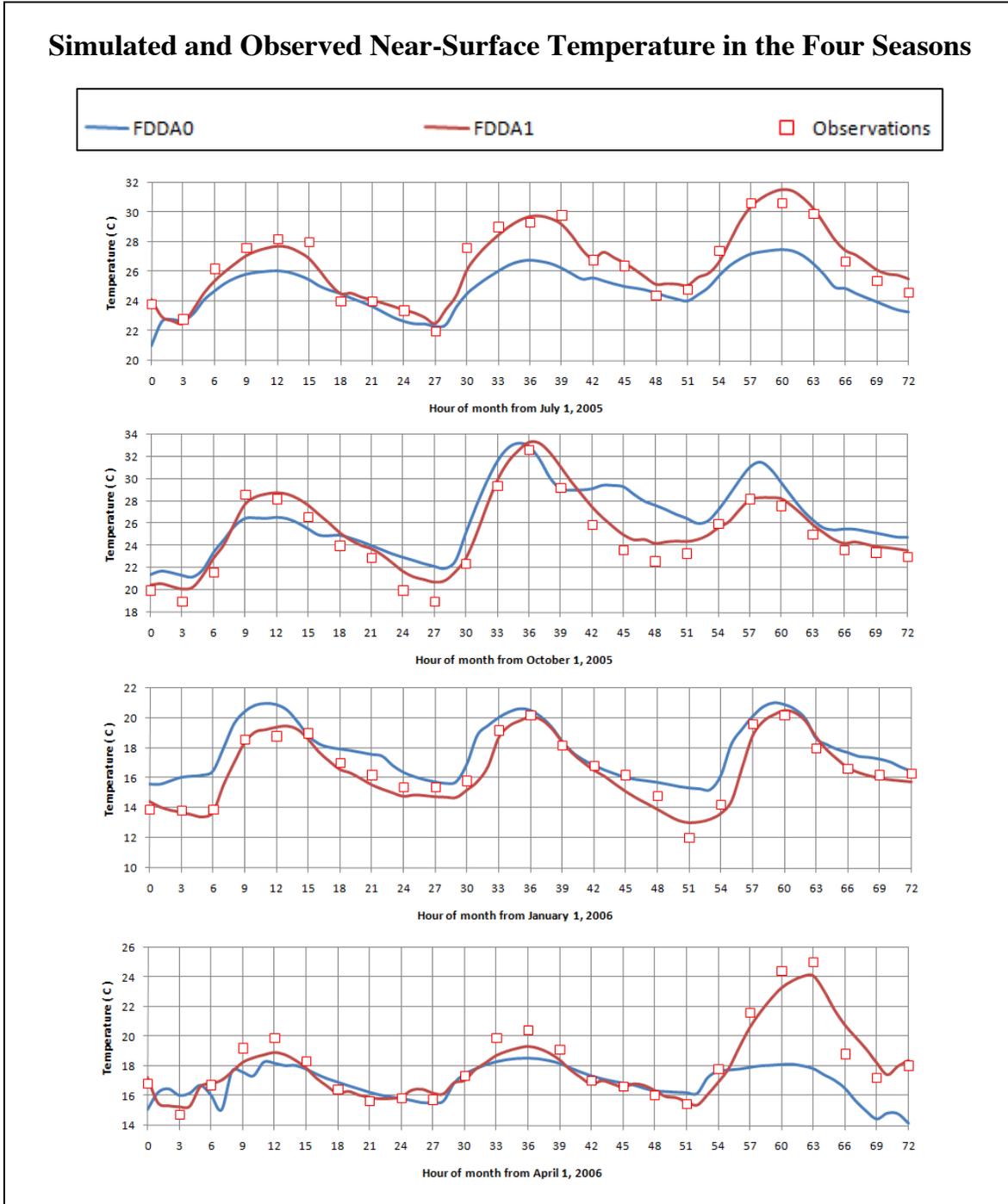


Figure F-1 Near-surface temperature in the four seasons at Alexandria/Nouzha station



Figure F-2 Total and daily gains of RMSE and MABE at Alexandria/Nouzha station

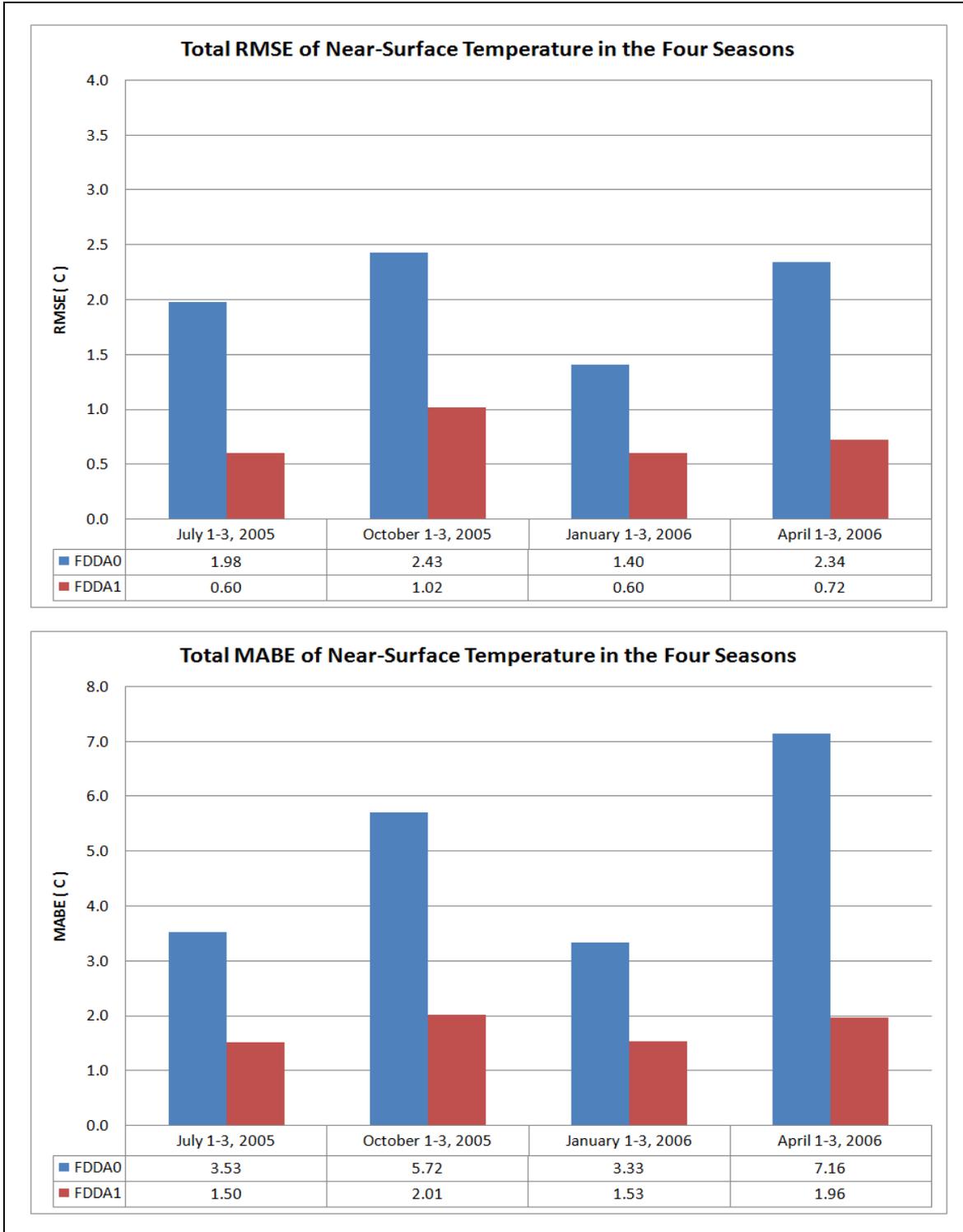


Figure F-3 Total RMSE and MABE in the four seasons at Alexandria/Nouzha Station

Daily RMSE and MABE of Near-Surface Temperature in the Four Seasons

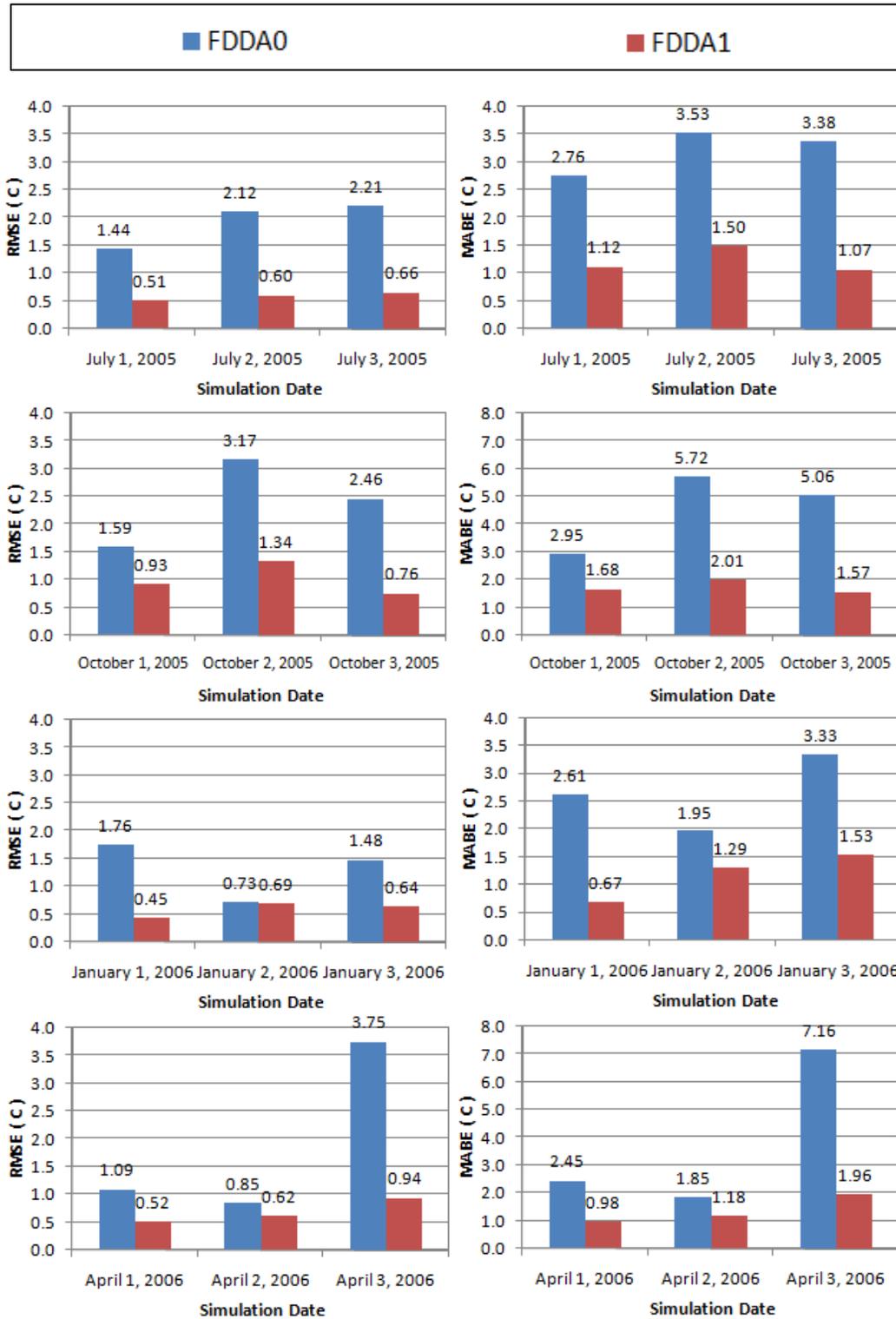


Figure F-4 Daily RMSE and MABE in the four seasons at Alexandria/Nouzha station

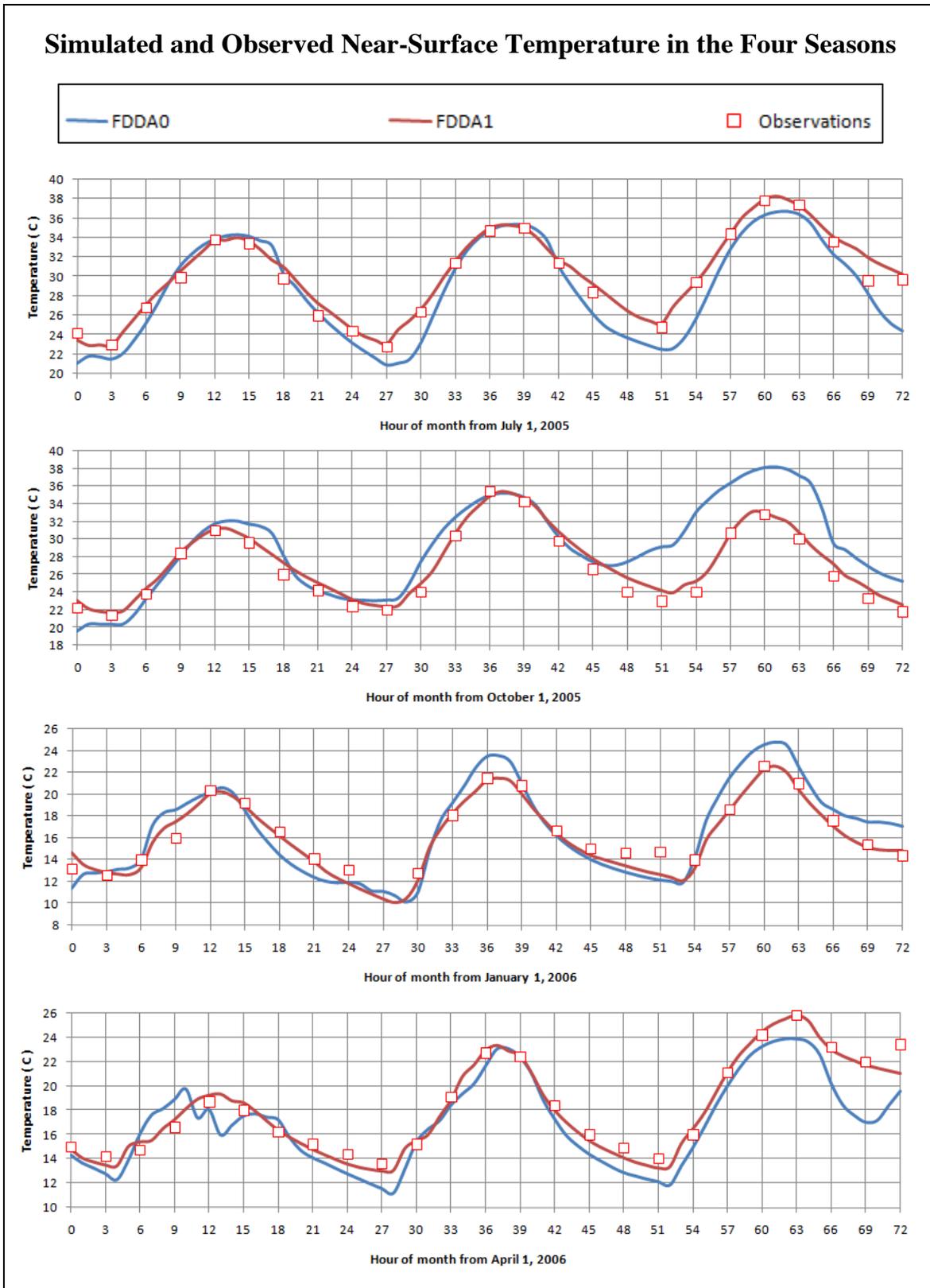


Figure F-5 Near-surface temperature in the four seasons at Cairo International Airport

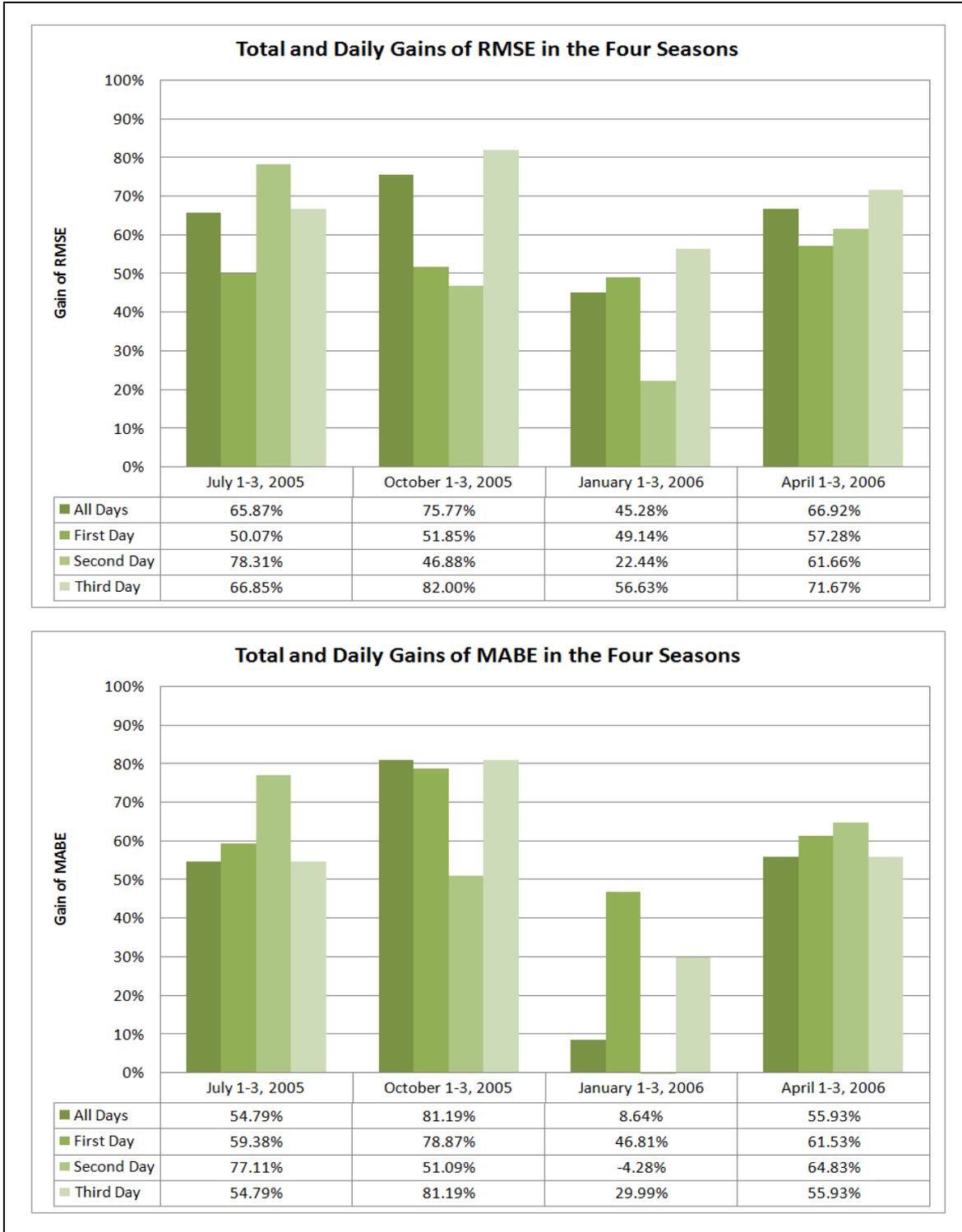


Figure F-6 Total and daily gains of RMSE and MABE at Cairo International Airport

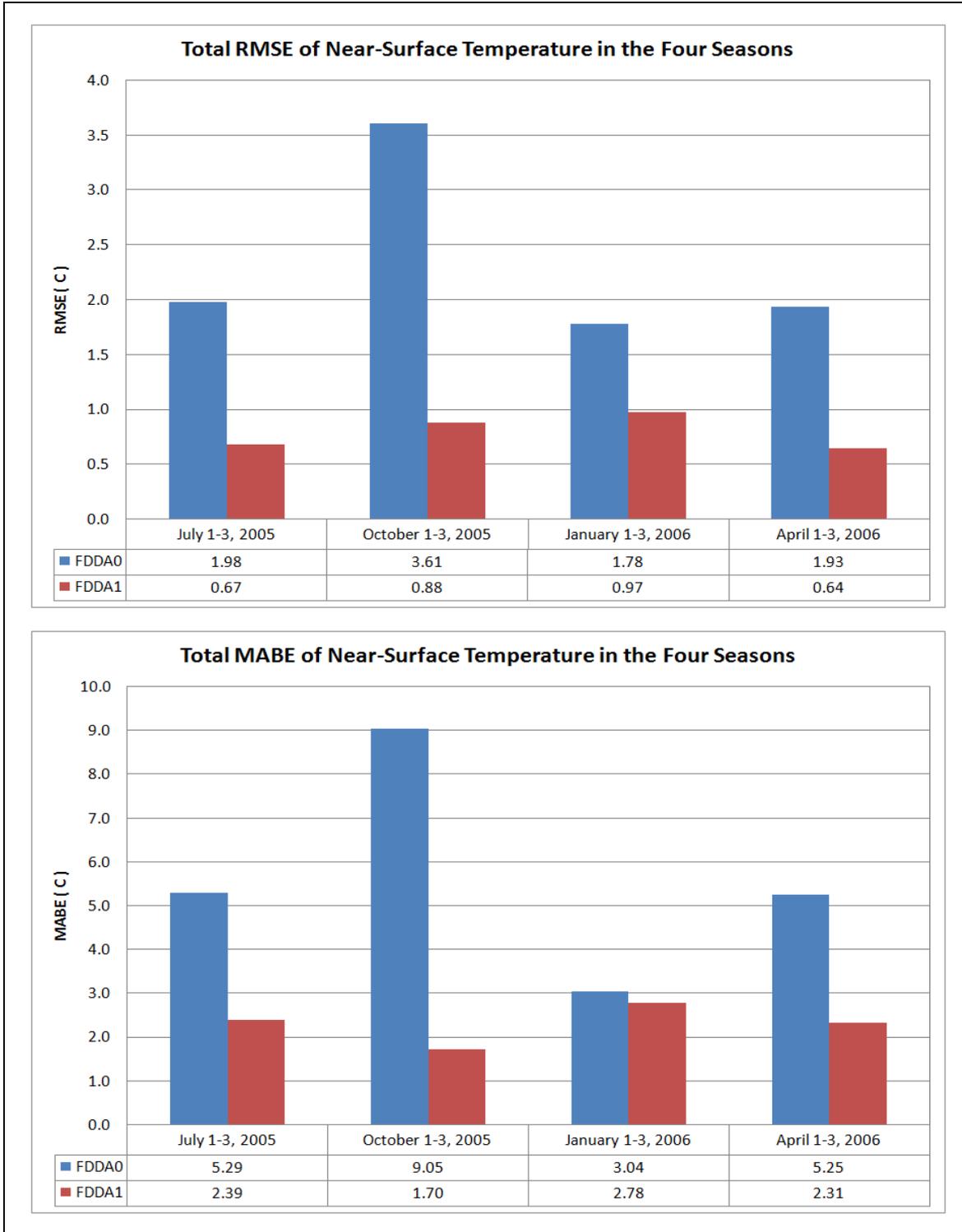


Figure F-7 Total RMSE and MABE in the four seasons at Cairo International Airport

Daily RMSE and MABE of Near-Surface Temperature in the Four Seasons

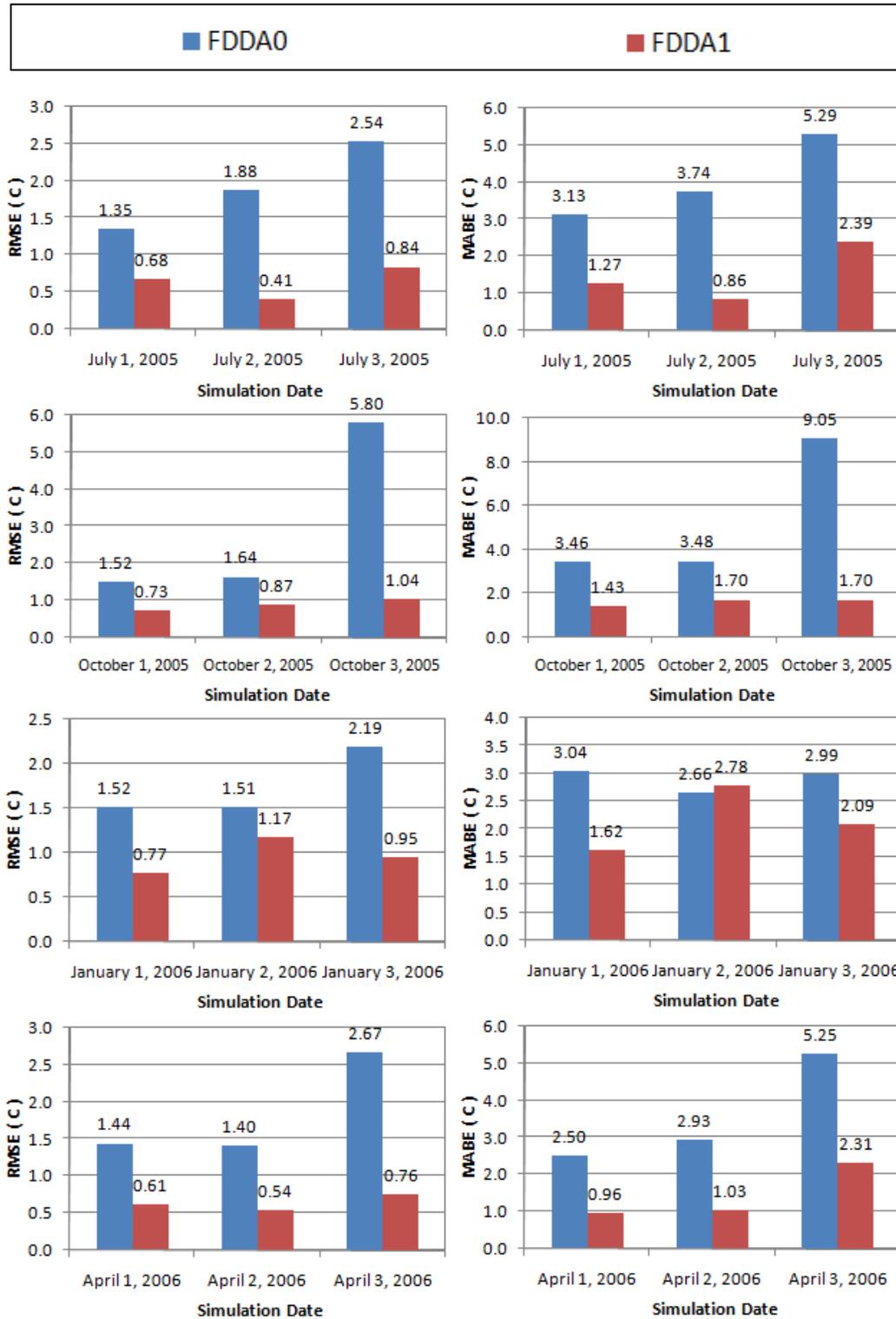


Figure F-8 Daily RMSE and MABE in the four seasons at Cairo International Airport

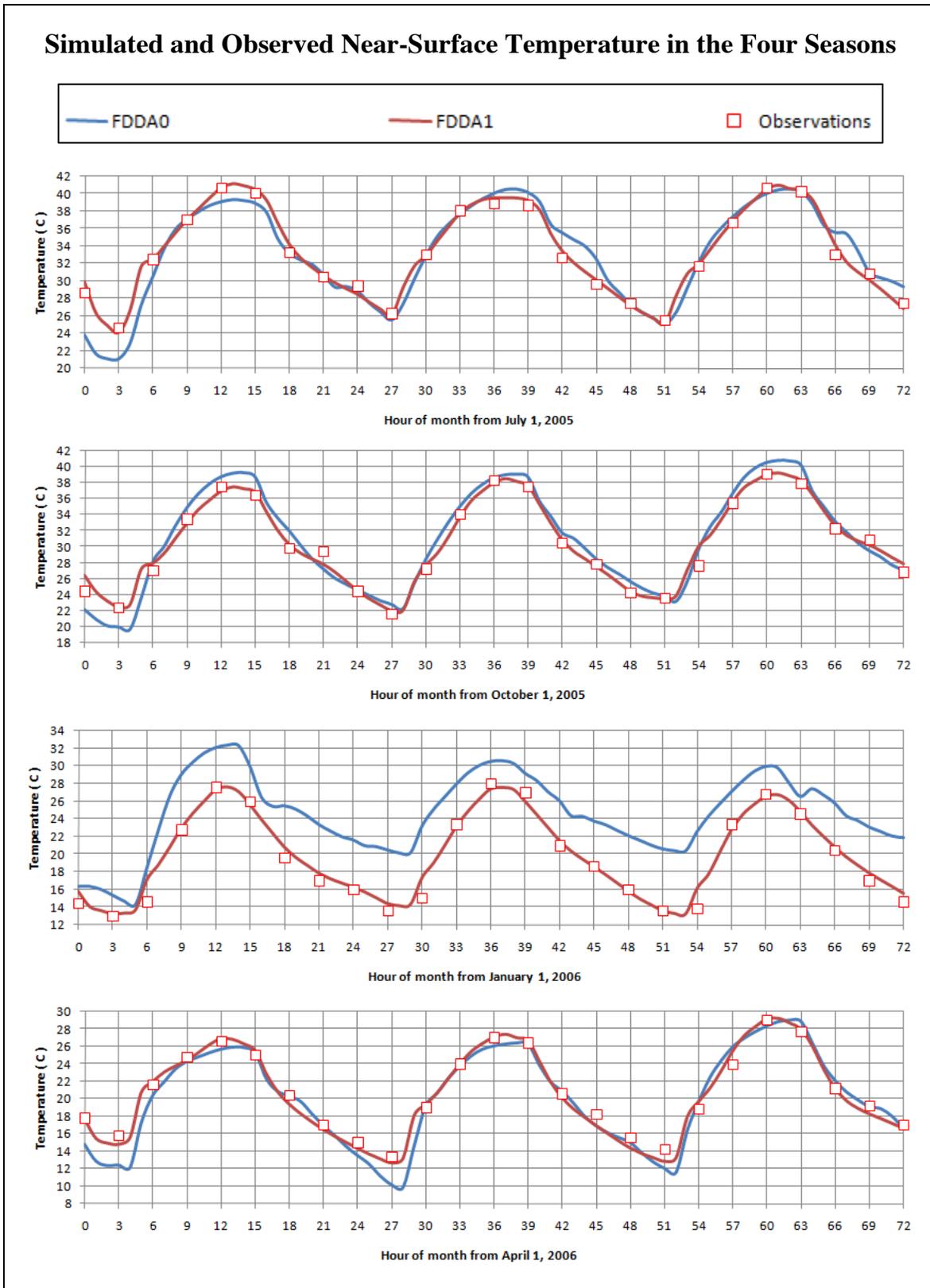


Figure F-9 Simulated and observed near-surface temperature at Luxor station



Figure F-10 Total and daily gains of RMSE and MABE at Luxor station

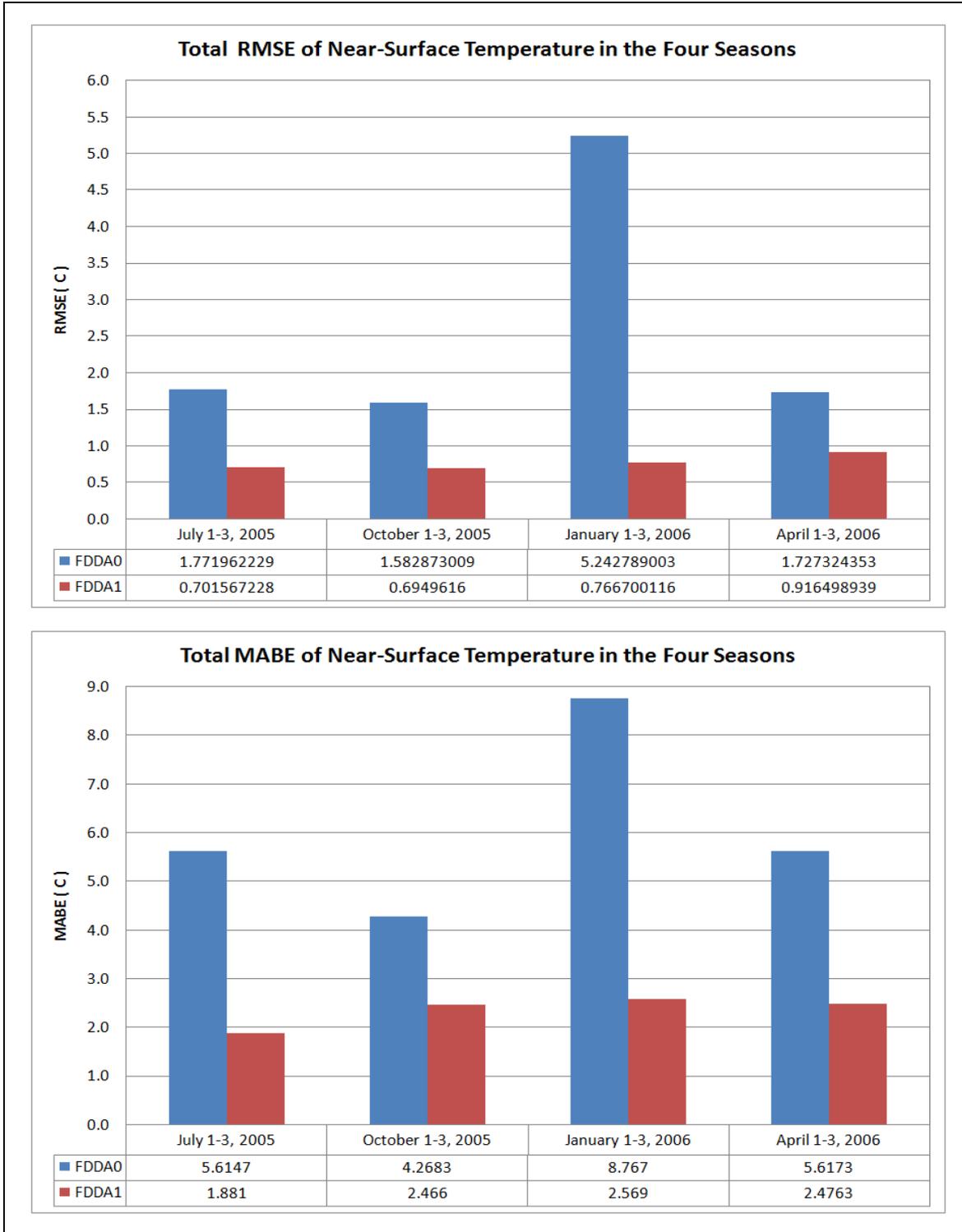


Figure F-11 Total RMSE and MABE in the four seasons at Luxor station

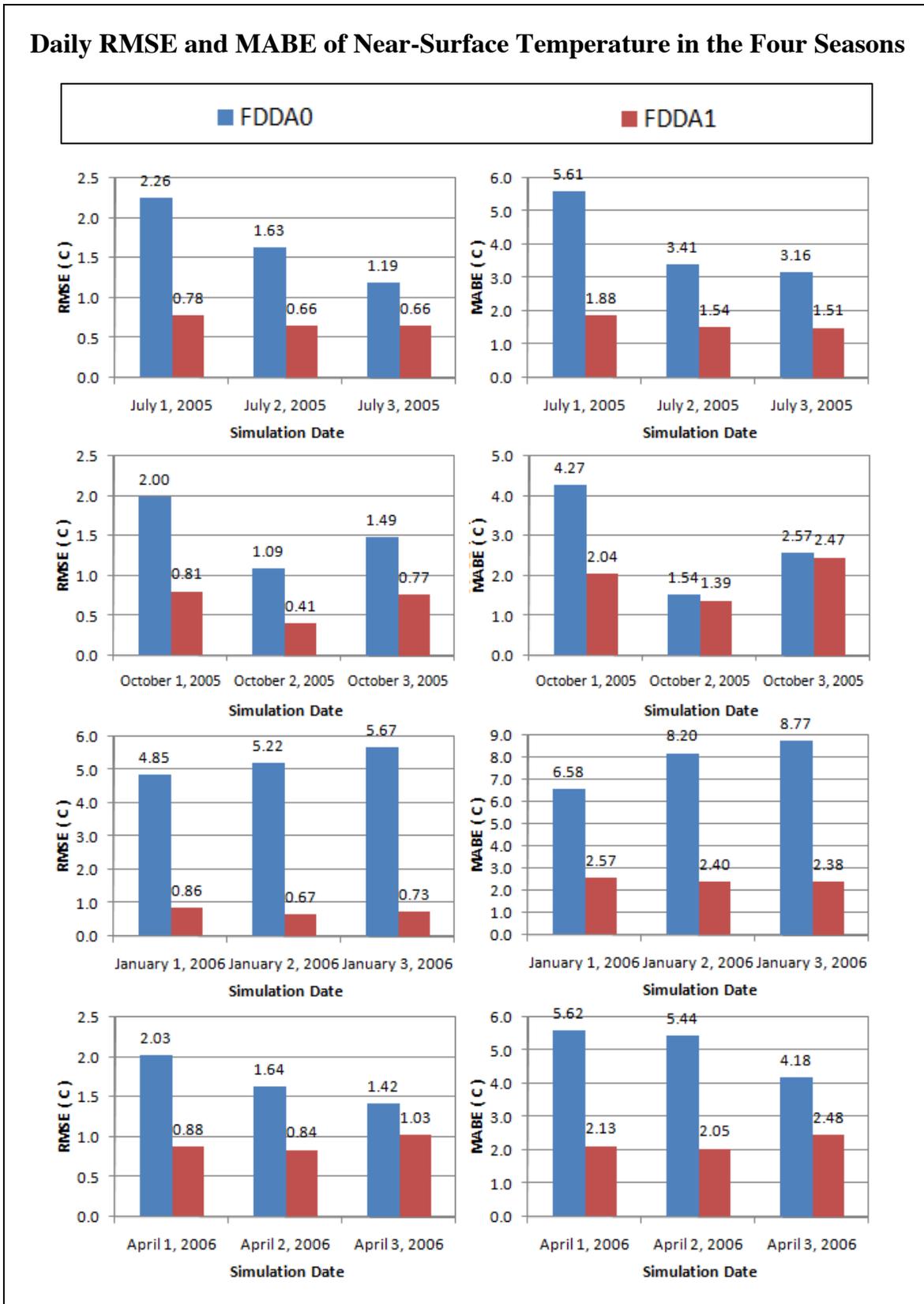


Figure F-12 Daily RMSE and MABE in the four seasons at Luxor station

Table F-1 RMSE and MABE in the four seasons at Alexandria/Nouzha station

Simulation		All Days		First Day		Second Day		Third Day	
		RMSE C	MABE C	RMSE C	MABE C	RMSE C	MABE C	RMSE C	MABE C
07/2005	FDDA0	1.976	3.529	1.436	2.760	2.117	3.529	2.210	3.379
	FDDA1	0.596	1.503	0.511	1.119	0.604	1.503	0.658	1.069
	Gain	69.85 %	57.41 %	64.43 %	59.46 %	71.46 %	57.41 %	70.23 %	69.35 %
10/2005	FDDA0	2.434	5.716	1.594	2.952	3.173	5.716	2.462	5.058
	FDDA1	1.016	2.008	0.931	1.675	1.337	2.008	0.756	1.572
	Gain	58.26 %	64.88 %	41.63 %	68.47%	57.87 %	64.88 %	69.29 %	68.92 %
01/2006	FDDA0	1.403	3.330	1.760	2.609	0.728	1.951	1.480	3.330
	FDDA1	0.596	1.534	0.728	0.667	0.689	1.287	0.639	1.534
	Gain	57.51 %	53.93 %	74.70 %	74.42 %	5.33 %	34.04 %	56.78 %	53.93 %
04/2006	FDDA0	2.340	7.156	1.093	2.447	0.853	1.854	3.751	7.156
	FDDA1	0.723	1.964	0.520	0.977	0.616	1.181	0.939	1.964
	Gain	69.11 %	72.55 %	52.40 %	60.08 %	27.82 %	36.30 %	74.97 %	72.55 %

Table F-2 RMSE and MABE in the four seasons at Cairo International Airport

Simulation		All Days		First Day		Second Day		Third Day	
		RMSE C	MABE C	RMSE C	MABE C	RMSE C	MABE C	RMSE C	MABE C
07/2005	FDDA0	1.977	5.293	1.355	3.134	1.881	3.736	2.537	5.293
	FDDA1	0.675	2.393	0.676	1.273	0.408	0.855	0.841	2.393
	Gain	65.87 %	54.79 %	50.07 %	59.38 %	78.31 %	77.11 %	66.85 %	54.79 %
10/2005	FDDA0	3.612	9.053	1.518	3.458	1.641	3.482	5.796	9.053
	FDDA1	0.875	1.703	0.731	1.425	0.872	1.703	1.043	1.703
	Gain	75.77 %	81.19 %	51.85 %	78.87 %	46.88 %	51.09 %	82.00 %	81.19 %
01/2006	FDDA0	1.777	3.040	1.516	3.040	1.510	2.663	2.190	2.992
	FDDA1	0.972	2.777	0.771	1.617	1.171	2.777	0.950	2.094
	Gain	45.28 %	8.64 %	49.14 %	46.81 %	22.44 %	-4.28 %	56.63 %	29.99 %
04/2006	FDDA0	1.931	5.249	1.438	2.498	1.405	2.926	2.667	5.249
	FDDA1	0.639	2.313	0.614	0.961	0.539	1.029	0.756	2.313
	Gain	66.92 %	55.93 %	57.28 %	61.53 %	61.66 %	64.83 %	71.67 %	55.93 %

Table F-3 RMSE and MABE in the four seasons at Luxor station

Simulation		All Days		First Day		Second Day		Third Day	
		RMSE C	MABE C	RMSE C	MABE C	RMSE C	MABE C	RMSE C	MABE C
07/2005	FDDA0	1.772	5.615	2.256	5.6147	1.635	3.411	1.190	3.157
	FDDA1	0.702	1.881	0.784	1.881	0.655	1.538	0.657	1.507
	Gain	60.41 %	66.50 %	65.25 %	66.50 %	59.92 %	54.91 %	44.78 %	52.27 %
10/2005	FDDA0	1.583	4.268	1.997	4.268	1.093	1.536	1.489	2.571
	FDDA1	0.695	2.466	0.805	2.043	0.412	1.387	0.772	2.466
	Gain	56.09 %	42.23 %	59.69 %	81.14 %	62.36 %	9.70 %	48.16 %	4.08 %
01/2006	FDDA0	5.243	8.767	4.855	6.580	5.224	8.197	5.670	8.767
	FDDA1	0.767	2.569	0.860	2.569	0.668	2.404	0.732	2.384
	Gain	85.38 %	70.70 %	82.28 %	60.96 %	87.22 %	70.67 %	87.08 %	72.81 %
04/2006	FDDA0	1.727	5.617	2.028	5.617	1.635	5.436	1.424	4.178
	FDDA1	0.916	2.476	0.876	2.130	0.838	2.055	1.029	2.476
	Gain	46.94 %	55.92 %	56.81 %	62.08 %	48.77 %	62.20 %	27.76 %	40.73 %

F.2. EVALUATION OF FDDA ON NWP FOR EGYPT

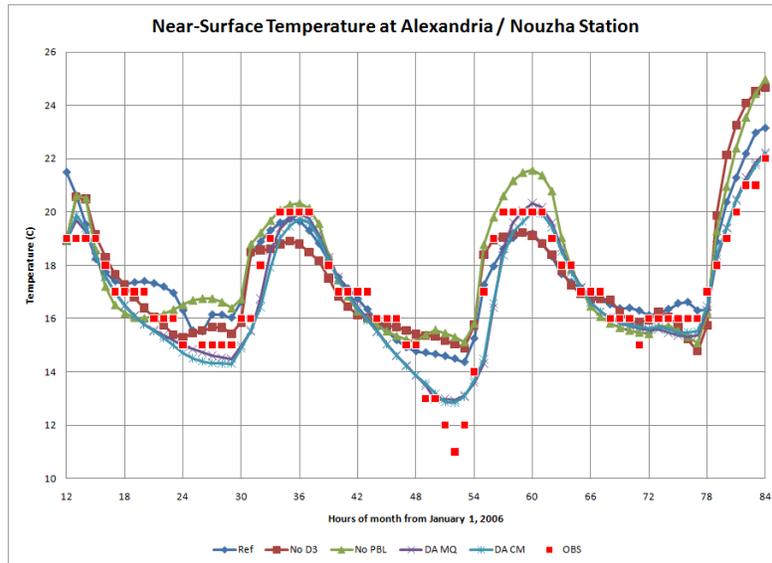


Figure F-13 The different sets of FDDA options at Alexandria/Nouzha station in winter

Table F-4 RMSE and different gains at Alexandria/Nouzha station in winter

Simulation	Ref	No D3	No PBL	DA MQ	DA CM
RMSE (C)	1.088	1.394	1.453	0.787	0.813
Gain (%)	0	-28.11	-33.54	27.66	25.25

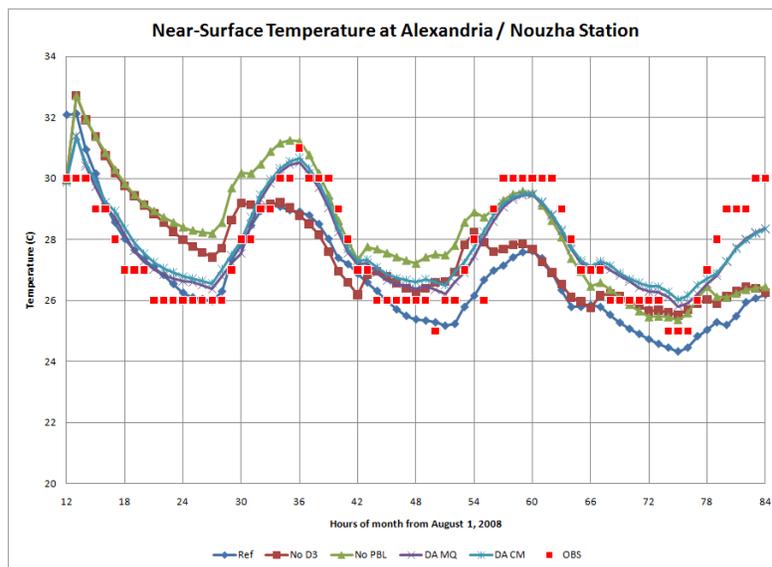


Figure F-14 The different sets of FDDA options at Alexandria/Nouzha station in summer

Table F-5 RMSE and different gains at Alexandria/Nouzha station in summer

Simulation	Ref	No D3	No PBL	DA MQ	DA CM
RMSE (C)	1.617	1.756	1.721	0.783	0.867
Gain (%)	0	-8.57	-6.39	51.62	46.42

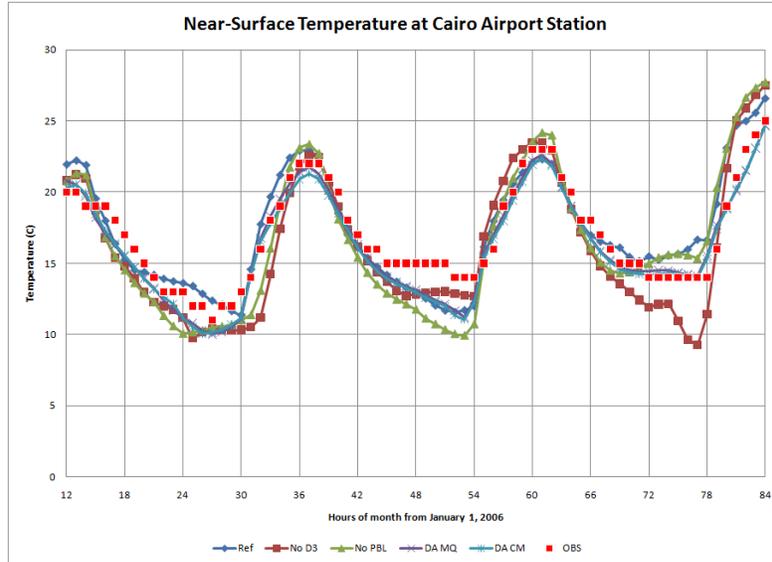


Figure F-15 The different sets of FDDA options at Cairo station in winter

Table F-6 RMSE and different gains at Cairo station in winter

Simulation	Ref	No D3	No PBL	DA MQ	DA CM
RMSE (C)	1.648	2.090	2.226	1.204	1.264
Gain (%)	0	-26.81	-35.09	26.95	23.34

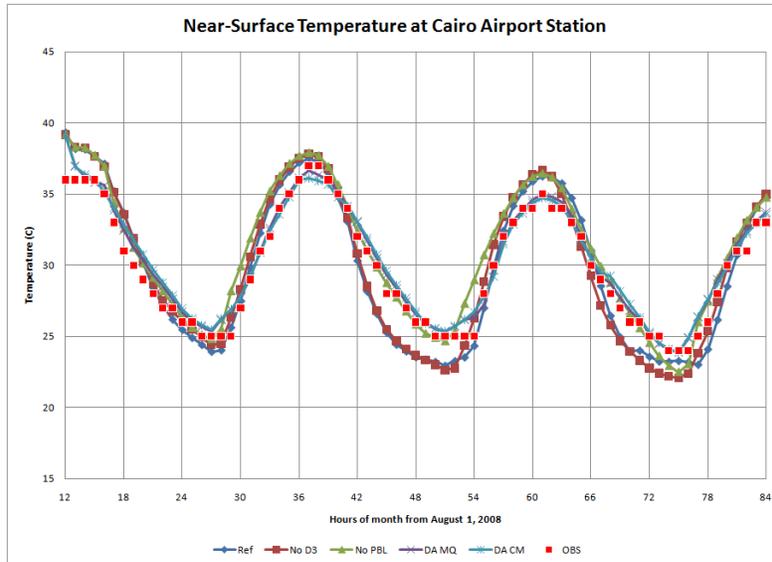


Figure F-16 The different sets of FDDA options at Cairo station in summer

Table F-7 RMSE and different gains at Cairo station in summer

Simulation	Ref	No D3	No PBL	DA MQ	DA CM
RMSE (C)	1.681	1.776	1.571	0.842	0.951
Gain (%)	0	-5.64	6.52	49.90	43.42

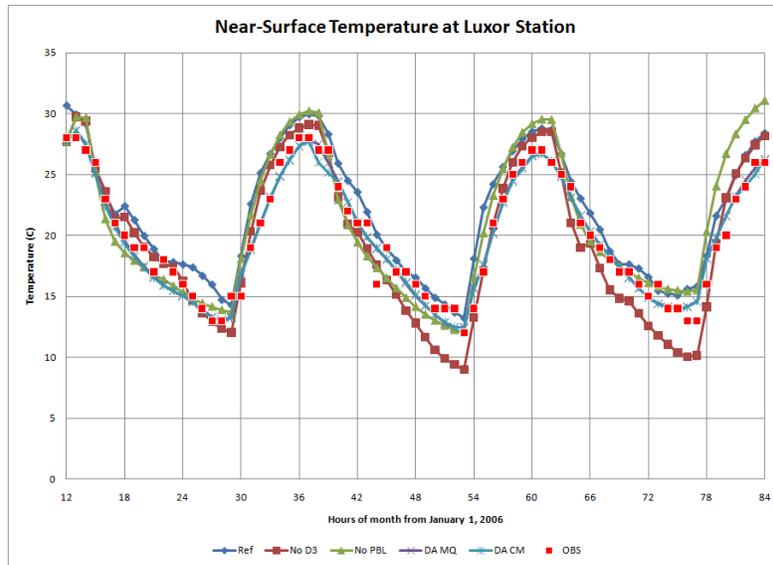


Figure F-17 The different sets of FDDA options at Luxor station in winter

Table F-8 RMSE and different gains at Luxor station in winter

Simulation	Ref	No D3	No PBL	DA MQ	DA CM
RMSE (C)	2.150	2.044	2.390	0.924	0.954
Gain (%)	0	4.94	-11.16	57.04	55.62

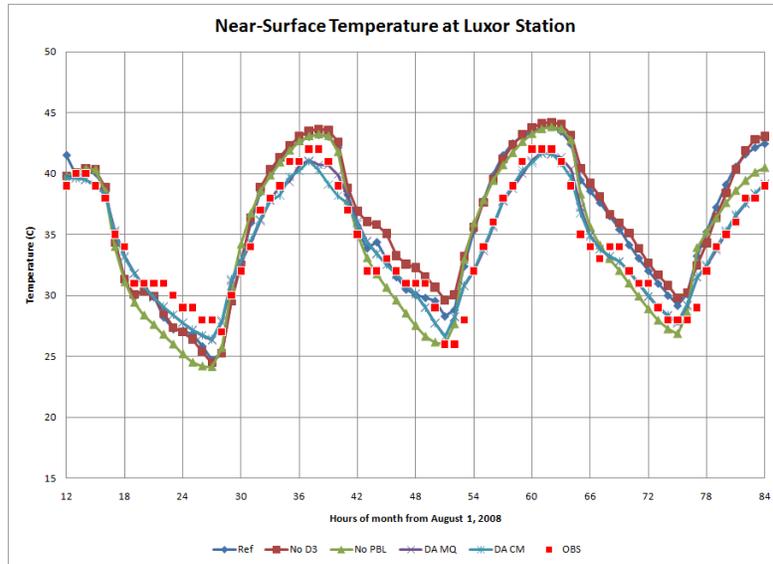


Figure F-18 The different sets of FDDA options at Luxor station in summer

Table F-9 RMSE and different gains at Luxor station in summer

Simulation	Ref	No D3	No PBL	DA MQ	DA CM
RMSE (C)	2.489	2.789	2.452	1.062	1.063
Gain (%)	0	-12.05	1.50	57.33	57.30

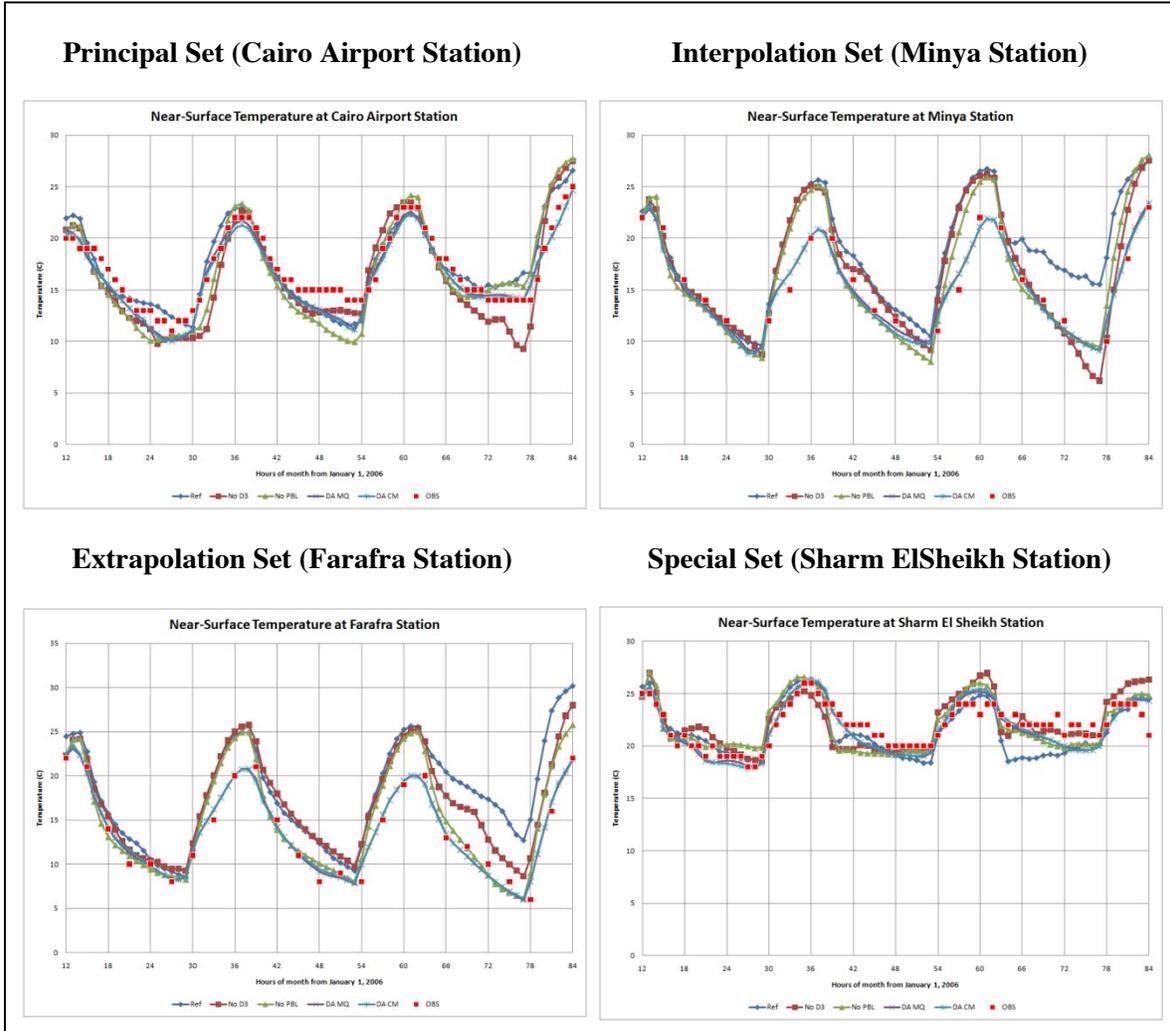


Figure F-19 Sample results for the different observation sets in winter

Table F-10 RMSE of near-surface temperature in winter 2006 at all stations

#	Station	Ref	No D3	No PBL	DA MQ	DA CM
1	Mersa Matruh	1.332	2.419	1.750	1.239	1.250
2	Dabaa	1.349	2.034	1.217	0.748	0.905
3	Port Alexandria	1.855	2.505	2.544	1.249	1.261
4	Alexandria / Nouzha	1.088	1.394	1.453	0.787	0.813
5	Baltim	1.669	1.943	1.879	1.083	1.156
6	Port Said / El Gamil	1.298	1.820	1.699	0.827	0.793
7	El Arish	3.635	2.661	2.641	1.089	1.123
8	Wadi El Natroon	1.911	2.078	1.775	1.491	1.336
9	Cairo Airport	1.648	2.090	2.226	1.204	1.263
10	Minya	4.322	3.119	2.990	1.131	1.126
11	Asyut	3.674	2.691	2.467	1.293	1.343
12	Luxor	2.150	2.044	2.390	0.924	0.954
13	Aswan	2.174	2.030	2.245	0.797	0.767
14	Siwa	2.774	1.734	1.407	1.068	0.949
15	Abu Simbel					
16	Baharia	2.724	1.013	1.360	0.698	0.776
17	Farafra	5.227	3.685	2.533	0.949	1.021
18	Dakela					
19	Kharga	4.213	2.212	1.755	1.271	1.930
20	Ras Sedr	2.053	1.839	2.003	1.028	1.058
21	Taba Airport	2.288	2.079	2.096	1.404	1.405
22	El Tor	3.489	4.010	3.877	2.250	2.187
23	Sharm El Sheikh	1.646	1.655	1.598	1.112	1.150
24	Hurguada	1.668	1.651	1.978	0.915	0.902
	Average	2.463	2.214	2.086	1.116	1.158
	Minimum	1.088	1.013	1.217	0.698	0.767
	Maximum	5.227	4.010	3.877	2.250	2.187

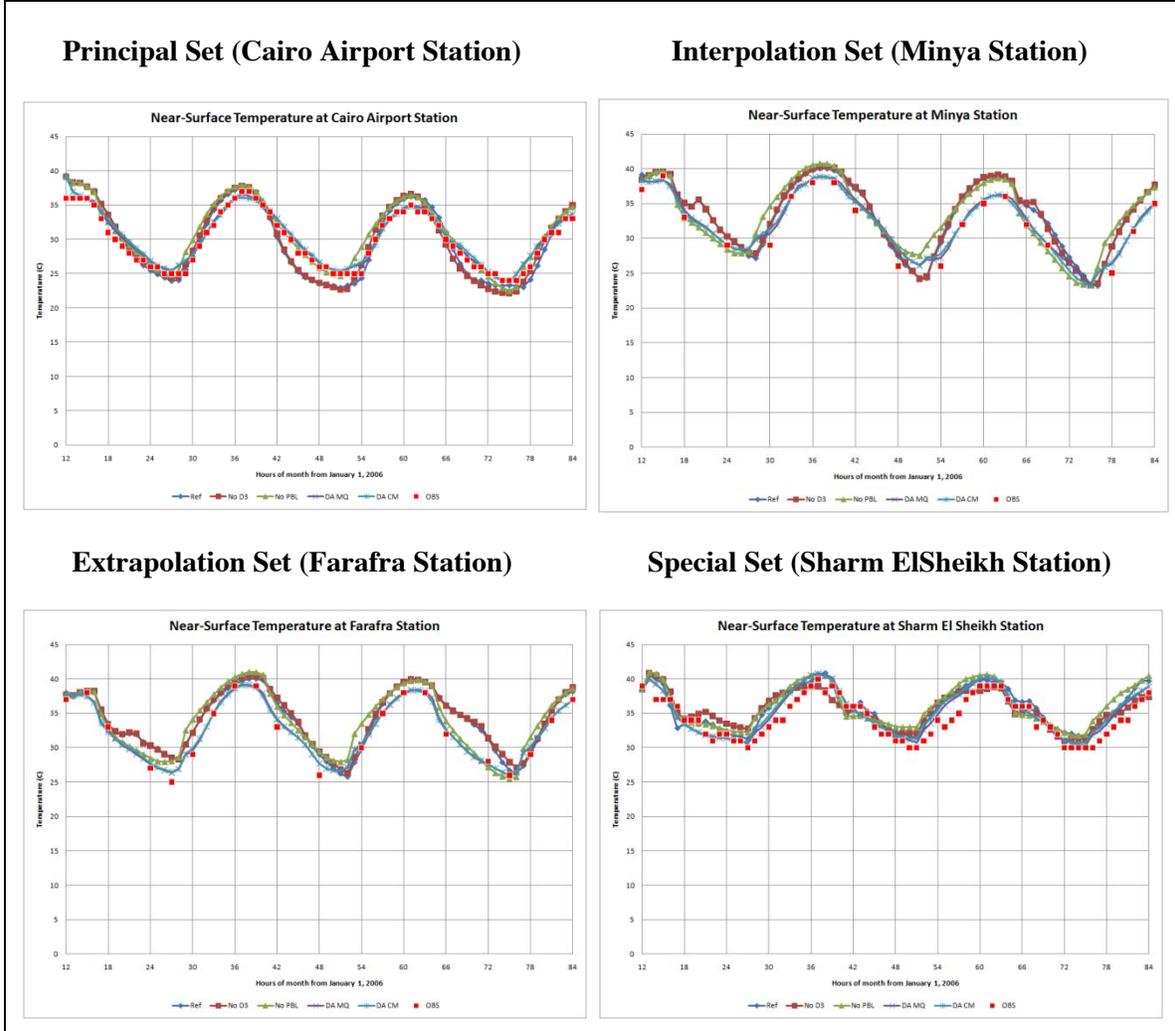


Figure F-20 Sample results for the different observation sets in winter

Table F-11 RMSE of near-surface temperature in summer 2008 at all stations

#	Station	Ref	No D3	No PBL	DA MQ	DA CM
1	Mersa Matruh	2.031	2.415	1.840	0.985	1.043
2	Dabaa	1.228	1.752	1.613	0.906	0.852
3	Port Alexandria	2.430	3.033	2.886	2.067	2.098
4	Alexandria / Nouzha	1.617	1.756	1.721	0.782	0.867
5	Baltim	3.024	2.819	2.518	1.626	1.667
6	Port Said / El Gamil	2.165	2.267	2.089	0.950	1.025
7	El Arish	2.951	2.820	2.214	1.164	1.266
8	Wadi El Natroon	3.443	3.598	3.915	2.981	2.929
9	Cairo Airport	1.681	1.776	1.571	0.842	0.951
10	Minya	2.644	2.754	3.020	1.029	1.076
11	Asyut	2.308	2.598	2.057	0.785	0.884
12	Luxor	2.489	2.789	2.452	1.062	1.063
13	Aswan	2.078	2.482	3.075	1.086	1.221
14	Siwa	1.633	1.587	1.745	1.014	0.826
15	Abu Simbel	2.631	3.053	2.840	0.886	1.153
16	Baharia	1.673	1.749	2.078	1.235	1.218
17	Farafra	2.171	2.352	2.309	0.697	0.738
18	Dakela	2.444	2.446	1.853	0.830	1.195
19	Kharga	2.209	2.665	1.912	0.964	0.962
20	Ras Sedr	1.252	1.920	2.276	1.801	1.870
21	Taba Airport	3.322	3.853	2.361	3.388	3.538
22	El Tor	3.922	3.821	3.021	1.539	1.657
23	Sharm El Sheikh	1.930	2.000	2.444	1.328	1.528
24	Hurguada	2.007	2.194	2.550	1.103	1.151
	Average	2.303	2.521	2.348	1.294	1.366
	Minimum	1.228	1.587	1.571	0.697	0.738
	Maximum	3.922	3.853	3.915	3.388	3.538

F.3. MODEL SENSITIVITY TO PHYSICS OPTIONS

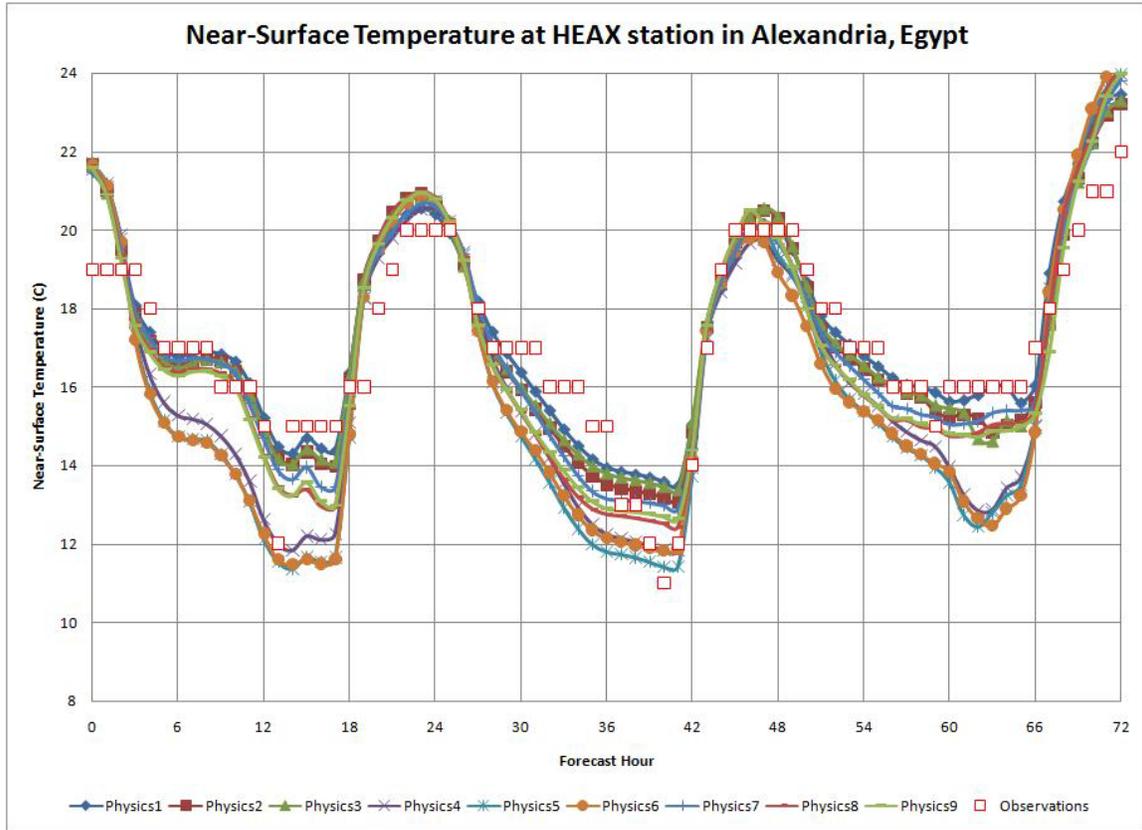


Figure F-21 Near-surface temperature at Alexandria/Nouzha station in winter

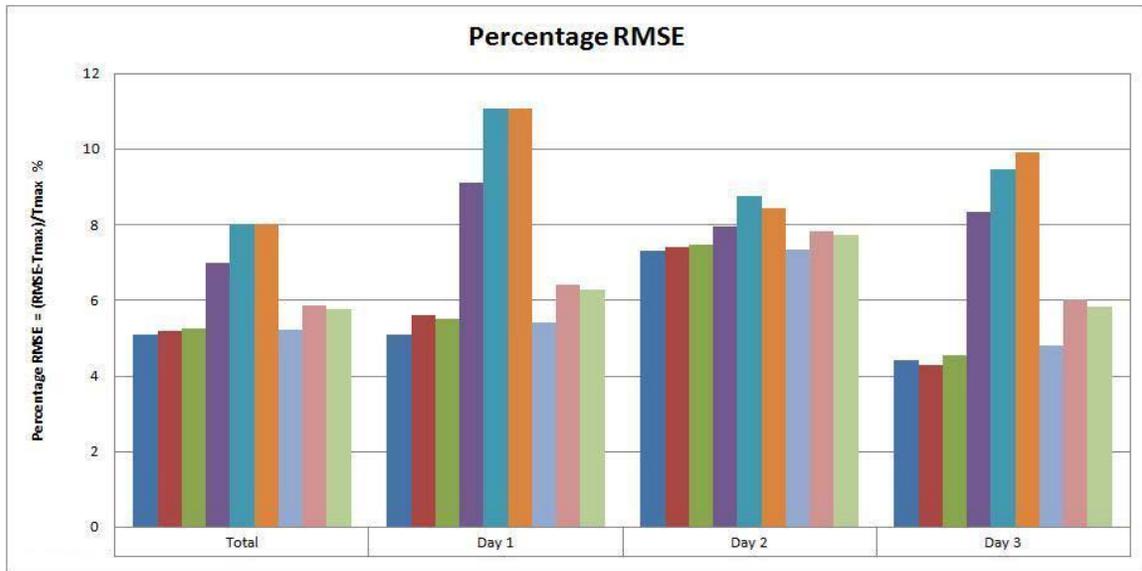


Figure F-22 RMSE of near-surface temperature at Alexandria/Nouzha station in winter

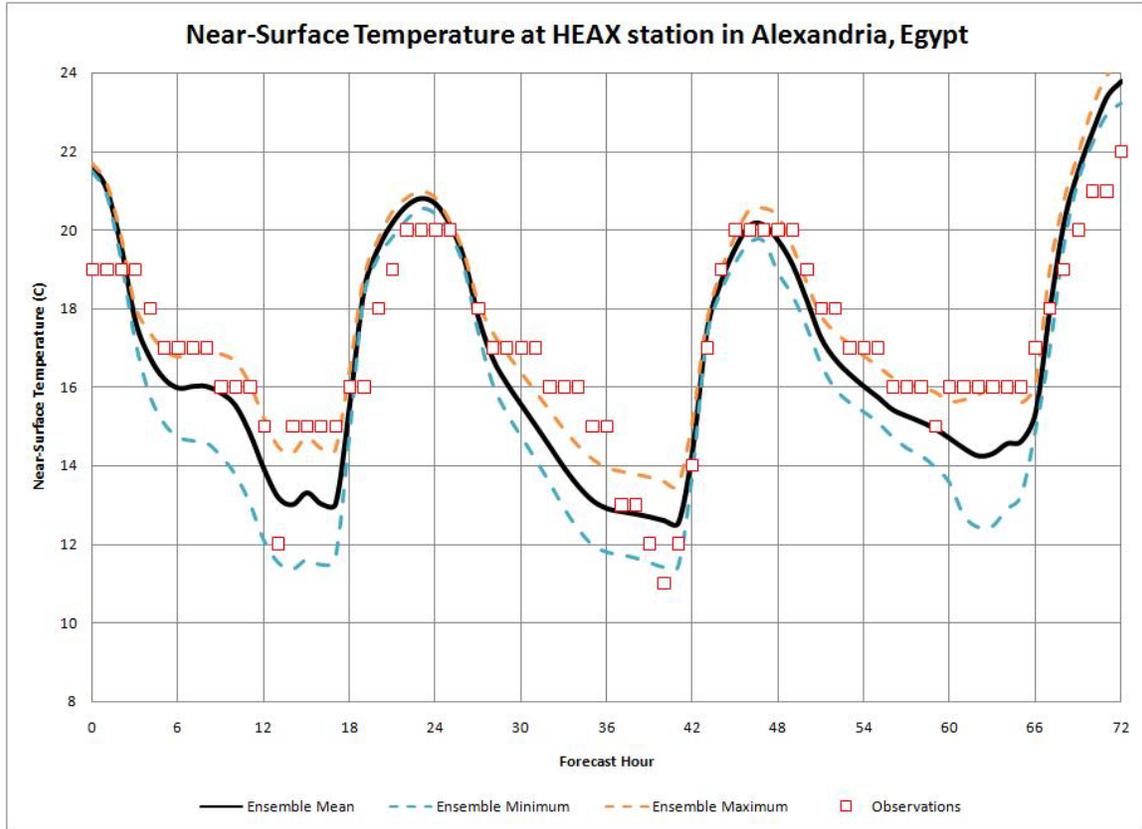


Figure F-23 Min, Max and mean temperature at Alexandria/Nouzha station in winter

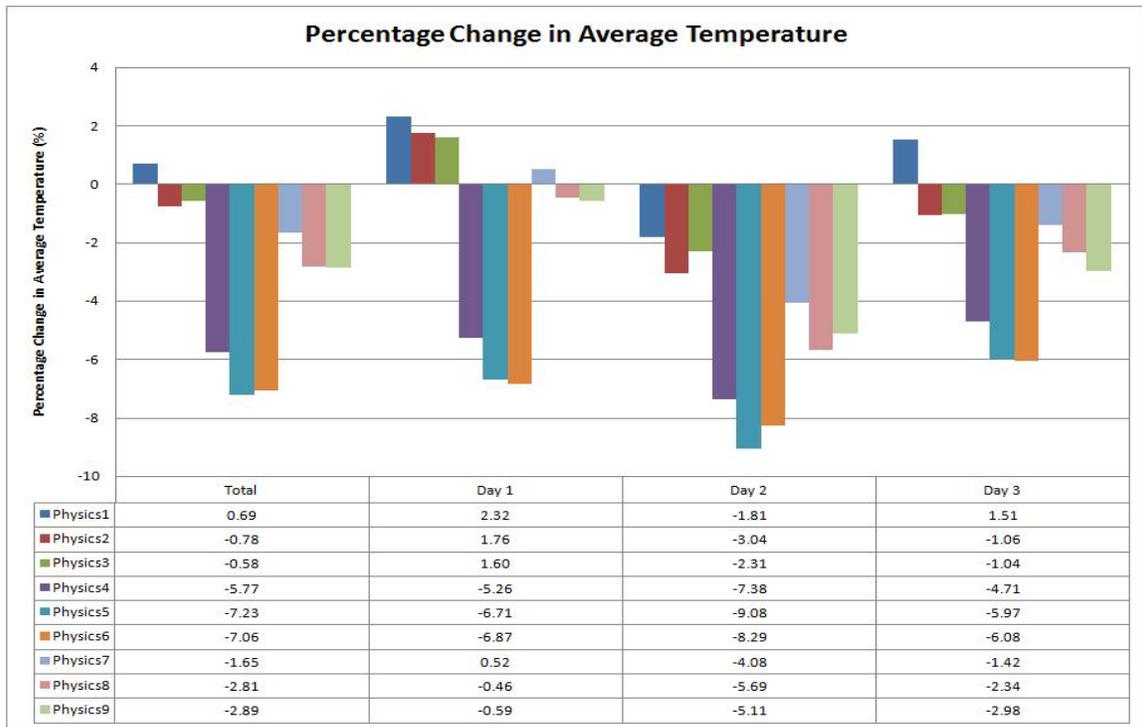


Figure F-24 Change in average temperature at Alexandria/Nouzha station in winter

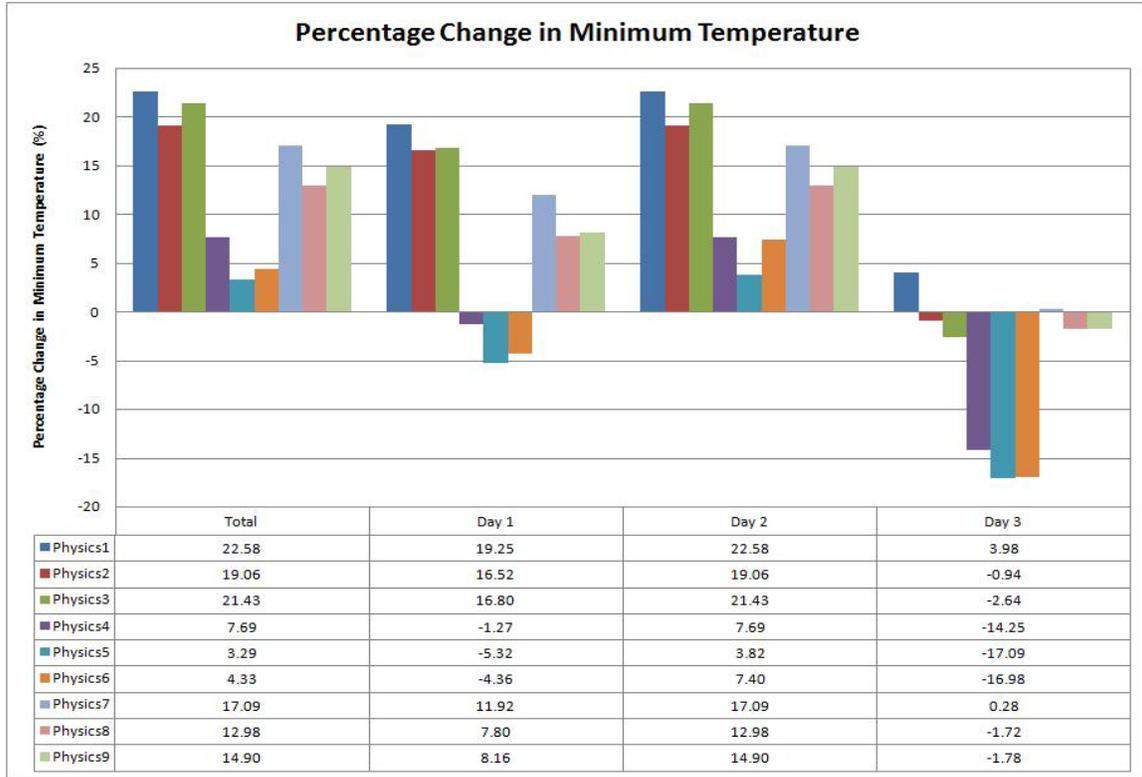


Figure F-25 Change in minimum temperature at Alexandria/Nouzha station in winter

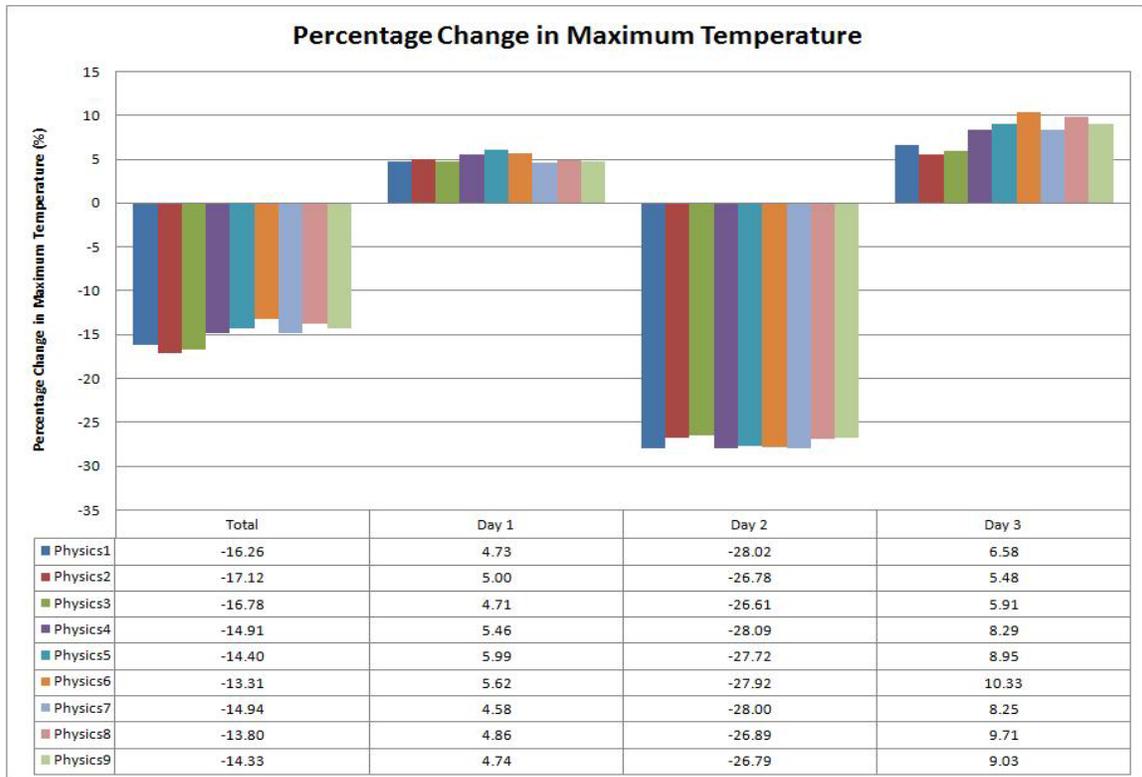


Figure F-26 Change in maximum temperature at Alexandria/Nouzha station in winter

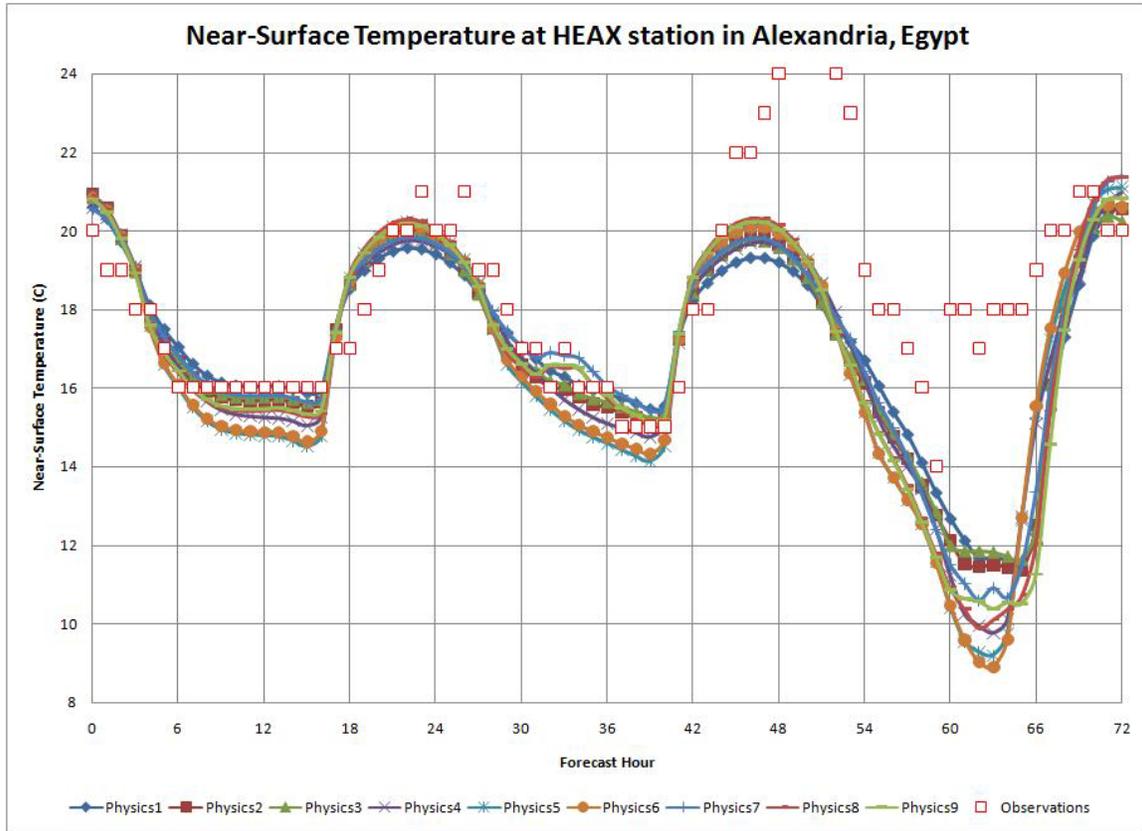


Figure F-27 Near-surface temperature at Alexandria/Nouzha station in spring

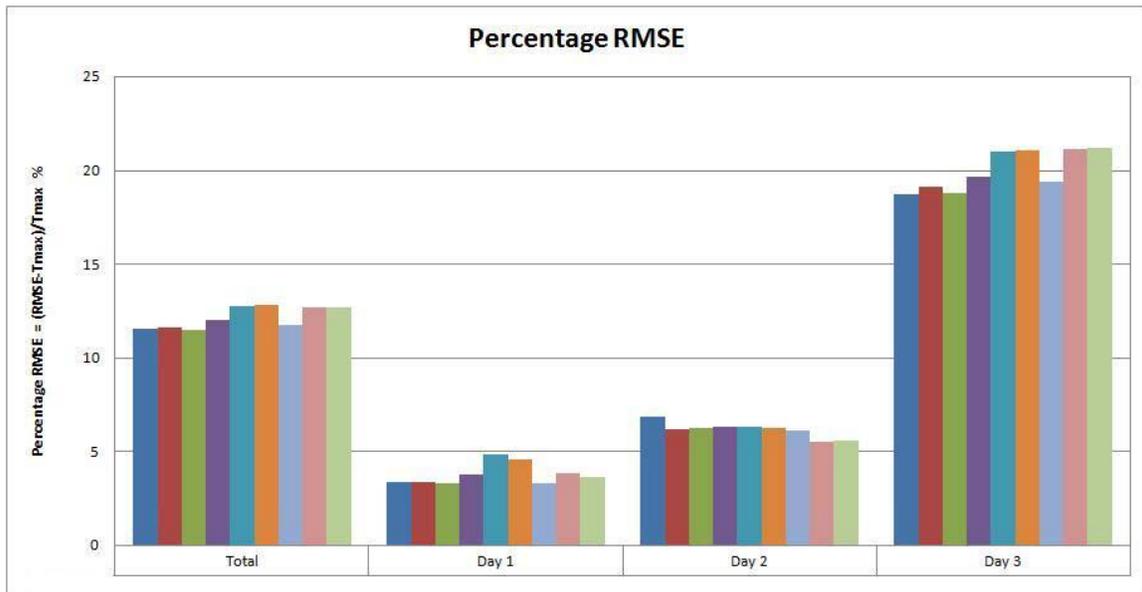


Figure F-28 RMSE of near-surface temperature at Alexandria/Nouzha station in spring

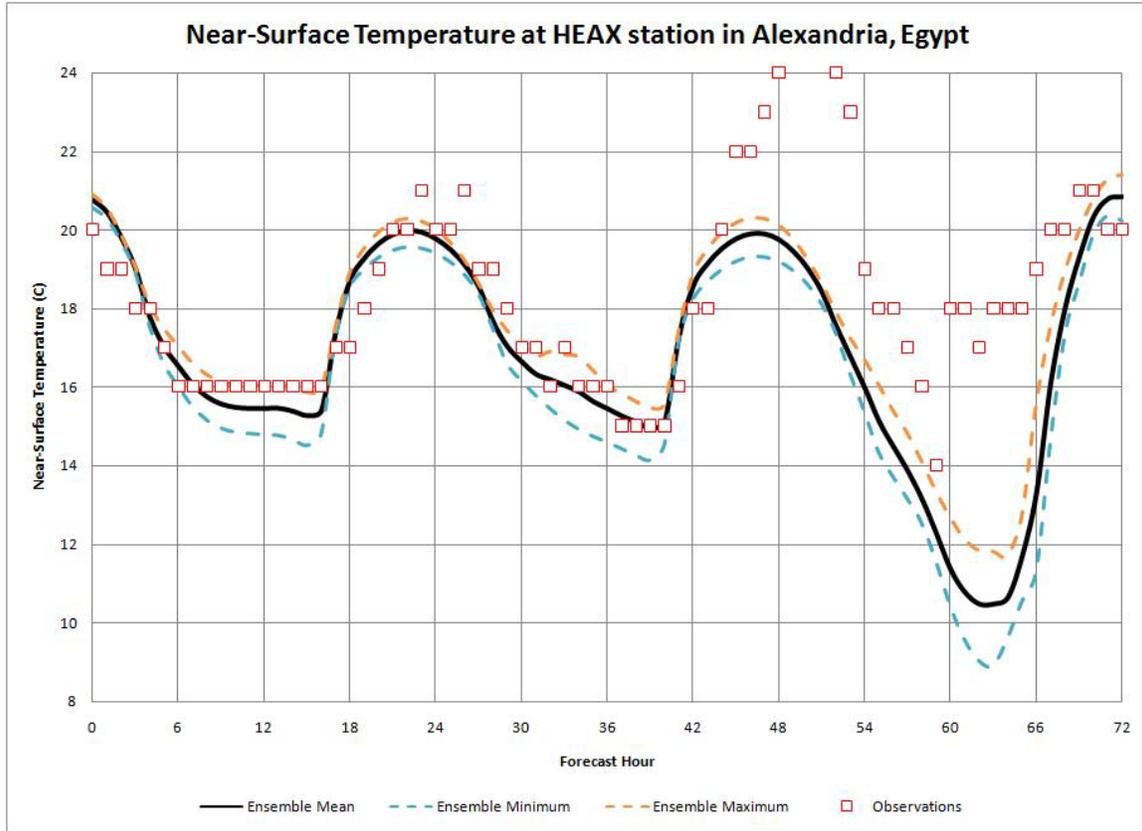


Figure F-29 Min, Max and mean temperature at Alexandria/Nouzha station in spring

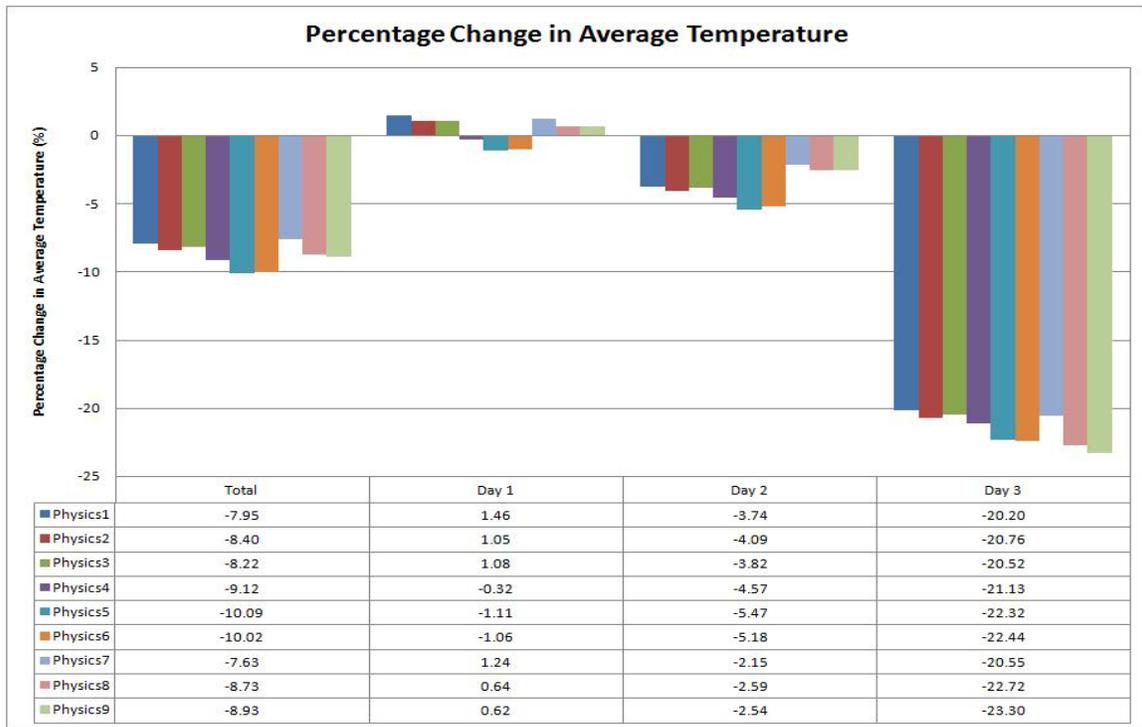


Figure F-30 Change in average temperature at Alexandria/Nouzha station in spring

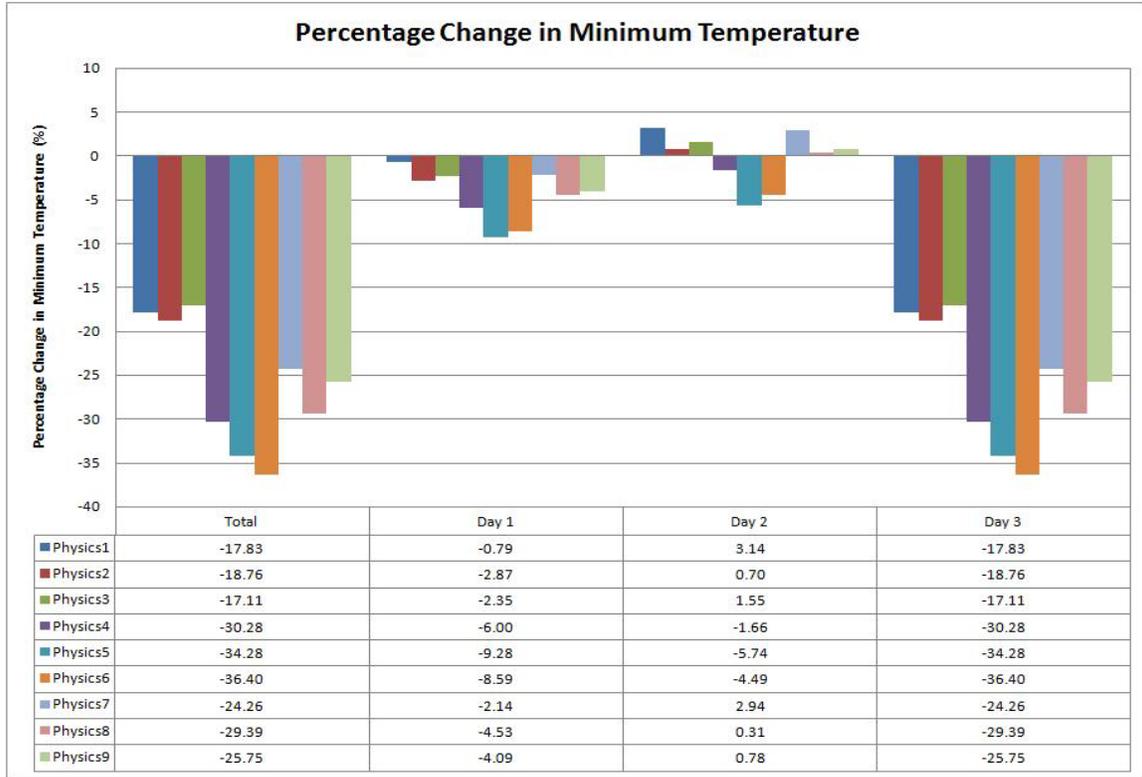


Figure F-31 Change in minimum temperature at Alexandria/Nouzha station in spring

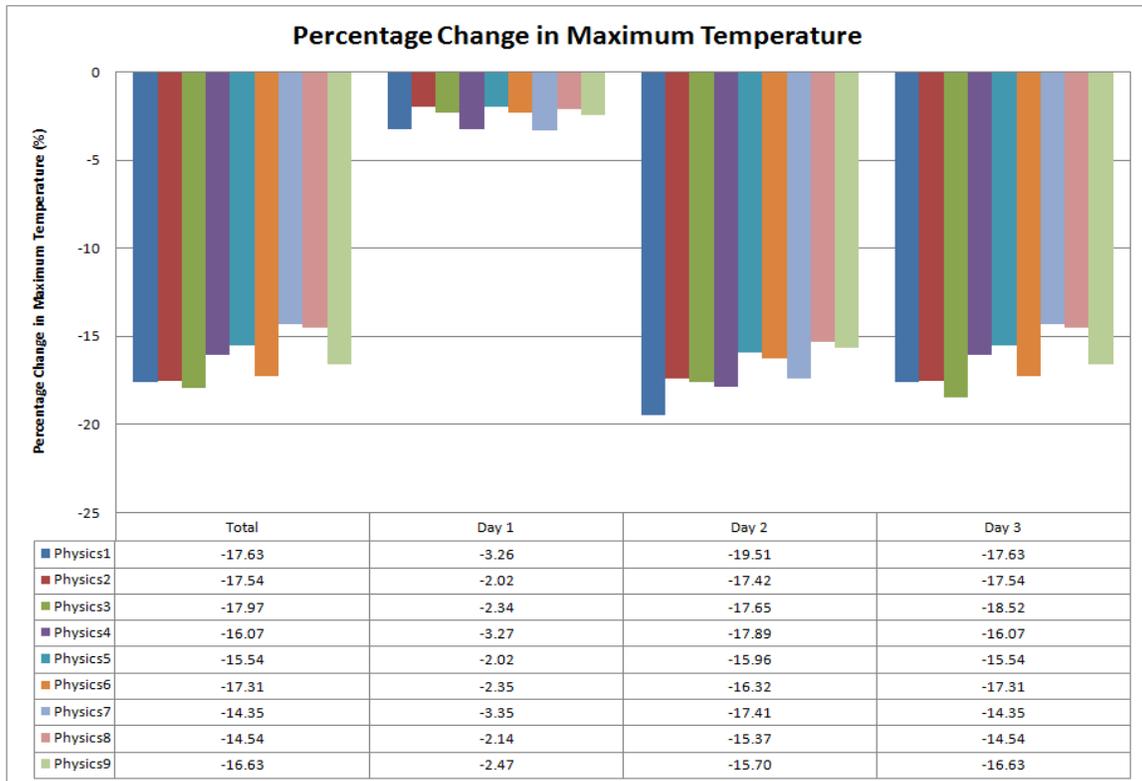


Figure F-32 Change in maximum temperature at Alexandria/Nouzha station in spring

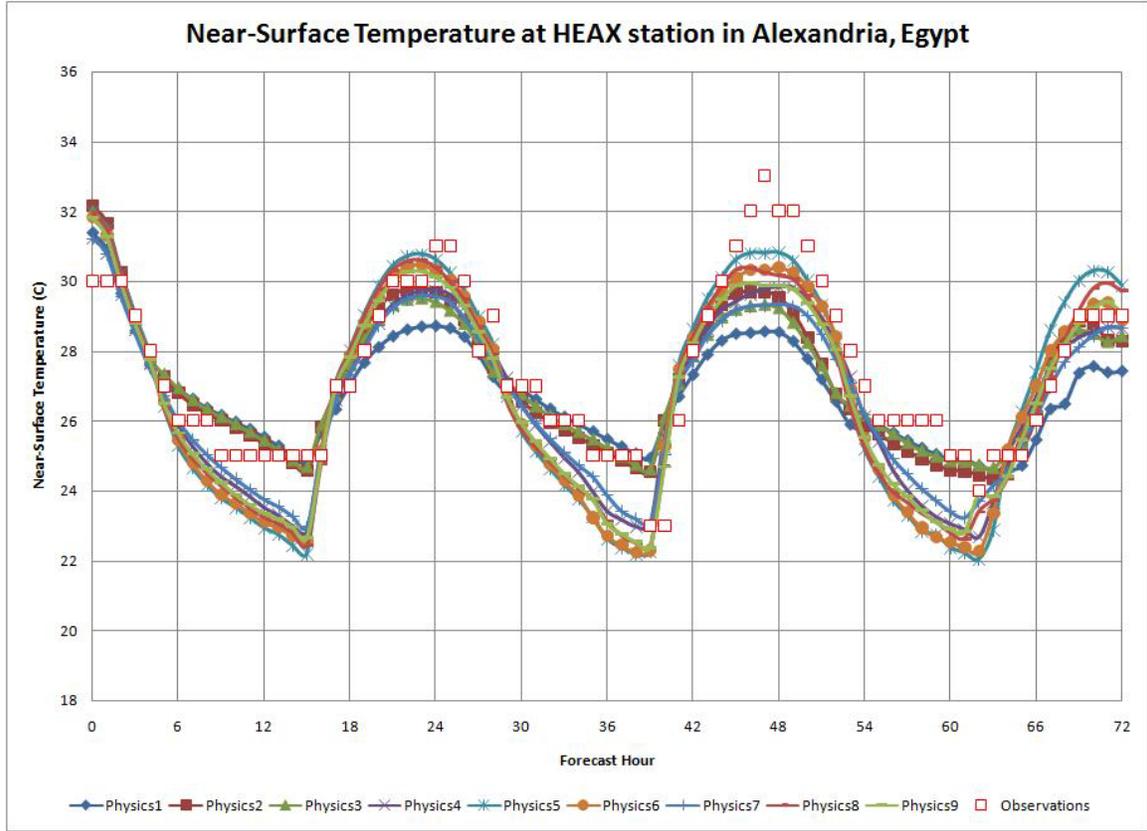


Figure F-33 Near-surface temperature at Alexandria/Nouzha station in summer

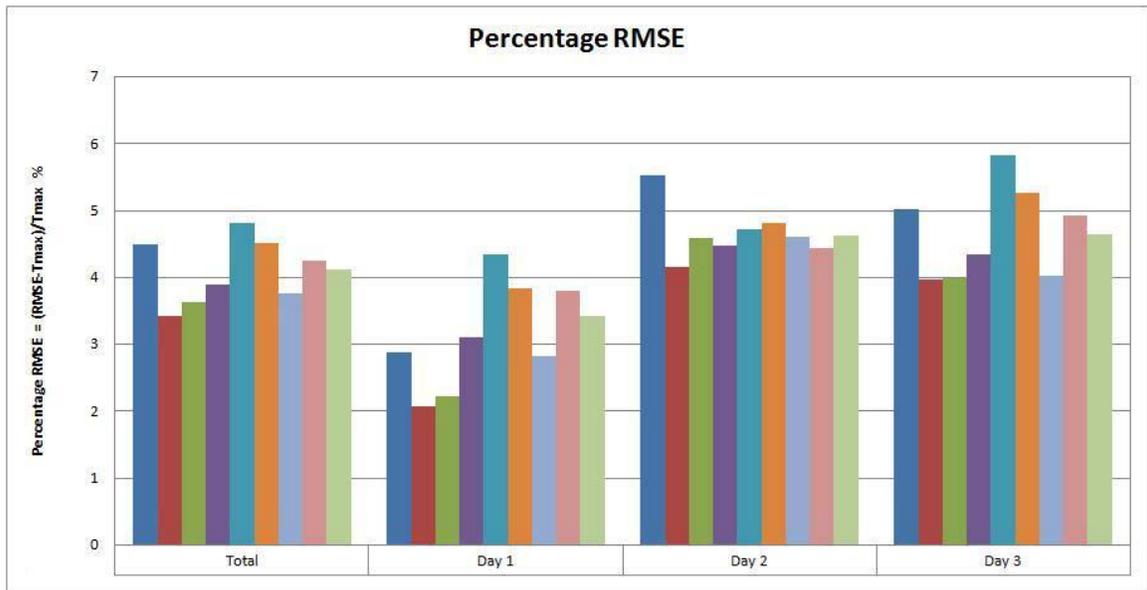


Figure F-34 RMSE of near-surface temperature at Alexandria/Nouzha station in summer

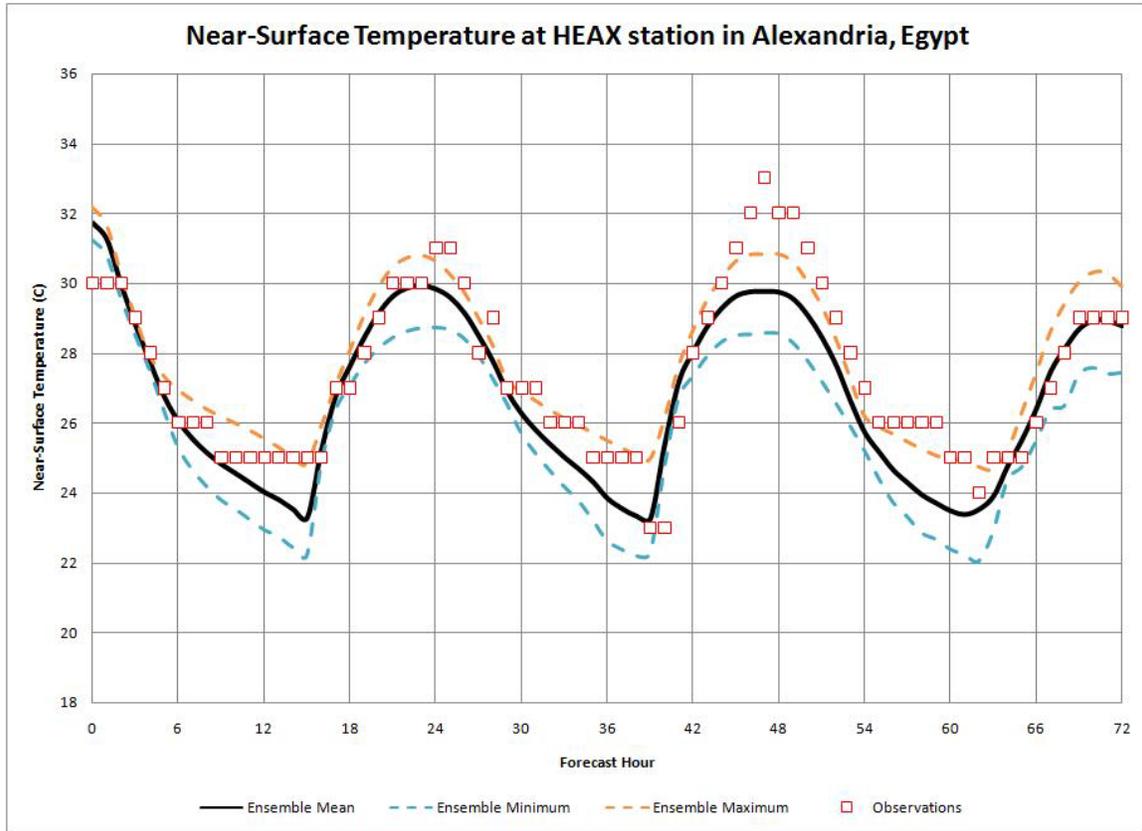


Figure F-35 Min, Max and mean temperature at Alexandria/Nouzha station in summer

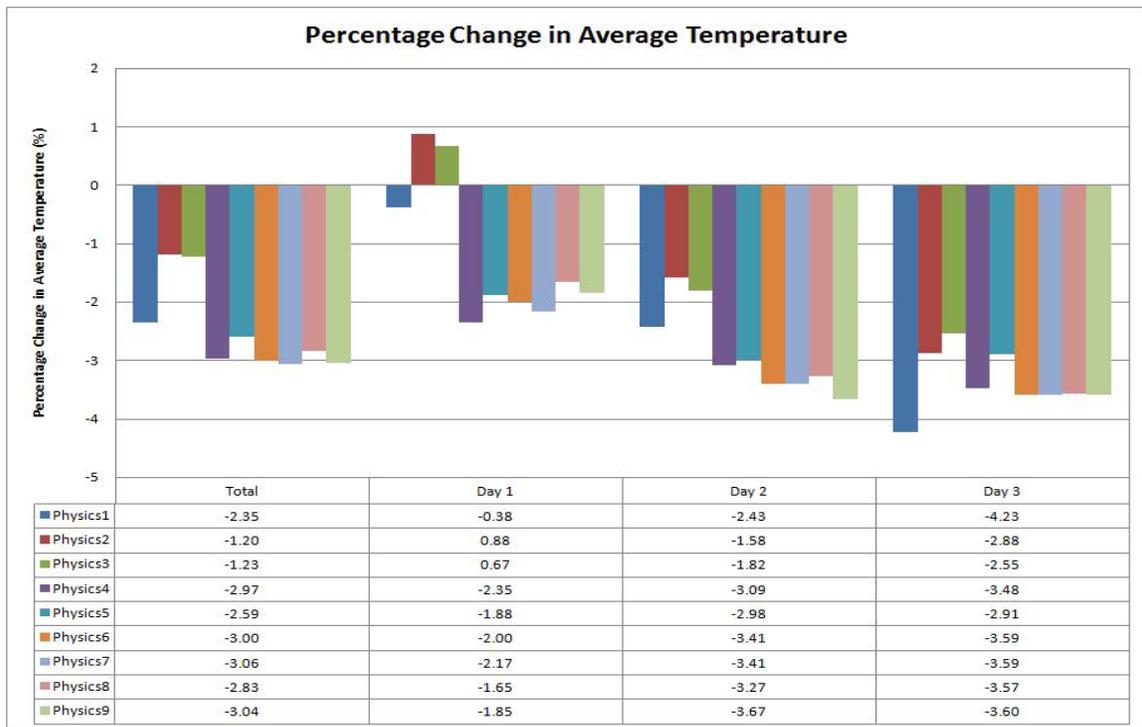


Figure F-36 Change in average temperature at Alexandria/Nouzha station in summer

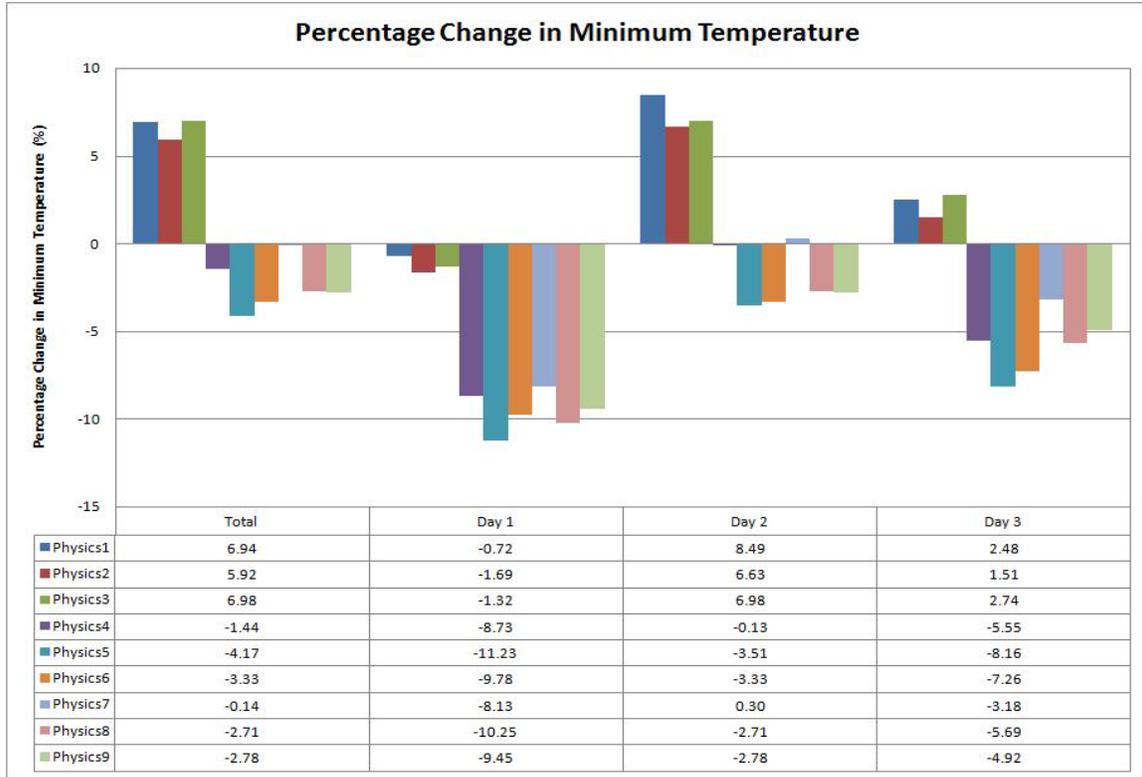


Figure F-37 Change in minimum temperature at Alexandria/Nouzha station in summer

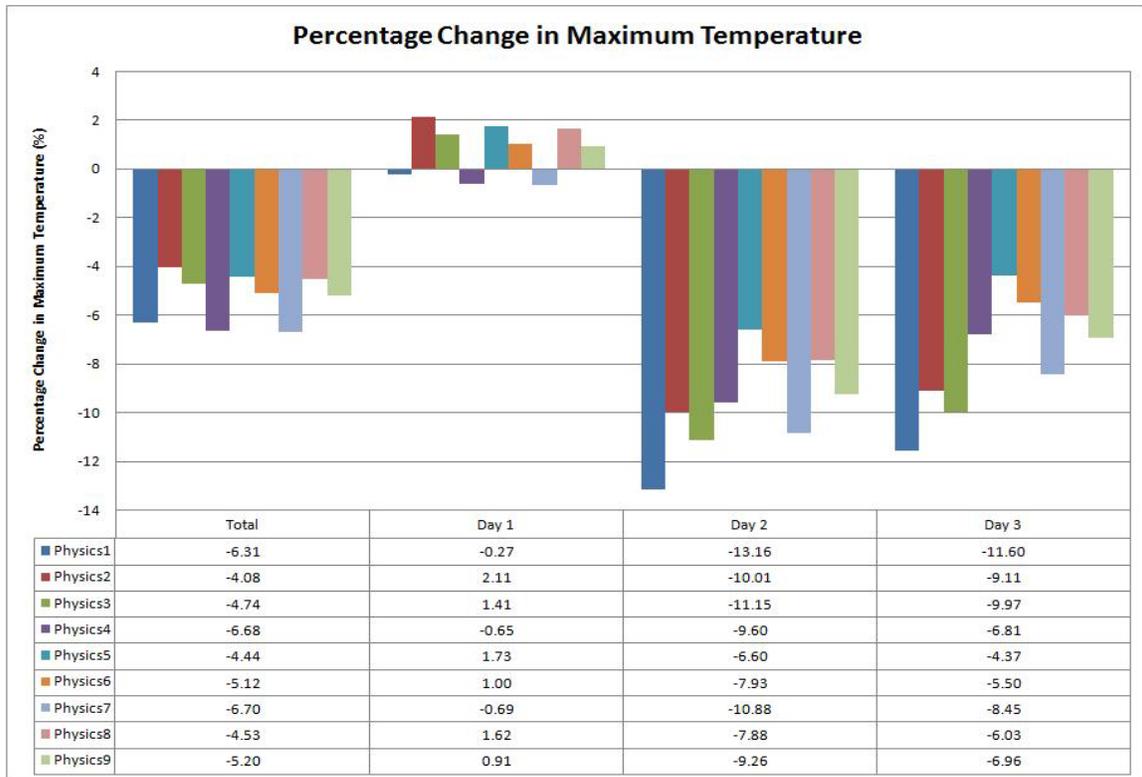


Figure F-38 Change in Maximum temperature at Alexandria/Nouzha station in summer

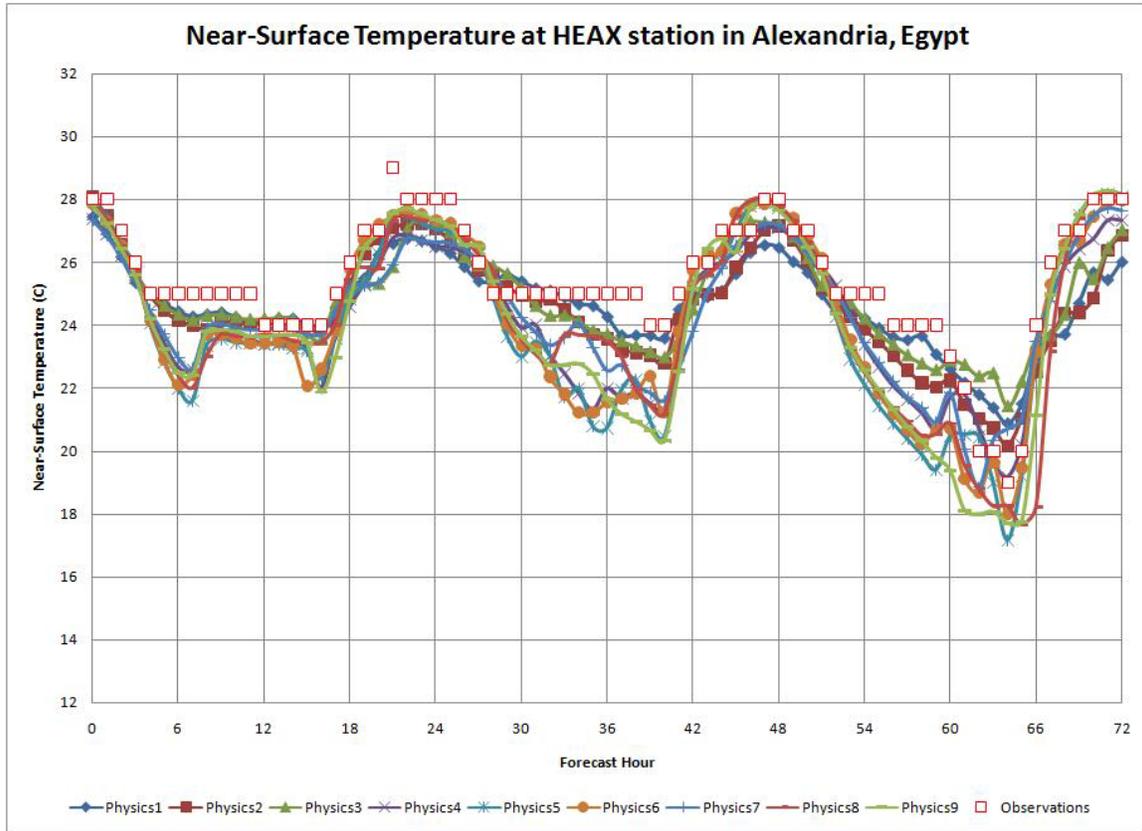


Figure F-39 Near-surface temperature at Alexandria/Nouzha station in autumn

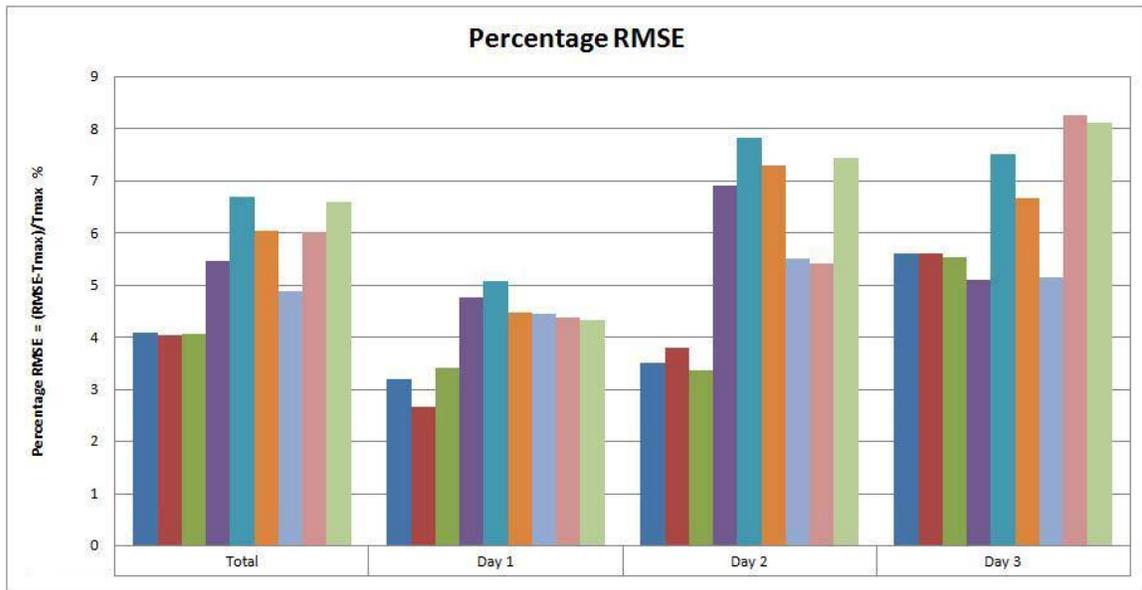


Figure F-40 RMSE of near-surface temperature at Alexandria/Nouzha station in autumn

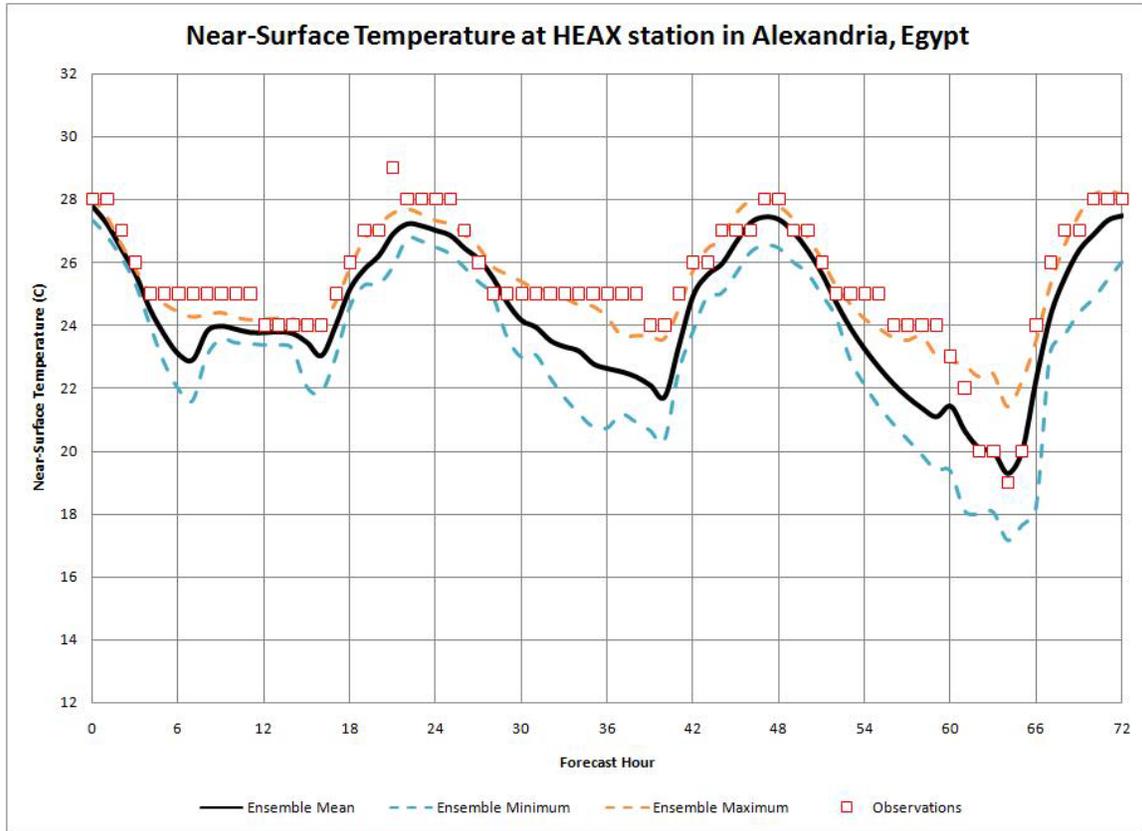


Figure F-41 Min, Max and mean temperature at Alexandria/Nouzha station in autumn

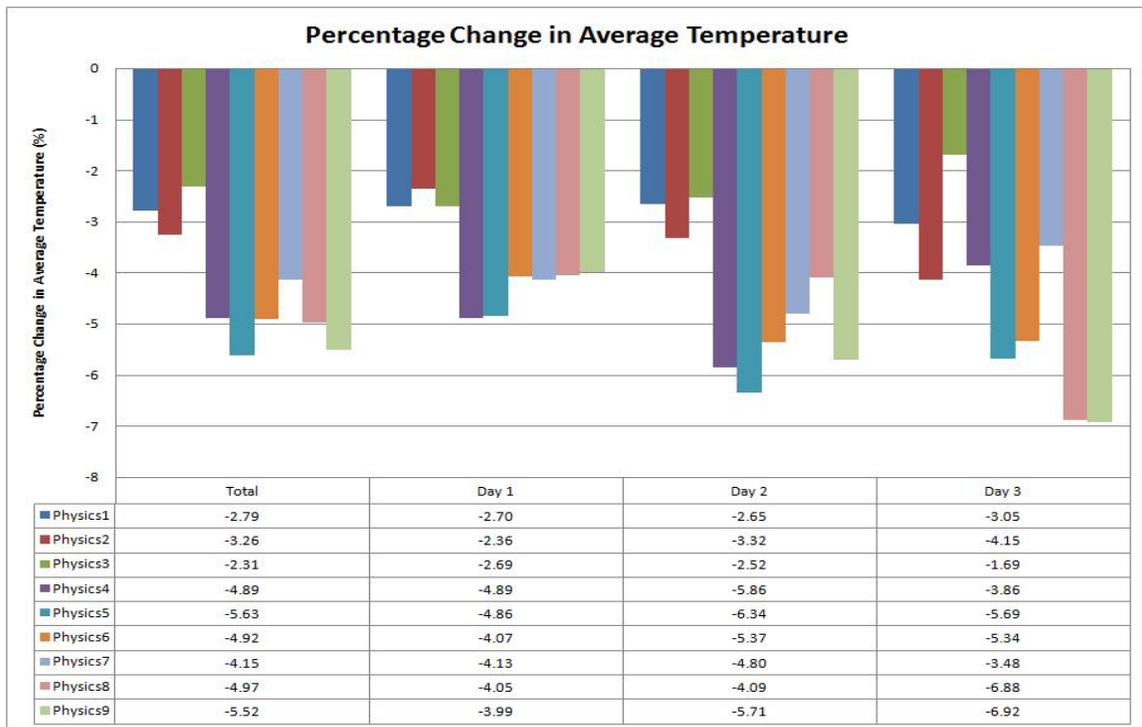


Figure F-42 Change in average temperature at Alexandria/Nouzha station in autumn

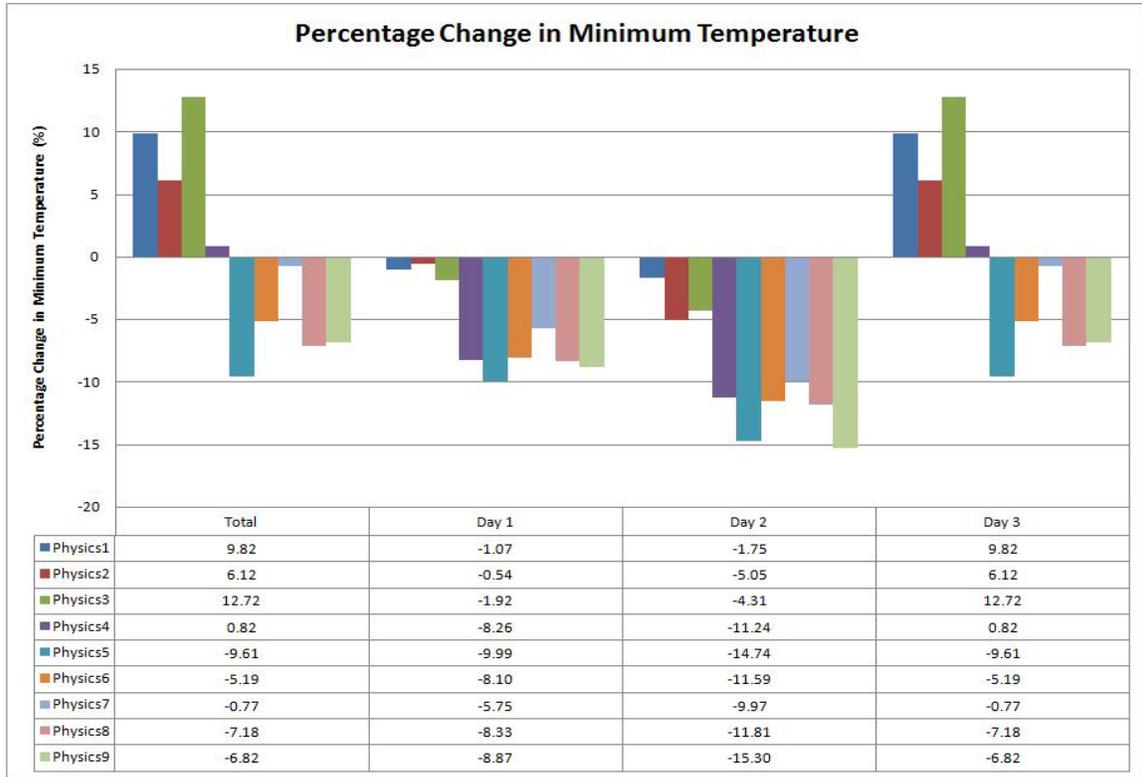


Figure F-43 Change in minimum temperature at Alexandria/Nouzha station in autumn

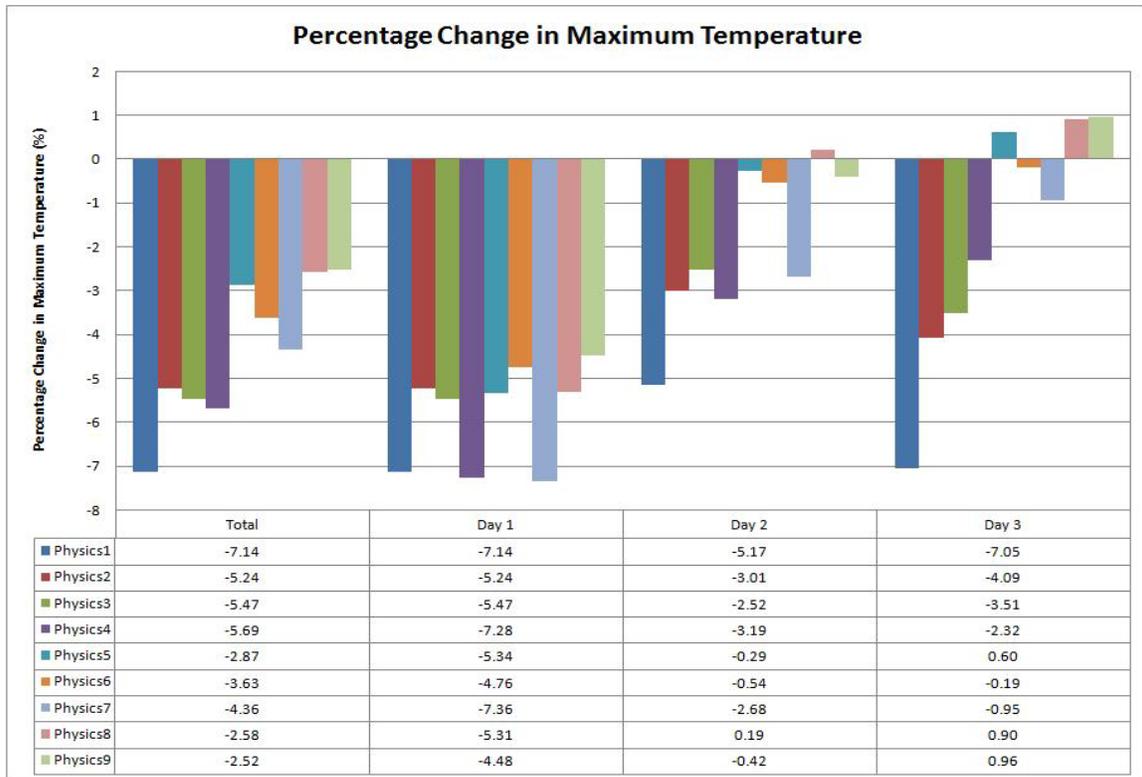


Figure F-44 Change in maximum temperature at Alexandria/Nouzha station in autumn

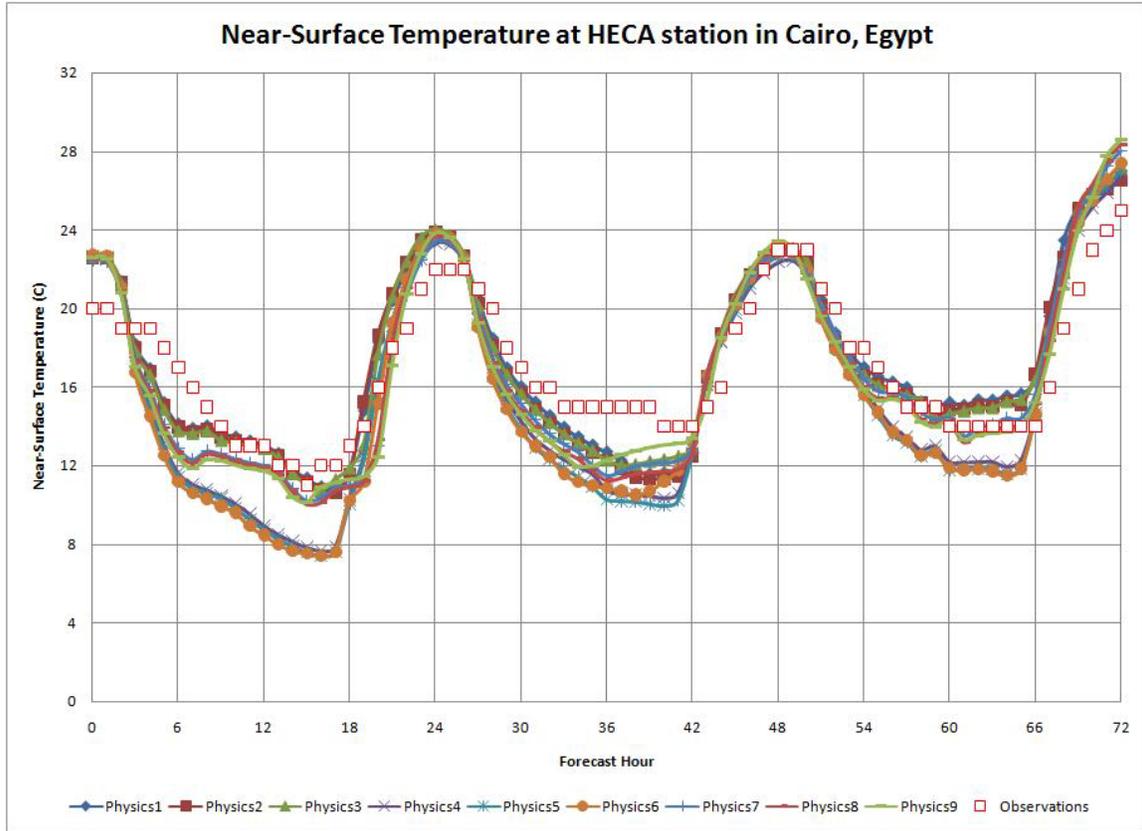


Figure F-45 Near-surface temperature at Cairo Airport station in winter

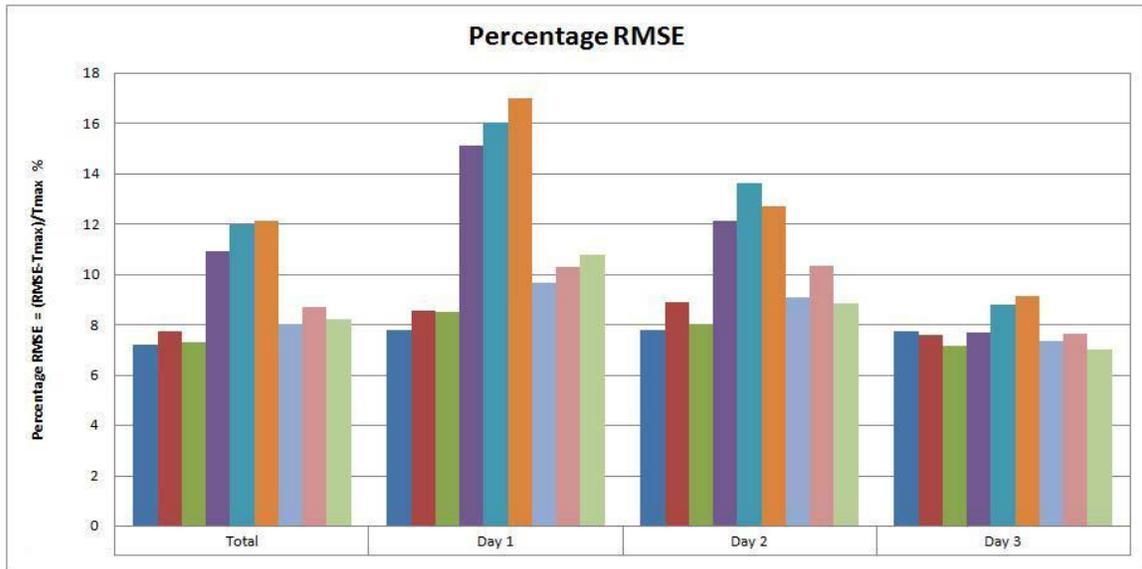


Figure F-46 RMSE of near-surface temperature at Cairo Airport station in winter

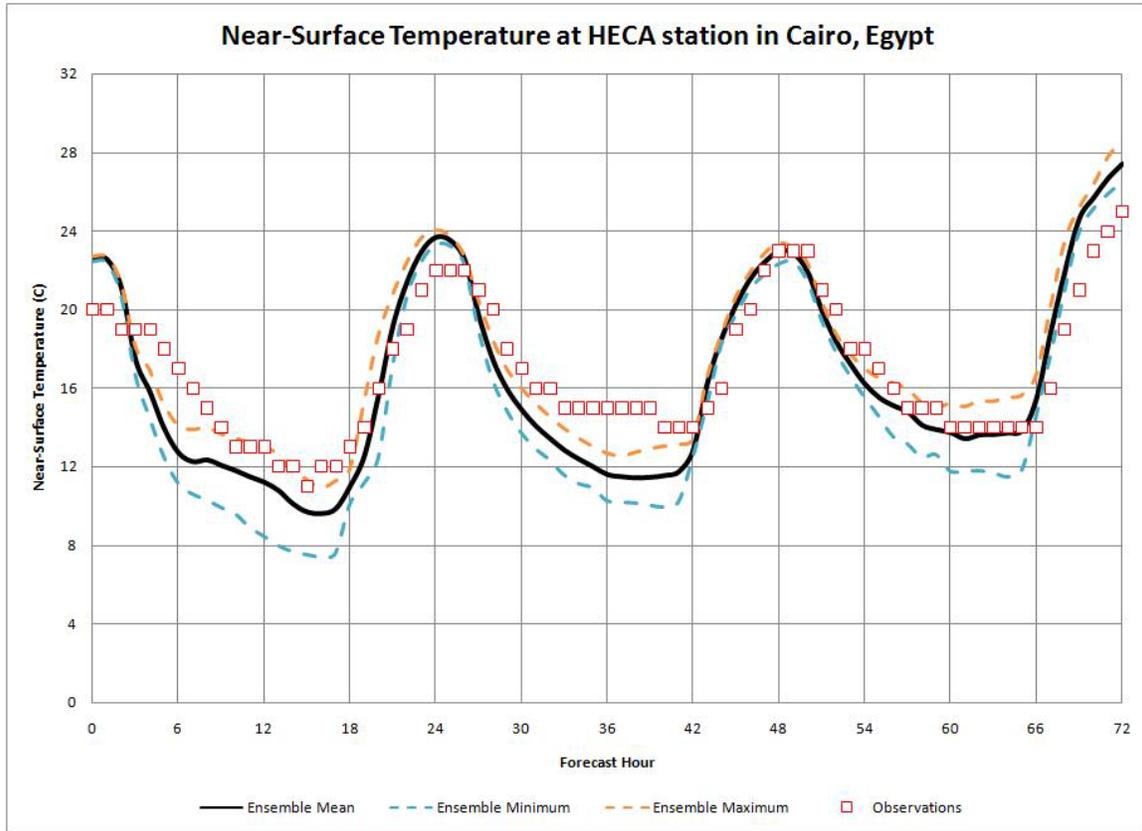


Figure F-47 Min, Max and mean temperature at Cairo Airport station in winter

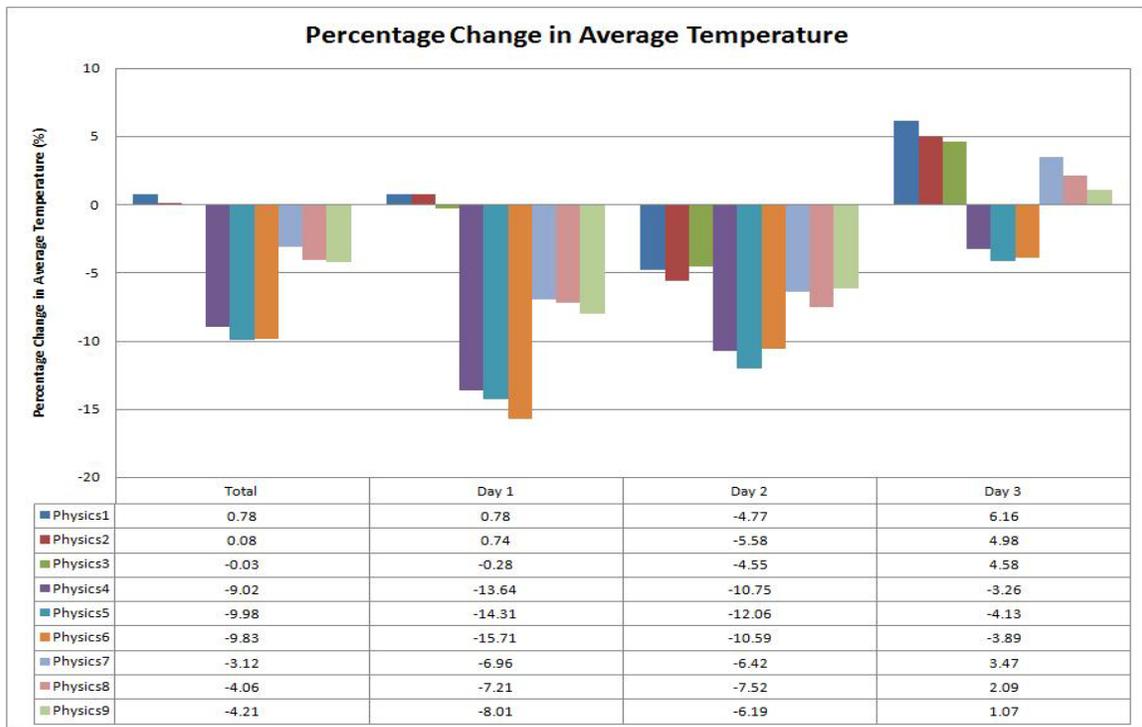


Figure F-48 Change in average temperature at Cairo station in winter

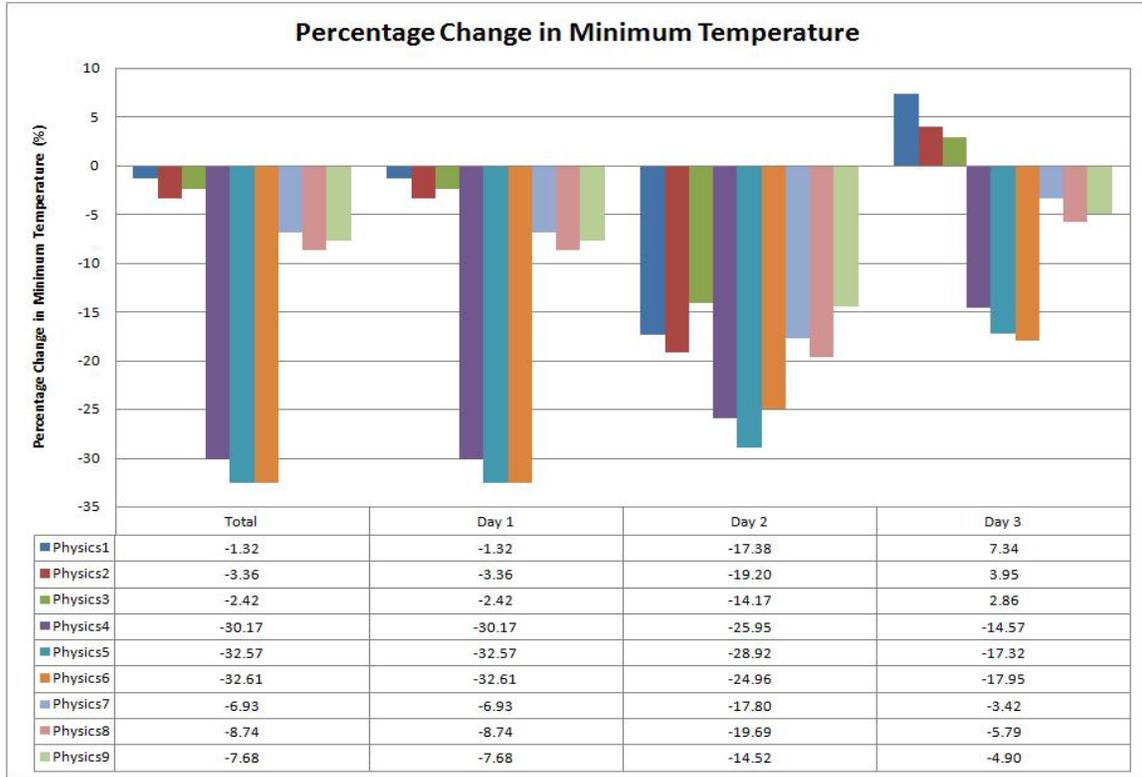


Figure F-49 Change in minimum temperature at Cairo Airport station in winter

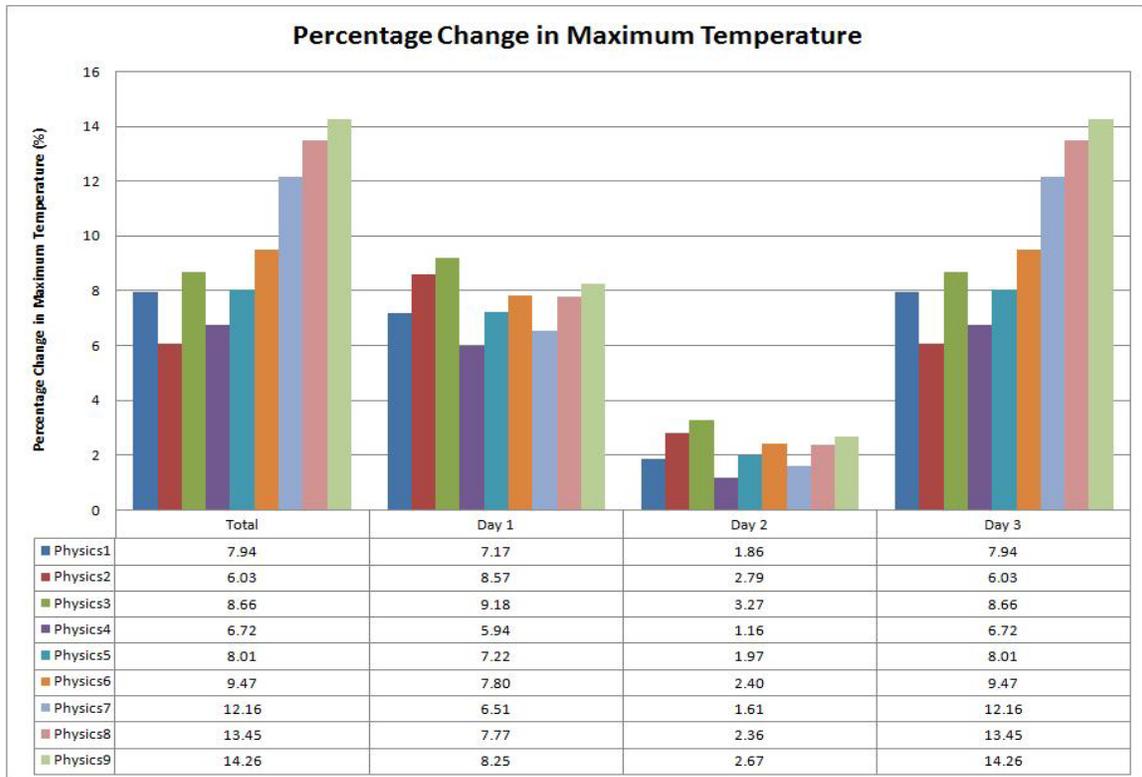


Figure F-50 Change in maximum temperature at Cairo Airport station in winter

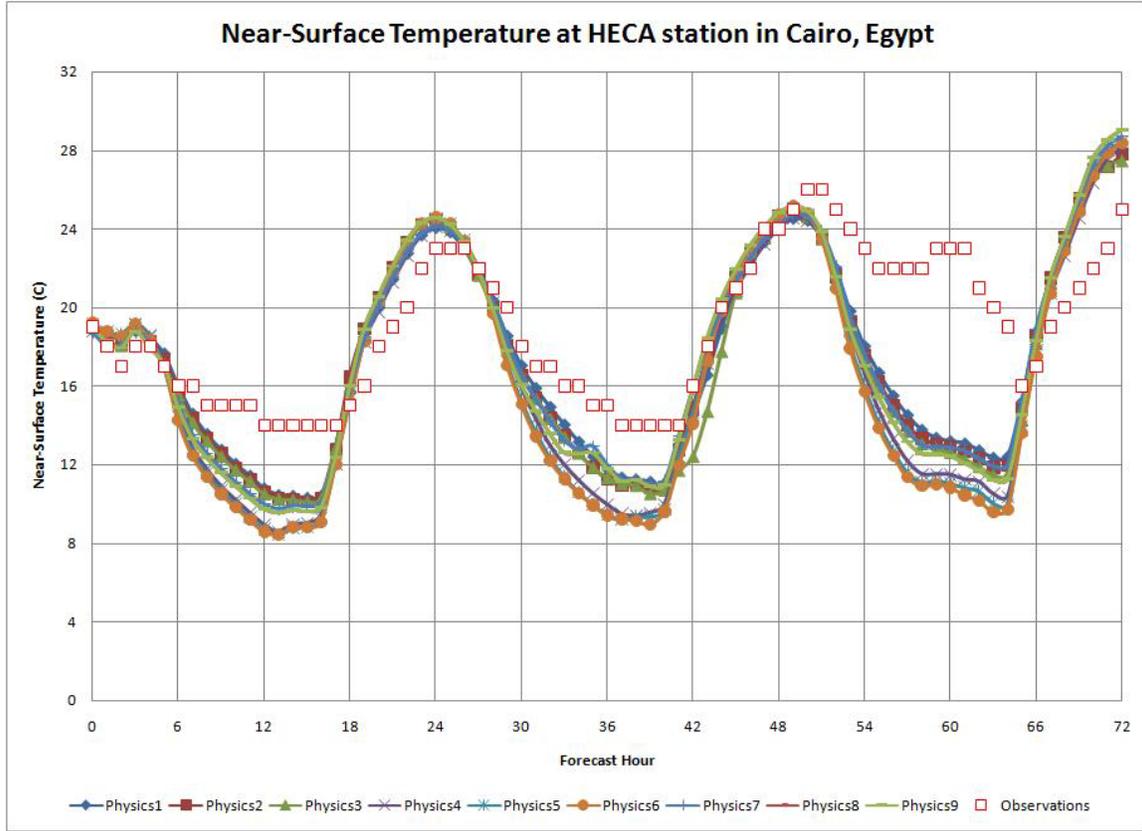


Figure F-51 Near-surface temperature at Cairo International station in spring

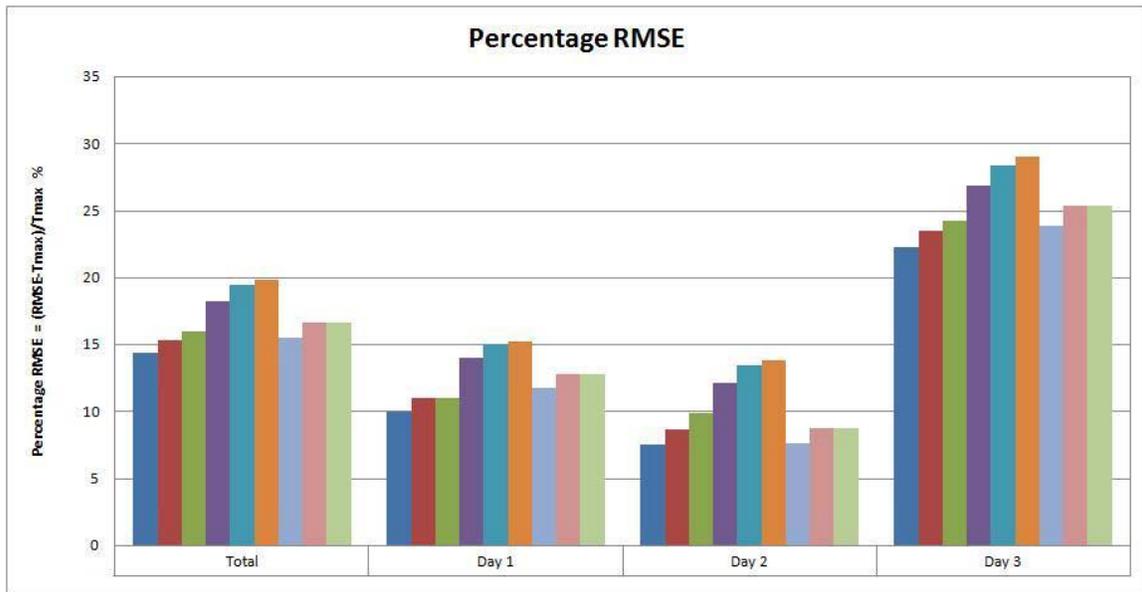


Figure F-52 RMSE of near-surface temperature at Cairo Airport station in spring

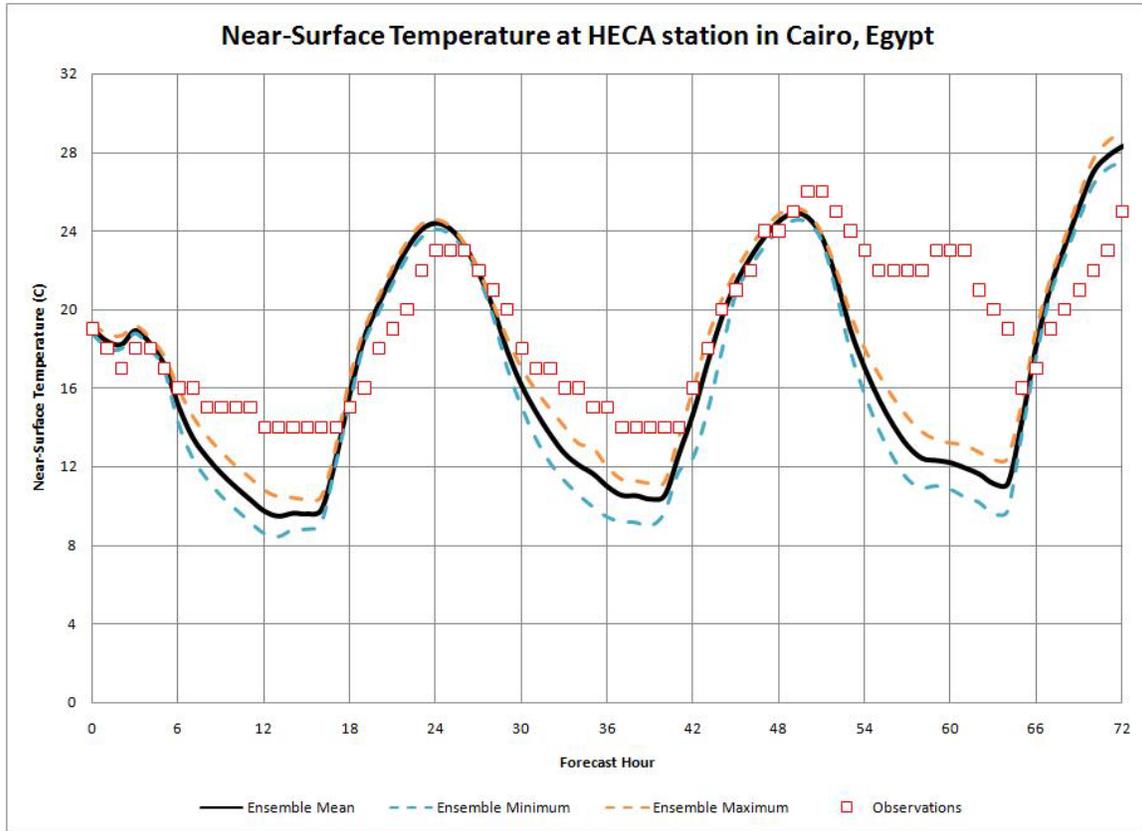


Figure F-53 Min, Max and mean temperature at Cairo Airport station in spring

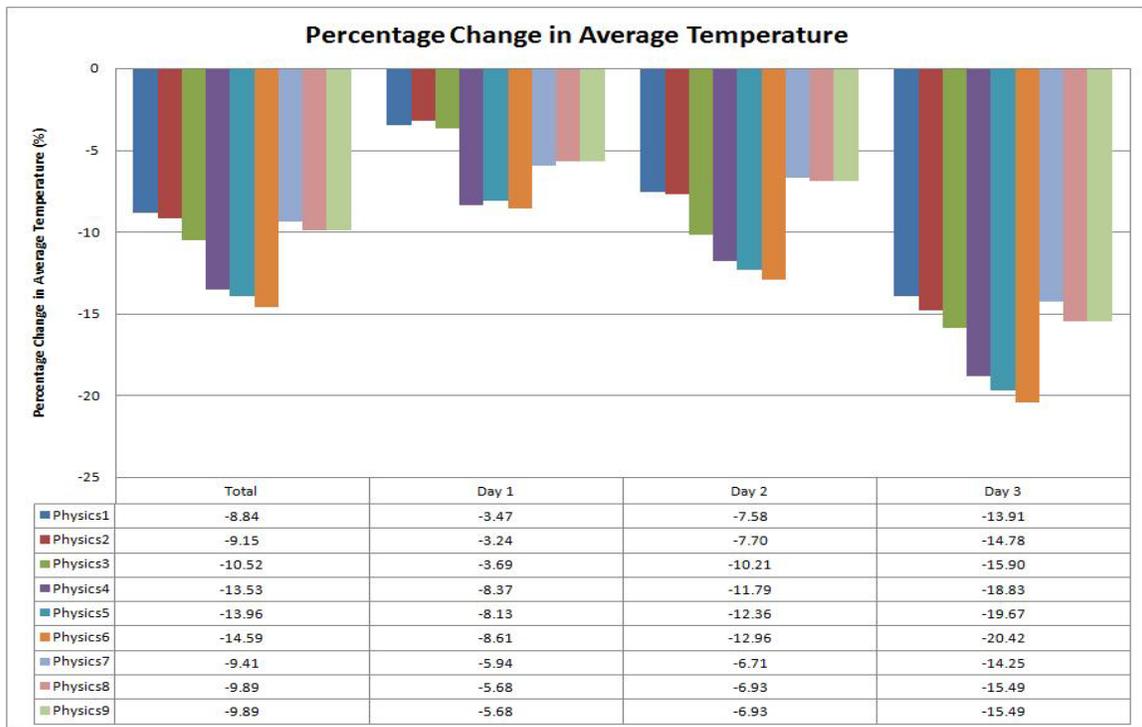


Figure F-54 Change in average temperature at Cairo International station in spring

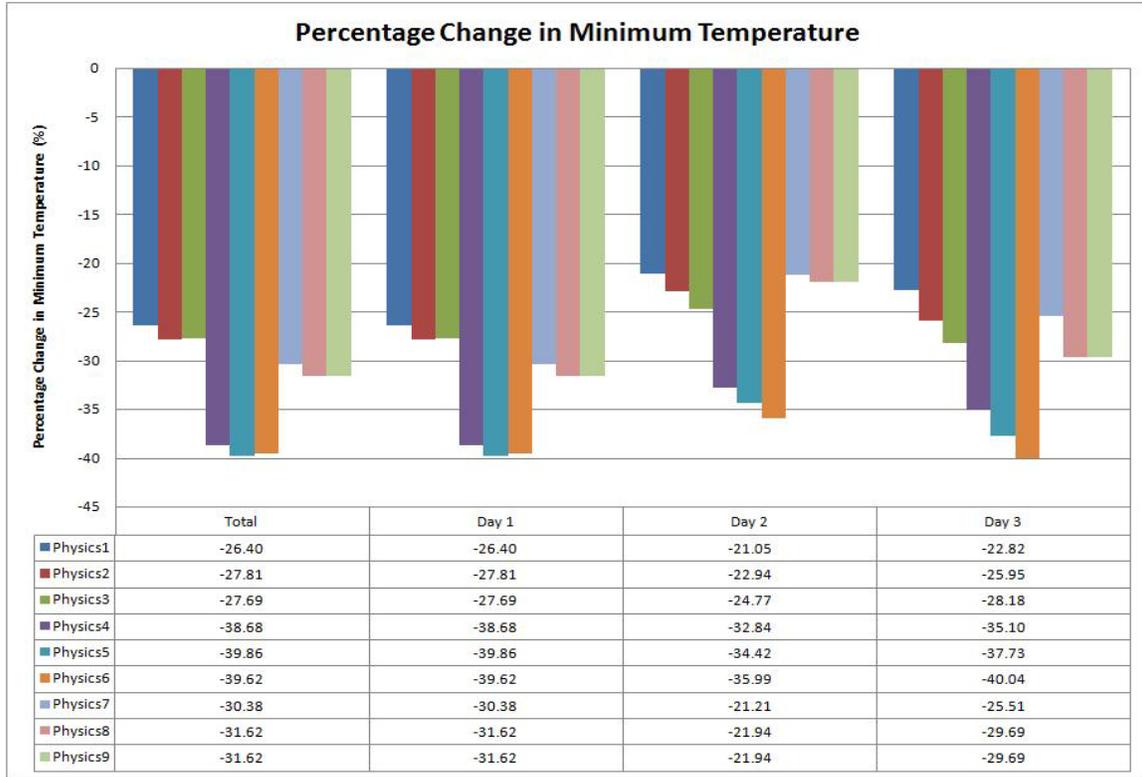


Figure F-55 Change in minimum temperature at Cairo Airport station in spring

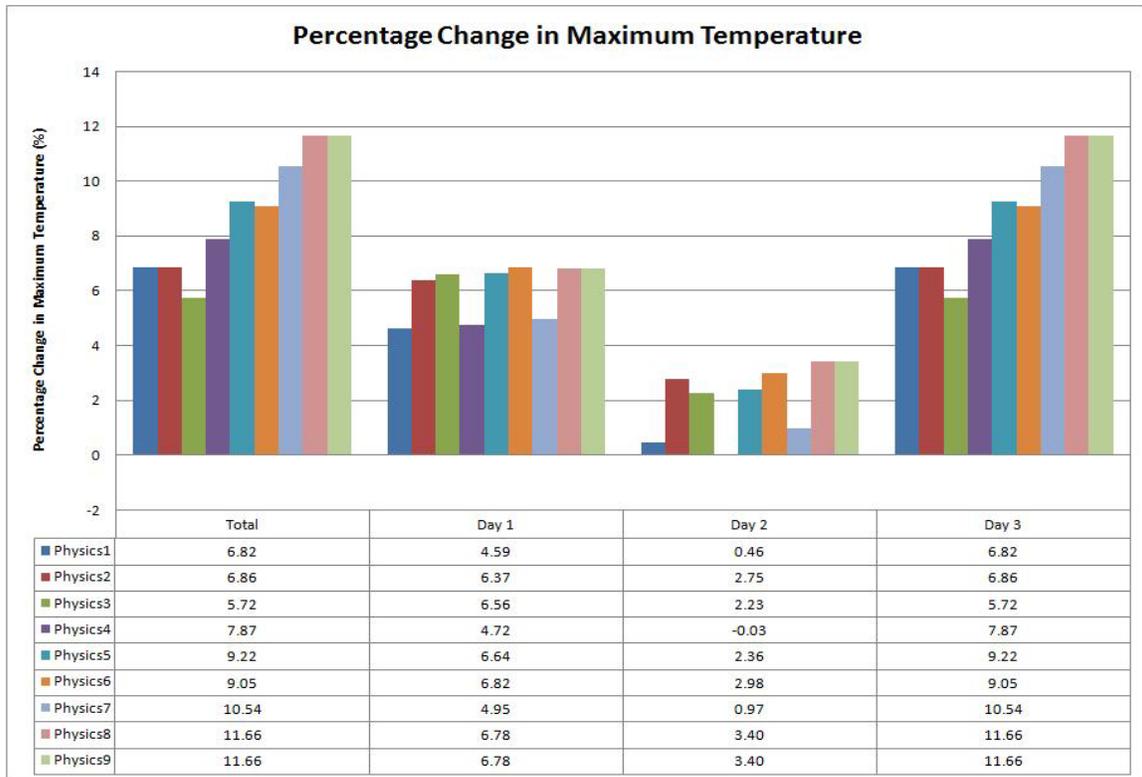


Figure F-56 Change in maximum temperature at Cairo Airport station in spring

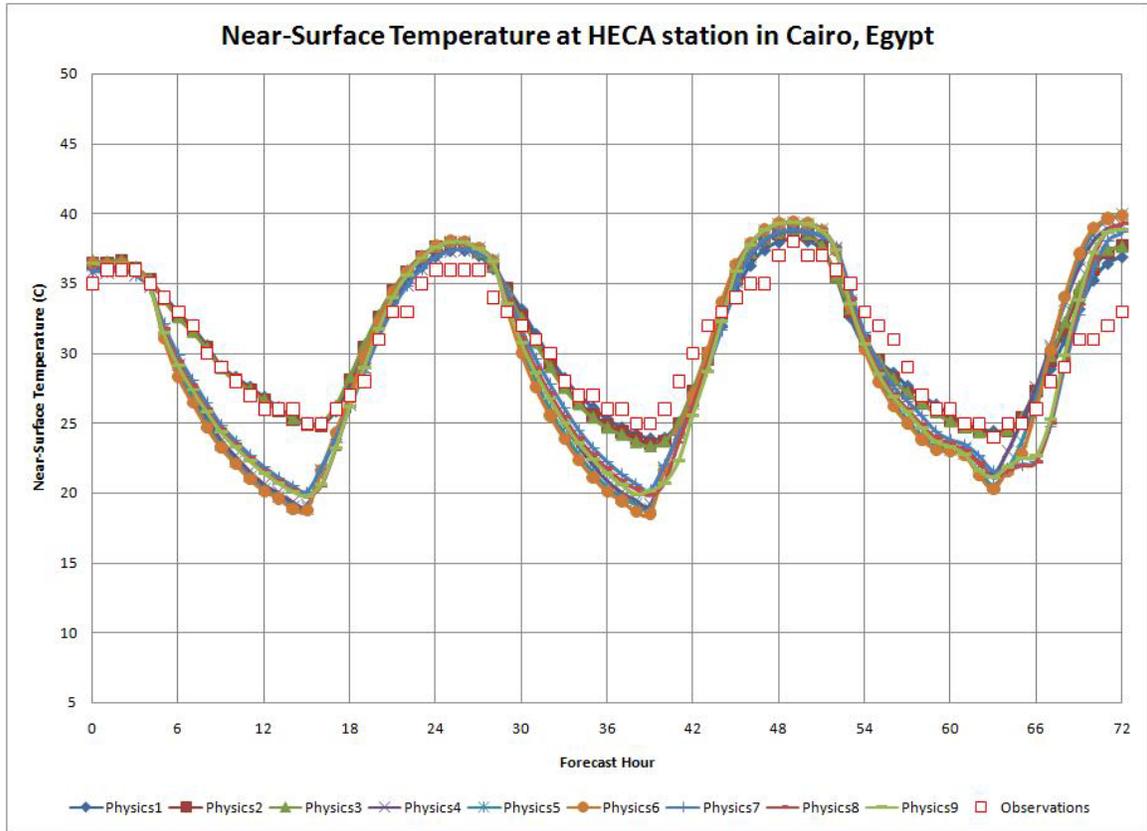


Figure F-57 Near-surface temperature at Cairo Airport station in summer

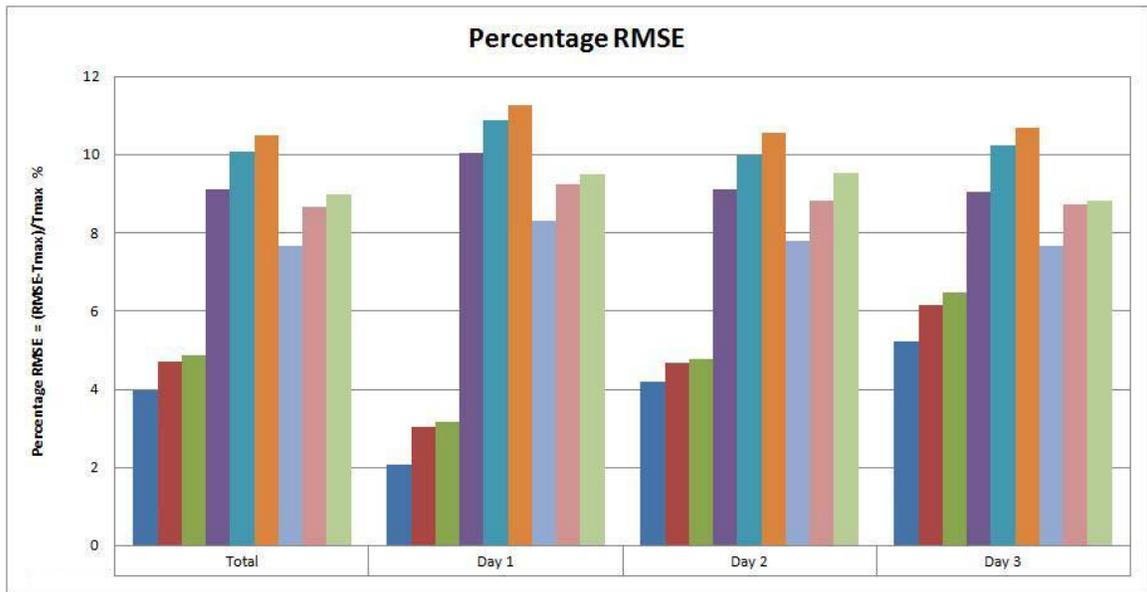


Figure F-58 RMSE of near-surface temperature at Cairo Airport station in summer

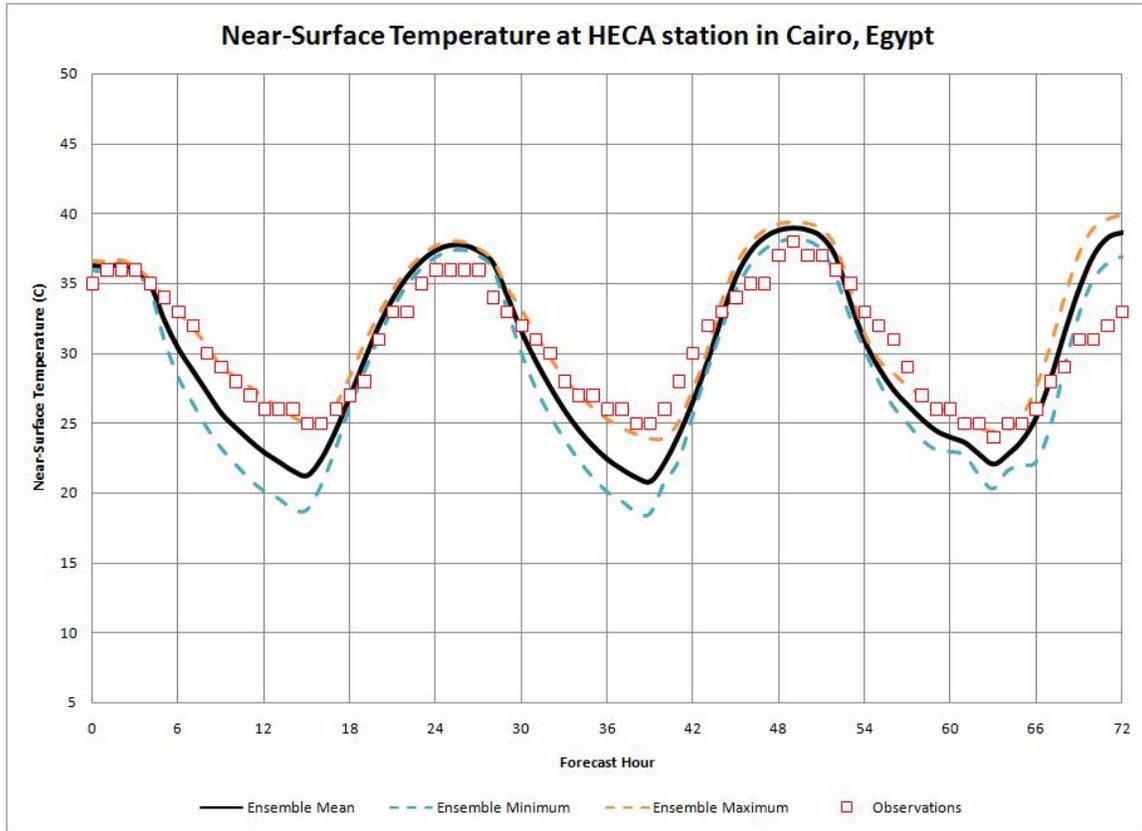


Figure F-59 Min, Max and mean temperature at Cairo Airport station in summer

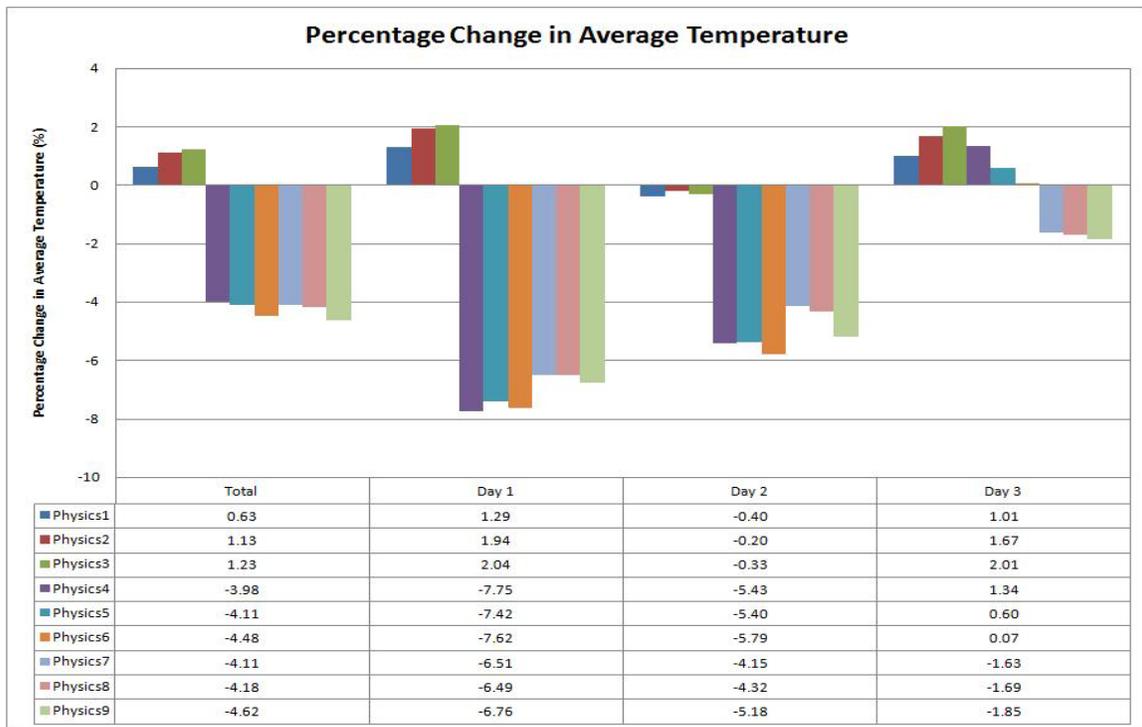


Figure F-60 Change in average temperature at Cairo Airport station in summer

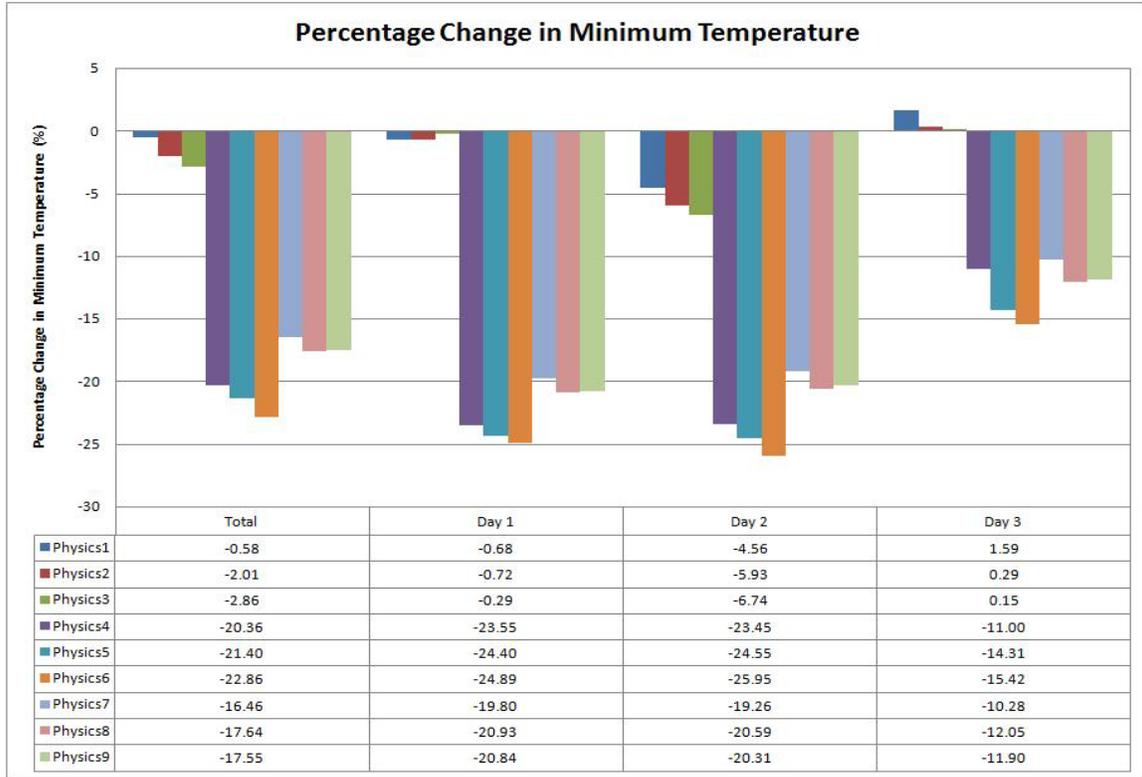


Figure F-61 Change in minimum temperature at Cairo Airport station in summer

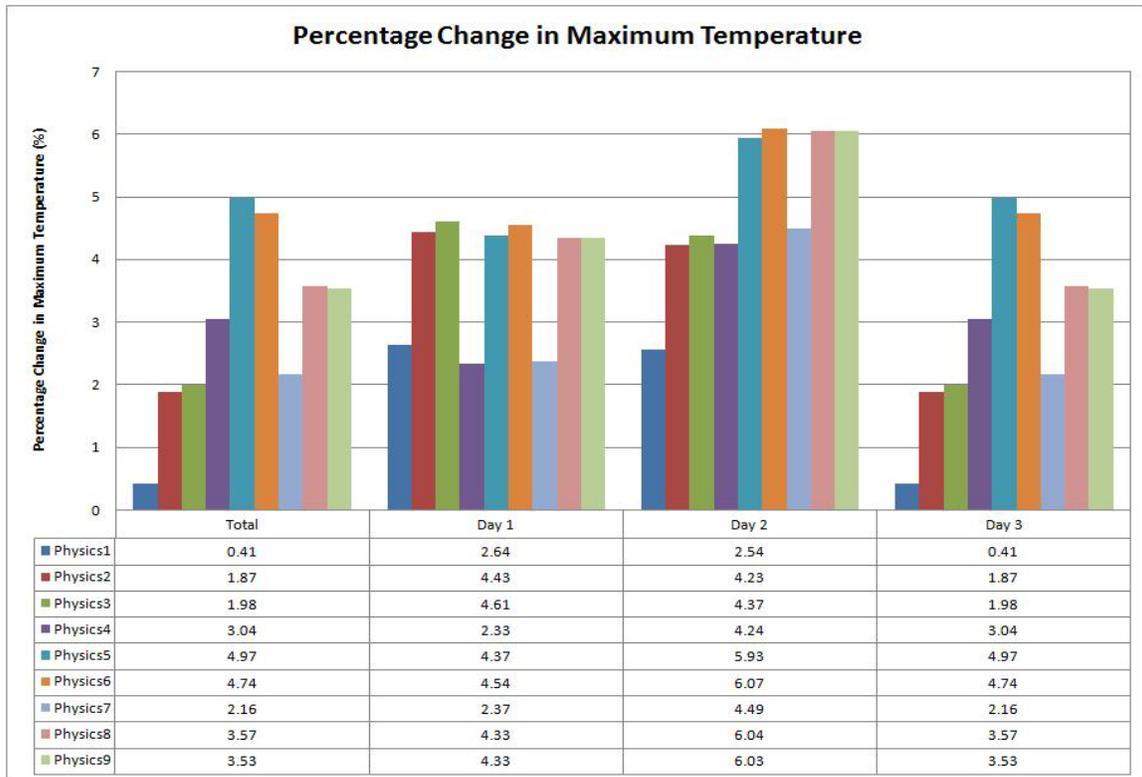


Figure F-62 Change in maximum temperature at Cairo Airport station in summer

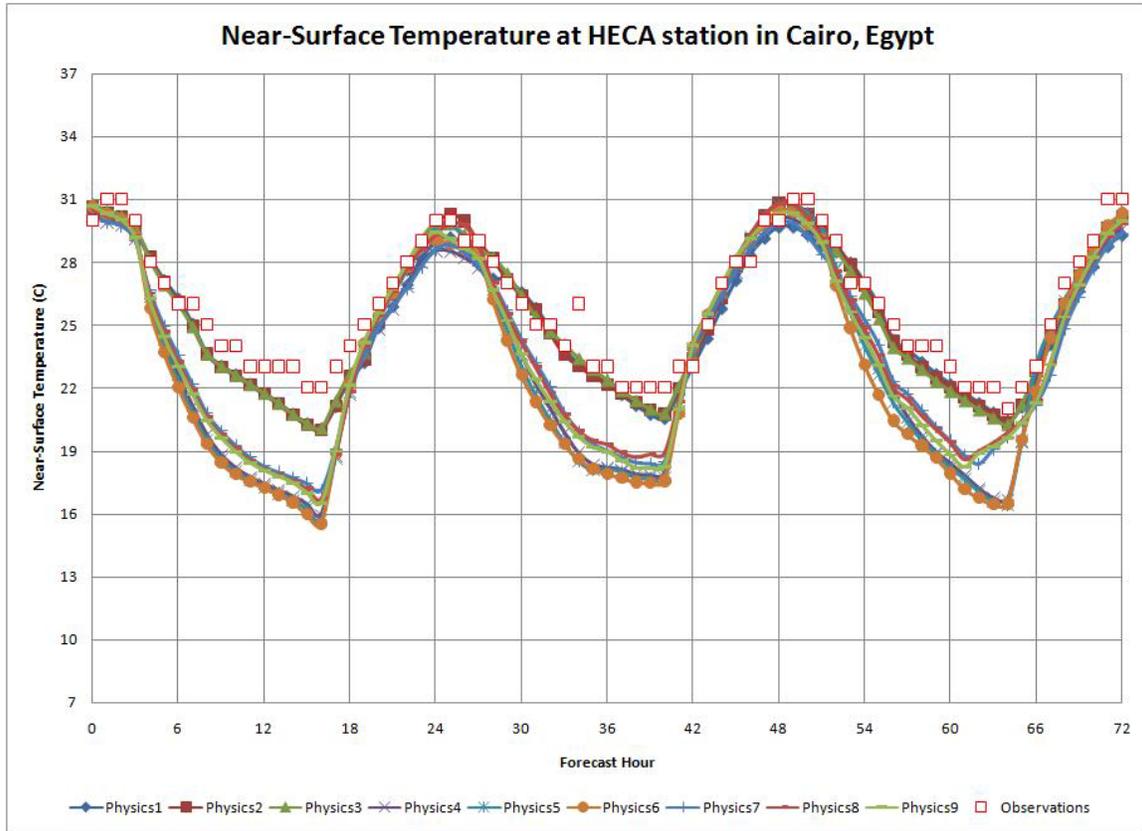


Figure F-63 Near-surface temperature at Cairo Airport station in autumn

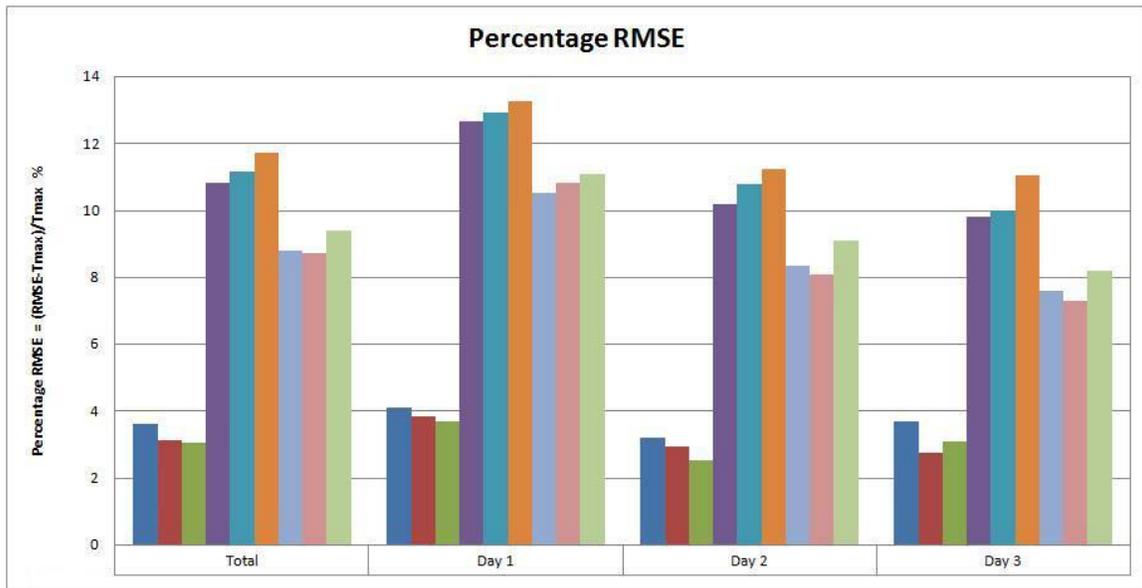


Figure F-64 RMSE of near-surface temperature at Cairo Airport station in autumn

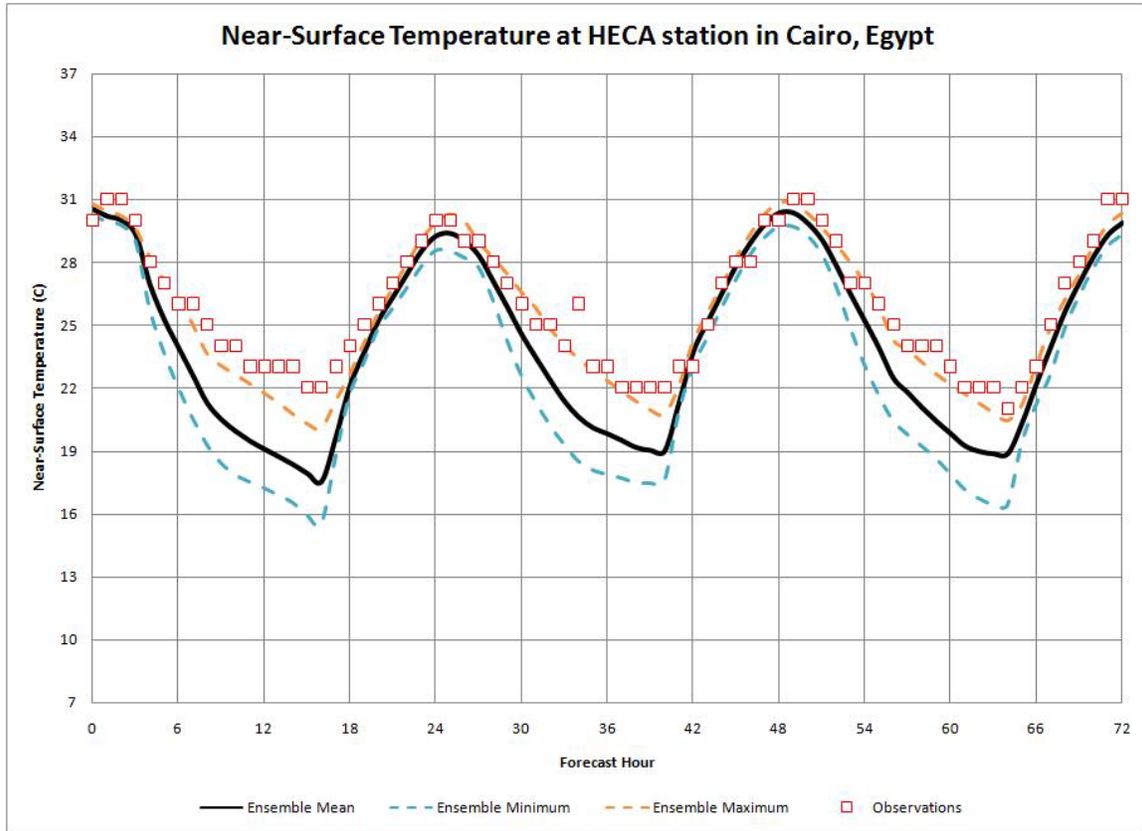


Figure F-65 Min, Max and mean temperature at Cairo Airport station in autumn

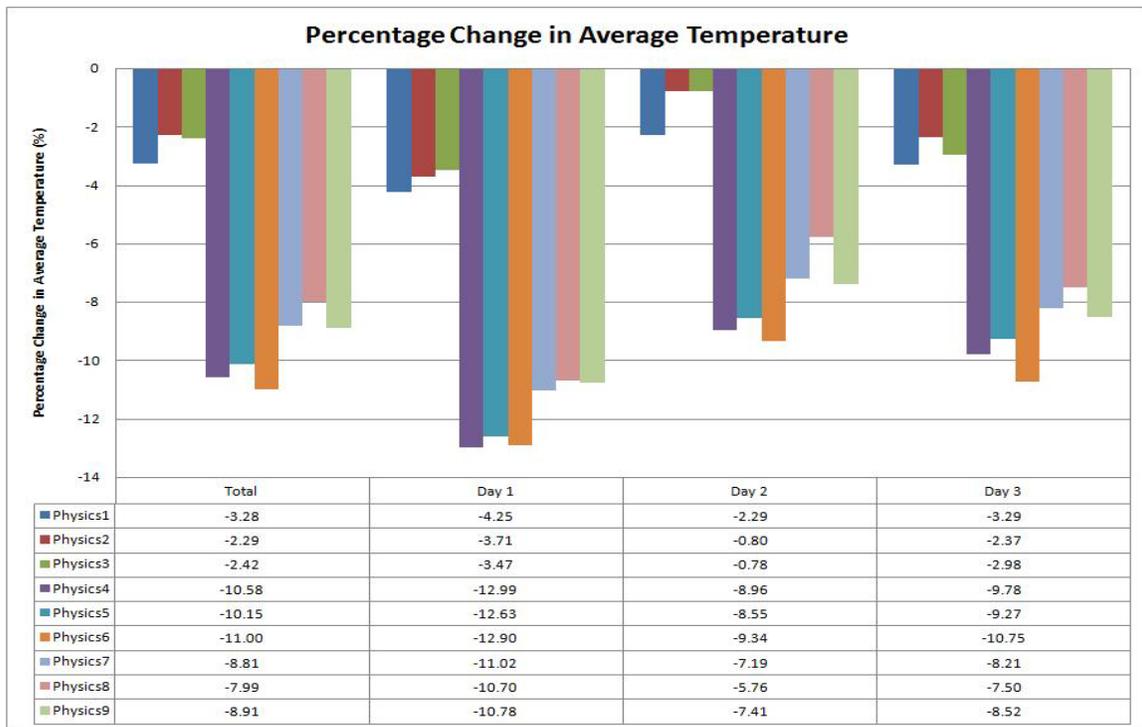


Figure F-66 Change in average temperature at Cairo Airport station in autumn

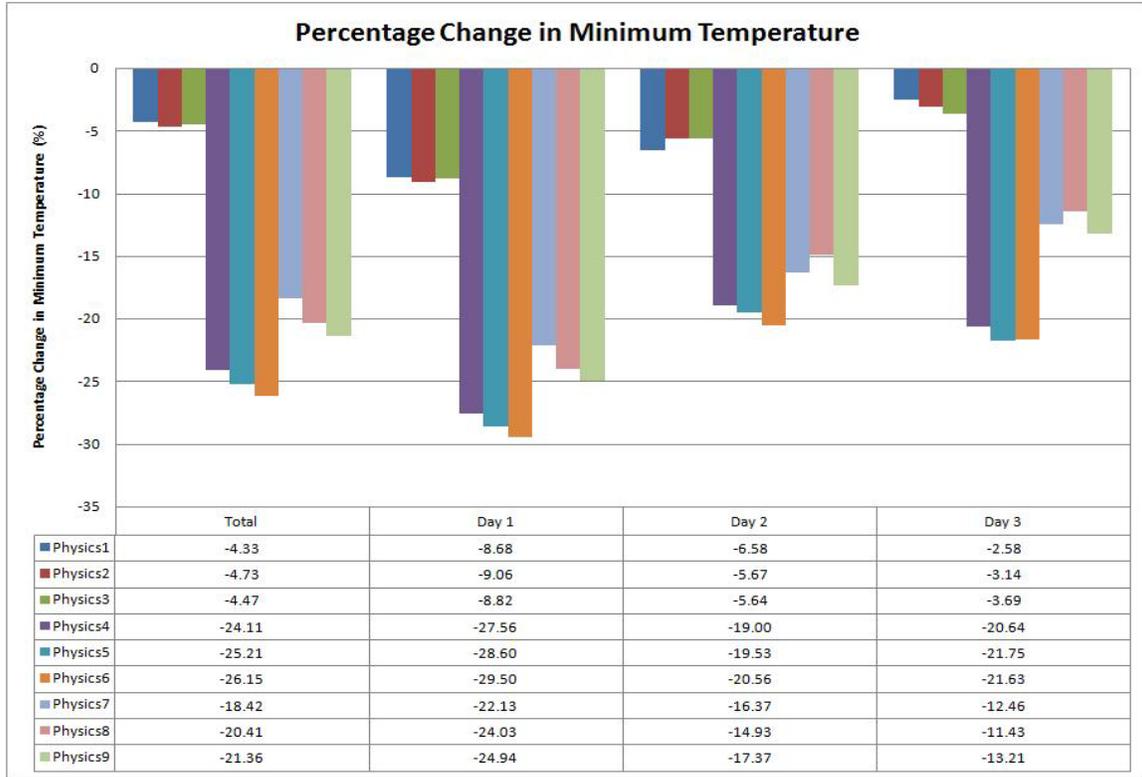


Figure F-67 Change in minimum temperature at Cairo Airport station in autumn

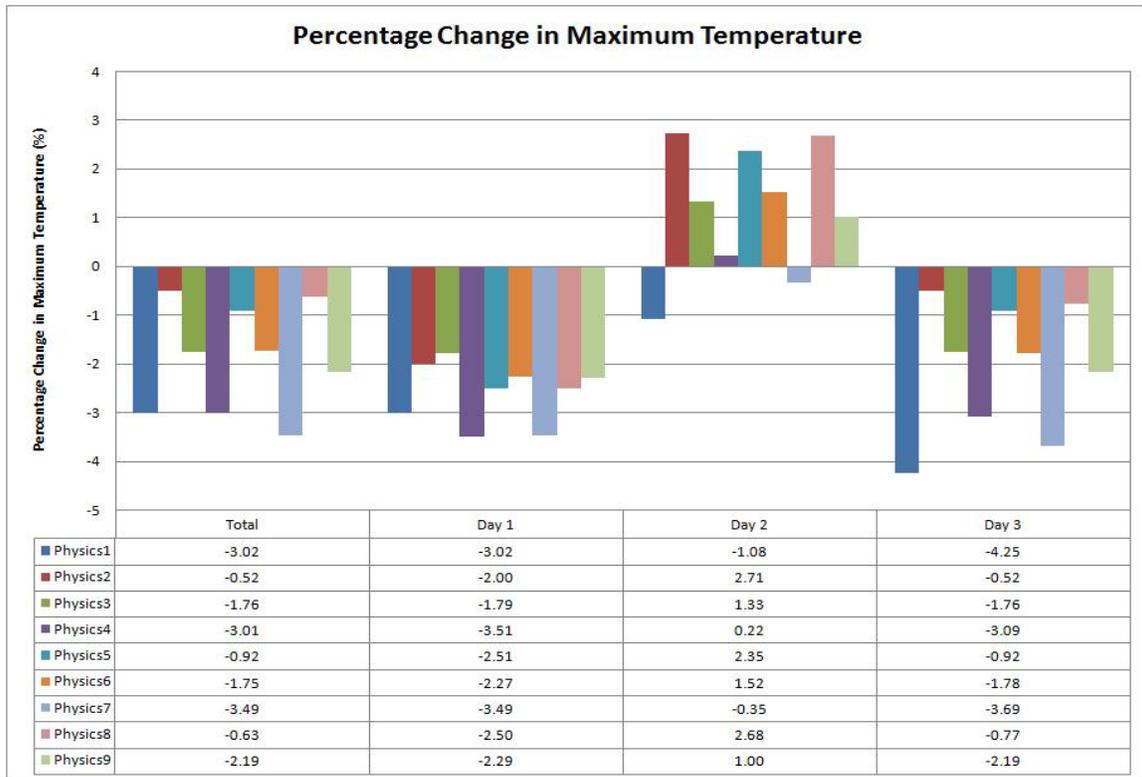


Figure F-68 Change in maximum temperature at Cairo Airport station in autumn

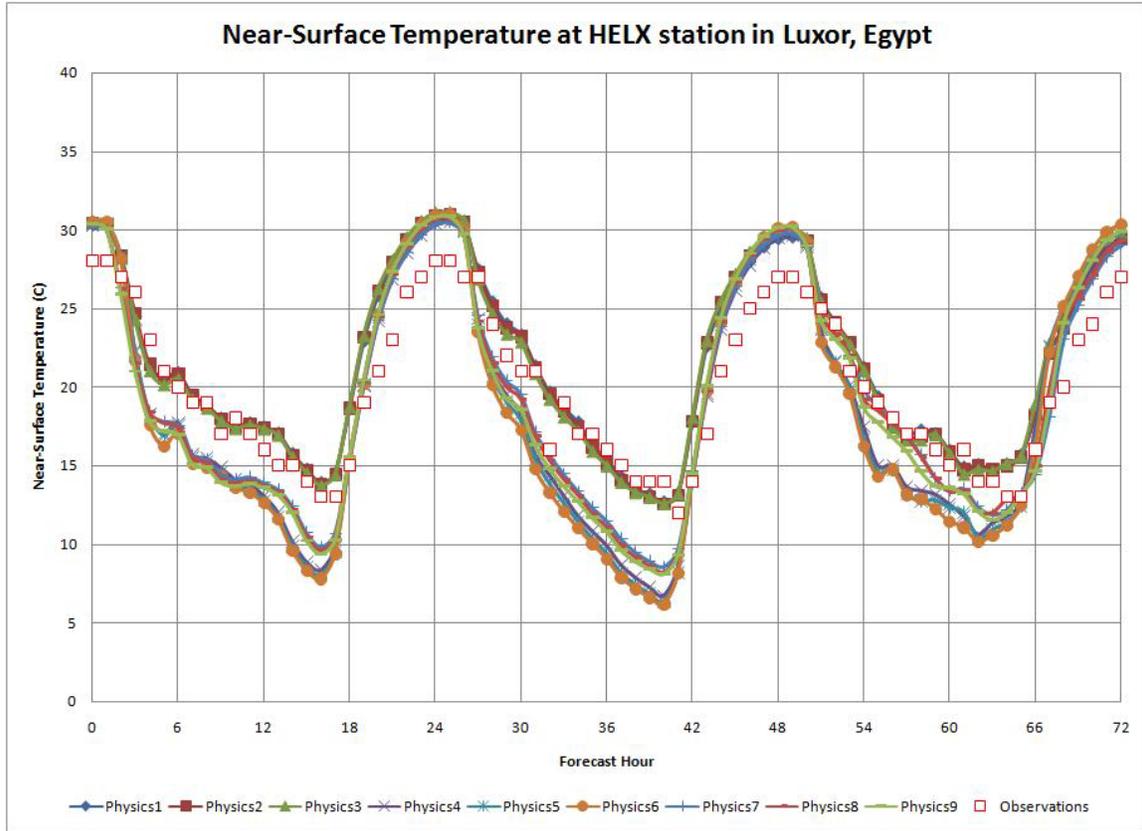


Figure F-69 Near-surface temperature at Luxor station in winter

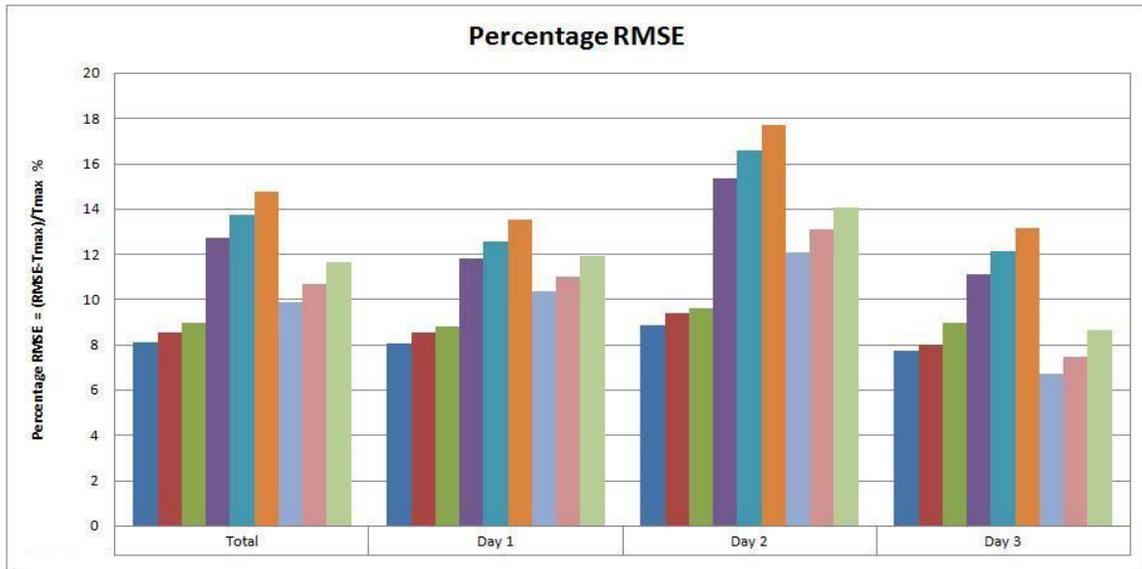


Figure F-70 RMSE of near-surface temperature at Luxor station in winter

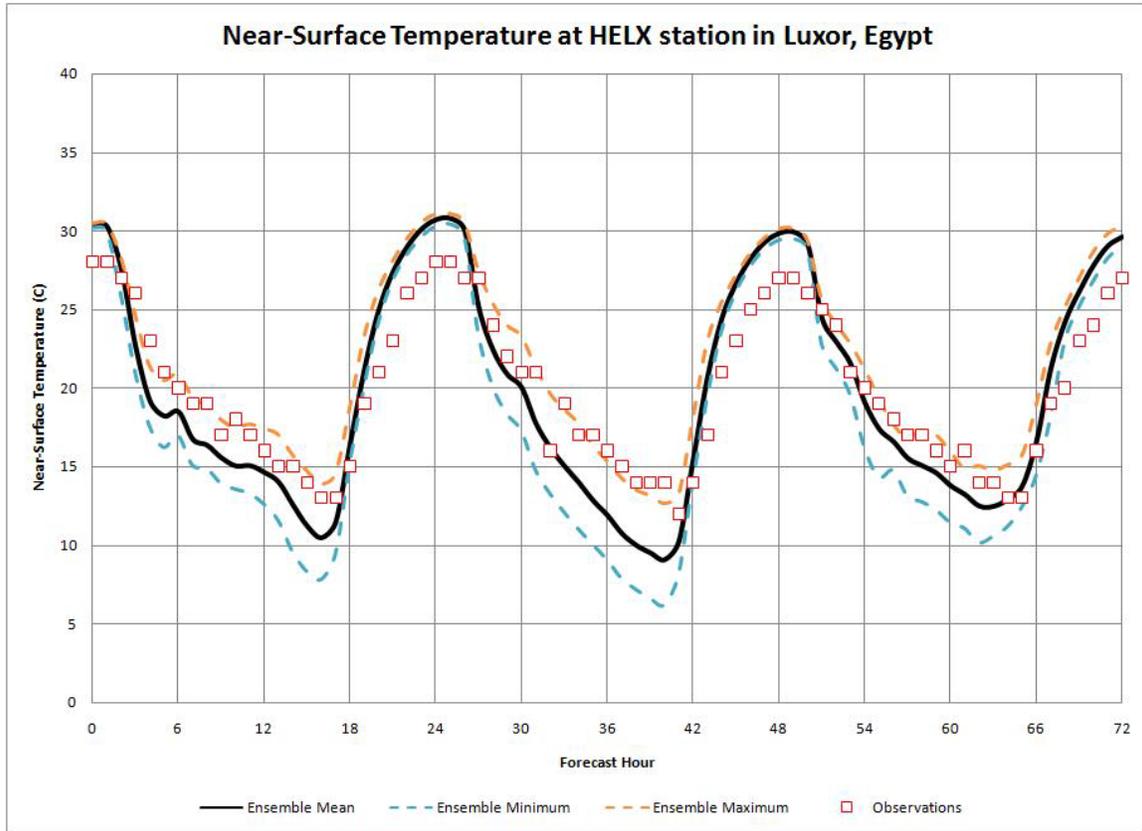


Figure F-71 Min, Max and mean temperature at Luxor station in winter

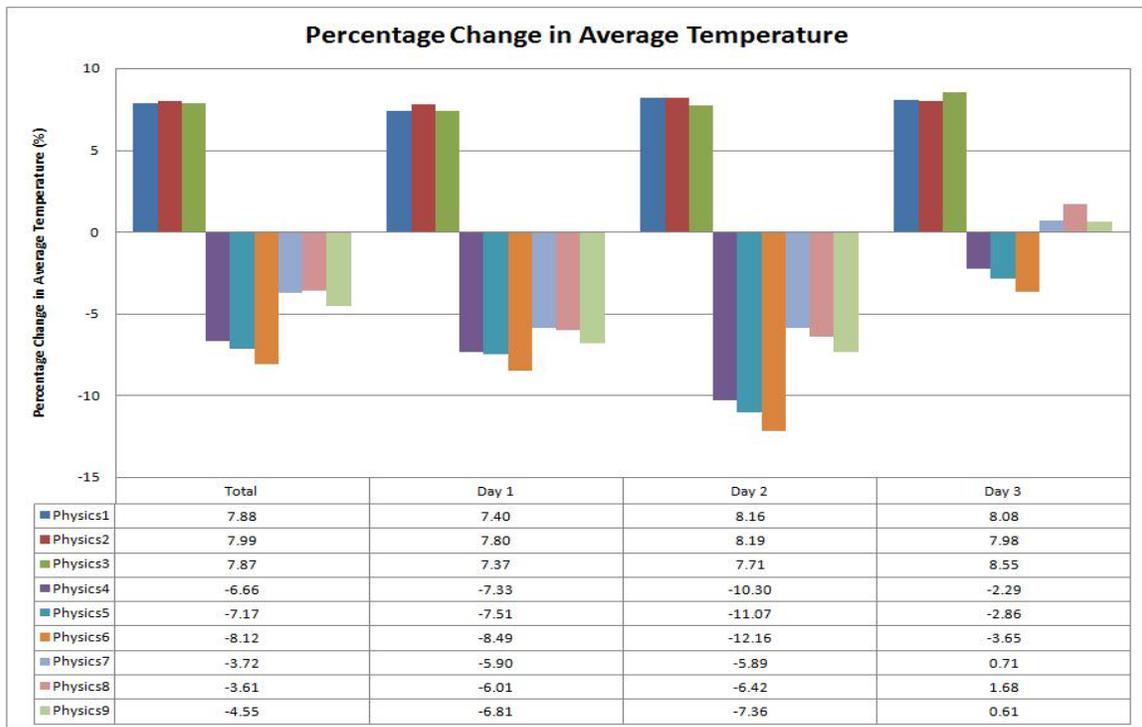


Figure F-72 Change in average temperature at Luxor station in winter

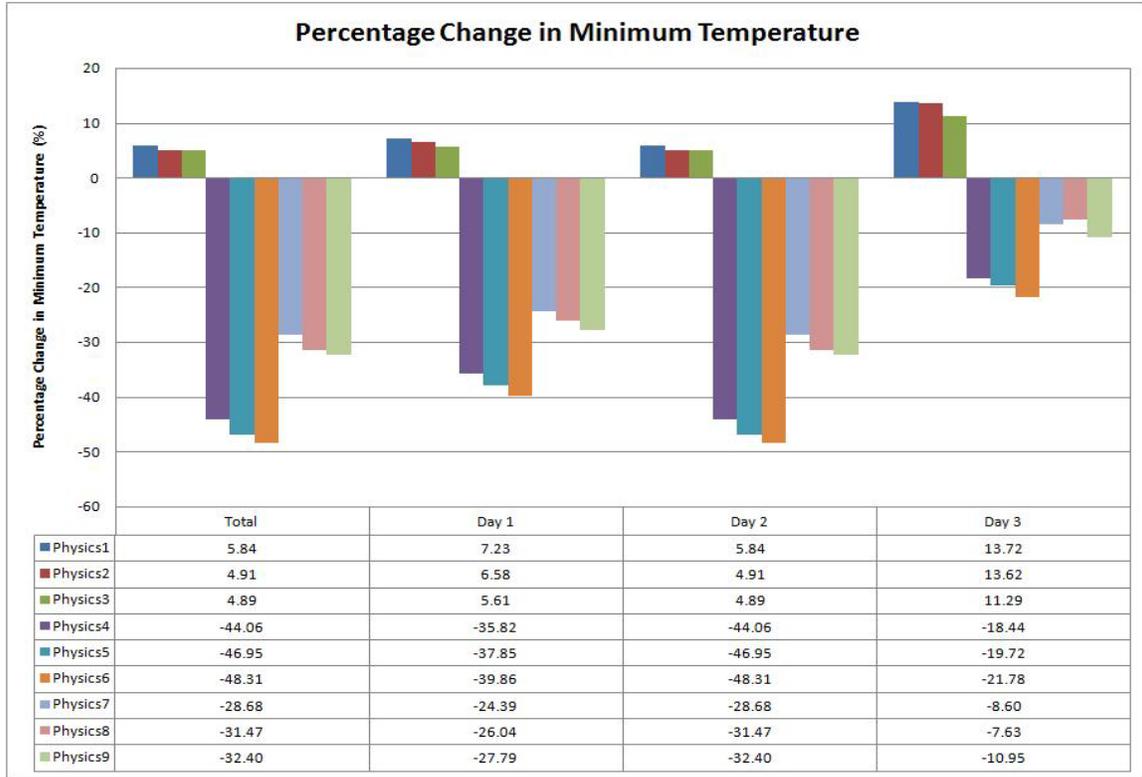


Figure F-73 Change in minimum temperature at Luxor station in winter

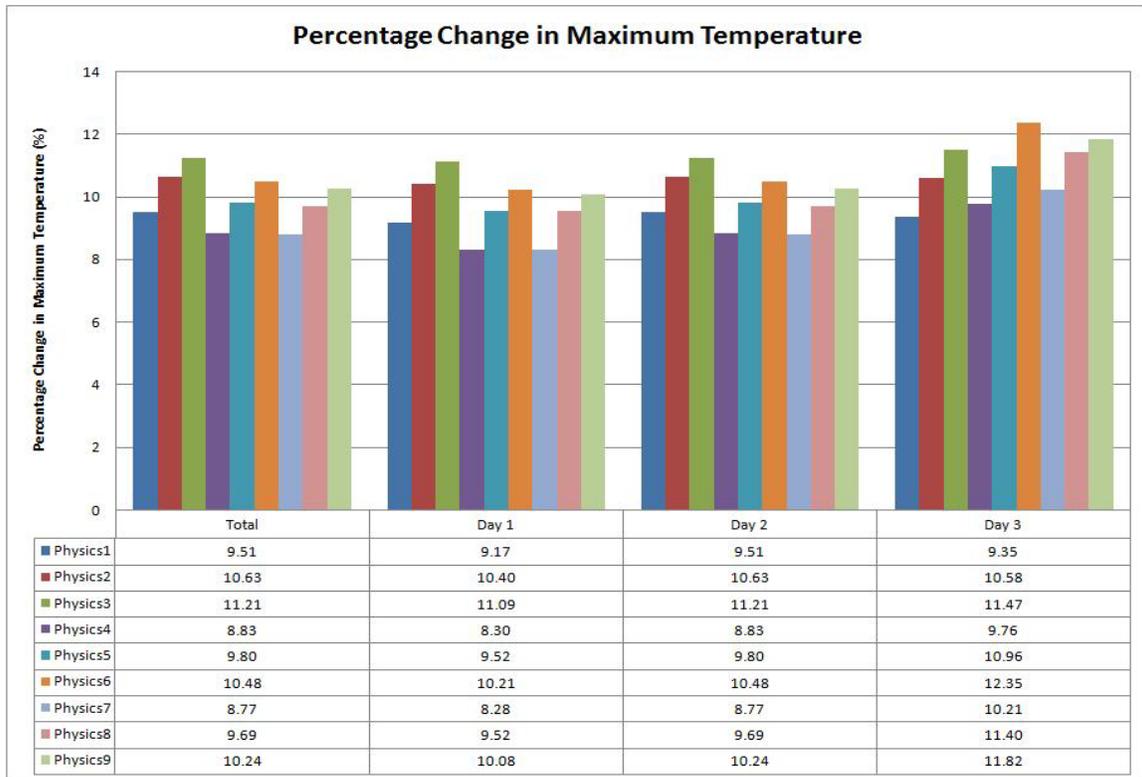


Figure F-74 Change in maximum temperature at Luxor station in winter

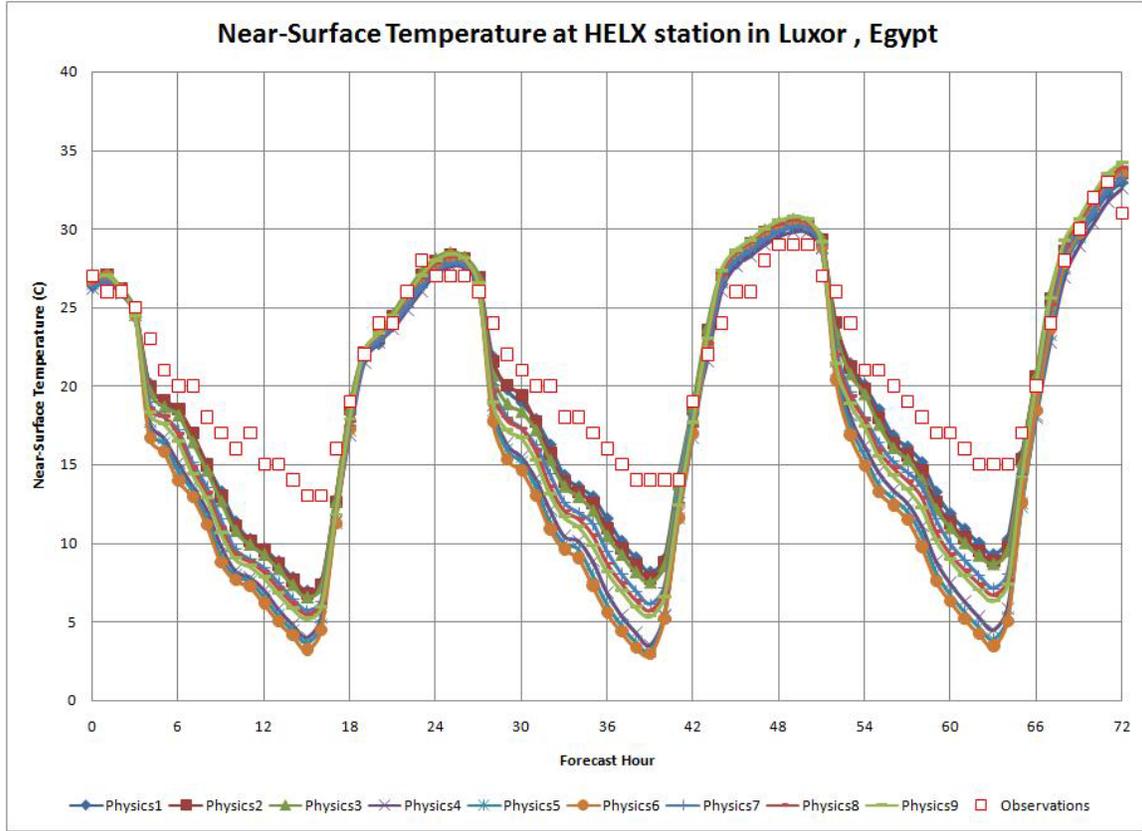


Figure F-75 Near-surface temperature at Luxor station in spring

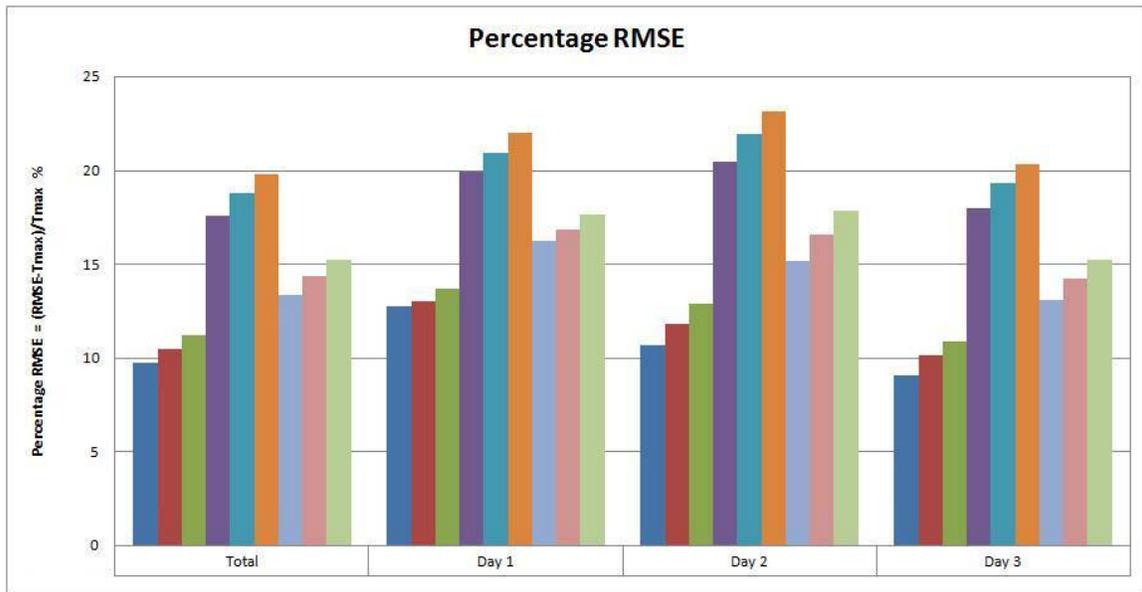


Figure F-76 RMSE of near-surface temperature at Luxor station in spring

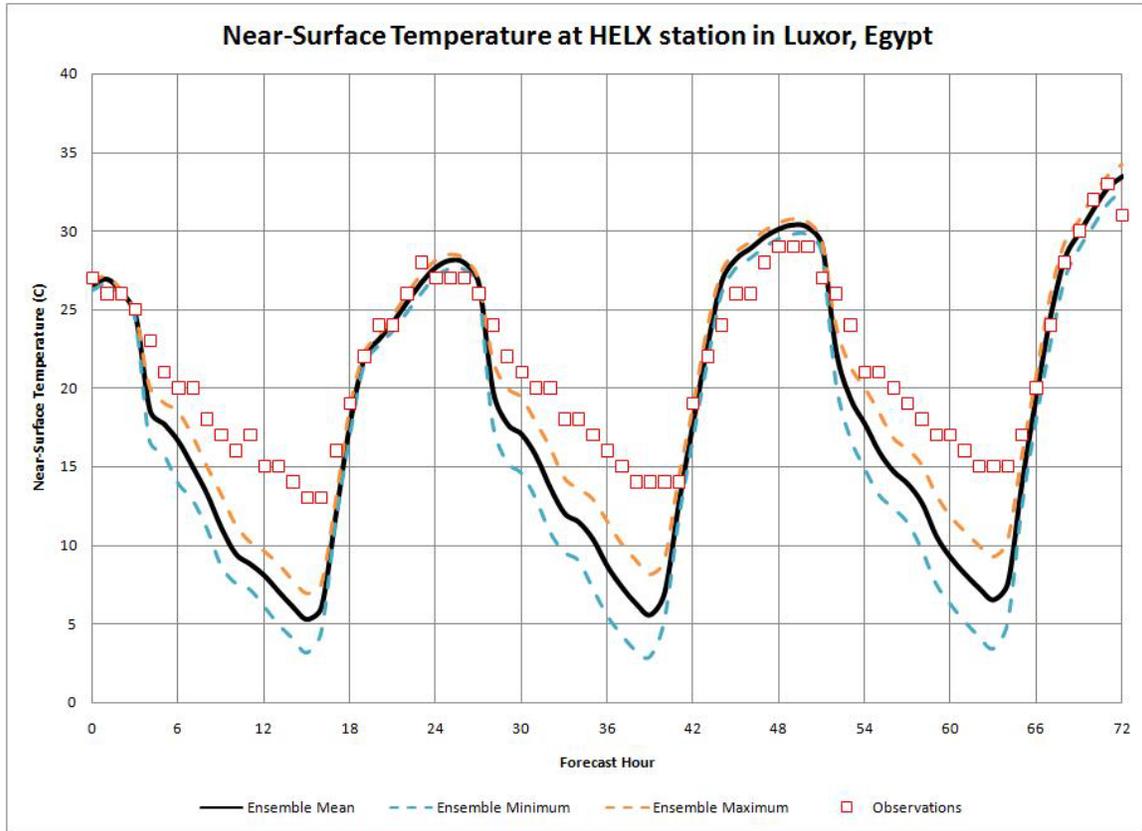


Figure F-77 Min, Max and mean temperature at Luxor station in spring

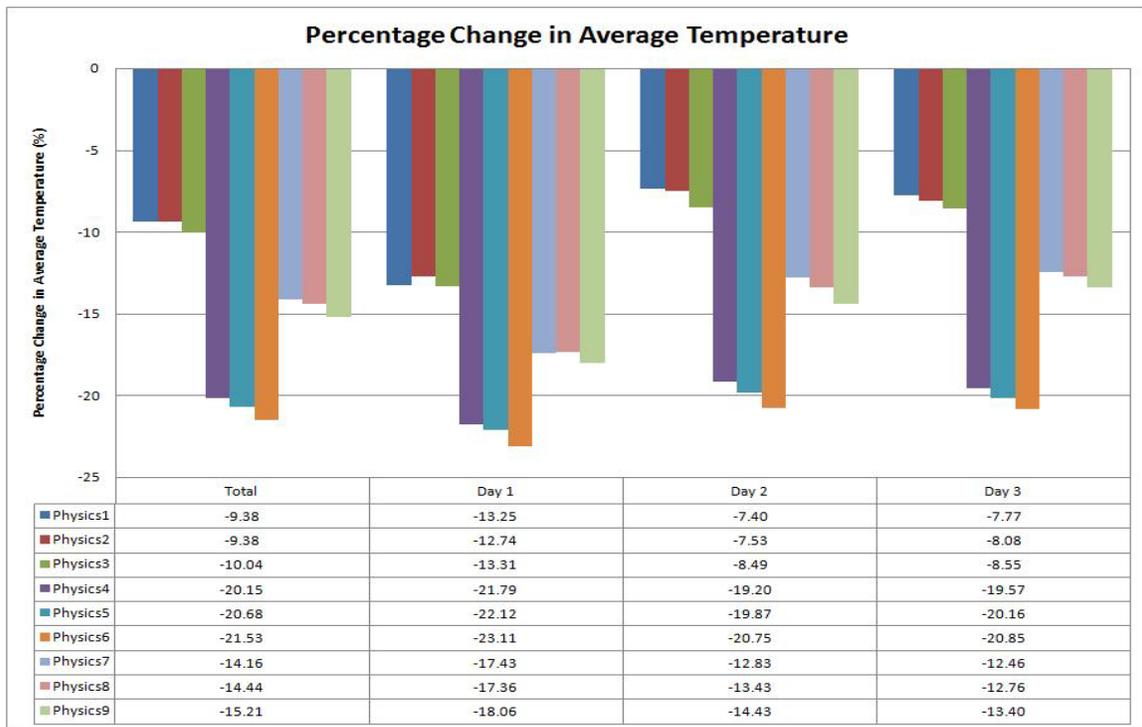


Figure F-78 Change in average temperature at Luxor station in spring

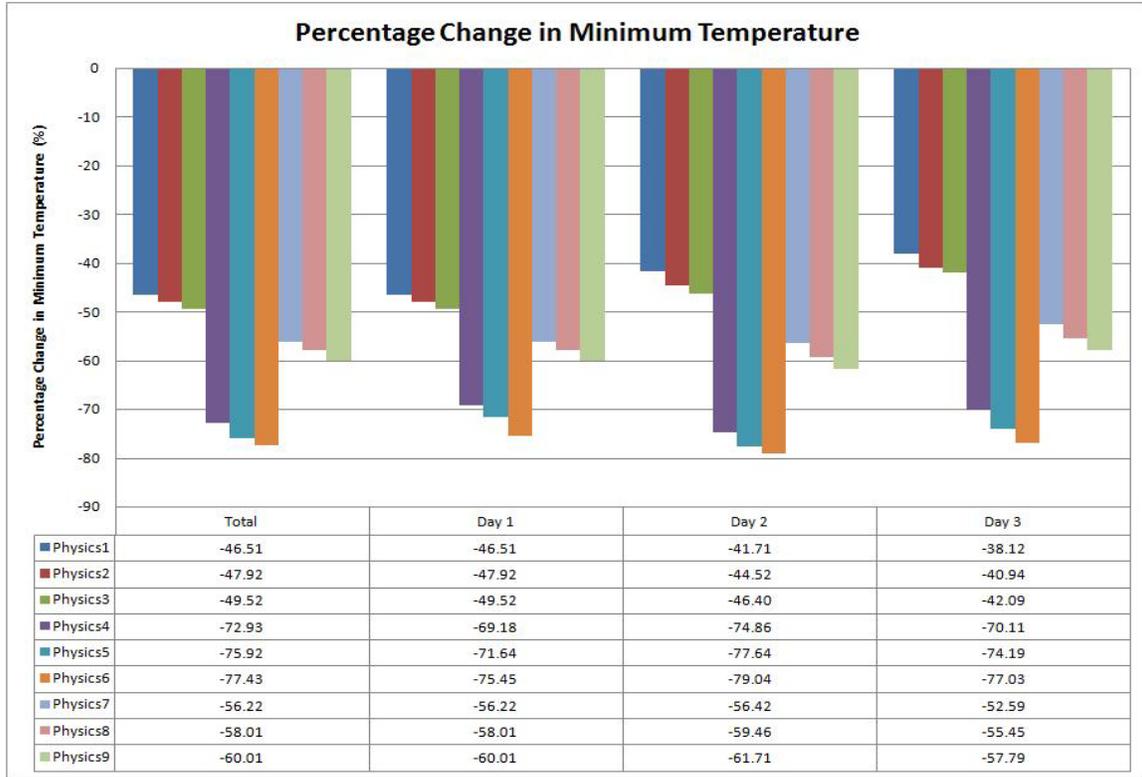


Figure F-79 Change in minimum temperature at Luxor station in spring

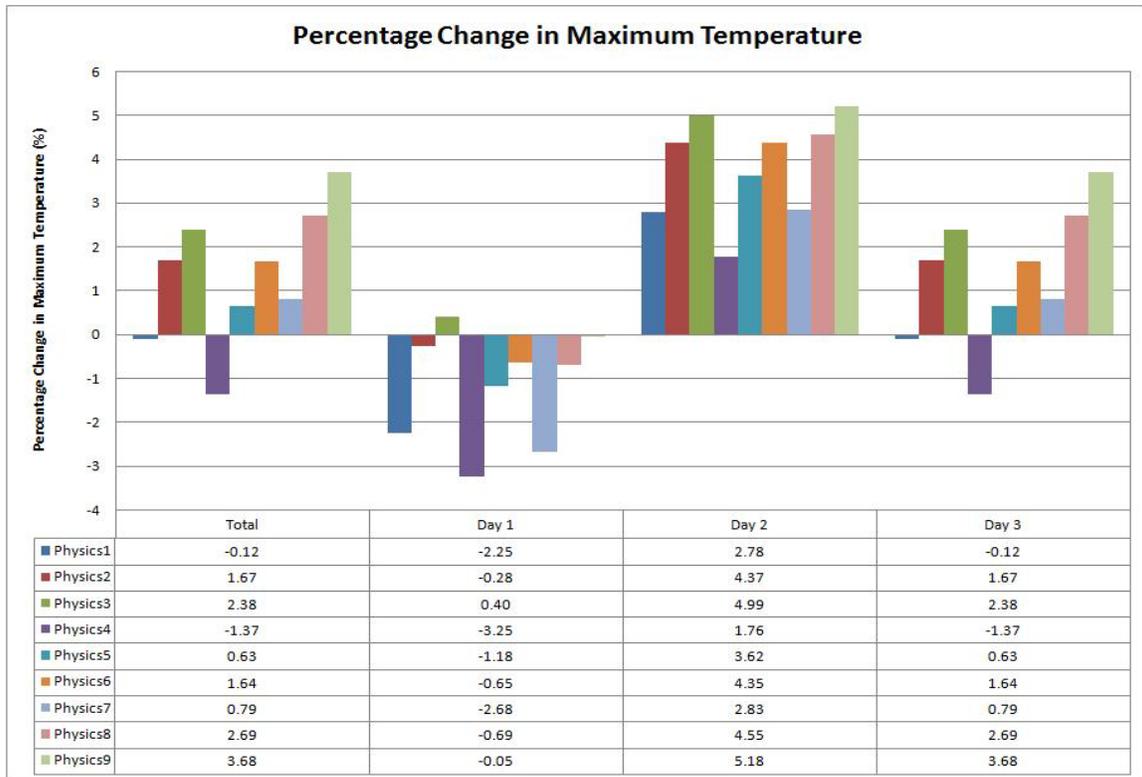


Figure F-80 Change in maximum temperature at Luxor station in spring

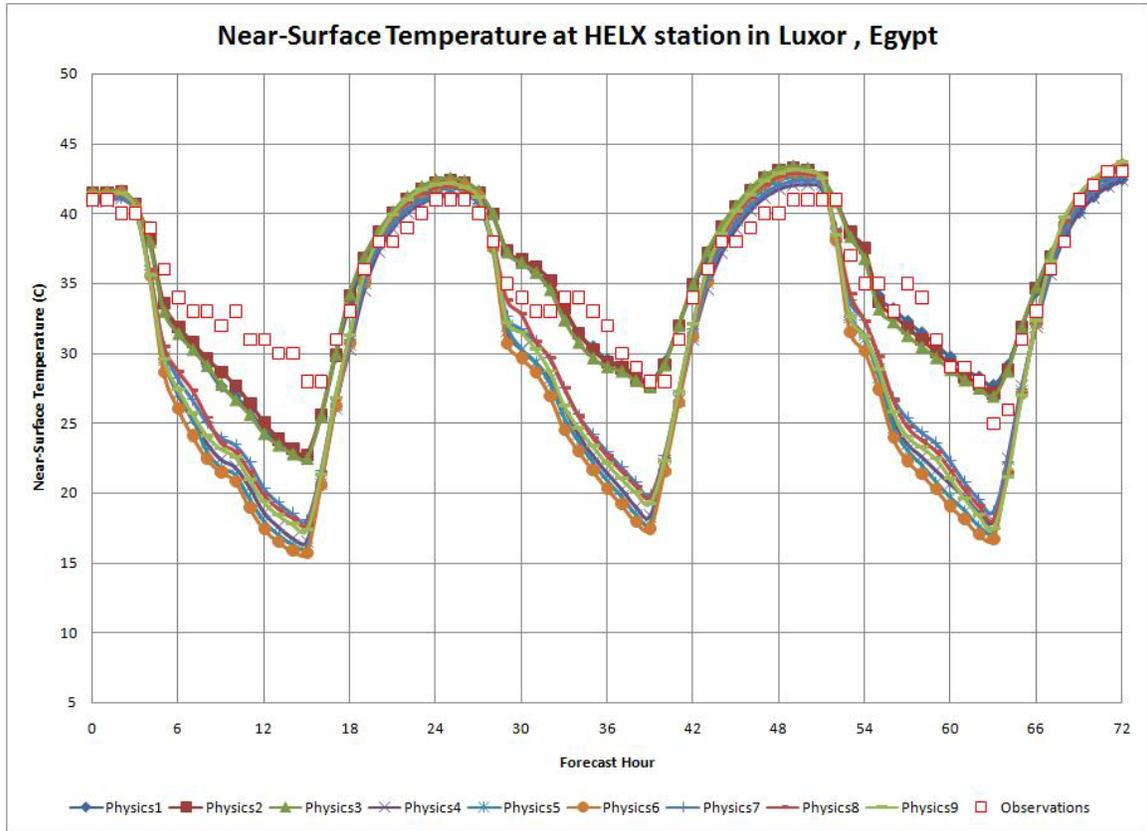


Figure F-81 Near-surface temperature at Luxor station in summer

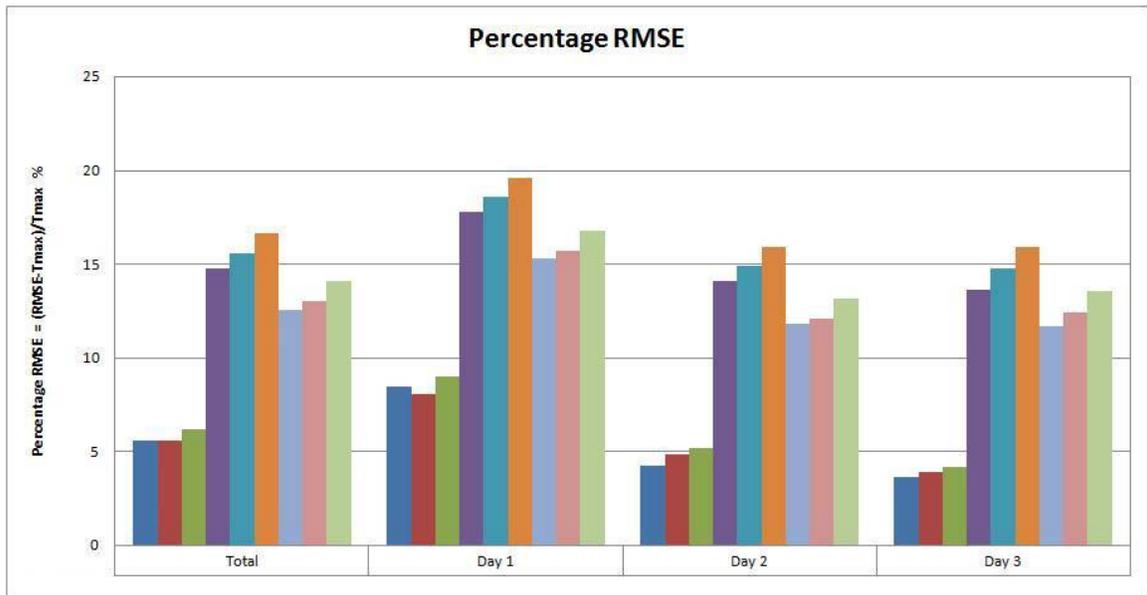


Figure F-82 RMSE of near-surface temperature at Luxor station in summer

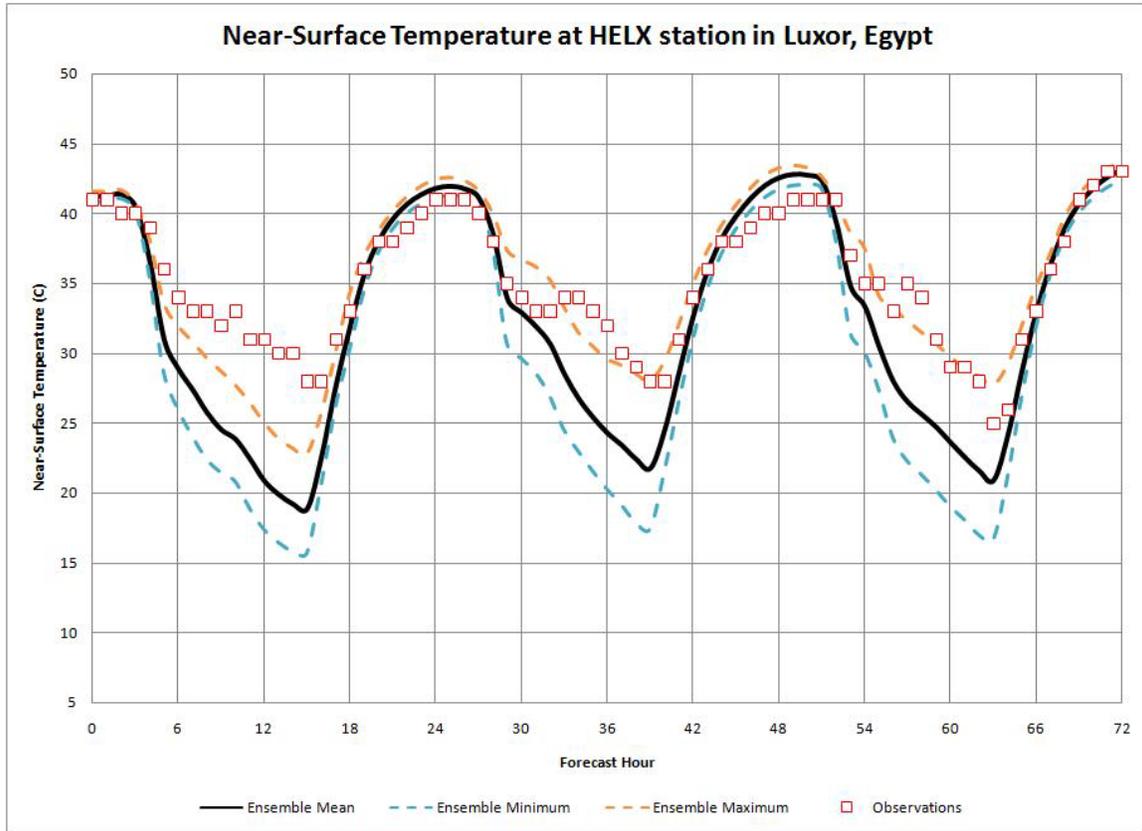


Figure F-83 Min, Max and mean temperature at Luxor station in summer

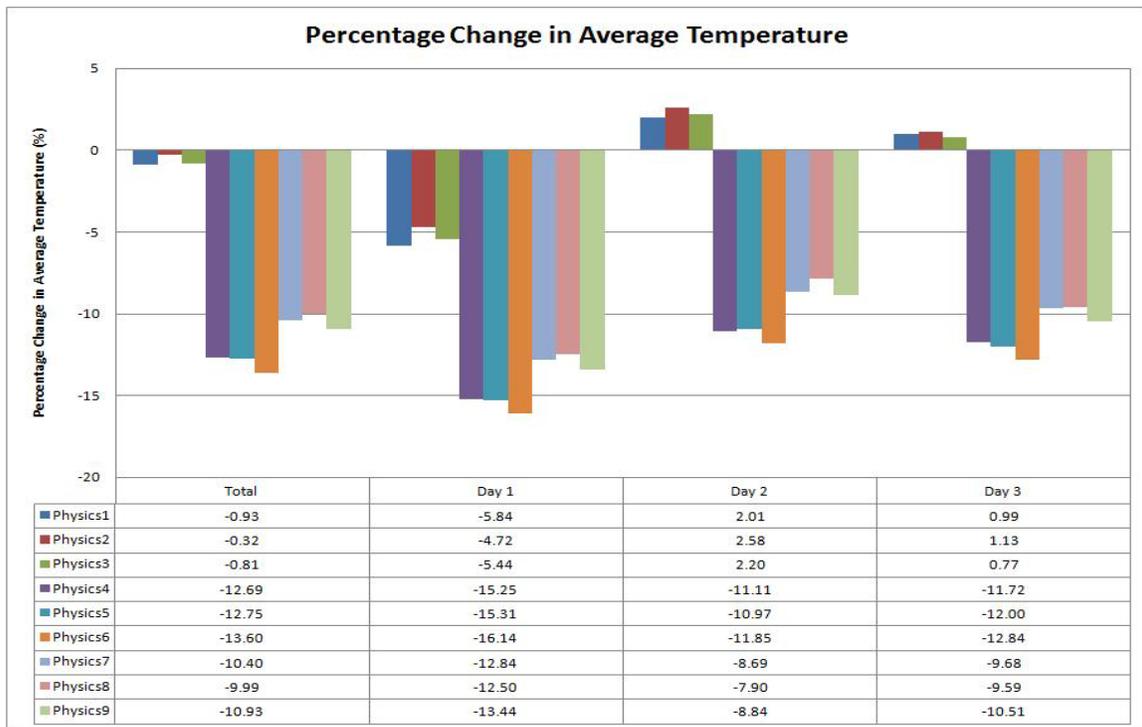


Figure F-84 Change in average temperature at Luxor station in summer

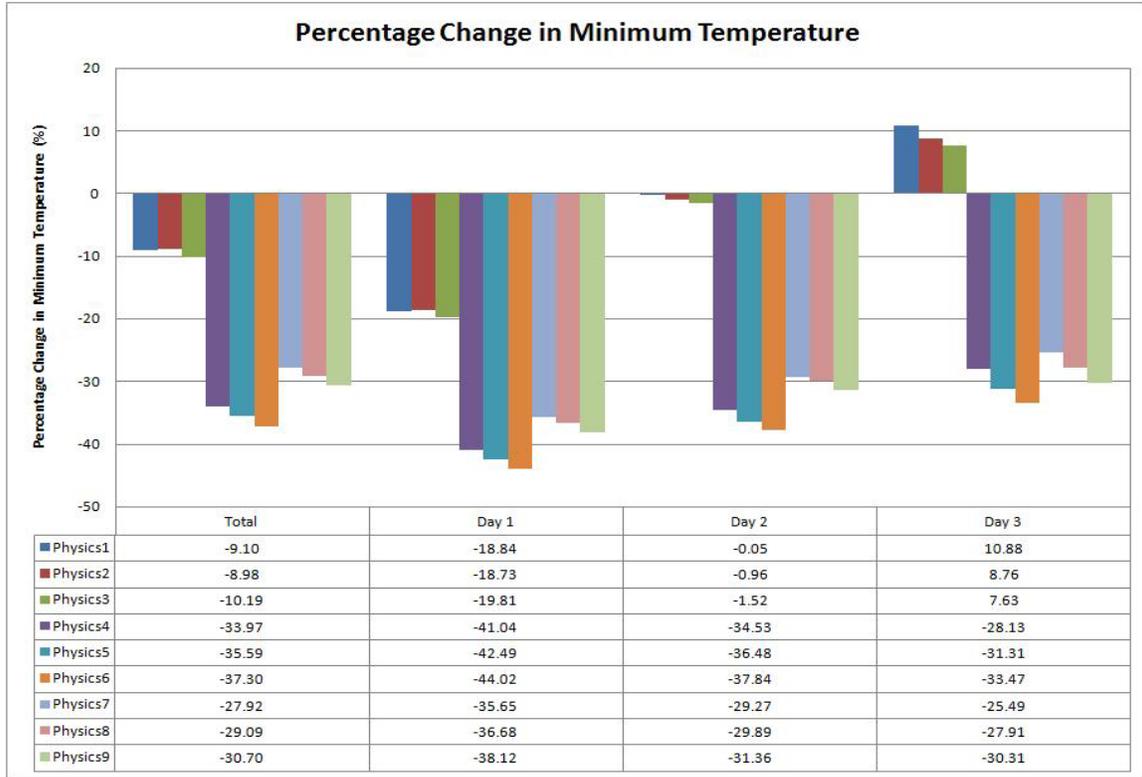


Figure F-85 Change in minimum temperature at Luxor station in summer

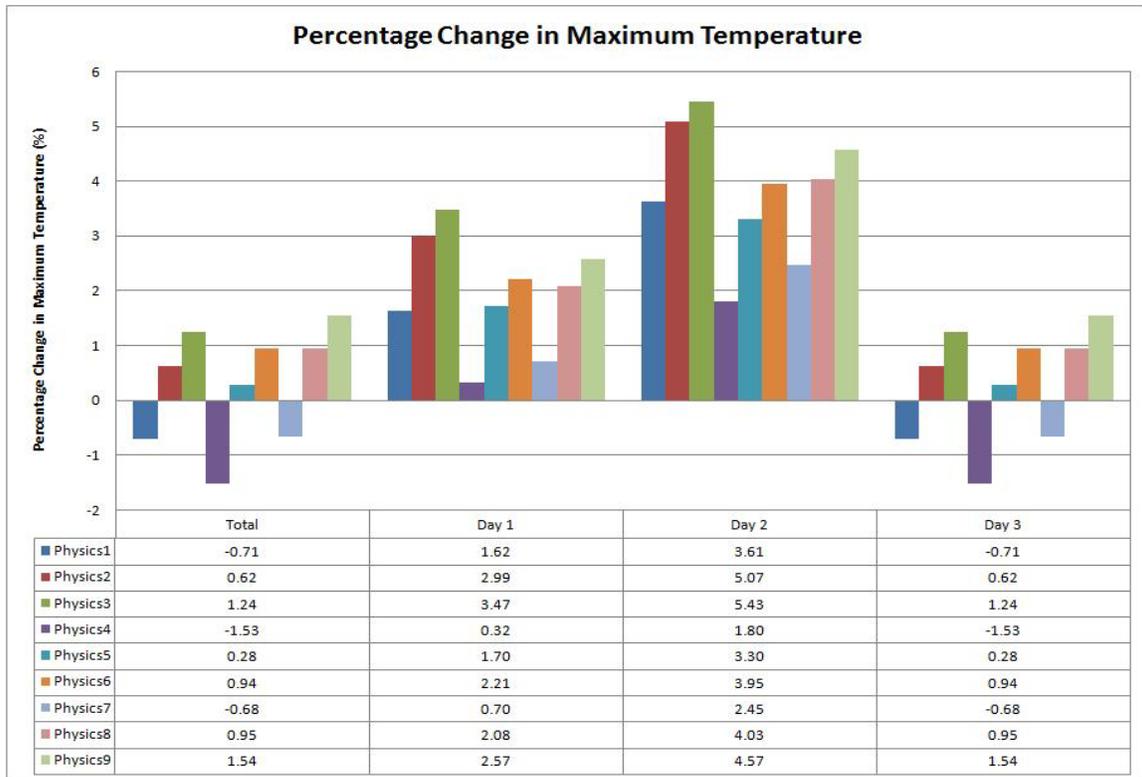


Figure F-86 Change in maximum temperature at Luxor station in summer

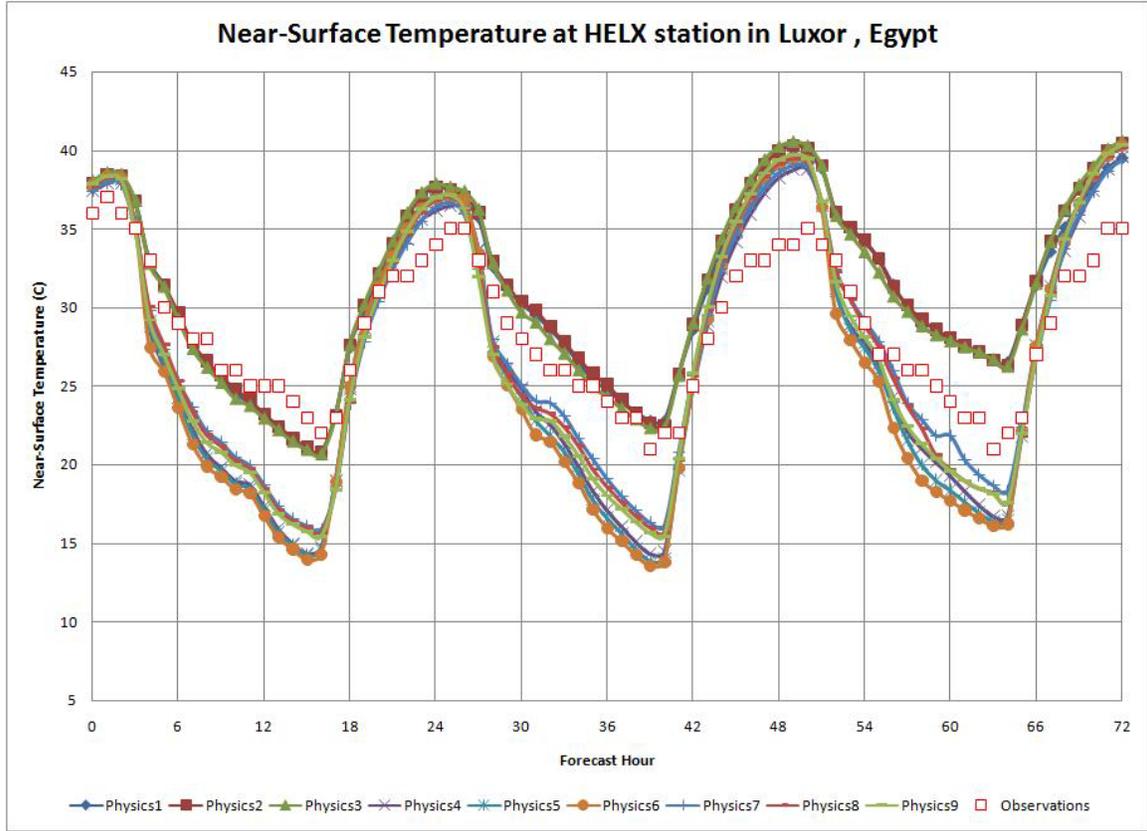


Figure F-87 Near-surface temperature at Luxor station in autumn

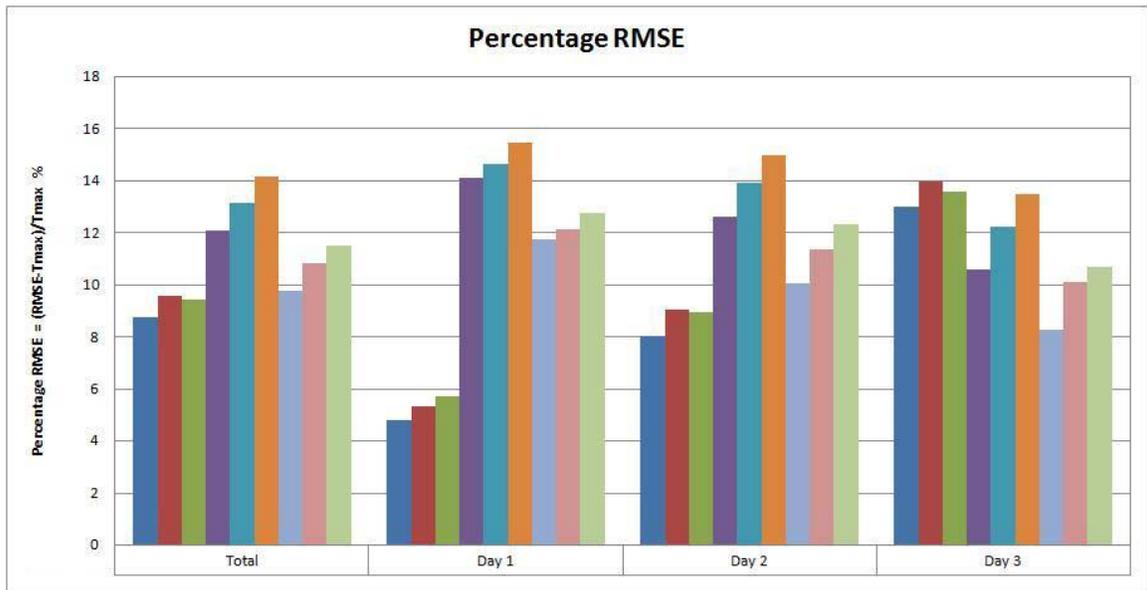


Figure F-88 RMSE of near-surface temperature at Luxor station in autumn

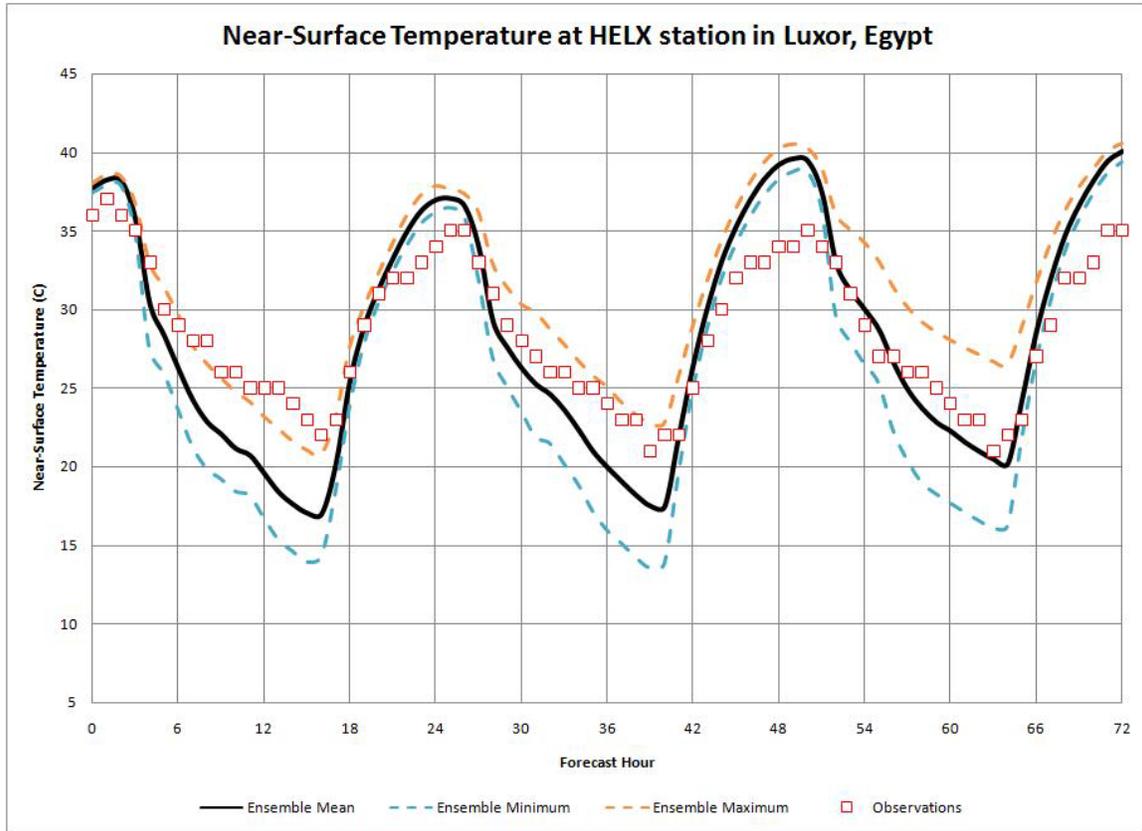


Figure F-89 Min, Max and mean temperature at Luxor station in autumn

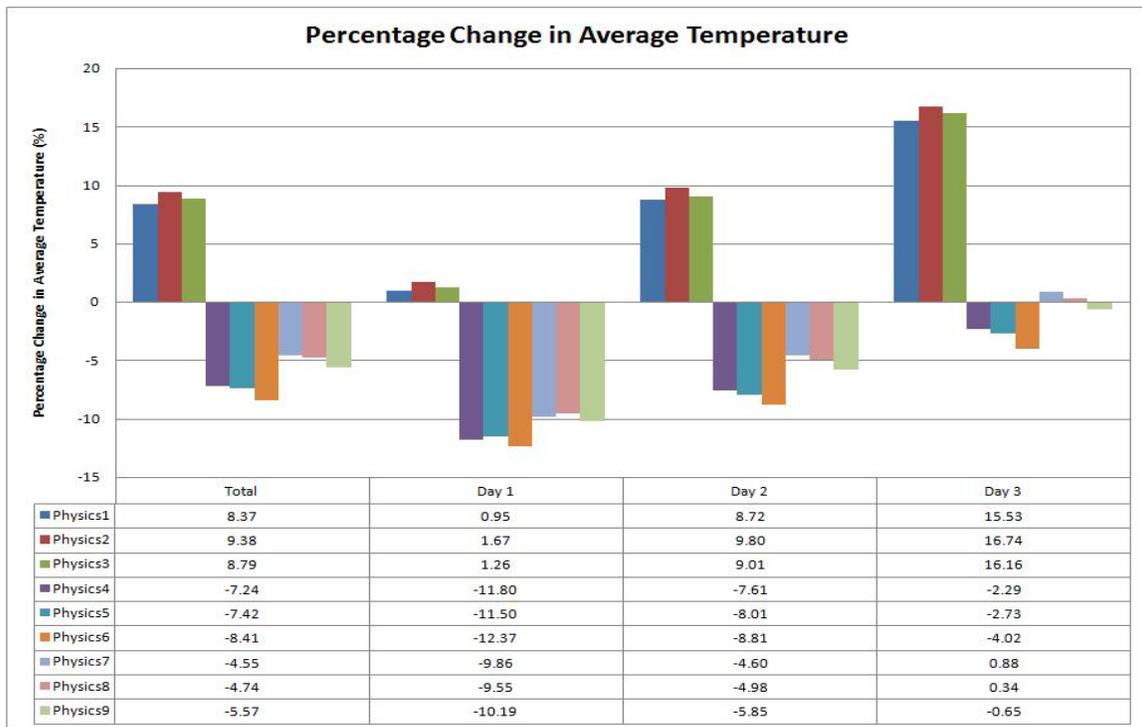


Figure F-90 Change in average temperature at Luxor station in autumn

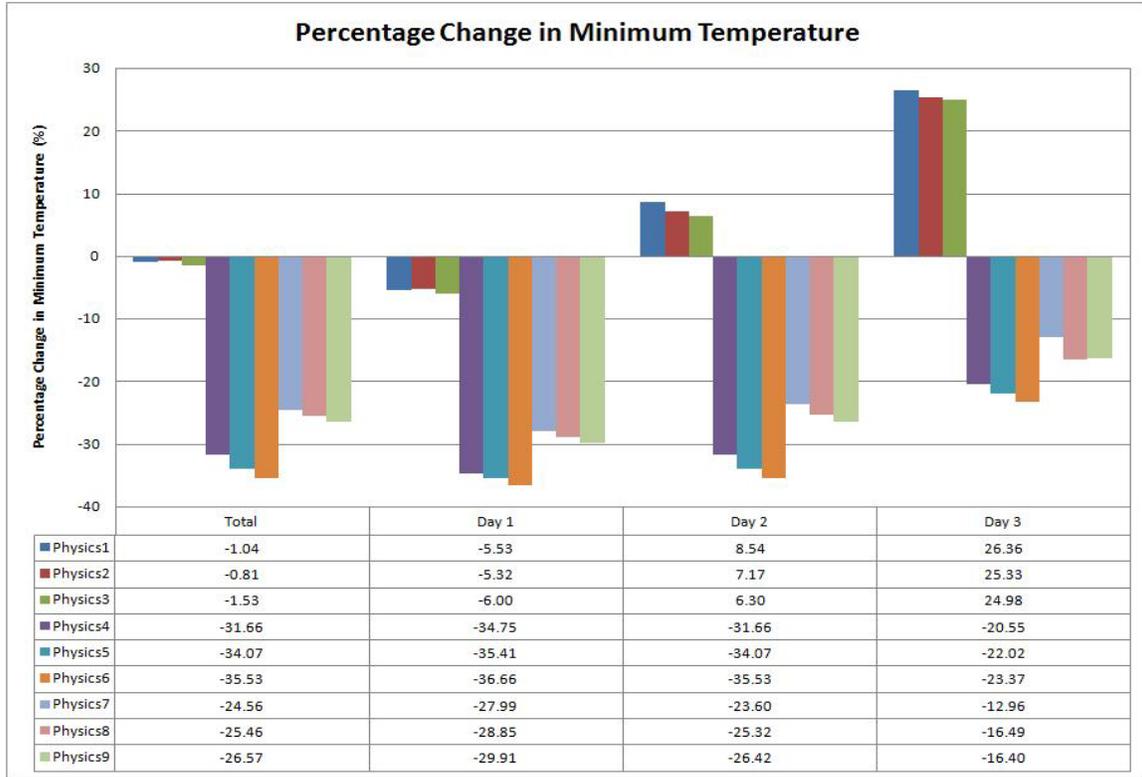


Figure F-91 Change in minimum temperature at Luxor station in autumn

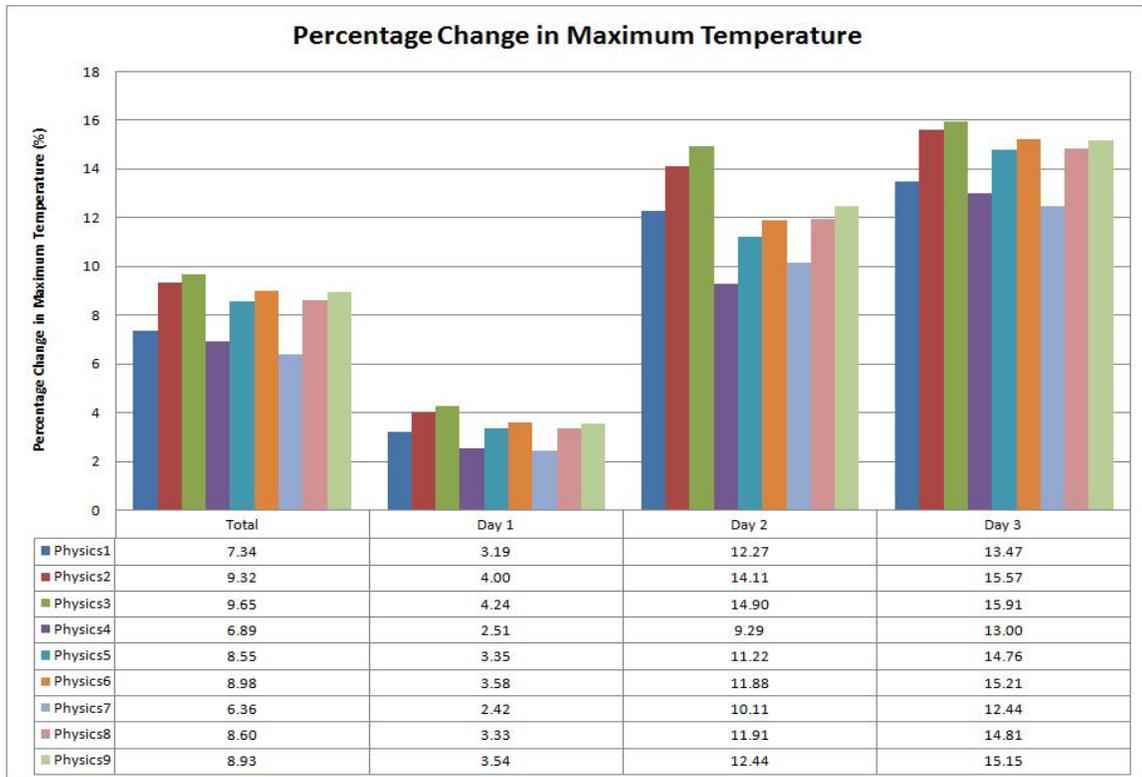


Figure F-92 Change in maximum temperature at Luxor station in autumn

F.4. ENSEMBLE FORECASTING EXPERIMENT

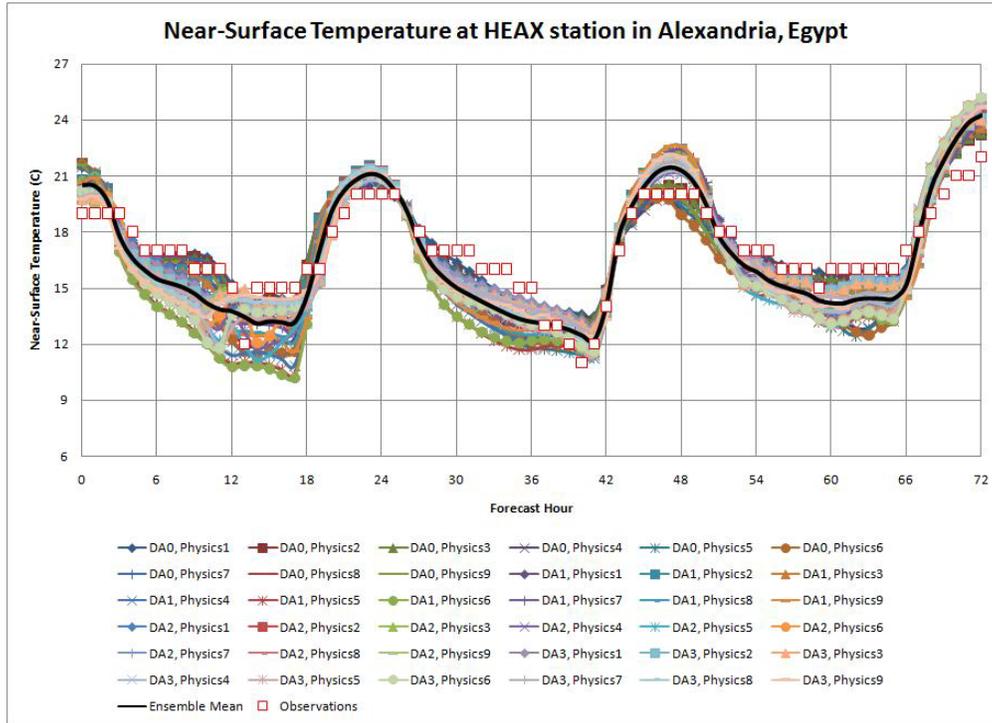


Figure F-93 Near-surface temperature at Alexandria/Nouzha station in winter

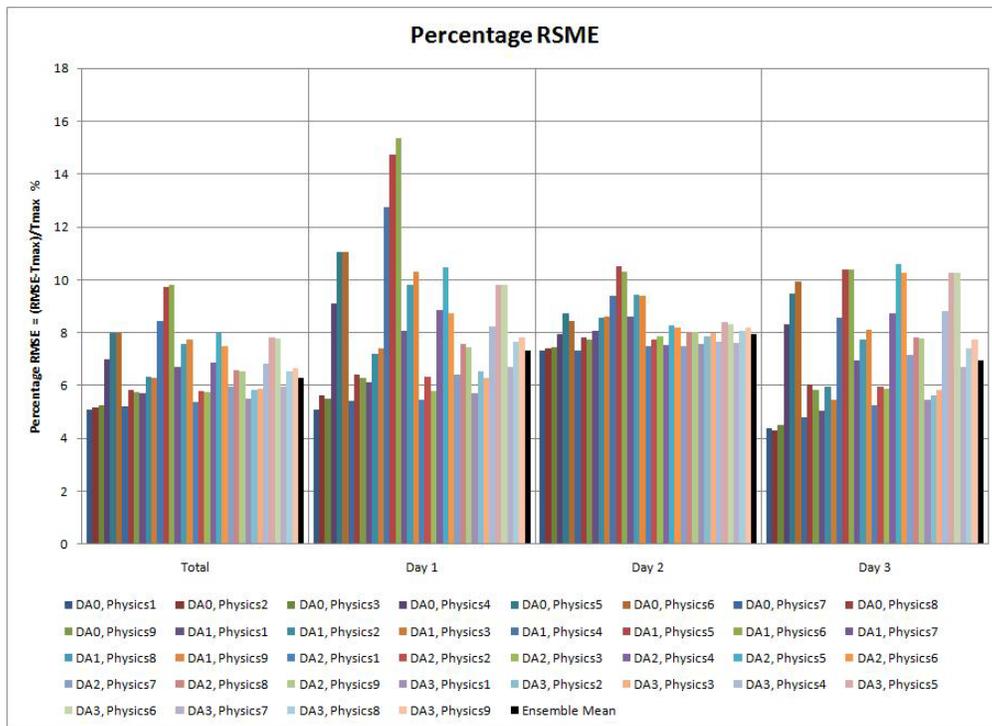


Figure F-94 RMSE of near-surface temperature at Alexandria/Nouzha station in winter

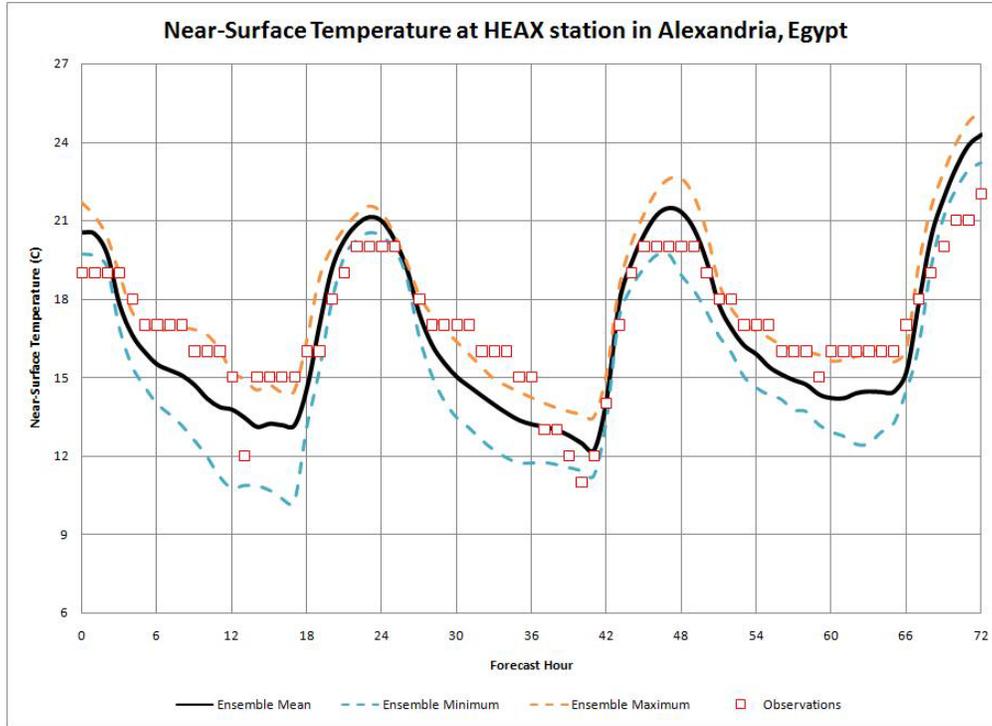


Figure F-95 Min, Max and mean temperature at Alexandria/Nouzha station in winter

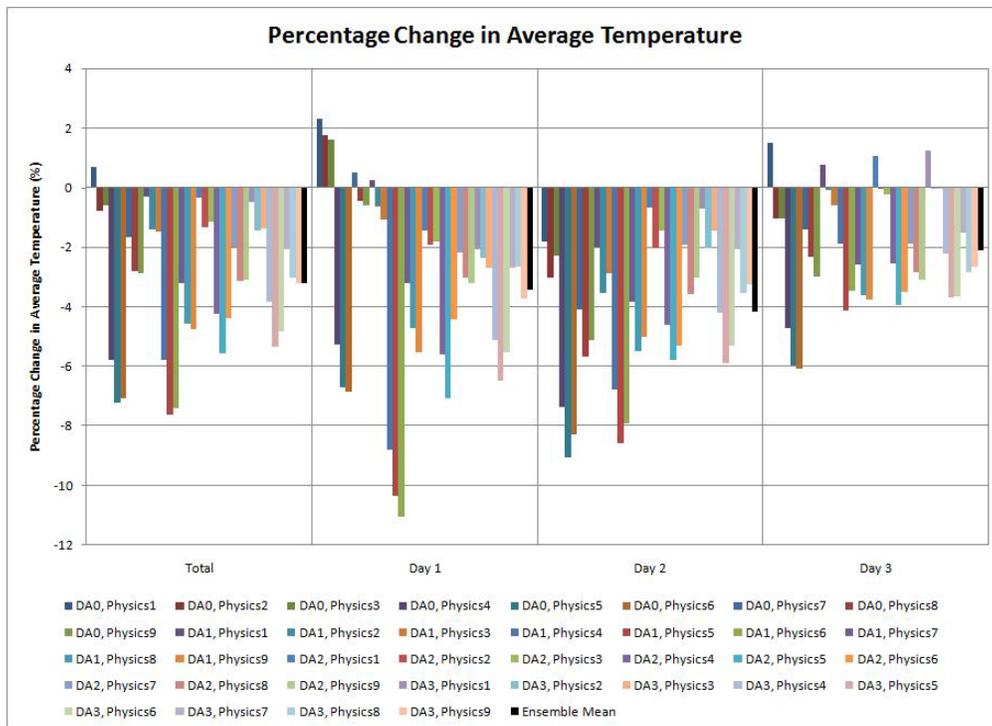


Figure F-96 Change in average temperature at Alexandria/Nouzha station in winter

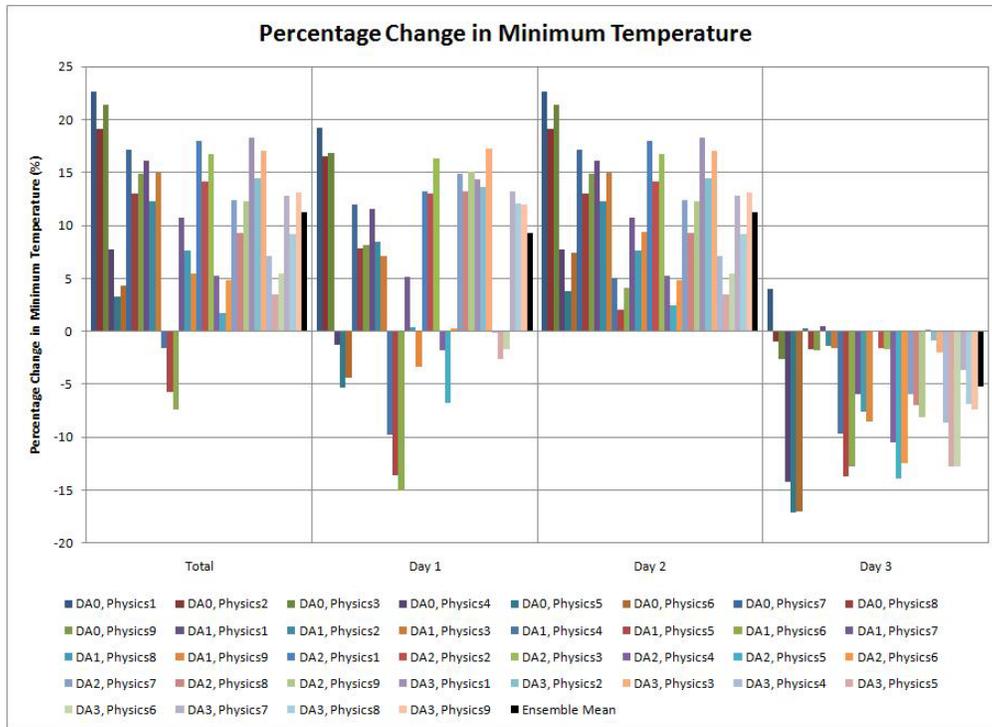


Figure F-97 Change in minimum temperature at Alexandria/Nouzha station in winter

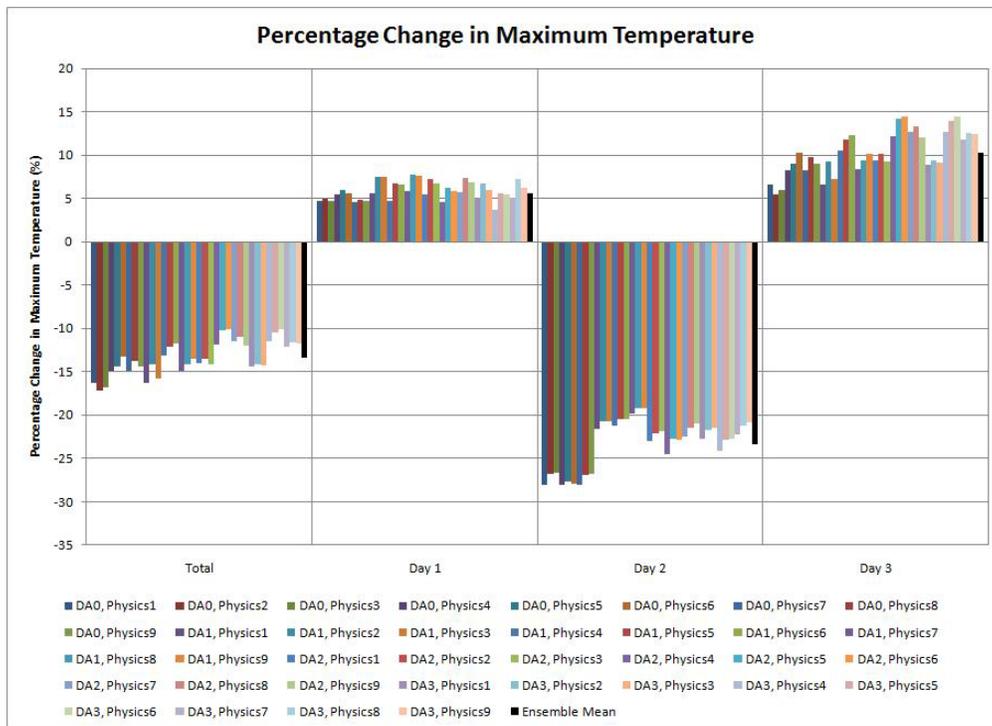


Figure F-98 Change in maximum temperature at Alexandria/Nouzha station in winter

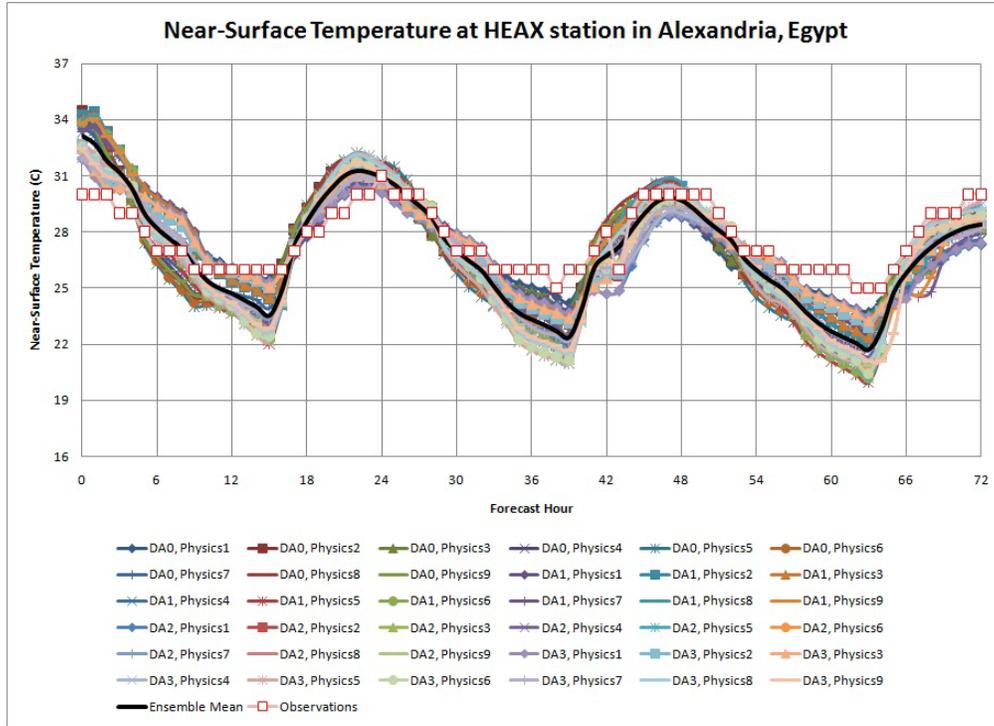


Figure F-99 Near-surface temperature at Alexandria/Nouzha station in summer

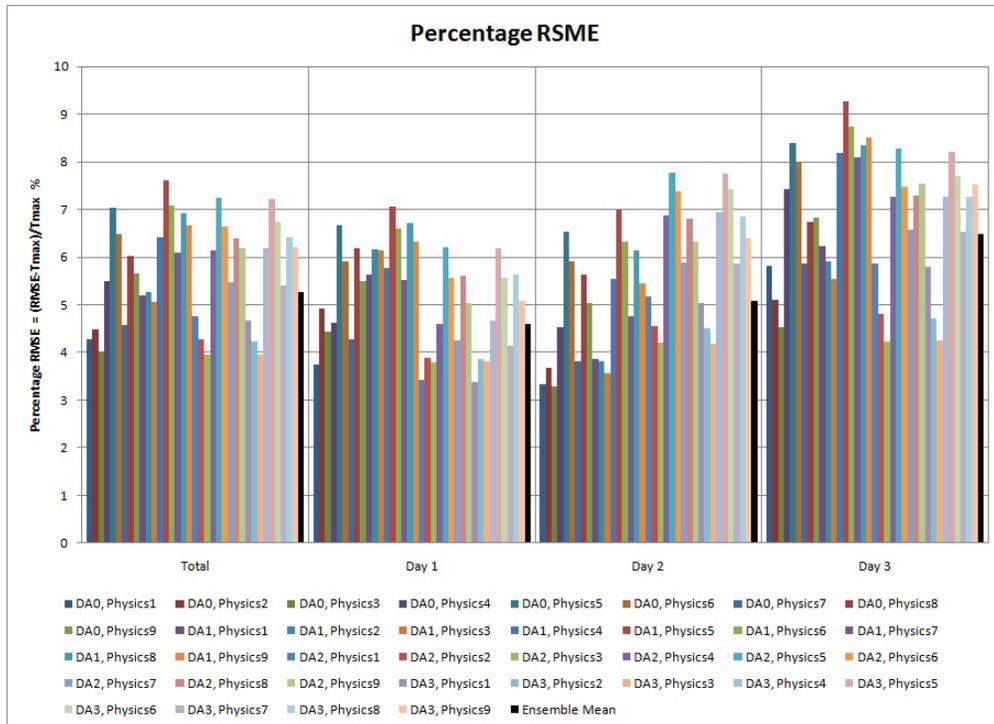


Figure F-100 RMSE of near-surface temperature at Alexandria/Nouzha station in summer

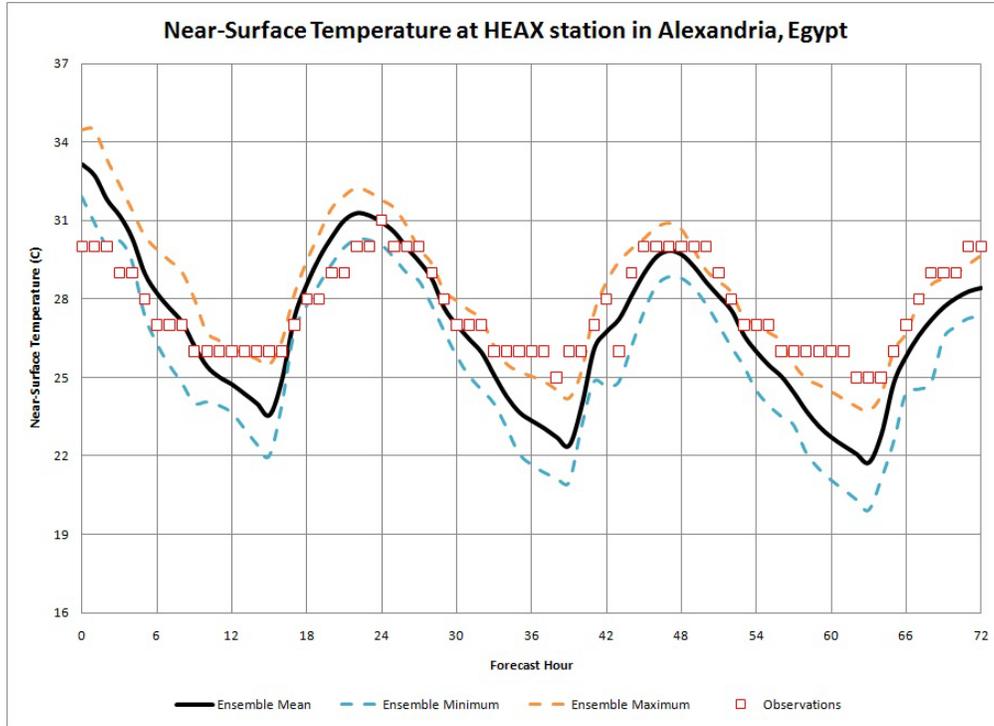


Figure F-101 Min, Max and Mean temperature at Alexandria/Nouzha station in summer

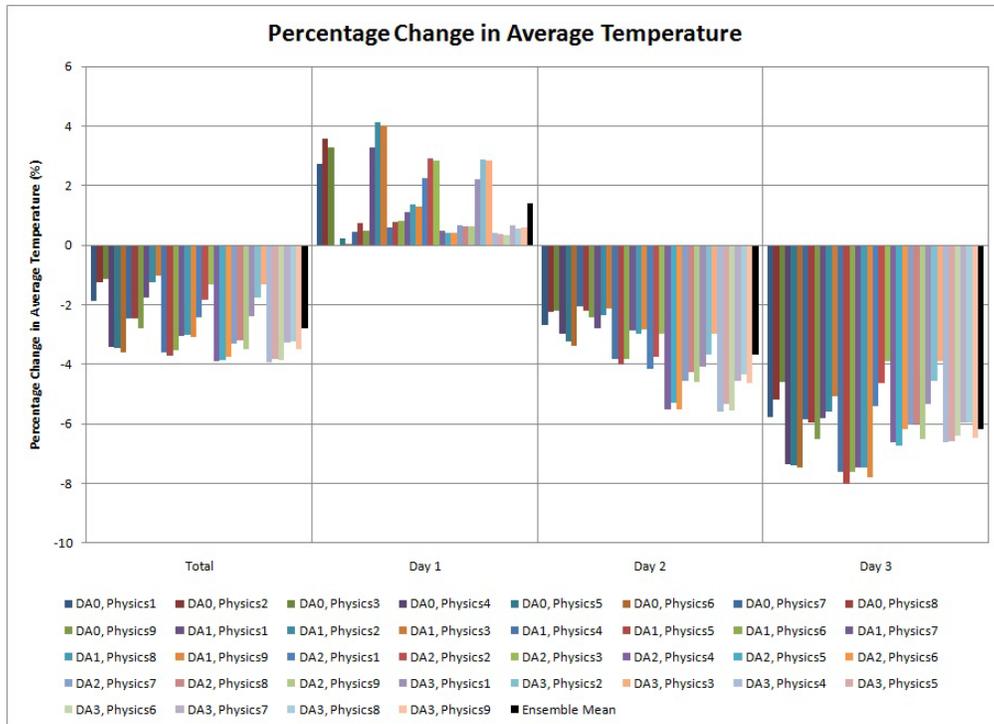


Figure F-102 Change in average temperature at Alexandria/Nouzha station in summer

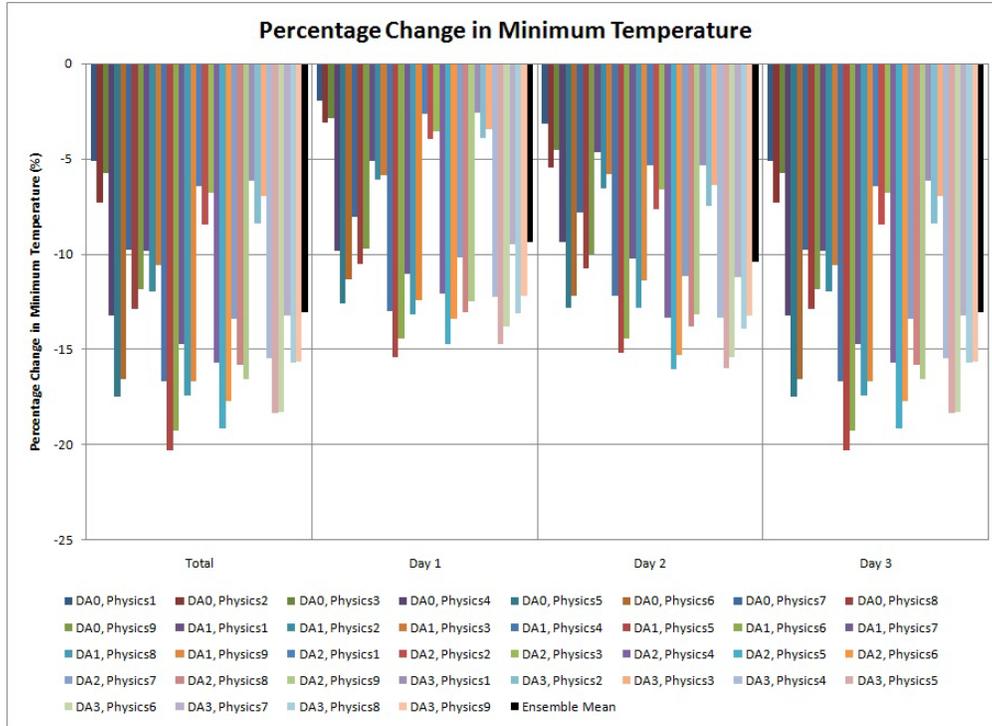


Figure F-103 Change in minimum temperature at Alexandria/Nouzha station in summer

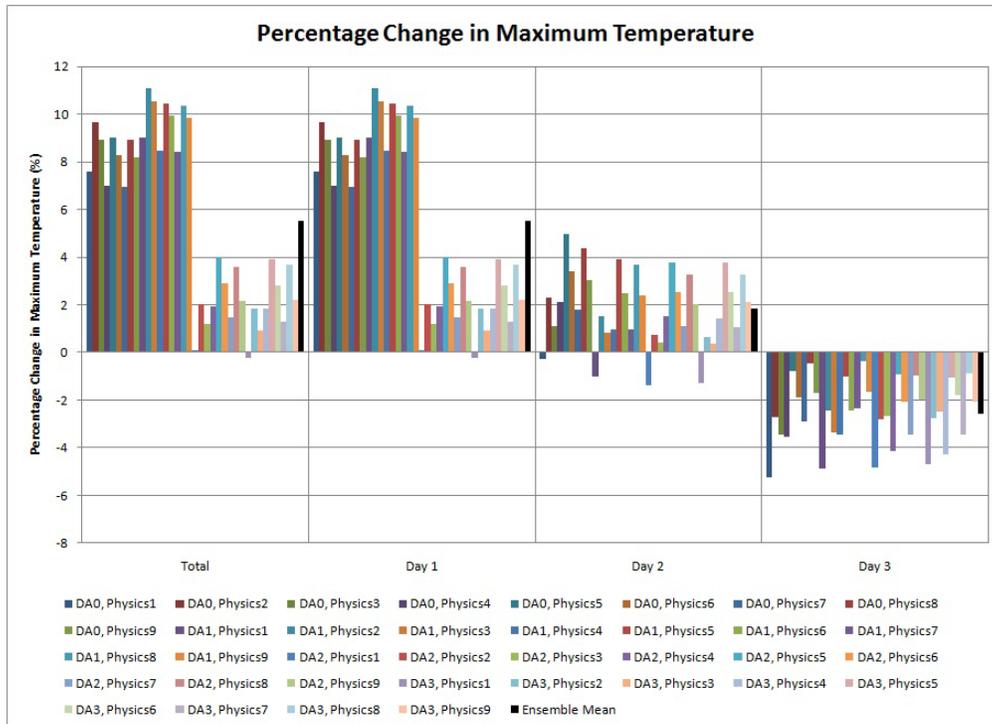


Figure F-104 Change in maximum temperature at Alexandria/Nouzha station in summer

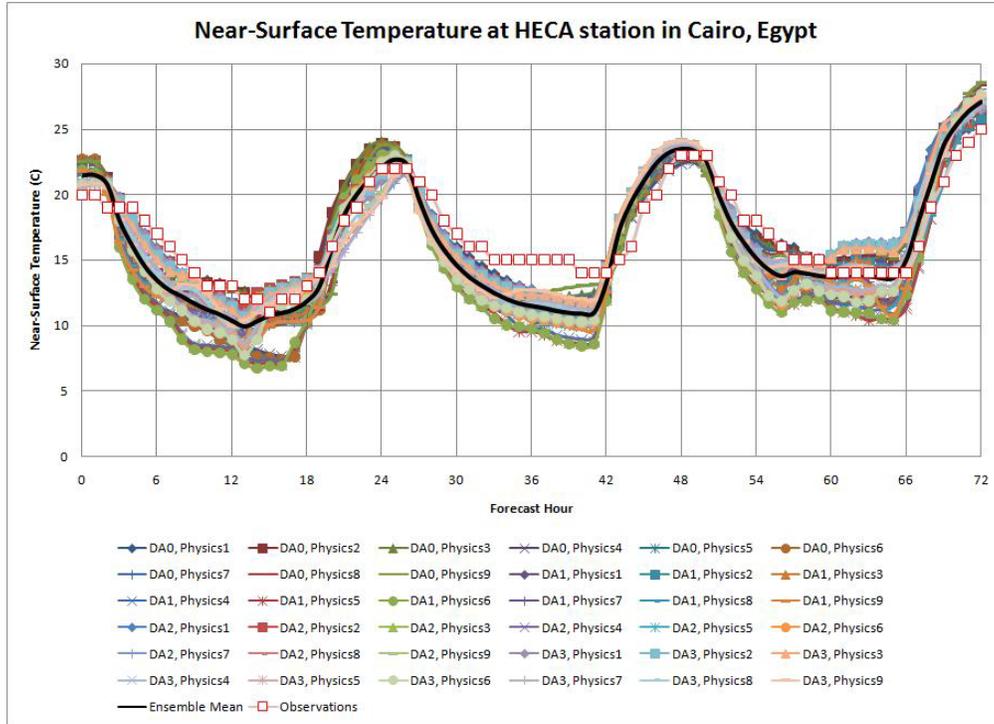


Figure F-105 Near-surface temperature at Cairo station in winter

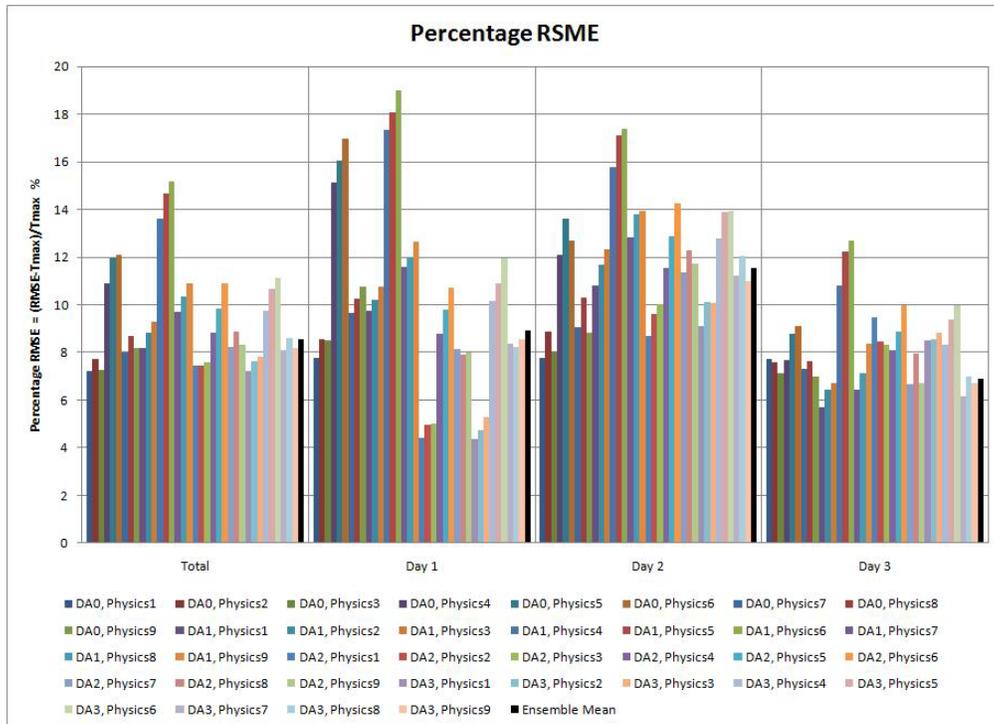


Figure F-106 RMSE of near-surface temperature at Cairo station in winter

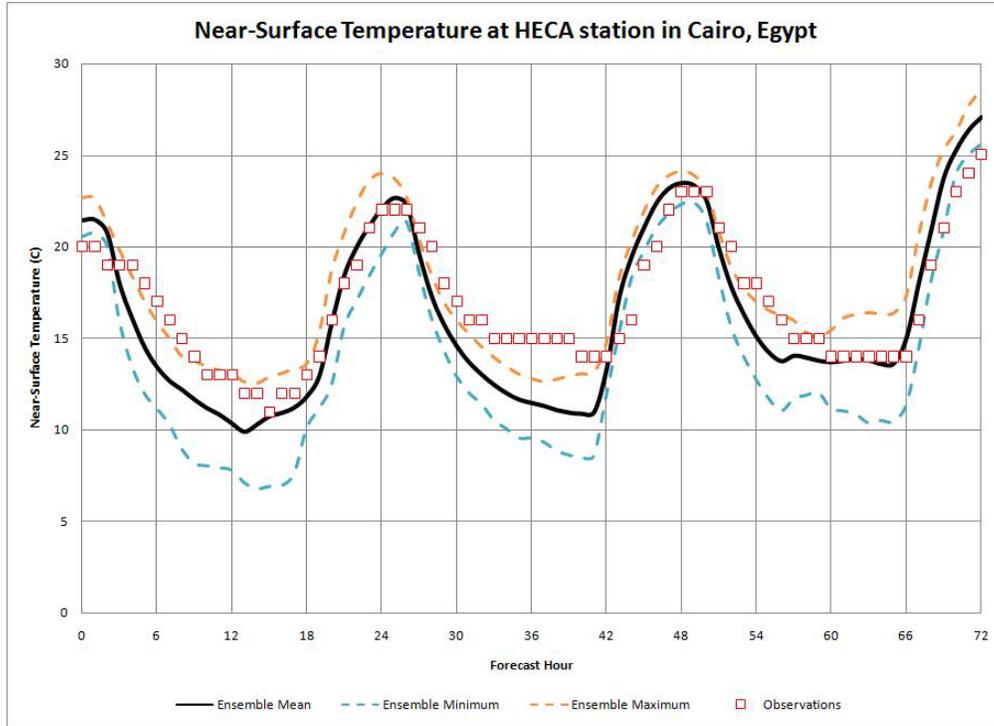


Figure F-107 Min, Max and Mean temperature at Cairo station in winter

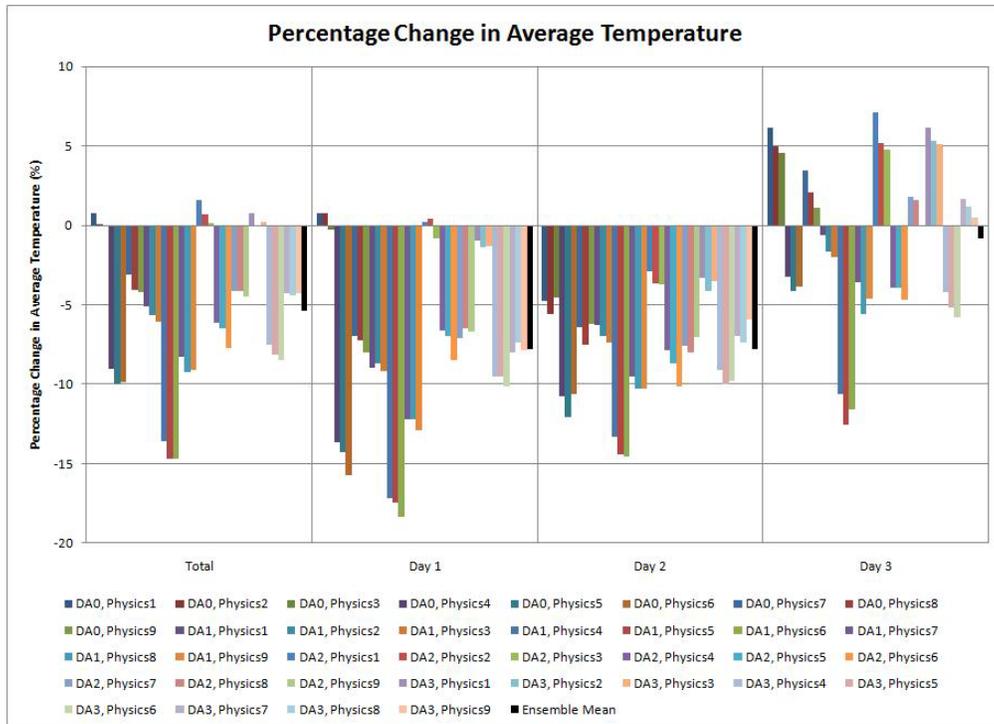


Figure F-108 Change in average temperature at Cairo station in winter

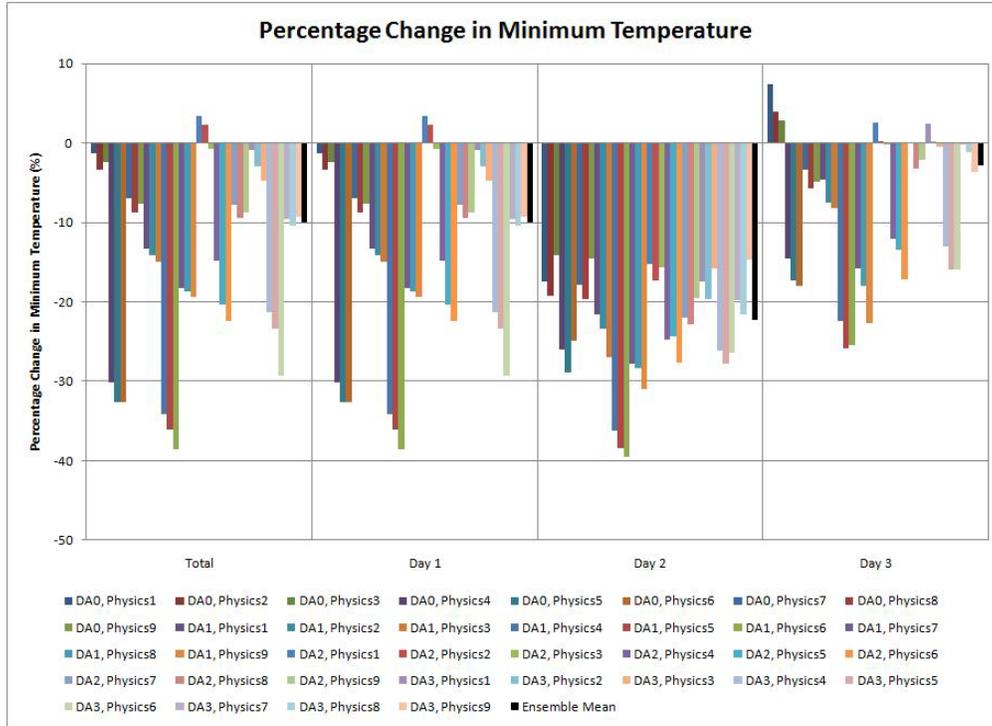


Figure F-109 Change in minimum temperature at Cairo station in winter

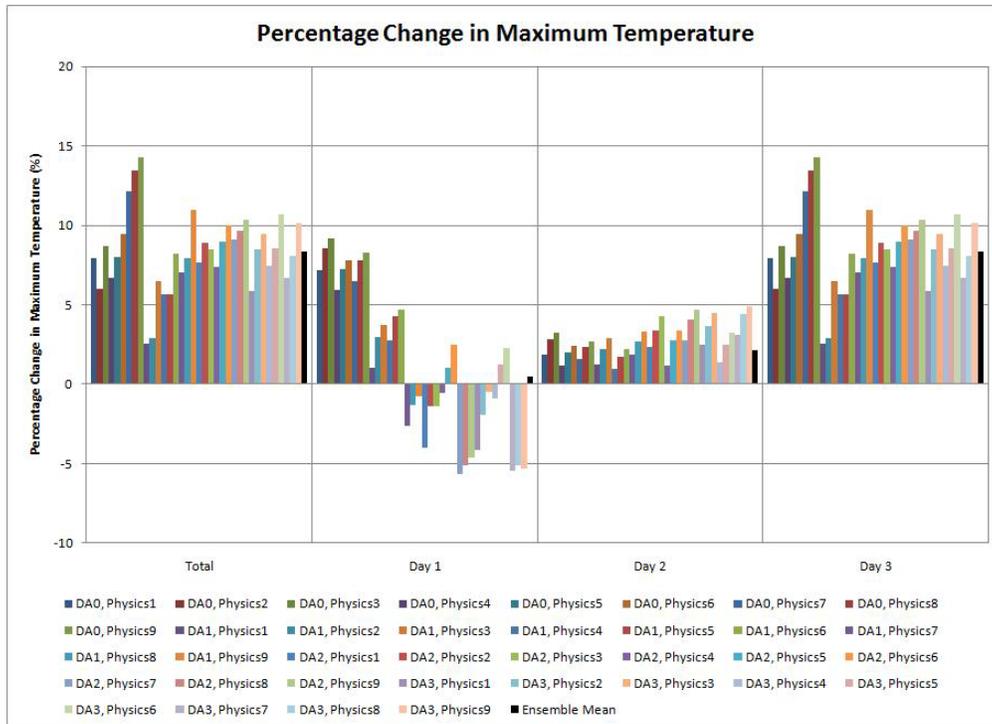


Figure F-110 Change in maximum temperature at Cairo station in winter

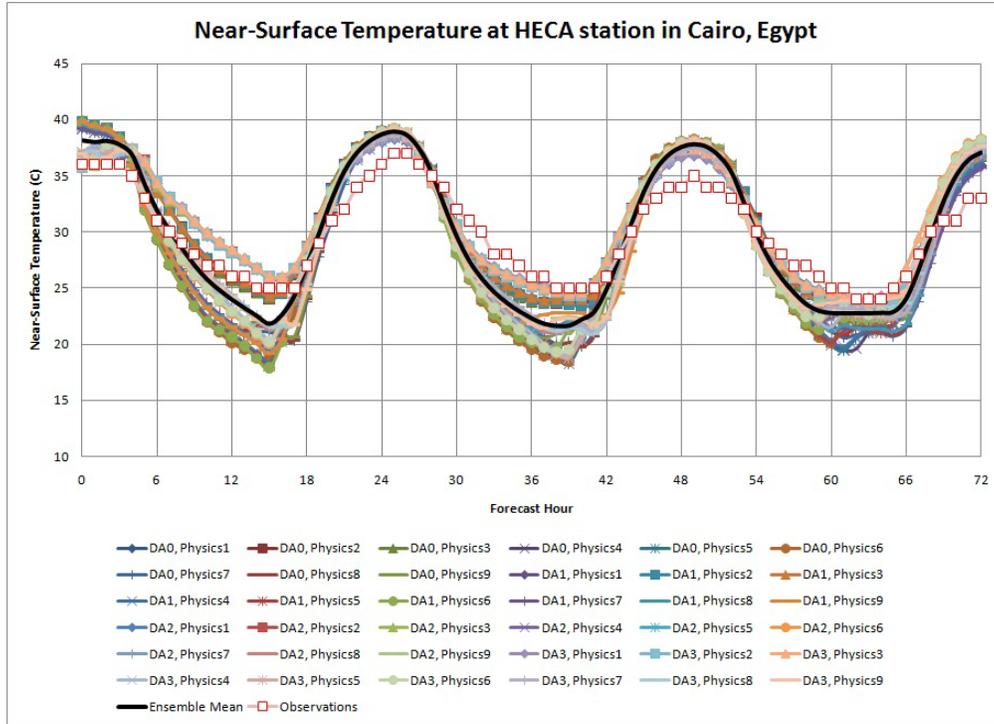


Figure F-111 Near-surface temperature at Cairo station in summer

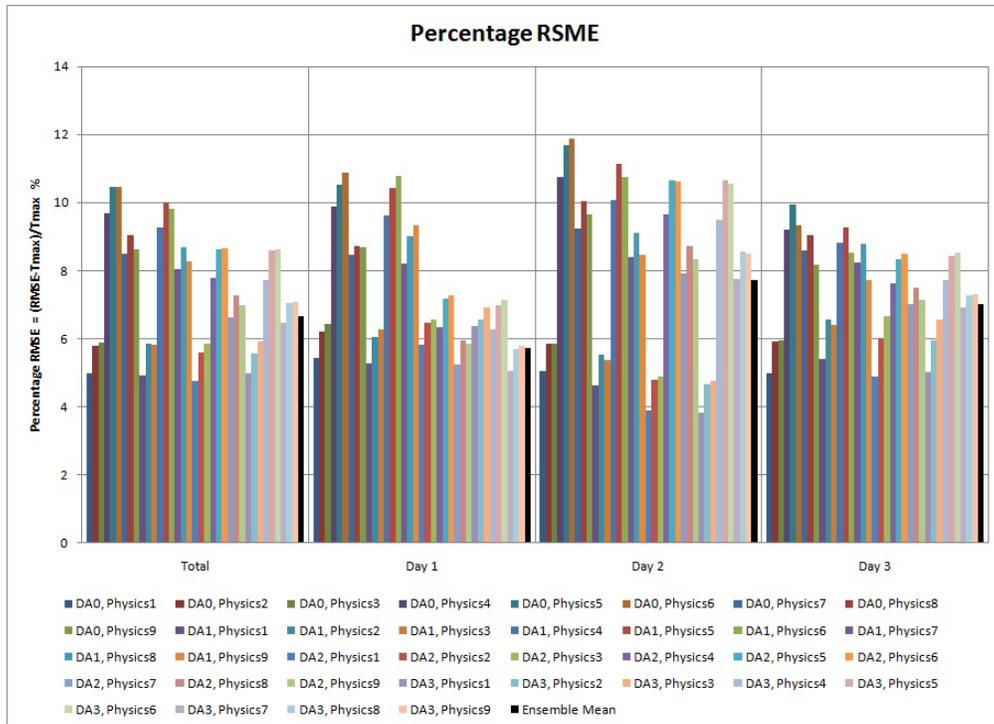


Figure F-112 RMSE of near-surface temperature at Cairo station in summer

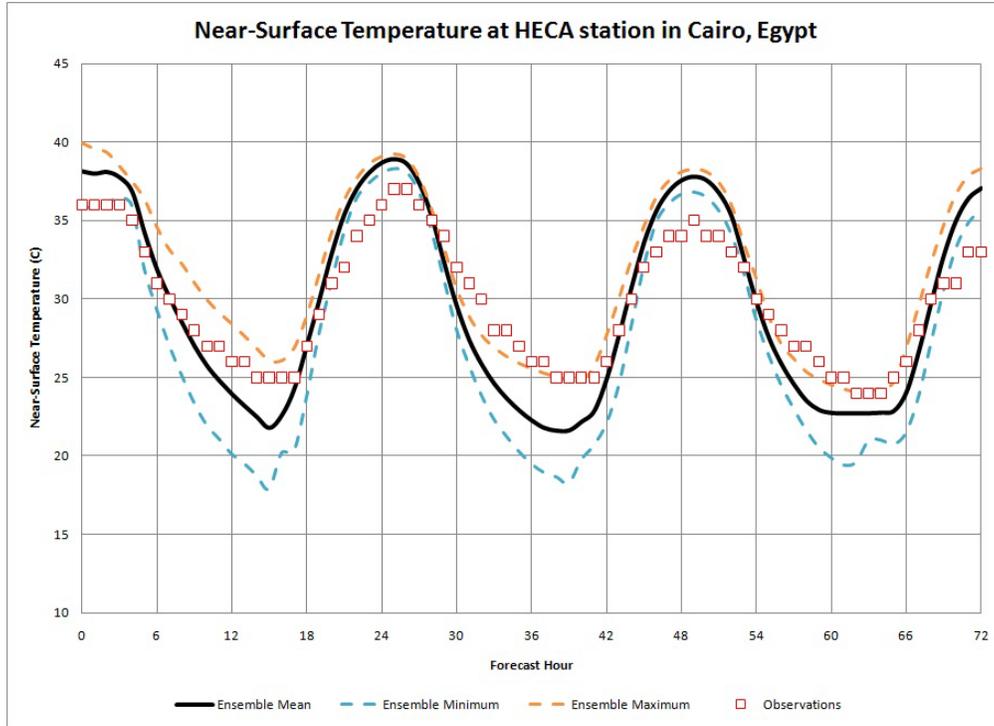


Figure F-113 Min, Max and Mean temperature at Cairo station in summer

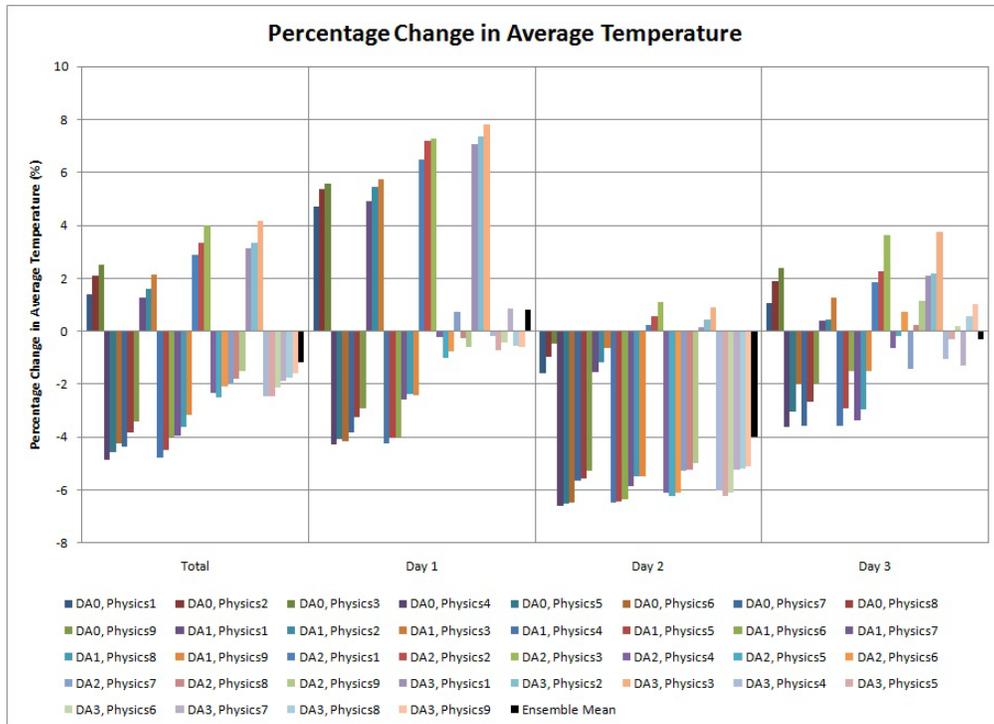


Figure F-114 Change in average temperature at Cairo station in summer

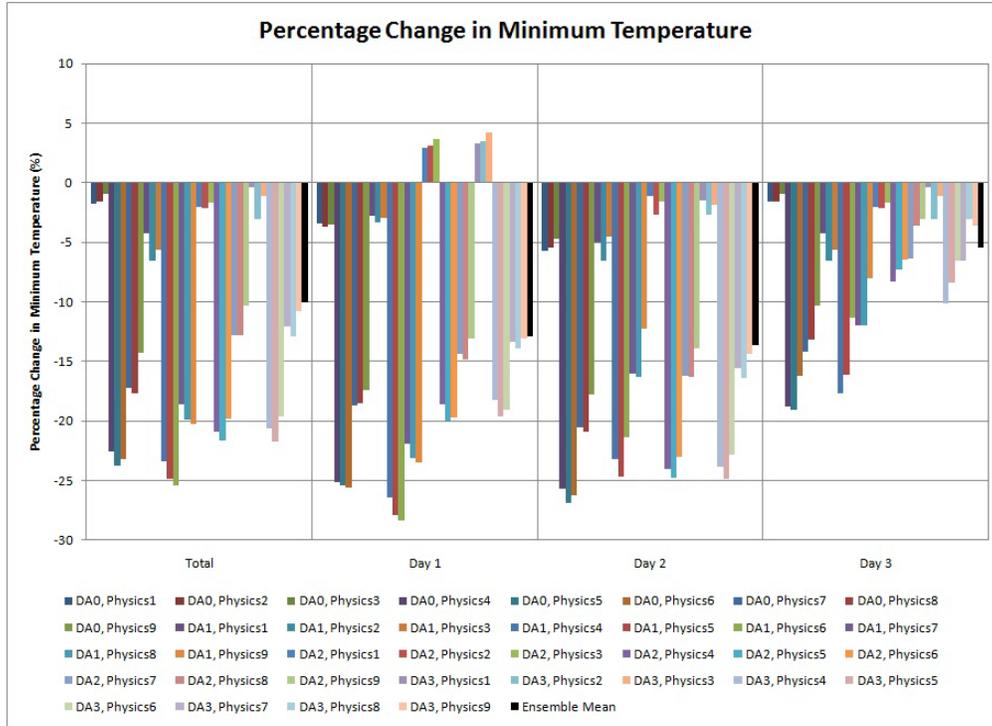


Figure F-115 Change in minimum temperature at Cairo station in summer

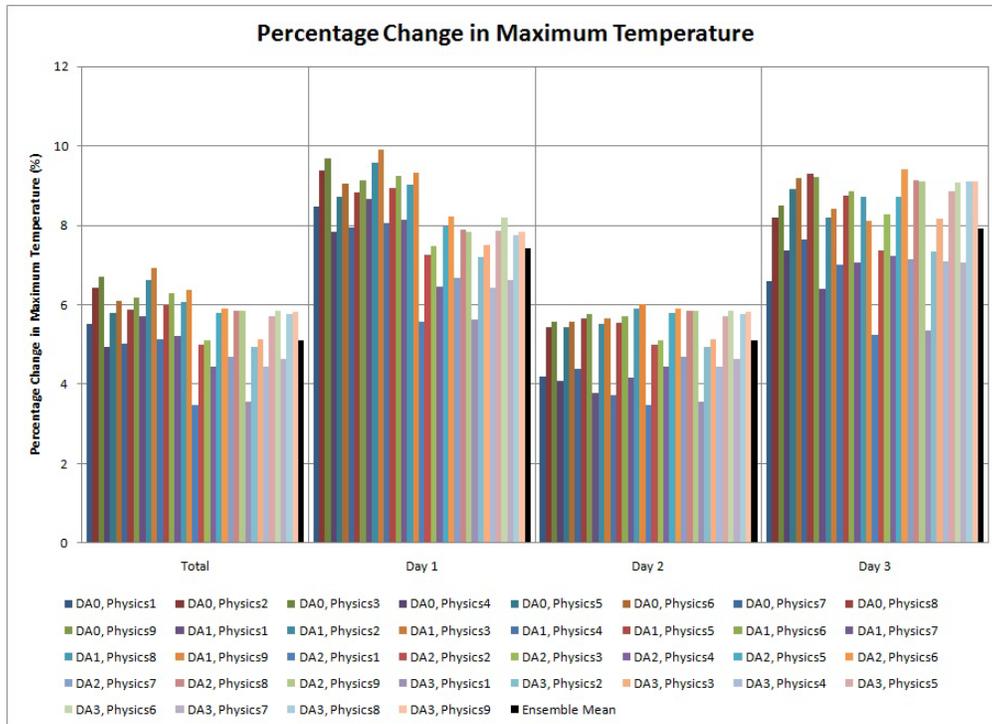


Figure F-116 Change in maximum temperature at Cairo station in summer

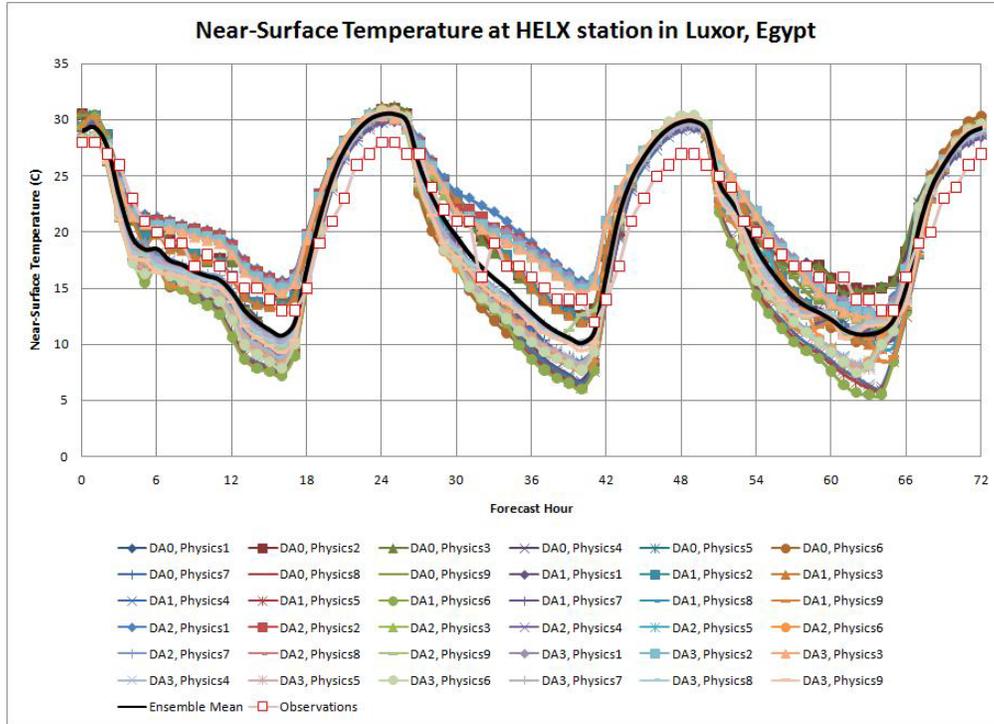


Figure F-117 Near-surface temperature at Luxor station in winter

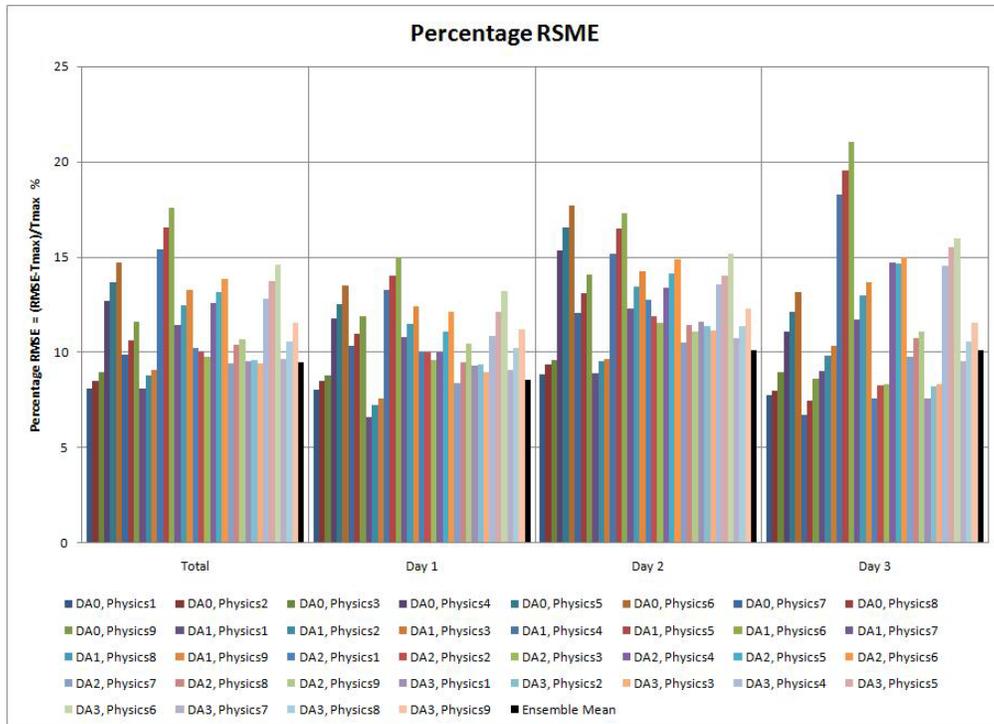


Figure F-118 RMSE of near-surface temperature at Luxor station in winter

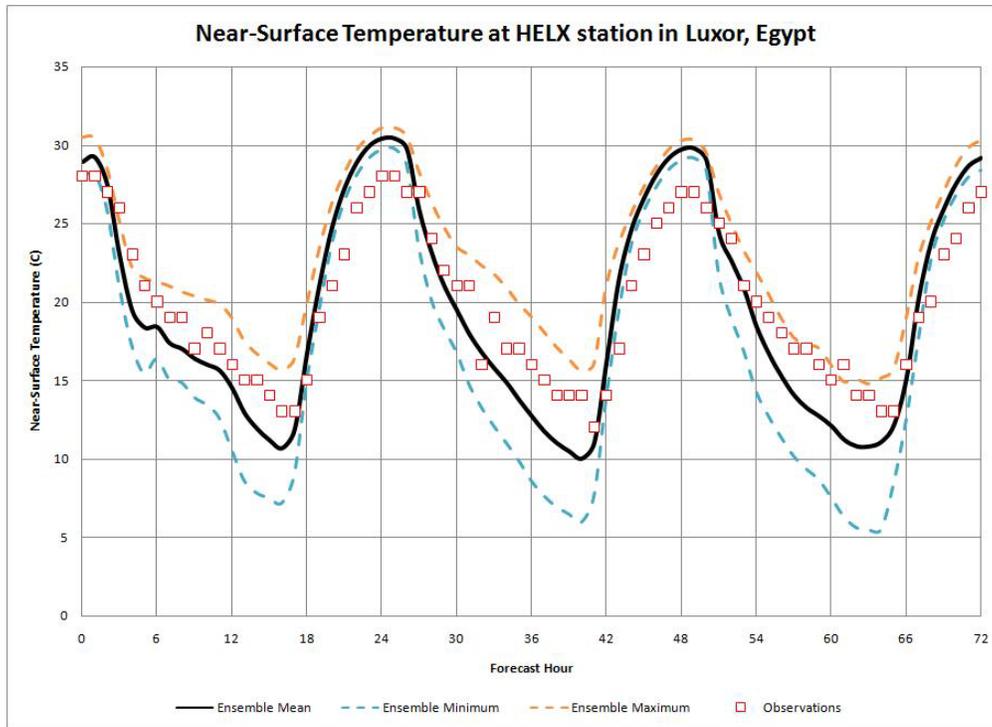


Figure F-119 Min, Max and Mean temperature at Luxor station in winter

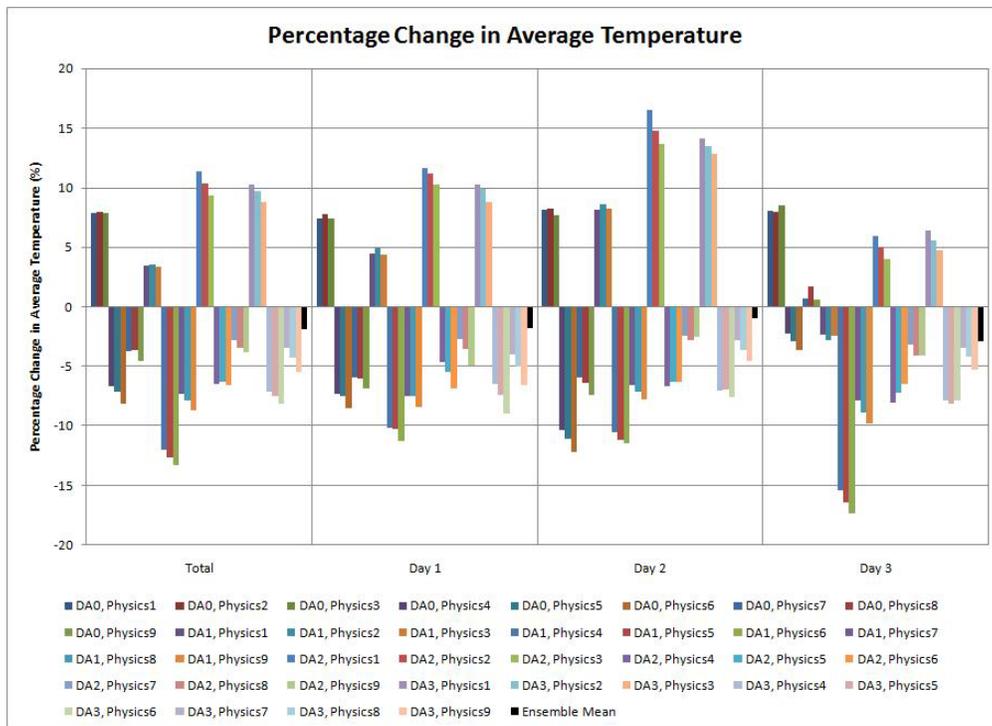


Figure F-120 Change in average temperature at Luxor station in winter

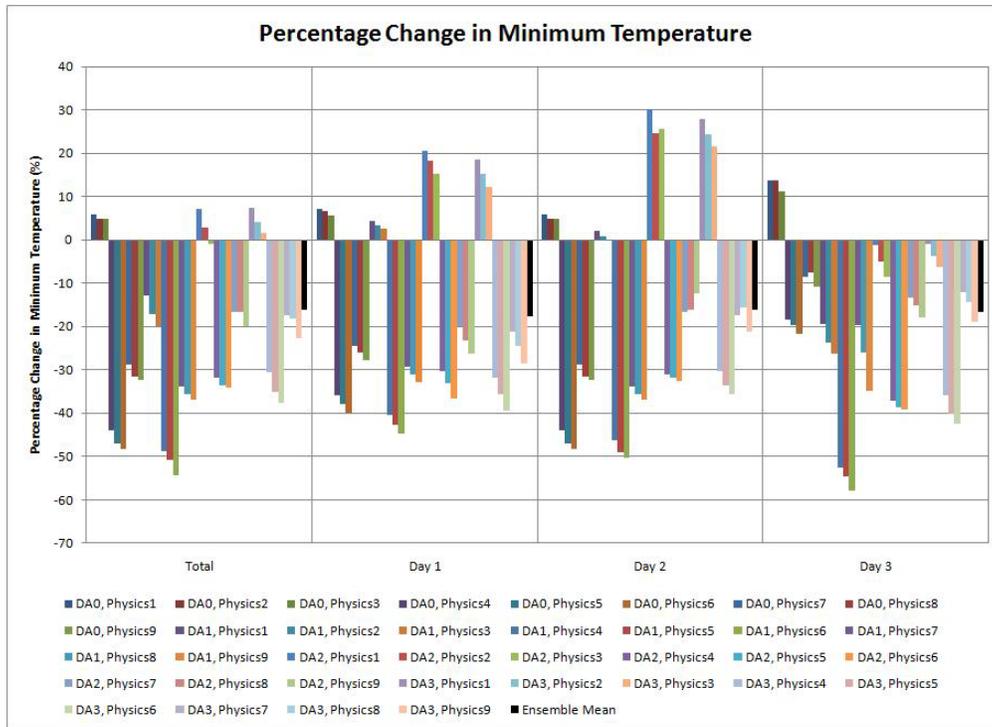


Figure F-121 Change in minimum temperature at Luxor station in winter

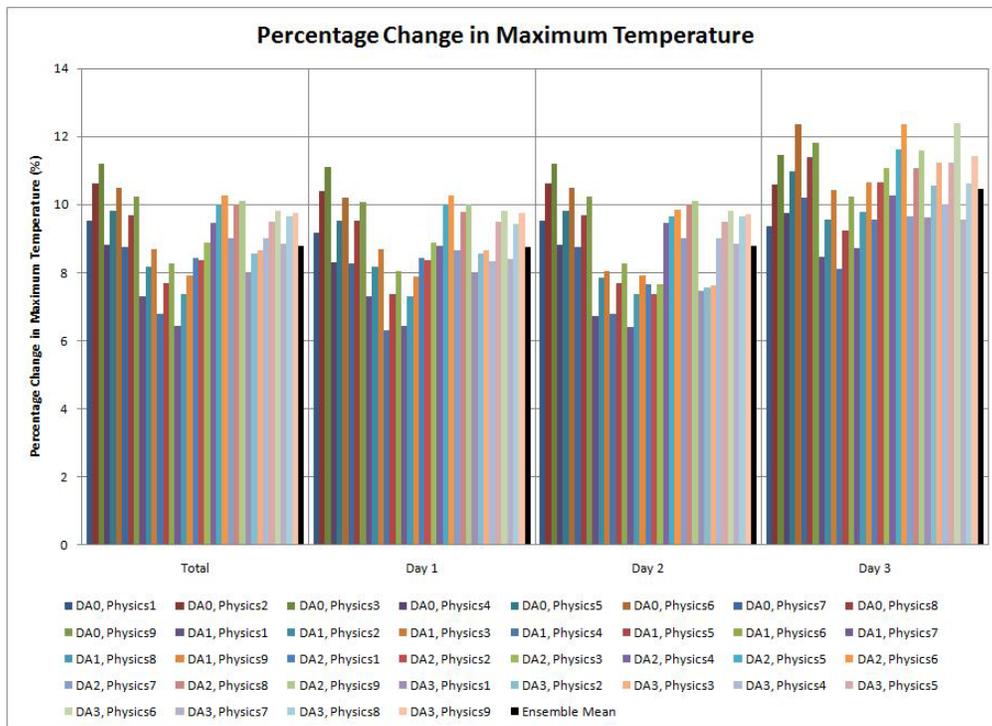


Figure F-122 Change in maximum temperature at Luxor station in winter

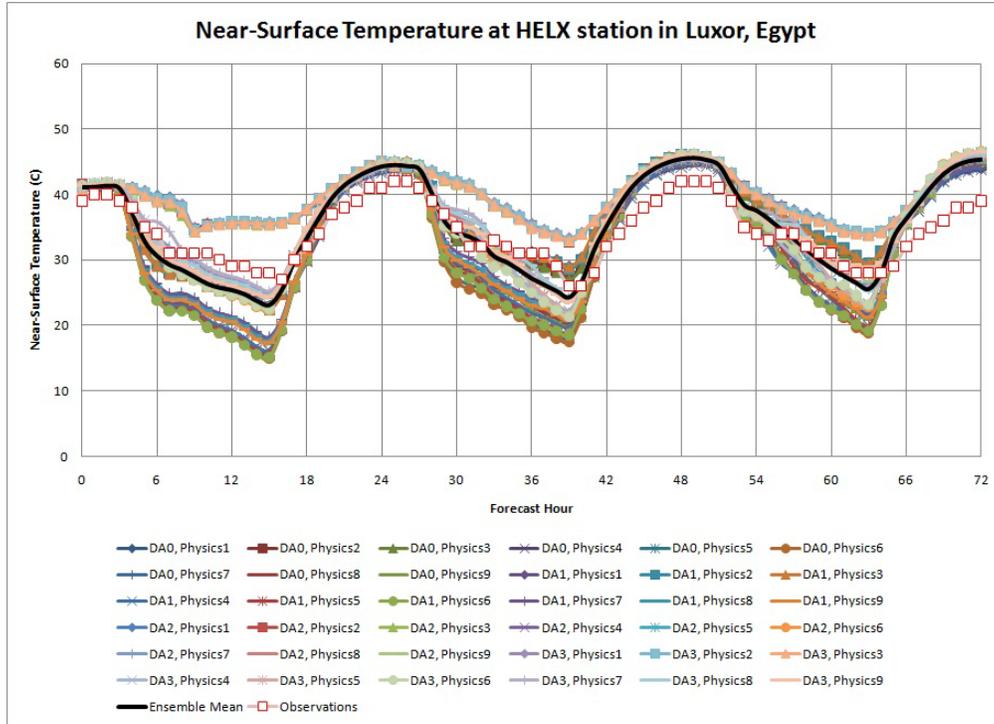


Figure F-123 Near-surface temperature at Luxor station in summer

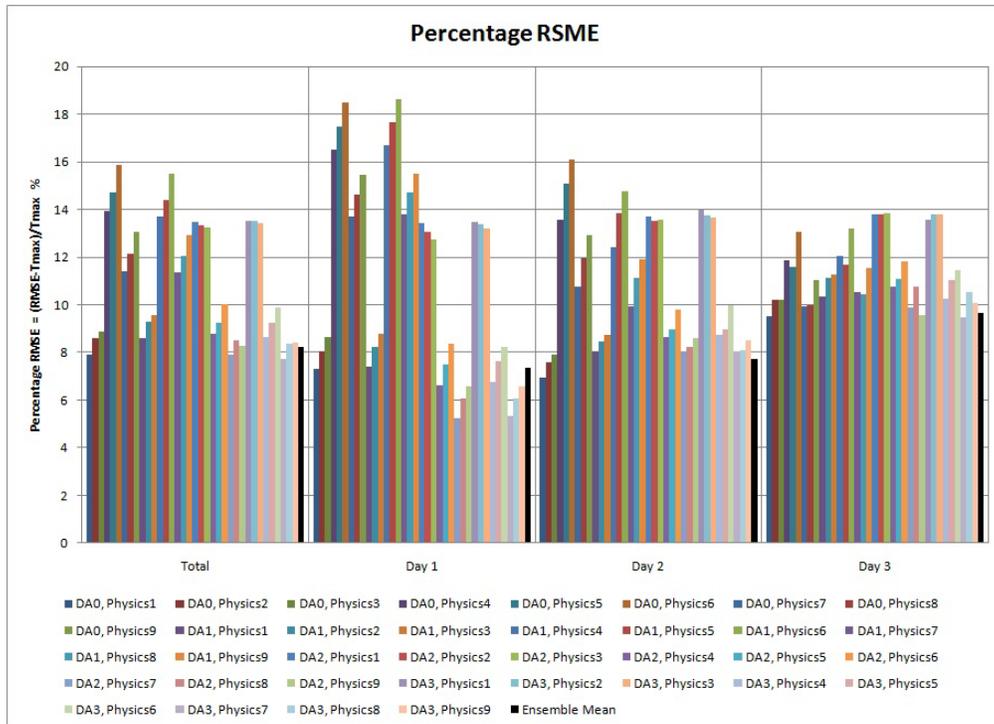


Figure F-124 RMSE of near-surface temperature at Luxor station in summer

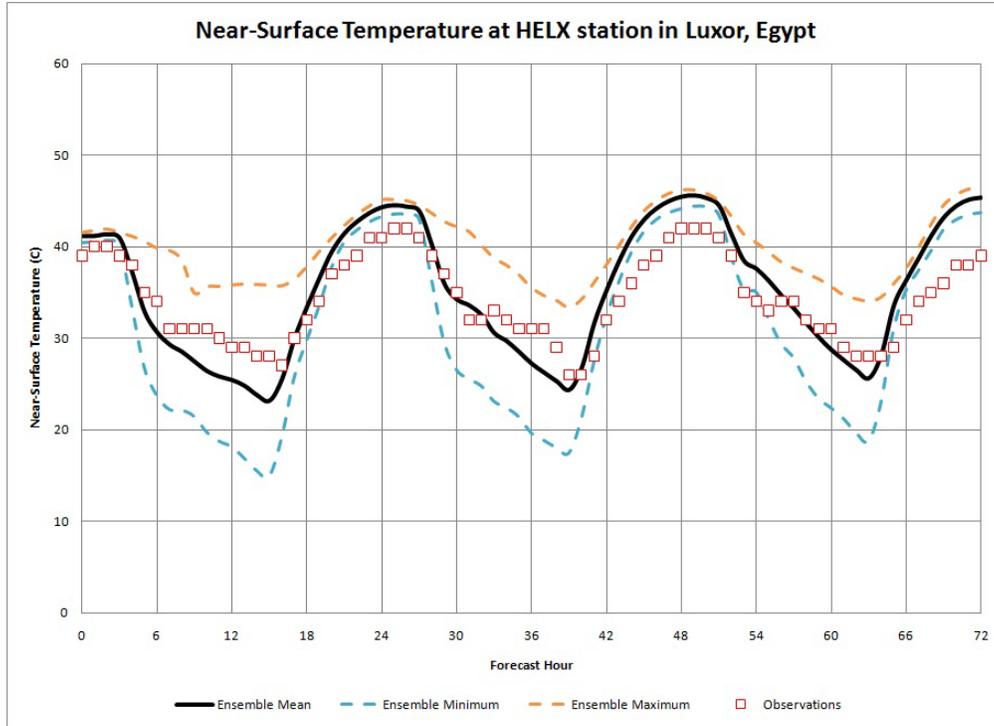


Figure F-125 Min, Max and Mean temperature at Luxor station in summer

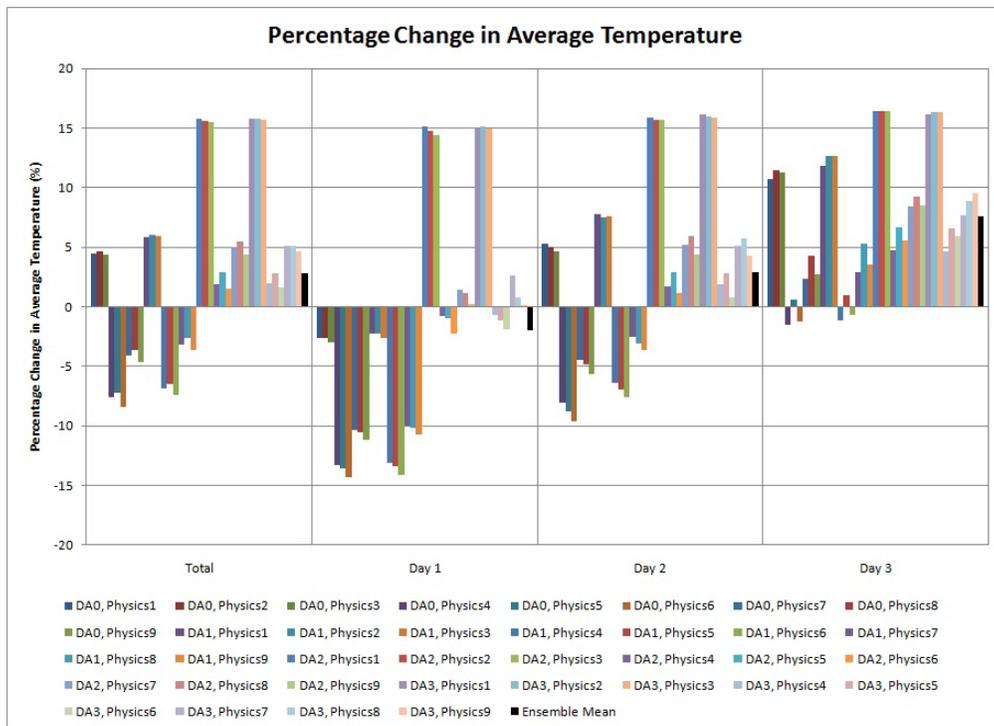


Figure F-126 Change in average temperature at Luxor station in summer

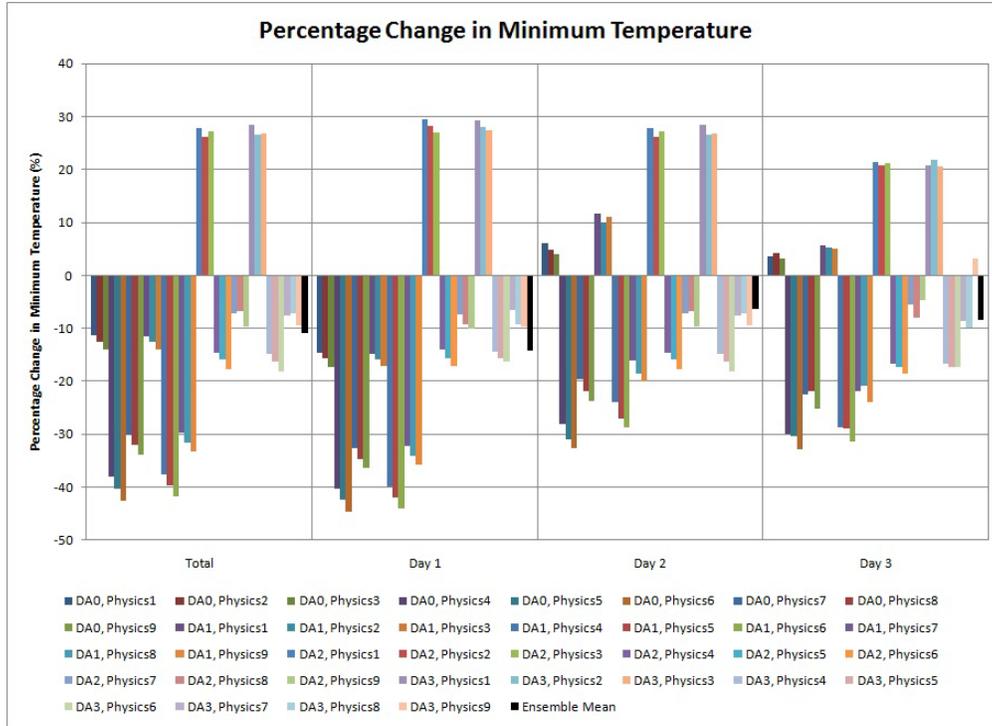


Figure F-127 Change in minimum temperature at Luxor station in summer

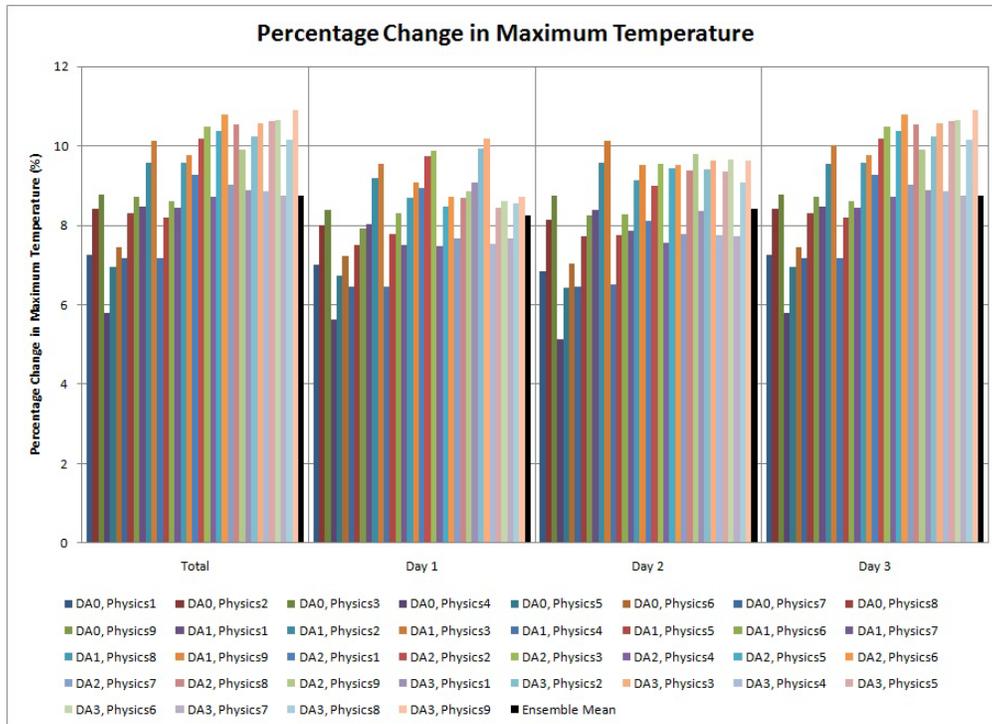


Figure F-128 Change in maximum temperature at Luxor station in summer

G AUTOMATION SCRIPTS

G.1. INITIALIZATION SCRIPT

```
#!/bin/sh
#$Id: MM5.INIT, v 3.0 2008/07/12 12:07:00 narss EGY $
#-----
# This script is used for initialization of MM5 inputs for analysis/forecasting.
#-----
# HISTORY:
# Version      Date          Comment          Author
# 1.0          .            12/07/06        Initial         Hamada Sultan
# 2.0          .            12/07/07        Hamada Sultan
# 3.0          .            12/07/08        Hamada Sultan
#-----
# COPYRIGHT (C) National Authority for Remote Sensing and Space Sciences (NARSS)
#-----

usage(){
  echo
  echo 'Usage:
  echo ' MM5.INIT YYYYMMDDHH TIME DA PHYSICS [COMMENT]
  echo
  echo '      DA = 0 : NO Data Assimilation
  echo '          = 1 : OA Initialization
  echo '          = 2 : FDDA-OA Dynamic Initialization 06 hr
  echo '          = 3 : FDDA-OA Dynamic Initialization 12 hr
  echo '          = 4 : 3DVAR Initialization
  echo '          = 5 : FDDA-3DVAR Dynamic Initialization 06 hr
  echo '          = 6 : FDDA-3DVAR Dynamic Initialization 12 hr
  echo '          = 7 : FDDA-OA Dynamic Analysis during run time
  echo '          = 8 : FDDA-3DVAR Dynamic Analysis during run time
  echo
  echo '      PHYSICS = 1 : MRF PBL with default thermal roughness length      & Cloud-Radiation
  echo '               = 2 : MRF PBL with Garratt thermal roughness length    & Cloud-Radiation
  echo '               = 3 : MRF PBL with Zilitunkevich thermal roughness length & Cloud-Radiation
  echo '               = 4 : MRF PBL with default thermal roughness length    & CCM2 Radiation
  echo '               = 5 : MRF PBL with Garratt thermal roughness length    & CCM2 Radiation
  echo '               = 6 : MRF PBL with Zilitunkevich thermal roughness length & CCM2 Radiation
  echo '               = 7 : MRF PBL with default thermal roughness length    & RRTM Longwave
  echo '               = 8 : MRF PBL with Garratt thermal roughness length    & RRTM Longwave
  echo '               = 9 : MRF PBL with Zilitunkevich thermal roughness length & RRTM Longwave
  echo
  echo '      ( YYYYMMDDHH is the starting time and TIME is the analysis/forecast time in hours )
  echo
}

if [ $# -lt 4 ]; then
  usage
  exit 1
fi
```

```
if [ $# = 5 ]; then
  COMMENT=_$5
fi
```

```
YYYYMMDDHHs=$1 TIME=$2 DA=$3 PHYSICS=$4
```

```
MODIFICATIONS(){
```

```
# User modifications:
```

```
MM5_ROOT=$HOME/mm5 # MM5 root directory
MM5_RUN=$HOME/mm5/INPUT # MM5 working directory
BUFR_ROOT=$HOME/bufr # BUFR root directory
BUFR_RUN=$HOME/bufr/run # BUFR working directory
AAPP_PREFIX=$AAPP_PREFIX # IAPP root directory
ATOVS_ENV=${AAPP_PREFIX}/ATOVS_ENV6 # Environment variables for AAPP
IAPP_ROOT=$HOME/iapp # IAPP root directory
IAPP_RUN=$HOME/iapp/run # IAPP working directory
IAPP Ancillary=/storage/data/iapp.ancillary # IAPP ancillary data
```

```
Where30sTer=/storage/data/mm5.terrain/GTOPO30 # TERRAIN 30 second Data
FNLDataDir=/storage/data/mm5.fnl # Global gridded data in grib format FNL
GFSDDataDir=/storage/data/mm5.gfs # Global gridded data in grib format GFS
GFSDDataDir2=/storage/data/ncep.gfs # Global gridded data in grib2 format GFS
GRIBDataDir=/storage/data/mm5.grib # Global gridded data in grib format FNL/GFS
ON29DataDir=/storage/data/mm5.adp # NCEP ADP global surface and upperair observations subsets in ON29 format
BUFRDataDir=/storage/data/mm5.bufr # NCEP ADP global surface and upperair observations BUFR/PRE PBufr
ATOVSDataDir=/storage/data/atovs.class # Advanced TIROS Operational Vertical Sounder NOAA/ATOV S; atovs.class/hrpt.hmf
```

```
GetGRIB=0 # 0: Do not download GRIB data, 1: Download GRIB data
GRIB=FNL # FNL or GFS will be used for MM5 REGRID program
GFS=1 # 1: GRIB format, 2: GRIB2 format
```

```
PrepareOBS=0 # 0: Do not prepare observation (prepare your data manually), 1: Prepare observations manually
```

```
GetConventionalOBS=0 # 0: Do not download conventional observational data, 1: Download conventional observational data
```

```
ConventionalOBS=2 # 0: None, 1: NCEP ADP BUFR, 2: NCEP ADP PREPBufr and 3: NCEP ADP ON29
```

```
GetSatelliteOBS=0 # 0: Do not download Satellite observational data, 1: Download Satellite observational data
```

```
SatelliteOBS=2 # 0: None, 1: ATOVS HRPT HMF, 2: ATOVS NOAA level 1b and 3: Other satellite observations
```

```
FDDA=1 # 0: NO FDDA, 1: FDDAGD, 2: FDDAOB and 3: Combined FDDAGD and FDDAOB... FDDA=2 f or resolution < 9 km
```

```
# South and North are from -90 to +90 & West and East are from -180 to +180
```

```
South=0
```

```
North=60
```

```
West=0
```

```
East=60
```

```
MAXNES=3 # Domain3 is the domain of interest (Egypt)
```

```

PHIC=27.255
XLONC=31.036
IMX=071097175999999 #  $N+2*((n2-1)+(n3-1)/3), 3*N-2+2*(n3-1), 9*N-8+0$ ; "N=61/3, n2=?,n3=19: 33+2*(n2-1),97,175" n2=9,20
JMX=071097175999999
DIS=081027009003001
NEST1=001020020999999 # 1, n2, n3, n4, n5
NESTJ=001020020999999 # 1, n2, n3, n4, n5

TISTEP=180060020006002
# TISTEP=060030020006002 # Time step for stability but more computational cost

RINBLW=81
RINXY=81
RINBLW=162
RINXY=162

}

GFS2GRIB(){
# Converting GFS data from GRIB2 to GRIB format:

InDataDir=$GFSDDataDir2 OutDataDir=$GFSDDataDir

Tmax=$TIME

if [ -d $InDataDir/gfs.$YYYYMMDDHH ]
then
cd $InDataDir/gfs.$YYYYMMDDHH
YYYYMMDD=$(expr substr $YYYYMMDDHH 1 8) EE=$(expr substr $YYYYMMDDHH 9 2) HH=$EE
for (( TT = 0 ; TT <= $Tmax; TT = TT + 6 ))
do
Ltt=$(expr length $TT)
if [ $Ltt -eq 1 ]
then
InFileName="gfs.t"$EE"z.pgrb2f0"$TT
else
InFileName="gfs.t"$EE"z.pgrb2f"$TT
fi
Lhh=$(expr length $HH)
if [ $Lhh -eq 1 ]
then
OutFileName="gfs_"$(expr substr $YYYYMMDD 3 6)"_0"$HH"_$EE"
else
OutFileName="gfs_"$(expr substr $YYYYMMDD 3 6)"_"$HH"_$EE"
fi
if [ -e $InDataDir/gfs.$YYYYMMDDHH/$InFileName ]
then
echo $OutFileName
cvgrib -g2i $InDataDir/gfs.$YYYYMMDDHH/$InFileName $OutDataDir/$OutFileName
else
echo "$InDataDir/gfs.$YYYYMMDDHH/$InFileName file does not exist !"
fi
let "HH += 3"
if [ $HH -eq 24 ]

```

```

    then
        HH="0"
        YYYYDDD=$(JulianDate.exe $YYYYMMDD)
        let "YYYYDDD += 1"
        YYYYMMDD=$(JulianDate.exe $YYYYDDD)
    fi
done
else
    echo "$InDataDir/gfs.$YYYYMMDDHH directory does not exist !"
fi
}

ATOVS2NC(){
# Converting ATOVS data to NetCDF format:

cd $FNLDataDir
for filename in fnl*
do
    if [ -e $filename ]
    then
        echo "  $filename"
        iapp_grib2nc.ksh $filename $IAPP_ROOT/decoders/files/iapp_ancillary.cdl > ancillary_$filename.lis 2>&1
    fi
done

rm *.lis
for filename in iapp_ancillary*
do
    if [ -e $filename ]
    then
        YY=$(expr substr $filename 18 2)
        MM=$(expr substr $filename 20 2)
        DD=$(expr substr $filename 22 2)
        YYMMDD=$YY$MM$DD
        FNL=$(expr substr $filename 24 2)
        mv $filename IAPP.ANCILLARY.D$YYMMDD.FNL$FNL.NC
    fi
done

mv *.NC $IAPP Ancillary

cd $IAPP_RUN

for filename in *_NOAA*_HMF_to_AAPP.hmf
do
    if [ -e $filename ]
    then
        . $ATOVS_ENV
        if [ $(expr substr $filename 3 2) -gt 80 ]
        then
            YYYY='19'$(expr substr $filename 3 2)
        else
            YYYY='20'$(expr substr $filename 3 2)
        fi
    fi

```

```

AAPP_RUN_NOAA -D -Y $YYYY $IAPP_RUN/$filename > aapp.log
mv *.ltd hirs.ltd
select_hirsld hirs.ltd $South $North $West $East > select_hirsld.log
rm $filename *.hrp *.lib *.lic hirs.ltd *.txt *.err *.par *.log
fi
done

for filename in *.N*.D*.S*.E*.B*
do
if [ -e $filename ]
then
. $ATOVS_ENV
if [ -f *HIR$(expr substr $filename 8 35)* ]
then
mv *HIR$(expr substr $filename 8 35)* hrsn.lib
HRSFLAG='True'
else
HRSFLAG='False'
fi
if [ -f *AMA$(expr substr $filename 8 35)* ]
then
mv *AMA$(expr substr $filename 8 35)* aman.lib
AMAFLAG='True'
else
AMAFLAG='False'
fi
if [ -f *AMB$(expr substr $filename 8 35)* ]
then
mv *AMB$(expr substr $filename 8 35)* ambn.lib
AMBFLAG='True'
elif [ -f *MHS$(expr substr $filename 8 35)* ]
then
mv *MHS$(expr substr $filename 8 35)* ambn.lib
AMBFLAG='True'
else
AMBFLAG='False'
fi
if [ "${HRSFLAG}" = "True" -a "${AMAFLAG}" = "True" -a "${AMBFLAG}" = "True" ]
then
atovin HIRS AMSU-A AMSU-B
atovpp ATOVS
select_hirsld hirs.ltd $South $North $West $East > select_hirs.log
rm *.lic hirs.ltd *.log *.lib
fi
if [ "${HRSFLAG}" = "False" -o "${AMAFLAG}" = "False" -o "${AMBFLAG}" = "False" ]
then
rm *.lib
fi
fi
done

for filename in *.ltd
do
if [ -e $filename ]
then
mv $filename HIRS.N$(expr substr $filename 7 2).D$(JulianDate.exe $(expr substr $filename 11 5)).S$

```

```
(expr substr $filename 25 6).E$(expr substr $filename 33 6).LID
fi
done

for filename in *.LID
do
  if [ -e $filename ]
  then
    N=$(expr substr $filename 7 2)
    D=$(expr substr $filename 11 6)
    S=$(expr substr $filename 19 6)
    E=$(expr substr $filename 27 6)
    FNL=$S
    let "FNL /= 60000"
    let "FNL *= 6"
    Lfml=$(expr length $FNL)
    if [ $Lfml -eq 1 ]
    then
      FNL='0'$FNL
    fi
    ln -sf $filename iapp_hirs.lid
    ln -sf $IAPP Ancillary/IAPP.ANCILLARY.D$D.FNL$FNL.NC iapp_ancillary.nc
    ln -sf ../filenames/FILENAMES.N$N iapp.fileNames
    ncgen -b $IAPP_ROOT/cdlfiles/uwretrievals.cdl
    iapp_main > iapp.log
    mv uwretrievals.nc IAPP.RETRIEVALS.N$N.D$D.S$S.E$E.NC
    rm iapp_hirs.lid iapp_ancillary.nc iapp.fileNames *.log
  fi
done
}
```

```
ATOVS2FDDAOB(){
```

```
# Converting ATOVS data to FDDAOB format (MM5OBS-DOMAIN):
```

```
IMX1=$(expr substr $IMX 1 3)
JMX1=$(expr substr $JMX 1 3)
DIS1=$(expr substr $DIS 1 3)
IMX2=$(expr substr $IMX 4 3)
JMX2=$(expr substr $JMX 4 3)
DIS2=$(expr substr $DIS 4 3)
IMX3=$(expr substr $IMX 7 3)
JMX3=$(expr substr $JMX 7 3)
DIS3=$(expr substr $DIS 7 3)
IMX4=$(expr substr $IMX 10 3)
JMX4=$(expr substr $JMX 10 3)
DIS4=$(expr substr $DIS 10 3)
IMX5=$(expr substr $IMX 13 3)
JMX5=$(expr substr $JMX 13 3)
DIS5=$(expr substr $DIS 13 3)
```

```
cd $IAPP_RUN
```

```
for filename in MM5OBS_DOMAIN*
do
```

```

if [ -e $filename ]
then
  rm ./$filename
fi
done

for (( NEST = 1 ; NEST <= $MAXNES; NEST = NEST + 1 ))
do
  if [ $NEST -eq "1" ]
  then
    IMX=$IMX1
    JMX=$JMX1
    DIS=$DIS1
  elif [ $NEST -eq "2" ]
  then
    IMX=$IMX2
    JMX=$JMX2
    DIS=$DIS2
  elif [ $NEST -eq "3" ]
  then
    IMX=$IMX3
    JMX=$JMX3
    DIS=$DIS3
  elif [ $NEST -eq "4" ]
  then
    IMX=$IMX4
    JMX=$JMX4
    DIS=$DIS4
  elif [ $NEST -eq "5" ]
  then
    IMX=$IMX5
    JMX=$JMX5
    DIS=$DIS5
  fi

cat << End_Of_Namelist | sed -e 's/#.*//; s/ *$// ' > ./MM5OBS.NCL$NEST
begin

; Inputs:
; -----

PHIC = $PHIC
XLONC = $XLONC

IMX = $IMX
JMX = $JMX
Pfull = (/ 1000,998,997,994,992,988,985,980,976,970,964,955,945,930,910,890,850,835,800,760,700,6
80,660,580,560,500,485,465,435,0425,375,325,275,225,175,160,140,130,100 /)
DIS = $DIS

; -----

diri = "."
fils = systemfunc ("ls "+diri+"IAPP.RETRIEVALS*.nc")
in = addfiles(fils,"r")
ListSetType (in, "cat")

```

```

DOY = (/ in[:]->Scanline_Day_of_Year /)
TOD = (/ in[:]->Scanline_UTC_Time_of_Day /)
PHI = (/ in[:]->Latitude /)
XLON = (/ in[:]->Longitude /)

P1 = (/ in[0]->Pressure_Levels /)
h = 0+(288.15/-6.5)*((P1/1013.25)^(-6.5*8.31432/28.9644/9.80665)-1)
P = 1000*(h/(275/-6.5)+1)^(1/(-6.5*8.31432/28.9644/9.80665))
KMX = dimsizes(Pfull)
Phalf = new(KMX-1,float)
do k = 0 , KMX - 2
  Phalf(k)=(Pfull(KMX-k-1)+Pfull(KMX-k-2))/2
end do

T1 = (/ in[:]->Temperature_Retrieval /)
Q1 = (/ in[:]->WaterVapor_Retrieval /)

DIMS = dimsizes(T1)

Along_Track = DIMS(0)
Across_Track = DIMS(1)
Pres_Levels = DIMS(2)

TIME = DOY * 100 + TOD / 1000 / 60 / 60
ip = dim_pqsort(TIME,1)

conv = 57.29578
A = 6370
PHI1 = 0
C2 = A * cos(PHI1 / conv)
YC = C2 * log((1 + sin(PHIC / conv)) / cos(PHIC / conv))

ratio = 1
Iz = 1
Jz = 1

IC0 = (IMX + 1) / 2
JC0 = (JMX + 1) / 2

IC = (IC0 - Iz) * ratio + 0.5
JC = (JC0 - Jz) * ratio + 0.5

MM5OBS_Number = 0
K_OBS = KMX - 1
MISSINGFLAG = ismissing(PHI)
do i = 0 , Along_Track - 1
  do j = 0 , Across_Track - 1
    if ( MISSINGFLAG(ip(i),j) .ne. "True" ) then
      T = int2p (P,T1(ip(i),j,:),Phalf,2)
      Q = int2p (P,Q1(ip(i),j,:),Phalf,2)
      do k = 0 , K_OBS - 1
        Y = C2 * log((1 + sin(PHI(ip(i),j) /conv)) / cos(PHI(ip(i),j) / conv))
        X = C2 * (XLON(ip(i),j) - XLONC) / conv
        TIMEOBS = TIME(ip(i))
        RIO = (IC0 + (Y - YC) / DIS - Iz) * ratio + 1.0
      end do
    end if
  end do
end do

```

```

RJO = (JC0 + X / DIS - Jz) * ratio + 1.0
RKO = k + 1
IVAR1 = 99999.0
IVAR2 = 99999.0
IVAR3 = T(k)
IVAR4 = Q(k) / 1000
IVAR5 = 99999.0
if ( RIO .gt. 1 ) .and. ( RIO .lt. IMX ) .and. ( RJO .gt. 1 ) .and. ( RJO .lt. JMX ) then
  setfileoption("bin","WriteByteOrder","BigEndian")
  fbinrecwrite("./MM5OBS_DOMAIN$NEST",-1,(/ TIMEOBS, RIO, RJO, RKO, IVAR1, IVAR2, IVAR3,
IVAR4, IVAR5 /))
  MM5OBS_Number = MM5OBS_Number + 1
  end if
end do
end if
end do
end do
print(MM5OBS_Number)

end
End_Of_Namelist

for filename in IAPP.RETRIEVALS*.NC
do
  if [ -e $filename ]
  then
    ln -sf $filename $filename.nc
  fi
done

  ncl MM5OBS.NCL$NEST
  rm *.nc MM5OBS.NCL*

done

}

# Applying user modifications:
MODIFICATIONS

# Calculating the date and time parameters:
YYYYs=$(expr substr $YYYYMMDDHHs 1 4)
MMs=$(expr substr $YYYYMMDDHHs 5 2)
DDs=$(expr substr $YYYYMMDDHHs 7 2)
HHs=$(expr substr $YYYYMMDDHHs 9 2)
YYYYJJJs=$(JulianDate.exe $YYYYs$MMs$DDs)
YYYYJJe=$TIME
let "YYYYJJe /= 24"
let "YYYYJJe += YYYYJJJs"
DHH=$YYYYJJe
let "DHH -= YYYYJJJs"
let "DHH *= 24"
let "DHH -= TIME"
HHe=$HHs
let "HHe -= DHH"
if [ $HHe -ge 24 ]

```

```

then
  let "HHe -= 24"
  let "YYYYJJ)e += 1"
elif [ $HHe -lt 0 ]
then
  let "HHe += 24"
  let "YYYYJJ)e -= 1"
fi
if [ $HHe -lt 10 ]
then
  HHe=0$HHe
fi
YYYYMMDDHHe=$(JulianDate.exe $YYYYJJ)e$HHe
YYYYe=$(expr substr $YYYYMMDDHHe 1 4)
MMe=$(expr substr $YYYYMMDDHHe 5 2)
DDe=$(expr substr $YYYYMMDDHHe 7 2)
HHe=$(expr substr $YYYYMMDDHHe 9 2)
if [ $DA -eq 0 -o $DA -eq 1 -o $DA -eq 4 ]
then
  FDDA=0
elif [ $DA -eq 2 -o $DA -eq 5 ]
then
  let "HHs -= 6"
elif [ $DA -eq 3 -o $DA -eq 6 ]
then
  let "HHs -= 12"
fi
if [ $HHs -lt 0 ]
then
  let "DDs -= 1"
  let "HHs += 24"
fi
TIMAX=$YYYYJJ)e
let "TIMAX -= YYYYJJ)s"
let "TIMAX *= 24"
let "TIMAX += HHe"
let "TIMAX -= HHs"
let "TIMAX *= 60"

# Preparing the GRIB data for initialization:
for (( DAY = $YYYYJJ)s ; DAY <= $YYYYJJ)e; DAY = DAY + 1 ))
do
  YYYYMMDD=$(JulianDate.exe $DAY)
  YYYY=$(expr substr $YYYYMMDD 1 4)
  MM=$(expr substr $YYYYMMDD 5 2)
  DD=$(expr substr $YYYYMMDD 7 2)
  HH1="00"
  HH2="18"
  if [ $DAY -eq $YYYYJJ)s ]
  then
    HH1=$HHs
    let "HH1 /= 6"
    let "HH1 *= 6"
    if [ $HH1 -lt $HHs ]
    then
      let "HH1 += 6"
    fi
  fi

```

```

fi
fi
if [ $DAY -eq $YYYYJJJe ]
then
  HH2=$HHe
  let "HH2 /= 6"
  let "HH2 *= 6"
  HRSFC2=$HHFDDAe
  let "HH2 /= 3"
  let "HH2 *= 3"
fi
for (( HH = $HHi ; HH <= $HH2; HH = HH + 6 ))
do
  if [ $HH -lt 12 ]
  then
    HH="0"$HH
  else
    HH=$HH
  fi
  if [ $GRIB = GFS ]
  then
    GFSFileName=$GFSDataDir/"gfs_"$(expr substr $YYYYMMDD 3 6)"_"$HH"_00"
    GRIBFileName=$GRIBDataDir/"grib_"$(expr substr $YYYYMMDD 3 6)"_"$HH"_00"
    ln -sf $GFSFileName $GRIBFileName
  elif [ $GRIB = FNL ]
  then
    FNLFileName=$FNLDataDir/"fnl_"$(expr substr $YYYYMMDD 3 6)"_"$HH"_00"
    GRIBFileName=$GRIBDataDir/"grib_"$(expr substr $YYYYMMDD 3 6)"_"$HH"_00"
    ln -sf $FNLFileName $GRIBFileName
  fi
done
done

# Running MM5 TERRAIN program:
cd $MM5_ROOT/TERRAIN
rm -f ./terrain.deck
cp $MM5_ROOT/DECK/terrain.deck ./
chmod 777 terrain.deck
rm -f TERRAIN_DOMAIN*
./terrain.deck $Where30sTer $MAXNES $PHIC $XLONC $IMX $JMX $DIS $NESTI $NESTJ
mv TERRAIN_DOMAIN* $MM5_ROOT/TEMP

# Running MM5 REGRID program:
cd $MM5_ROOT/REGRID/pregrid
rm -f ./pregrid.csh
cp $MM5_ROOT/DECK/pregrid.csh ./
chmod 777 pregrid.csh
rm -f $MM5_ROOT/REGRID/pregrid/grib.misc/GRIBFILE* $MM5_ROOT/REGRID/pregrid/grib.misc/*FILE:*
rm -f $MM5_ROOT/REGRID/pregrid/*FILE:*
./pregrid.csh $GRIBDataDir $YYYYs $MMs $DDs $HHs $YYYYe $MMe $DDe $HHe
rm -f $MM5_ROOT/REGRID/pregrid/grib.misc/GRIBFILE* $MM5_ROOT/REGRID/pregrid/grib.misc/*FILE:*
cd $MM5_ROOT/REGRID/regridder
rm -f ./regridder.deck
cp $MM5_ROOT/DECK/regridder.deck ./
chmod 777 regridder.deck
./regridder.deck $YYYYs $MMs $DDs $HHs $YYYYe $MMe $DDe $HHe

```

```

rm -f $MM5_ROOT/REGRID/pregrid/*FILE:*
mv REGRID_DOMAIN1 $MM5_ROOT/TEMP
ln -sf $MM5_ROOT/TEMP/REGRID_DOMAIN1 $MM5_ROOT/TEMP/INTERPF_DOMAIN1

# Running MM5 INTERPF program:
cd $MM5_ROOT/INTERPF
rm -f $MM5_ROOT/INTERPF/*_DOMAIN*
rm -f ./interp.deck
cp $MM5_ROOT/DECK/interp.deck1 ./interp.deck
chmod 777 interp.deck
./interp.deck $YYYYs $MMs $DDs $HHs $YYYYe $MMe $DDe $HHe
mv *_DOMAIN1 $MM5_ROOT/TEMP
rm -f $MM5_ROOT/INTERPF/*_DOMAIN*

# Preparing Observational data:
mkdir $MM5_ROOT/TEMP/OBS
if [ $PrepareOBS -eq 1 -a $DA -gt 0 ]
then
  if [ $ConventionalOBS -eq 1 -o $ConventionalOBS -eq 2 ]
  then
    cd $BUFR_RUN
    BUFR2OBS
  elif [ $ConventionalOBS -eq 3 ]
  then
    cd $MM5_ROOT/FETCH
    rm f ./fetch.deck
    cp $MM5_ROOT/DECK/fetch.deck ./
    mkdir $MM5_ROOT/TEMP/OBS
    chmod 777 fetch.deck
    ./fetch.deck $MM5_ROOT $MM5_ROOT/TEMP $ON29DataDir $YYYYs$MMs$DDs$HHs $YYYYe$MMe
$DDe$HHe $East $West $South $North
  fi
  if [ $SatelliteOBS -eq 1 -o $SatelliteOBS -eq 2 ]
  then
    cd $IAPP_RUN
    atovs2nc
    atovs2mm5obs
    mv $IAPP_RUN/MM5OBS_DOMAIN* $MM5_ROOT/TEMP
  elif [ $SatelliteOBS -eq 3 ]
  then
    echo "Unknown satellite observations !"
  fi
else
  echo "Prepare the required observations manually !"
fi

# Creating the run directory and moving the inputs to it:
if [ $TIME -lt 10 ]
then
  RUNDIRPATH=$MM5_RUN/MM5.RUN.$YYYYMMDDHHs.00$TIME'HR'.DA$DA.PHYSICS$PHYSICS$
COMMENT
  elif [ $TIME -gt 10 -a $TIME -lt 100 ]
  then
    RUNDIRPATH=$MM5_RUN/MM5.RUN.$YYYYMMDDHHs.0$TIME'HR'.DA$DA.PHYSICS$PHYSICS$C
OMMENT
  else

```

```

RUNDIRPATH=$MM5_RUN/MM5.RUN.$YYYYMMDDHHs.$TIME'HR'.DA$DA.PHYSICS$PHYSICS$C
COMMENT
fi
mkdir $RUNDIRPATH
cd $MM5_ROOT/TEMP
mkdir $RUNDIRPATH/INPUT
mv TERRAIN_DOMAIN* MMINPUT_DOMAINi LOWBDY_DOMAINi BDYOUT_DOMAINi $RUNDIRPATH/I
NPUT
if [ $FDDA -eq 1 -o $FDDA -eq 3 ]
then
mv $MM5_ROOT/TEMP/OBS $RUNDIRPATH
fi
if [ $FDDA -eq 2 -o $FDDA -eq 3 ]
then
mv MM5OBS_DOMAIN* $RUNDIRPATH/INPUT
fi

# Creating the required decks for running:
mkdir $RUNDIRPATH/DECK
TISTEP1=$(expr substr $TISTEP 1 3)
RINBLW1=$RINBLW
RINXY1=$RINXY
for (( NEST = 1 ; NEST <= $MAXNES; NEST = NEST + 1 ))
do
$MM5_ROOT/DECK/mm5.deck.make $NEST $IMX $JMX $TIMAX $TISTEP1 $PHYSICS $RINBLW1 $RINX
Y1 $DA
mv $MM5_ROOT/TEMP/mm5.deck$NEST $RUNDIRPATH/DECK/mm5.deck$NEST
if [ $NEST -lt $MAXNES ]
then
NESTPLUS1=$NEST
let "NESTPLUS1 += 1"
cp $MM5_ROOT/DECK/nestdown.deck$NESTPLUS1 $RUNDIRPATH/DECK/nestdown.deck$NESTPLUS1
fi

if [ $FDDA -gt 0 ]
then
cp $MM5_ROOT/DECK/interp.f.deck$NEST $RUNDIRPATH/DECK/interp.f.deck$NEST
cp $MM5_ROOT/DECK/little_r.deck$NEST $RUNDIRPATH/DECK/little_r.deck$NEST
cp $MM5_ROOT/DECK/interp.b.deck$NEST $RUNDIRPATH/DECK/interp.b.deck$NEST
fi
start=$NEST
let "start *= 3"
let "start += 1"
TISTEP1=$(expr substr $TISTEP $start 3)
let "RINBLW1 /= 3"
let "RINXY1 /= 3"
done

# Copying the MM5 models to running directory:
mkdir $RUNDIRPATH/MODEL
for (( NEST = 1 ; NEST <= $MAXNES; NEST = NEST + 1 ))
do
cp $MM5_ROOT/MODEL/MM5.PHYSICS$PHYSICS.TAR.GZ$NEST $RUNDIRPATH/MODEL
done

```

```
# Creating the output directories:
mkdir $RUNDIRPATH/OUTPUT
mkdir $RUNDIRPATH/LOG
```

```
# Cleaning the TEMP directory
rm -r $MM5_ROOT/TEMP/*
```

```
exit
```

G.2. RUNNING SCRIPT

```
#!/bin/sh
#$Id: MM5.RUN, v 3.0 2008/08/12 12:07:00 narss EGY $
#-----
# This script is used for running MM5 modeling system for analysis/forecasting.
#-----
# HISTORY:
# Version      Date          Comment          Author
# 1.0          .            12/07/06        Initial          Hamada Sultan
# 2.0          .            12/07/07        .                Hamada Sultan
# 3.0          .            12/07/08        .                Hamada Sultan
#-----
# COPYRIGHT (C) National Authority for Remote Sensing and Space Sciences (NARSS)
#-----

# Working Directories: (edit here according to the system directory structure)
MM5_ROOT=$HOME/mm5 # MM5 root directory; contains NESTDOWN, INTERPB, LITTLE_R and INTER
PF programs.
MM5_RUN=/storage/nwp # MM5 running directory; contains input runs directories.
MM5_DIR=$HOME/mm5/MM5 # MM5 model directory; models will be extracted from input directories t
o it.
# STORAGE=/storage # The storage path; to move completed runs to it (uncomment with line 399).

# Parallel Processing: (edit here according to the type of processing system)
PP=1 # 0: Serial(Single-Core),1: Parallel OpenMP(Single-Core/Multi-Core),2: Parallel MPICH2(Single-
Core/Multi-Core/Cluster)
CORE=2 # 1: Single-Core processor, 2: Dual-Core Processor, 4: Quad-Core Processor, ...etc

# Cleaning working directories:
echo
echo "Cleaning working directories..."
echo
rm -f $MM5_ROOT/NESTDOWN/*_DOMAIN*
rm -f $MM5_ROOT/INTERPB/*_DOMAIN*
rm -f $MM5_ROOT/LITTLE_R/*_DOMAIN*
rm -f $MM5_ROOT/INTERPF/*_DOMAIN*
rm -rf $MM5_ROOT/TEMP/*
rm -rf $MM5_DIR

cd $MM5_RUN

for directory in MM5.RUN*HR*DA*PHYSICS*
do
```

```

if [ -d $directory ]
then

echo "=====
echo "# WORKING WITH DIRECTORY: $MM5_RUN/$directory" :
echo "=====

# Calculating the date and time parameters:
YYYYMMDDHHS=$(expr substr $directory 9 10)
TIME=$(expr substr $directory 20 3)
DA=$(expr substr $directory 28 1)
PHYSICS=$(expr substr $directory 37 1)
if [ $TIME -lt 100 ]
then
    TIME=$(expr substr $TIME 2 2)
fi

YYYYs=$( expr substr $YYYYMMDDHHS 1 4)
MMs=$( expr substr $YYYYMMDDHHS 5 2)
DDs=$( expr substr $YYYYMMDDHHS 7 2)
HHs=$( expr substr $YYYYMMDDHHS 9 2)
YYYYJJs=$(JulianDate.exe $YYYYs$MMs$DDs)
YYYYJJe=$TIME
let "YYYYJJe /= 24"
let "YYYYJJe += YYYYJJs"
DHH=$YYYYJJe
let "DHH -= YYYYJJs"
let "DHH *= 24"
let "DHH -= TIME"
HHe=$HHs
let "HHe -= DHH"
if [ $HHe -ge 24 ]
then
    let "HHe -= 24"
    let "YYYYJJe += 1"
elif [ $HHe -lt 0 ]
then
    let "HHe += 24"
    let "YYYYJJe -= 1"
fi
if [ $HHe -lt 10 ]
then
    HHe=0$HHe
fi
YYYYMMDDHHe=$(JulianDate.exe $YYYYJJe)$HHe
YYYYe=$(expr substr $YYYYMMDDHHe 1 4)
MMe=$(expr substr $YYYYMMDDHHe 5 2)
DDe=$(expr substr $YYYYMMDDHHe 7 2)
HHe=$(expr substr $YYYYMMDDHHe 9 2)
if [ $DA -eq 0 ]
then
    FDDA=0
elif [ $DA -eq 2 -o $DA -eq 5 ]
then
    let "HHs -= 6"
elif [ $DA -eq 3 -o $DA -eq 6 ]

```

```

then
  let "HHs -= 12"
fi
if [ $HHs -lt 0 ]
then
  let "DDs -= 1"
  let "HHs += 24"
fi
TIMAX=${YYYYJJ}e
let "TIMAX -= YYYYJJ)s"
let "TIMAX *= 24"
let "TIMAX += HHe"
let "TIMAX -= HHs"
let "TIMAX *= 60"

# Calculating the starting and ending nests:
MINNES=1
cd $directory/OUTPUT
for filename in MMOUT_DOMAIN*
do
  if [ -e $filename ]
  then
    MINNES=$(expr substr $filename 13 1)
    let "MINNES += 1"
  fi
done
MAXNES=1
cd $MM5_RUN
cd $directory/INPUT
for filename in TERRAIN_DOMAIN*
do
  if [ -e $filename ]
  then
    MAXNES=$(expr substr $filename 15 1)
  fi
done
cd $MM5_RUN

# Checking if it is a completed run:
if [ $MINNES -gt $MAXNES ]
then
  echo
  echo "COMPLETED RUN..."
  echo
else

# Linking the input files to the working directory:
cd $MM5_ROOT/TEMP
rm -f $MM5_ROOT/TEMP/TERRAIN_DOMAIN*
ln -sf $MM5_RUN/$directory/INPUT/TERRAIN_DOMAIN* ./
for filename in $MM5_RUN/$directory/INPUT/MM5OBS_DOMAIN*
do
  if [ -e $filename ]
  then
    rm -f $MM5_ROOT/TEMP/$filename
    ln -sf $filename ./

```

```

    fi
done
echo
echo "Starting Nest = $MINNES   Ending Nest = $MAXNES"
echo
if [ $MINNES -eq 1 ]
then
    rm -f $MM5_ROOT/TEMP/MMINPUT_DOMAINI
    rm -f $MM5_ROOT/TEMP/BDYOUT_DOMAINI
    rm -f $MM5_ROOT/TEMP/LOWBDY_DOMAINI
    ln -sf $MM5_RUN/$directory/INPUT/MMINPUT_DOMAIN$MINNES MMINPUT_DOMAINI
    ln -sf $MM5_RUN/$directory/INPUT/BDYOUT_DOMAIN$MINNES BDYOUT_DOMAINI
    ln -sf $MM5_RUN/$directory/INPUT/LOWBDY_DOMAIN$MINNES LOWBDY_DOMAINI
# Running NESTDOWN for interrupted runs:
elif [ $MINNES -gt 1 -a $MINNES -le $MAXNES ]
then
    NEST=$MINNES
    let "NEST -= 1"
    echo
    echo "Running NESTDOWN for interrupted runs..."
    echo
    rm -f $MM5_ROOT/TEMP/MMOUT_DOMAINI
    ln -sf $MM5_RUN/$directory/OUTPUT/MMOUT_DOMAIN$NEST MMOUT_DOMAINI
    cd $MM5_ROOT/NESTDOWN
    NESTPLUSI=$NEST
    let "NESTPLUSI += 1"
    if [ $NEST != 1 ]
    then
        mv $MM5_ROOT/TEMP/TERRAIN_DOMAIN$NEST $MM5_ROOT/TEMP/TERRAIN_DOMAINI
        mv $MM5_ROOT/TEMP/TERRAIN_DOMAIN$NESTPLUSI $MM5_ROOT/TEMP/TERRAIN_DOMAI
N2
    fi
    rm -f nestdown.deck
    ln -sf $MM5_RUN/$directory/DECK/nestdown.deck$NESTPLUSI nestdown.deck
    chmod 777 nestdown.deck
    ./nestdown.deck $YYYYs $MMs $DDs $HHs $YYYYe $MMe $DDe $HHe
    rm -f $MM5_RUN/$directory/INPUT/MMINPUT_DOMAIN$NESTPLUSI
    rm -f $MM5_RUN/$directory/INPUT/LOWBDY_DOMAIN$NESTPLUSI
    rm -f $MM5_RUN/$directory/INPUT/BDYOUT_DOMAIN$NESTPLUSI
    mv MMINPUT_DOMAIN* $MM5_RUN/$directory/INPUT/MMINPUT_DOMAIN$NESTPLUSI
    mv LOWBDY_DOMAIN* $MM5_RUN/$directory/INPUT/LOWBDY_DOMAIN$NESTPLUSI
    mv BDYOUT_DOMAIN* $MM5_RUN/$directory/INPUT/BDYOUT_DOMAIN$NESTPLUSI
    rm -f $MM5_ROOT/TEMP/MMINPUT_DOMAINI
    rm -f $MM5_ROOT/TEMP/LOWBDY_DOMAINI
    rm -f $MM5_ROOT/TEMP/BDYOUT_DOMAINI
    cd $MM5_ROOT/TEMP
    ln -sf $MM5_RUN/$directory/INPUT/MMINPUT_DOMAIN$NESTPLUSI MMINPUT_DOMAINI
    ln -sf $MM5_RUN/$directory/INPUT/LOWBDY_DOMAIN$NESTPLUSI LOWBDY_DOMAINI
    ln -sf $MM5_RUN/$directory/INPUT/BDYOUT_DOMAIN$NESTPLUSI BDYOUT_DOMAINI
    cd $MM5_RUN
fi

# Loop for the remaining nests:
for (( NEST = $MINNES ; NEST <= $MAXNES; NEST = NEST + 1 ))
do

```

```

echo "_____ "
echo "Running MM5 modeling system for DOMAIN$NEST..."
echo "_____ "
echo
echo "Checking for MM5 model..."
echo
if [ -e $MM5_RUN/$directory/MODEL/MM5.PHYSICS$PHYSICS.TAR.GZ$NEST ]
then
  echo
  echo "  Found MM5 model for DOMAIN$NEST"
  echo
  if [ -d $MM5_DIR ]
  then
    rm -r $MM5_DIR
  fi
  cd $MM5_ROOT
  tar -zxf $MM5_RUN/$directory/MODEL/MM5.PHYSICS$PHYSICS.TAR.GZ$NEST
  cd $MM5_RUN
else
  echo
  echo "Please copy MM5 model MM5.PHYSICS$PHYSICS.TAR.GZ$NEST to $MM5_RUN/$directory
/MODEL ! "
  echo
  exit
fi

# Checking for the observations to use OA (LITTLE_R):
if [ -d $MM5_RUN/$directory/OBS ]
then
  rm -rf $MM5_ROOT/TEMP/OBS
  mkdir $MM5_ROOT/TEMP/OBS
  cd $MM5_ROOT/TEMP/OBS
  ln -sf $MM5_RUN/$directory/OBS/* ./
  cd $MM5_RUN
  echo
  echo "Running INTERPB..."
  echo
  cd $MM5_ROOT/INTERPB
  rm -f interpb.deck
  ln -sf $MM5_RUN/$directory/DECK/interpb.deck$NEST interpb.deck
  chmod 777 interpb.deck
  ./interpb.deck $YYYYs $MMs $DDs $HHs $YYYYe $MMe $DDe $HHe
  rm -f FILE_MMOUTP:* REGRID_DOMAIN*
  rm -f $MM5_ROOT/TEMP/MMOUTP_DOMAIN*
  mv MMOUTP_DOMAIN* $MM5_ROOT/TEMP/MMOUTP_DOMAIN*
  echo
  echo "Running LITTLE_R..."
  echo
  cd $MM5_ROOT/LITTLE_R
  rm -f little_r.deck
  ln -sf $MM5_RUN/$directory/DECK/little_r.deck$NEST little_r.deck
  chmod 777 little_r.deck
  ./little_r.deck "$MM5_ROOT/TEMP" $YYYYs$MMs$DDs$HHs $YYYYe$MMe$DDe$HHe $DA
  rm -f $MM5_ROOT/TEMP/LITTLE_R_DOMAIN*
  mv LITTLE_R_DOMAIN* $MM5_ROOT/TEMP/LITTLE_R_DOMAIN*
  rm -f $MM5_ROOT/TEMP/INTERPF_DOMAIN*

```

```

mv $MM5_ROOT/TEMP/LITTLE_R_DOMAINI $MM5_ROOT/TEMP/INTERPF_DOMAINI
rm -f $MM5_RUN/$directory/INPUT/SFCFDDA_DOMAIN$NEST
mv SFCFDDA_DOMAIN* $MM5_RUN/$directory/INPUT/SFCFDDA_DOMAIN$NEST
rm -f $MM5_ROOT/TEMP/SFCFDDA_DOMAINI
cd $MM5_ROOT/TEMP
ln -sf $MM5_RUN/$directory/INPUT/SFCFDDA_DOMAIN$NEST SFCFDDA_DOMAINI
echo
echo "Running INTERPF..."
echo
cd $MM5_ROOT/INTERPF
rm -f interpf.deck
ln -sf $MM5_RUN/$directory/DECK/interpf.deck$NEST interpf.deck
chmod 777 interpf.deck
./interpf.deck $YYYYs $MMs $DDs $HHs $YYYYe $MMe $DDe $HHe
rm -f $MM5_ROOT/TEMP/MMINPUT_DOMAINI
rm -f $MM5_ROOT/TEMP/LOWBDY_DOMAINI
rm -f $MM5_ROOT/TEMP/BDYOUT_DOMAINI
mv MMINPUT_DOMAIN* $MM5_ROOT/TEMP/MMINPUT_DOMAINI
mv LOWBDY_DOMAIN* $MM5_ROOT/TEMP/LOWBDY_DOMAINI
mv BDYOUT_DOMAIN* $MM5_ROOT/TEMP/BDYOUT_DOMAINI
else
echo
echo "The observations directory $MM5_RUN/$directory/OBS does not exist !"
echo
fi

```

Determining the suitable FDDA type:

```

if [ $DA -eq 0 ]
then
  FDDA=0
elif [ $NEST -le 3 ]
then
  FDDA=1
elif [ $NEST -gt 3 ]
then
  FDDA=2
fi

```

Preparing inputs for running MM5:

```

cd $MM5_DIR/Run
rm -f MMINPUT_DOMAINI
rm -f LOWBDY_DOMAINI
rm -f BDYOUT_DOMAINI
ln -sf $MM5_ROOT/TEMP/MMINPUT_DOMAINI MMINPUT_DOMAINI
ln -sf $MM5_ROOT/TEMP/LOWBDY_DOMAINI LOWBDY_DOMAINI
ln -sf $MM5_ROOT/TEMP/BDYOUT_DOMAINI BDYOUT_DOMAINI

if [ $FDDA -eq 1 -o $FDDA -eq 3 ]
then
  if [ -e $MM5_ROOT/TEMP/SFCFDDA_DOMAINI ]
  then
    rm -f SFCFDDA_DOMAINI
    ln -sf $MM5_ROOT/TEMP/SFCFDDA_DOMAINI SFCFDDA_DOMAINI
  else
    echo "SFCFDDA_DOMAIN$NEST does not exist, please check LITTLE_R job !"
  fi
fi

```

```

if [ -e $MM5_ROOT/TEMP/MMINPUT2_DOMAIN$NEST ]
then
  rm -f MMINPUT2_DOMAINi
  ln -sf $MM5_ROOT/TEMP/MMINPUT2_DOMAIN$NEST MMINPUT2_DOMAINi
else
  ln -sf MMINPUT_DOMAINi MMINPUT2_DOMAINi
fi
fi
if [ $FDDA -eq 2 -o $FDDA -eq 3 ]
then
  if [ -e $MM5_ROOT/TEMP/MM5OBS_DOMAIN$NEST ]
  then
    rm -f MM5OBS_DOMAINi
    ln -sf $MM5_ROOT/TEMP/MM5OBS_DOMAIN$NEST MM5OBS_DOMAINi
  fi
fi
echo
echo "Running MM5..."
echo

cd $MM5_DIR
rm -f mm5.deck
ln -sf $MM5_RUN/$directory/DECK/mm5.deck$NEST mm5.deck
chmod 777 mm5.deck
./mm5.deck $PP $CORE

# Moving the output files to the working directory:
rm -f $MM5_ROOT/TEMP/MMOUT_DOMAIN*
mv ./Run/MMOUT_DOMAIN* $MM5_ROOT/TEMP/MMOUT_DOMAINi
if [ $PP -eq 2 ]
then
  rm -f $MM5_ROOT/TEMP/rs1*
  mv ./Run/rs1* $MM5_ROOT/TEMP
else
  rm -f $MM5_ROOT/TEMP/MM5.LOG
  rm -f $MM5_ROOT/TEMP/fort.26
  mv ./Run/MM5.LOG $MM5_ROOT/TEMP
  mv ./Run/fort.26 $MM5_ROOT/TEMP
fi
rm -f $MM5_ROOT/TEMP/MM5.TIME
mv ./Run/MM5.TIME $MM5_ROOT/TEMP

if [ $NEST -lt $MAXNES ]
then
  echo
  echo "Running NESTDOWN..."
  echo
  cd $MM5_ROOT/NESTDOWN
  NESTPLUSi=$NEST
  let "NESTPLUSi += 1"
  if [ $NEST != 1 ]
  then
    mv $MM5_ROOT/TEMP/TERRAIN_DOMAIN$NEST $MM5_ROOT/TEMP/TERRAIN_DOMAINi
    mv $MM5_ROOT/TEMP/TERRAIN_DOMAIN$NESTPLUSi $MM5_ROOT/TEMP/TERRAIN_DOMA
IN2
  fi

```

```

rm -f nestdown.deck
ln -sf $MM5_RUN/$directory/DECK/nestdown.deck$NESTPLUS1 nestdown.deck
chmod 777 nestdown.deck
./nestdown.deck $YYYYs $MMs $DDs $HHs $YYYYe $MMe $DDe $HHe
rm -f $MM5_RUN/$directory/INPUT/MMINPUT_DOMAIN$NESTPLUS1
rm -f $MM5_RUN/$directory/INPUT/LOWBDY_DOMAIN$NESTPLUS1
rm -f $MM5_RUN/$directory/INPUT/BDYOUT_DOMAIN$NESTPLUS1
mv MMINPUT_DOMAIN* $MM5_RUN/$directory/INPUT/MMINPUT_DOMAIN$NESTPLUS1
mv LOWBDY_DOMAIN* $MM5_RUN/$directory/INPUT/LOWBDY_DOMAIN$NESTPLUS1
mv BDYOUT_DOMAIN* $MM5_RUN/$directory/INPUT/BDYOUT_DOMAIN$NESTPLUS1
rm -f $MM5_ROOT/TEMP/MMINPUT_DOMAIN1
rm -f $MM5_ROOT/TEMP/LOWBDY_DOMAIN1
rm -f $MM5_ROOT/TEMP/BDYOUT_DOMAIN1
cd $MM5_ROOT/TEMP
ln -sf $MM5_RUN/$directory/INPUT/MMINPUT_DOMAIN$NESTPLUS1 MMINPUT_DOMAIN1
ln -sf $MM5_RUN/$directory/INPUT/LOWBDY_DOMAIN$NESTPLUS1 LOWBDY_DOMAIN1
ln -sf $MM5_RUN/$directory/INPUT/BDYOUT_DOMAIN$NESTPLUS1 BDYOUT_DOMAIN1
cd $MM5_RUN
fi

# Moving the output files to the run directory:
mv $MM5_ROOT/TEMP/MMOUT_DOMAIN1 $MM5_RUN/$directory/OUTPUT/MMOUT_DOMAIN$
NEST
if [ $PP -eq 2 ]
then
for filename in rsl.out*
do
Lrslout=$(expr length $filename)
Lsuffix=$Lrslout
let "Lsuffix -= 7"
Suffix=$(expr substr $filename 8 $Lsuffix)
mv $MM5_ROOT/TEMP/rsl.out.* $MM5_RUN/$directory/LOG/RSL.OUT$suffix_"DOMAIN$NEST
done
for filename in rsl.error*
do
Lrslerror=$(expr length $filename)
Lsuffix=$Lrslerror
let "Lsuffix -= 9"
Suffix=$(expr substr $filename 10 $Lsuffix)
mv $MM5_ROOT/TEMP/rsl.error.* $MM5_RUN/$directory/LOG/RSL.ERROR$suffix_"DOMAIN$
NEST
done
else
mv $MM5_ROOT/TEMP/MM5.LOG $MM5_RUN/$directory/LOG/MM5.LOG_DOMAIN$NEST
mv $MM5_ROOT/TEMP/fort.26 $MM5_RUN/$directory/LOG/MM5.STN_DOMAIN$NEST
fi
mv $MM5_ROOT/TEMP/MM5.TIME $MM5_RUN/$directory/LOG/MM5.TIME_DOMAIN$NEST
done

# Moving the run to the storage:
# mv $MM5_RUN/$directory $STORAGE

# Cleaning working directory:
rm -r $MM5_DIR
rm -r $MM5_ROOT/TEMP/*
fi

```

```

fi
cd $MM5_RUN
done

exit

```

G.3. VISUALIZATION SCRIPT

```

% MATLAB
% $Id: readmm5nc.m, v 3.0 2008/07/12 12:07:00 narss EGY $
%-----
% This script is used for MM5 outputs in NetCDF format from MM5toNetCDF code.
%-----
% HISTORY:
% Version          Date          Comment          Author
% 1.0              .          12/07/06        Initial          Hamada Sultan
% 2.0              12/07/07        Hamada Sultan
% 3.0              12/07/08        Hamada Sultan
%-----
% COPYRIGHT (C) National Authority for Remote Sensing and Space Sciences (NARSS)
%-----
clear all;
clc;

% User Modifications:
%-----
% Inputs:
tobs=[25 25 25 25 25 25 25 25 25];
Analysis='ANALYSIS_2006010412.nc';
Run1='MM5.RUN.2006010112.072HR.DA7.PHYSICS1.nc';
Run2='MM5.RUN.2006010112.072HR.DA0.PHYSICS1.nc';
TimeIndex=73; % Analysis time w.r.t. the model output

% Selected Area for Statistics:
iStart=1;
iEnd=174; % lmax_crs
jStart=1;
jEnd=174; % jmax_crs

% Open NetCDF file:
%-----
ncid = netcdf.open(Analysis,'NC_NOWRITE');
ncid1 = netcdf.open(Run1,'NC_NOWRITE');
ncid2 = netcdf.open(Run2,'NC_NOWRITE');

% Inquire the number of dimensions, variables, and global attributes
%-----
[ndims,nvars,ngatts,unlimdimid] = netcdf.inq(ncid);
for varid = 0:nvars-1
    [varname,xtype,dimids,natts] = netcdf.inqVar(ncid1,varid);
    varnamesize=size(varname);
    for i = 1:varnamesize(2)
        varnames(varid+1,i)=varname(i);
    end
end

```

end

% Get Simulation Time ID:

%-----

timeid = netcdf.inqVarID(ncid1,'time');

%Get Simulation Time:

%-----

time=netcdf.getVar(ncid1,timeid);

Nmax=numel(time);

% Get Horizontal Grid ID's:

%-----

longicrsid = netcdf.inqVarID(ncid1,'longicrs');

latitcrsid = netcdf.inqVarID(ncid1,'latitcrs');

longidotid = netcdf.inqVarID(ncid1,'longidot');

latitdotid = netcdf.inqVarID(ncid1,'latitdot');

terrainid = netcdf.inqVarID(ncid1,'terrain');

% Get Horizontal Grid:

%-----

lon_crs=netcdf.getVar(ncid1,longicrsid);

lat_crs=netcdf.getVar(ncid1,latitcrsid);

lon_dot=netcdf.getVar(ncid1,longidotid);

lat_dot=netcdf.getVar(ncid1,latitdotid);

terrain=netcdf.getVar(ncid1,terrainid);

[lmax_crs,jmax_crs] = size(lon_crs);

[lmax_dot,jmax_dot] = size(lon_dot);

% Get Vertical Grid ID's:

%-----

ptopid = netcdf.inqVarID(ncid1,'ptop');

base_slpid = netcdf.inqVarID(ncid1,'base_slp');

base_sltid=netcdf.inqVarID(ncid1,'base_slt');

ppid = netcdf.inqVarID(ncid1,'pp');

base_lrid=netcdf.inqVarID(ncid1,'base_lr');

sigma_levelid=netcdf.inqVarID(ncid1,'sigma_level');

sigma_level_fullid=netcdf.inqVarID(ncid1,'sigma_level_full');

% Get Vertical Grid:

%-----

ptop=netcdf.getVar(ncid1,ptopid);

base_slp=netcdf.getVar(ncid1,base_slpid);

base_slt=netcdf.getVar(ncid1,base_sltid);

pp=netcdf.getVar(ncid1,ppid);

base_lr=netcdf.getVar(ncid1,base_lrid);

sigma_level=netcdf.getVar(ncid1,sigma_levelid);

sigma_level_full=netcdf.getVar(ncid1,sigma_level_fullid);

Kmax=numel(sigma_level);

p_level=ptop+(base_slp-ptop)*sigma_level;

p_level_full=ptop+(base_slp-ptop)*sigma_level_full;

z_level=0+(base_slt/-6.5)*((p_level/base_slp).^(-6.5*8.31432/28.9644/9.80665)-1)*1000;

z=zeros(lmax_crs,jmax_crs,Kmax);

P00=base_slp;

Ts0=base_slt;

A=base_lr;

```

Ptop=ptop;
R=287;
g=9.8066;
Ps0=zeros(lmax_crs,jmax_crs);
P0=zeros(lmax_crs,jmax_crs,Kmax);
for i=1:lmax_crs
    for j=1:jmax_crs
        Ps0(i,j)=P00*exp(-Ts0/A+((-Ts0/A)^2-2*g*terrain(i,j)/A/R)^0.5)-Ptop;
        for k=1:Kmax
            % Old Method:
            z(i,j,k)=z_level(k)+terrain(i,j);
            % New Method:
            z(i,j,k)=z_level(k)+sigma_level(k)*terrain(i,j);
            % Correct Method (using INTERPF Tutorial):
            P0(i,j,k)=Ps0(i,j)*sigma_level(k)+Ptop;
            z(i,j,k)=-1*(R*A/2/g*(log(P0(i,j,k)/P00))^2+R*Ts0/g*log(P0(i,j,k)/P00));
        end
    end
end

% Get variable ID (e.g., for temperature):
%-----
tid = netcdf.inqVarID(ncid,'t');
t1id = netcdf.inqVarID(ncid1,'t');
t2id = netcdf.inqVarID(ncid2,'t');

% Get variable data (e.g., for temperature):
%-----
t=netcdf.getVar(ncid,tid)-273.15;
t1=netcdf.getVar(ncid1,t1id)-273.15;
t2=netcdf.getVar(ncid2,t2id)-273.15;

% Analysis and Statistics:
%-----
dt1=t(:,:)-t1(:,,:;TimeIndex);
dt2=t(:,:)-t2(:,,:;TimeIndex);
dt=t1(:,,:;TimeIndex)-t2(:,,:;TimeIndex);
% dt1=(t(:,,:)-t1(:,,:;TimeIndex))./t(:,,:)*100;
% dt2=(t(:,,:)-t2(:,,:;TimeIndex))./t(:,,:)*100;
% dt=(t1(:,,:;TimeIndex)-t2(:,,:;TimeIndex))./t1(:,,:;TimeIndex)*100;

fid = fopen(strcat('STATISTICS',Analysis(9:19),'.log'),'w');
for k=1:Kmax
    Layer=Kmax-k+1;

    tobsmean=mean(tobs);
    tmean=mean(mean(t(iStart:iEnd,jStart:jEnd,Layer)));
    t1mean=mean(mean(t1(iStart:iEnd,jStart:jEnd,Layer,TimeIndex)));
    t2mean=mean(mean(t2(iStart:iEnd,jStart:jEnd,Layer,TimeIndex)));
    dt1mean=mean(mean(dt1(iStart:iEnd,jStart:jEnd,Layer)));
    dt2mean=mean(mean(dt2(iStart:iEnd,jStart:jEnd,Layer)));
    dtmean=mean(mean(dt(iStart:iEnd,jStart:jEnd,Layer)));

    tobsmedian=median(tobs);
    tmedian=median(median(t(iStart:iEnd,jStart:jEnd,Layer)));

```

```

t1median=median(median(t1(iStart:iEnd,jStart:jEnd,Layer,TimeIndex)));
t2median=median(median(t2(iStart:iEnd,jStart:jEnd,Layer,TimeIndex)));
dt1median=median(median(dt1(iStart:iEnd,jStart:jEnd,Layer)));
dt2median=median(median(dt2(iStart:iEnd,jStart:jEnd,Layer)));
dtmedian=median(median(dt(iStart:iEnd,jStart:jEnd,Layer)));

tobsstd=std(tobs);
tstd=std(std(t(iStart:iEnd,jStart:jEnd,Layer)));
t1std=std(std(t1(iStart:iEnd,jStart:jEnd,Layer,TimeIndex)));
t2std=std(std(t2(iStart:iEnd,jStart:jEnd,Layer,TimeIndex)));
dt1std=std(std(dt1(iStart:iEnd,jStart:jEnd,Layer)));
dt2std=std(std(dt2(iStart:iEnd,jStart:jEnd,Layer)));
dtstd=std(std(dt(iStart:iEnd,jStart:jEnd,Layer)));

tobsmin=min(tobs);
tmin=min(min(t(iStart:iEnd,jStart:jEnd,Layer)));
t1min=min(min(t1(iStart:iEnd,jStart:jEnd,Layer,TimeIndex)));
t2min=min(min(t2(iStart:iEnd,jStart:jEnd,Layer,TimeIndex)));
dt1min=min(min(dt1(iStart:iEnd,jStart:jEnd,Layer)));
dt2min=min(min(dt2(iStart:iEnd,jStart:jEnd,Layer)));
dtmin=min(min(dt(iStart:iEnd,jStart:jEnd,Layer)));

tobsmax=max(tobs);
tmax=max(max(t(iStart:iEnd,jStart:jEnd,Layer)));
t1max=max(max(t1(iStart:iEnd,jStart:jEnd,Layer,TimeIndex)));
t2max=max(max(t2(iStart:iEnd,jStart:jEnd,Layer,TimeIndex)));
dt1max=max(max(dt1(iStart:iEnd,jStart:jEnd,Layer)));
dt2max=max(max(dt2(iStart:iEnd,jStart:jEnd,Layer)));
dtmax=max(max(dt(iStart:iEnd,jStart:jEnd,Layer)));

tobsvar=var(tobs);
tvar=var(var(t(iStart:iEnd,jStart:jEnd,Layer)));
t1var=var(var(t1(iStart:iEnd,jStart:jEnd,Layer,TimeIndex)));
t2var=var(var(t2(iStart:iEnd,jStart:jEnd,Layer,TimeIndex)));
dt1var=var(var(dt1(iStart:iEnd,jStart:jEnd,Layer)));
dt2var=var(var(dt2(iStart:iEnd,jStart:jEnd,Layer)));
dtvar=var(var(dt(iStart:iEnd,jStart:jEnd,Layer)));

tobsmode=mode(tobs);
tmode=mode(mode(t(iStart:iEnd,jStart:jEnd,Layer)));
t1mode=mode(mode(t1(iStart:iEnd,jStart:jEnd,Layer,TimeIndex)));
t2mode=mode(mode(t2(iStart:iEnd,jStart:jEnd,Layer,TimeIndex)));
dt1mode=mode(mode(dt1(iStart:iEnd,jStart:jEnd,Layer)));
dt2mode=mode(mode(dt2(iStart:iEnd,jStart:jEnd,Layer)));
dtmode=mode(mode(dt(iStart:iEnd,jStart:jEnd,Layer)));

fprintf(fid,'k =');
fprintf(fid,'%3.0f\n',Layer);
fprintf(fid,'\n');
fprintf(fid,'Variables: ');
% fprintf(fid,'\n');
fprintf(fid,'Tobs T T1 T2 dT1 dT2 dT');
fprintf(fid,'\n');
fprintf(fid,'Mean: ');
% fprintf(fid,'\n');
fprintf(fid,'%9.9f %9.9f %9.9f %9.9f %9.9f %9.9f %9.9f\n',...

```

```

        tobsmean,tmean,tlmean,t2mean,dtlmean,dt2mean,dtmean);
fprintf(fid,'Median: ');
% fprintf(fid,'\n');
fprintf(fid,'%9.9f %9.9f %9.9f %9.9f %9.9f %9.9f %9.9f\n',...
        tobsmedian,tmedian,tlmedian,t2median,dtlmedian,dt2median,dtmedian);
fprintf(fid,'Standard_Deviation: ');
% fprintf(fid,'\n');
fprintf(fid,'%9.9f %9.9f %9.9f %9.9f %9.9f %9.9f %9.9f\n',...
        tobsstd,tstd,tlstd,t2std,dtlstd,dt2std,dtstd);
fprintf(fid,'Minimum: ');
% fprintf(fid,'\n');
fprintf(fid,'%9.9f %9.9f %9.9f %9.9f %9.9f %9.9f %9.9f\n',...
        tobsmin,tmin,tlmin,t2min,dtlmin,dt2min,dtmin);
fprintf(fid,'Maximum: ');
% fprintf(fid,'\n');
fprintf(fid,'%9.9f %9.9f %9.9f %9.9f %9.9f %9.9f %9.9f\n',...
        tobsmax,tmax,tlmax,t2max,dtlmax,dt2max,dtmax);
fprintf(fid,'Variance: ');
% fprintf(fid,'\n');
fprintf(fid,'%9.9f %9.9f %9.9f %9.9f %9.9f %9.9f %9.9f\n',...
        tobsvar,tvar,tlvar,t2var,dtlvar,dt2var,dtvar);
fprintf(fid,'Mode: ');
% fprintf(fid,'\n');
fprintf(fid,'%9.9f %9.9f %9.9f %9.9f %9.9f %9.9f %9.9f\n',...
        tobsmode,tmode,tlmode,t2mode,dtlmode,dt2mode,dtmode);
fprintf(fid,'\n');
end
fclose(fid);

% Contour variable data (e.g., for temperature):
%-----
% for p=1:38
%   contour(lon_crs,lat_crs,t(:,39-p,1));
%   drawnow;
% end
% for n=1:73
%   contourf(lon_crs,lat_crs,t(:,38,n),50);
%   drawnow;
% end

% Close NetCDF File:
%-----
netcdf.close(ncid1);
netcdf.close(ncid2);

% Write to Tecplot:
%-----

fid = fopen('temp.dat','w');
fprintf(fid,'Title="Temperature Error in MM5 Output 2"');
fprintf(fid,'\n');
fprintf(fid,'Variables="Longitude","Latitude","Temperature","Temperature1","Temperature2","Temperature
Error1","Temperature Error2","Temperature Difference12"');
fprintf(fid,'\n');
for k=1:Kmax
    fprintf(fid,'Zone i = ');

```

```

fprintf(fid,'%3.0f',lmax_crs);
fprintf(fid,' j =');
fprintf(fid,'%3.0f',jmax_crs);
fprintf(fid,'\n');
for j=1:jmax_crs
    for i=1:lmax_crs
        fprintf(fid,'%9.9f %9.9f %9.9f %9.9f %9.9f %9.9f %9.9f %9.9f\n',...
            lon_crs(i,j),lat_crs(i,j),t(i,j,Kmax-k+1),t1(i,j,Kmax-k+1,TimeIndex),t2(i,j,Kmax-
k+1,TimeIndex),dt1(i,j,Kmax-k+1),dt2(i,j,Kmax-k+1),dt(i,j,Kmax-k+1));
        end
    end
    fprintf(fid,'TEXT X=3 Y=95 C=blue T="Date: ');
    fprintf(fid,Analysis(16:17));
    fprintf(fid,');
    fprintf(fid,Analysis(14:15));
    fprintf(fid,');
    fprintf(fid,Analysis(10:13));
    fprintf(fid,'');
    fprintf(fid,'ZN=');
    fprintf(fid,'%3.0f',k);
    fprintf(fid,'H=');
    fprintf(fid,'%3.0f',3);
    fprintf(fid,'TEXT X=72 Y=95 C=blue T="Time: ');
    fprintf(fid,Analysis(18:19));
    fprintf(fid,':00:00 UTC');
    fprintf(fid,'ZN=');
    fprintf(fid,'%3.0f',k);
    fprintf(fid,'H=');
    fprintf(fid,'%3.0f',3);
    fprintf(fid,'TEXT X=3 Y=3 C=blue T="Pressure Level: ');
    fprintf(fid,'%3.0f',p_level(Kmax-k+1)/100);
    fprintf(fid,' hPa');
    fprintf(fid,'ZN=');
    fprintf(fid,'%3.0f',k);
    fprintf(fid,'H=');
    fprintf(fid,'%3.0f',3);
    fprintf(fid,'TEXT X=50 Y=1 C=red T="National Authority for Remote Sensing and Space Sciences');
    fprintf(fid,'ZN=');
    fprintf(fid,'%3.0f',k);
    fprintf(fid,'H=');
    fprintf(fid,'%3.0f',1.5);
end
fclose(fid);
dos "C:\Program Files\Tecplot\Tec360 2009\bin\preplot.exe" "temp.dat";
movefile('temp.plt',strcat('TEMPERATURE',Analysis(9:19),'.plt'));
delete temp.dat

```


ملخص الرسالة

الهدف الرئيسي من هذه الرسالة هو تنقيح وتحسين نتائج النمذجة العددية للطقس لمصر حيث أنها كدولة نامية لها إحتياجات خاصة لدمج هذه النتائج بالعديد من المشاريع الوطنية. ومن الأمثلة على ذلك الخطط الطموحة للطاقة المتجددة لعام 2020 وما بعده، ومعالجة تلوث الهواء في المدن الكبرى في مصر، وإدارة أنشطة الزراعة والري، والإستغلال الأمثل للموارد، وغيرها. هذا وقد بدأ العمل في هذه الرسالة بتجميع وتحليل وإظهار بيانات الأقمار الصناعية المتعلقة بالمناخ ودراسة تأثيرات استخدام تقنيات تمثيل البيانات على نتائج النمذجة العددية للطقس لمصر. تناولت الدراسة تأثيرات الخيارات المختلفة لتقنيات تمثيل البيانات في فصول السنة الأربعة، وعند أماكن ذات كثافات مختلفة من المواقع الرصدية. وكان الإهتمام الرئيسي بدراسة التأثيرات على دقة النتائج في المناطق الفقيرة بالبيانات الرصدية، وكذلك في المواقع ذات التضاريس المعقدة. أظهرت النتائج تحسناً في دقة النموذج وارتباط أفضل بين درجات الحرارة المحسوبة والمرصودة، لا سيما بالقرب من سطح الأرض، كنتيجة لتمثيل البيانات. كما تأثرت دقة النتائج تأثراً سلبياً في المناطق الخالية من البيانات الرصدية، وبالقرب من السواحل، وفي المواقع ذات التضاريس المعقدة. إلا أنه من الممكن الحد من ذلك عن طريق زيادة عدد ونوعية محطات الأرصاد في هذه المواقع.

بعد ذلك، تمت دراسة التأثيرات الناتجة عن تغيير وقت وطريقة التهيئة والخيارات الفيزيائية المختلفة للنموذج العددي. وقد أظهرت النتائج أن النموذج العددي ذي حساسية للتهيئة، وكذلك الخيارات الفيزيائية المختلفة. وقد تمت الإستفادة من ذلك في إجراء تجربة أولية لاستخدام تقنية تنبؤ المجموعات لمواجهة عدم التأكد المرتبط بالشروط الإبتدائية أو بالنموذج العددي. في إطار ذلك تم استخدام تسع خيارات فيزيائية مختلفة بناءً على ثلاث مخططات للطبقة المتاخمة و ثلاث مخططات للإشعاع. وكان إجمالي عدد أعضاء مجموعة التنبؤات هو ستة وثلاثون تنبؤ. تم استخدام تسع تنبؤات مرجعية بينما تم الإستفادة من تقنية تمثيل البيانات في السبع وعشرين تنبؤ الآخرين عن طريق إنتاج ثلاث مجموعات مختلفة من الشروط الإبتدائية. تم تقييم نتائج هذه التجربة على أساس النسبة المئوية للفرق في الحد الأقصى والحد الأدنى والمتوسط في درجات الحرارة مقارنةً بالبيانات الرصدية. أشارت النتائج إلى أن متوسط مجموعة التنبؤات يكون دائماً أفضل وأكثر ارتباطاً مع البيانات الرصدية من التنبؤ الواحد. كما أن دقة التنبؤ تتناسب طردياً مع عدد ونوعية أعضاء مجموعة التنبؤات. وختاماً، فإن استخدام تقنيات تمثيل البيانات وتنبؤ المجموعات قد ساهم في تحقيق الهدف المرجو من هذه الرسالة وتحسين نتائج المحاكاة والتنبؤ بالطقس لمصر تمهيداً لدمجها بالمشاريع الوطنية المختلفة.

تقنيات تنبؤ المجموعات وتمثيل البيانات في النمذجة العددية للطقس لمصر

إعداد

م/ حماده سلطان سليمان بدر

رسالة مقدمة إلى كلية الهندسة - جامعة القاهرة
كجزء من متطلبات الحصول على درجة الماجستير في
هندسة الطيران والفضاء

يعتمد من لجنة الممتحنين:



أ.د. **عاطف عمر شريف**، مشرف الرسالة الرئيسي
الأستاذ بقسم هندسة الطيران والفضاء بجامعة القاهرة



أ.د. **جيمي دوديا**، عضواً
الأستاذ بالمركز القومي الأمريكي لبحوث الغلاف الجوي



أ.د. **محمد عز الدين الراعي**، عضواً
العميد السابق لمعهد الدراسات العليا والبحوث بجامعة الإسكندرية

كلية الهندسة، جامعة القاهرة
الجيزة، جمهورية مصر العربية

فبراير 2011

تقنيات تنبؤ المجموعات وتمثيل البيانات في النمذجة العددية للطقس لمصر

إعداد

م/ حماده سلطان سليمان بدر

رسالة مقدمة إلى كلية الهندسة - جامعة القاهرة
كجزء من متطلبات الحصول على درجة الماجستير في
هندسة الطيران والفضاء

تحت إشراف

أ. د. عاطف عمر شريف

أستاذ بقسم هندسة الطيران والفضاء، كلية الهندسة - جامعة القاهرة

.....
عاطف عمر شريف

أ. م. د. بسمان محمد نبيل الحديدي

أستاذ مساعد بقسم هندسة الطيران والفضاء، كلية الهندسة - جامعة القاهرة

.....
بسمان محمد نبيل الحديدي

أ. م. د. حمدي عبدالله قنديل

أستاذ مساعد بقسم الهندسة الميكانيكية، كلية الهندسة - جامعة الإسكندرية،
كلية الهندسة وعلم المواد - الجامعة الألمانية بالقاهرة

.....
حمدي عبدالله قنديل

كلية الهندسة، جامعة القاهرة
الجيزة، جمهورية مصر العربية

فبراير 2011

تقنيات تنبؤ المجموعات وتمثيل البيانات في النمذجة العددية للطقس لمصر

إعداد

م/ حماده سلطان سليمان بدر

رسالة مقدمة إلى كلية الهندسة - جامعة القاهرة
كجزء من متطلبات الحصول على درجة الماجستير في
هندسة الطيران والفضاء

كلية الهندسة، جامعة القاهرة
الجيزة، جمهورية مصر العربية

فبراير 2011