HAROKOPIO UNIVERSITY OF ATHENS
DEPARTMENT OF GEOGRAPHY

ATMOSPHERE AND CLIMATE DYNAMICS GROUP (ACDG)

http://meteoclima.gr

Data Assimilation and Nowcasting

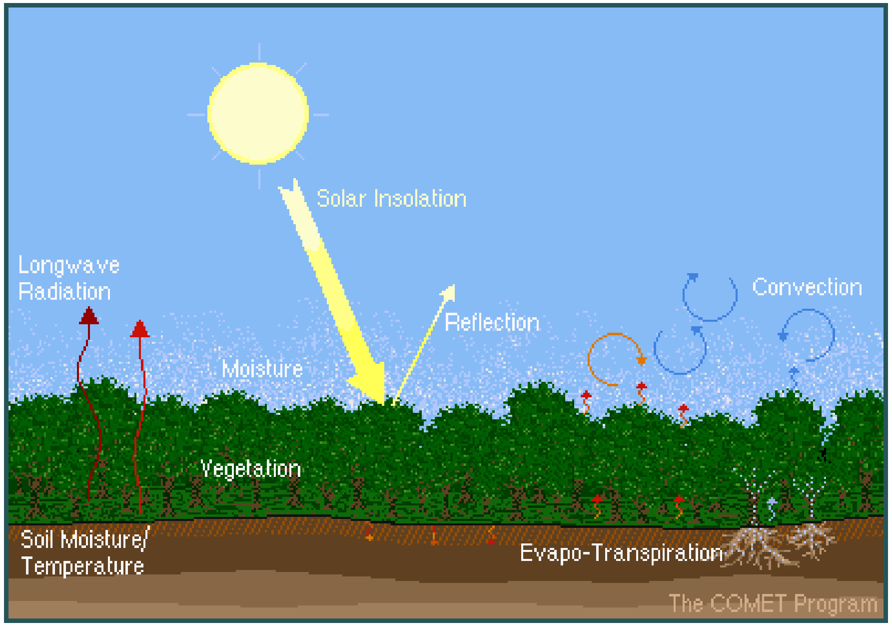
Petros Katsafados

pkatsaf@hua.gr

February 2022

Numerical Weather Prediction-NWP

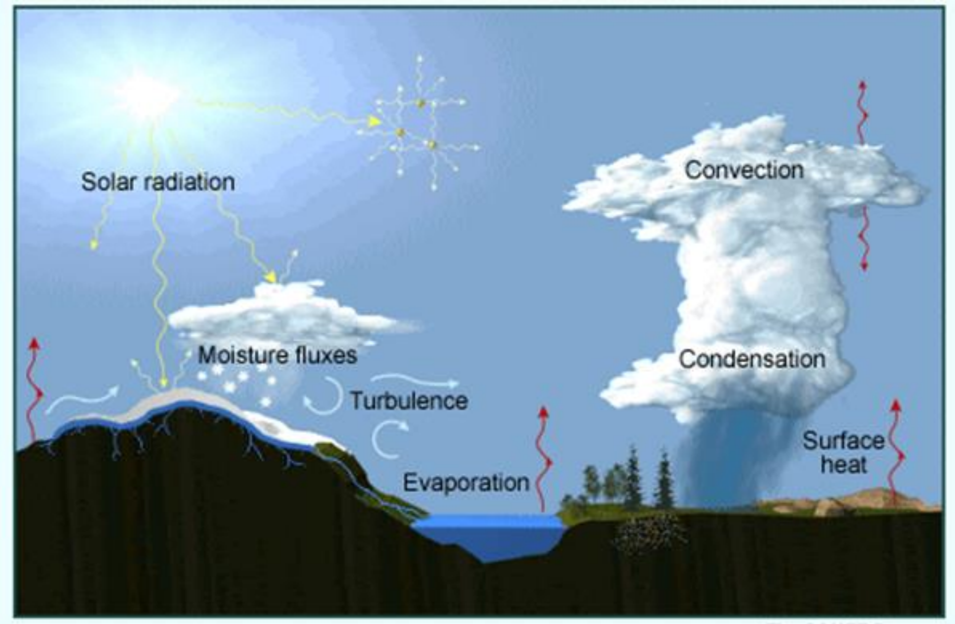
- The only way to predict the weather and the climate are the NWP models
- NWP models incorporate as many physically meaningful models of the atmospheric and surface processes as possible given computational constraints
- Coupled atmosphere/ocean/land models become even more sophisticated to incorporate the feedback between the atmosphere and underlying water surfaces



The source of this material is the Cooperative Program for Operational Meteorology, Education, and Training (COMET®) Website at http://meted.ucar.edu/ of the University Corporation for Atmospheric Research (UCAR) pursuant to a Cooperative Agreement with National Oceanic and Atmospheric Administration.

Atmospheric model definition

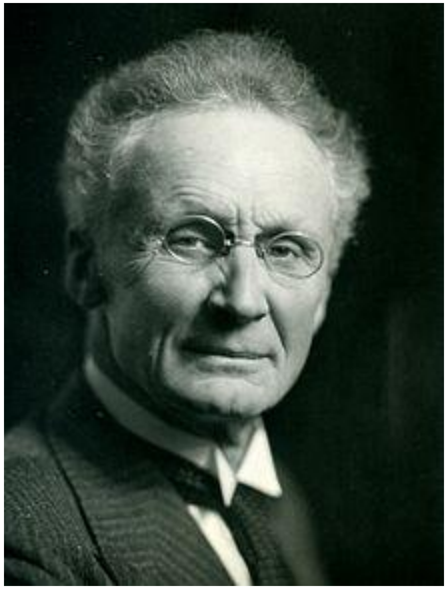
- Atmospheric models are set of partial differential equations defining the basic atmospheric variables (temperature, wind, geopotential, humidity, pressure)
- They describe the spatiotemporal state of the atmospheric conditions
- Atmospheric models incorporate as many physically meaningful models of surface processes as possible (given computational constraints) to make accurate forecasts



NWP principles

- Certain physical laws of motion and conservation of energy (for example, *Newton's Second Law of Motion* and the *First Law of Thermodynamics*) govern the evolution of the atmosphere
- Vilhelm Bjerknes first recognized that NWP was possible in principle in 1904
- He proposed that NWP could be seen as an initial value problem in mathematics:

⚠ since equations govern how meteorological variables change with time, if we know the initial condition of the atmosphere, we can solve the equations to obtain new values of those variables at a later time (i.e., make a forecast)



NWP fundamental concept

- To mathematically represent an NWP model in its simplest form, we write:

$$\frac{\Delta A}{\Delta t} = F(A)$$

- **ΔA** equals the change in a forecast variable at a particular point in space
- **Δt** equals the change in time (stands for the timestep)
- **F(A)** represents terms that can cause changes in the value of A

- In NWP, future values of meteorological variables are solved for by finding their initial values and then adding the physical forcing that acts on the variables over the time period of the forecast. This is stated as:

$$A^{forecast} = A^{initial} + F(A)\,\Delta t$$

- where F(A) stands for the combination of all of the kinds of forcing that can occur

NWP primitive equations

- The five equations govern changes in the motion and thermodynamics of the atmosphere and are derived from the set of conservation laws of momentum, mass, energy, and moisture
- This set of equations is considered to be closed and complete (meaning that we can forecast values of all terms by solving each of the equations in the proper sequence):
- All equations use the same basic forecast variables (u, v, ω, T, q, and z)

Wind Forecast Equations

1a. $$\frac{\partial u}{\partial t} = -u\frac{\partial u}{\partial x} - v\frac{\partial u}{\partial y} - \omega\frac{\partial u}{\partial p} + fv - g\frac{\partial z}{\partial x} + F_x$$

1b. $$\frac{\partial v}{\partial t} = -u\frac{\partial v}{\partial x} - v\frac{\partial v}{\partial y} - \omega\frac{\partial v}{\partial p} - fu - g\frac{\partial z}{\partial y} + F_y$$

Continuity Equation

2. $$\frac{\partial u}{\partial x} + \frac{\partial v}{\partial y} + \frac{\partial \omega}{\partial p} = 0$$

Temperature Forecast Equation

3. $$\frac{\partial T}{\partial t} = -u\frac{\partial T}{\partial x} - v\frac{\partial T}{\partial y} - \omega\left(\frac{\partial T}{\partial p} - \frac{RT}{c_p p}\right) + \frac{H}{c_p}$$

Moisture Forecast Equation

4. $$\frac{\partial q}{\partial t} = -u\frac{\partial q}{\partial x} - v\frac{\partial q}{\partial y} - \omega\frac{\partial q}{\partial p} + E - P$$

Hydrostatic Equation

5. $$\frac{\partial z}{\partial p} = -\frac{RT}{pg}$$

Finite differences approach

- Let the numerical solution of the moisture equation
- It only includes 1-D moisture advection terms
- It is a simplified form of the Eq. (4)
- Its numerical solution is based on the finite differences method
- How do the terms of this equation represent on a 3-D grid?

Derivative form of simplified moisture equation

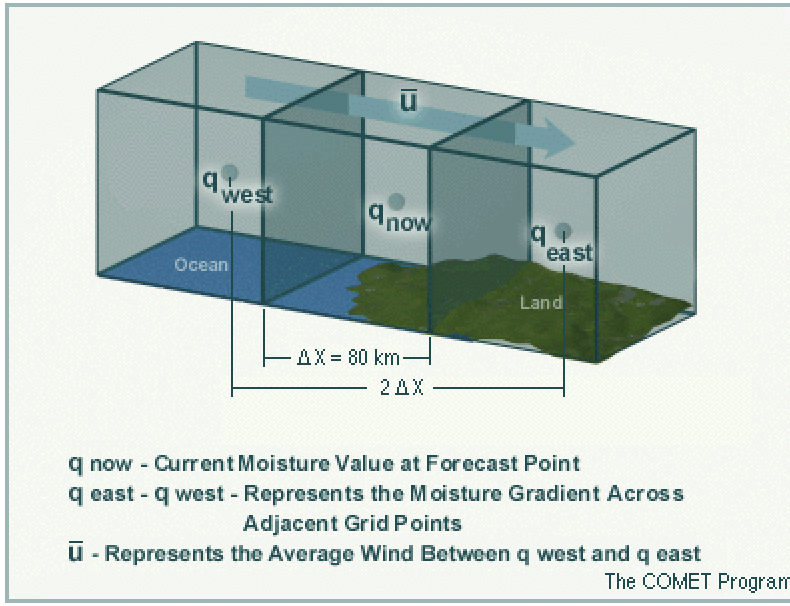
$$\frac{\partial q}{\partial t} = -\bar{u}\frac{\partial q}{\partial x}$$

Finite difference form of simplified moisture equation

$$\frac{\left(q^{t+1} - q^{t}\right)_{x,y}}{\Delta t} = -\bar{u}\,\frac{q^{t}_{x+1,y} - q^{t}_{x-1,y}}{2\Delta x}$$

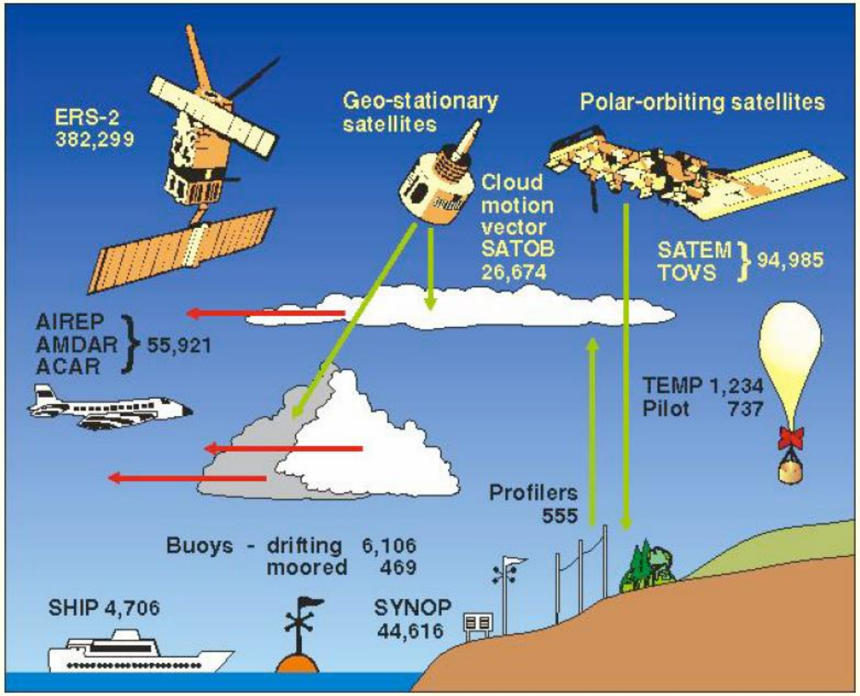
Written more conceptually

$$q^{forecast} = q^{now} - \bar{u}\,\frac{\Delta t}{2\Delta x}\left(q_{east} - q_{west}\right)^{now}$$



Data Assimilation (DA)

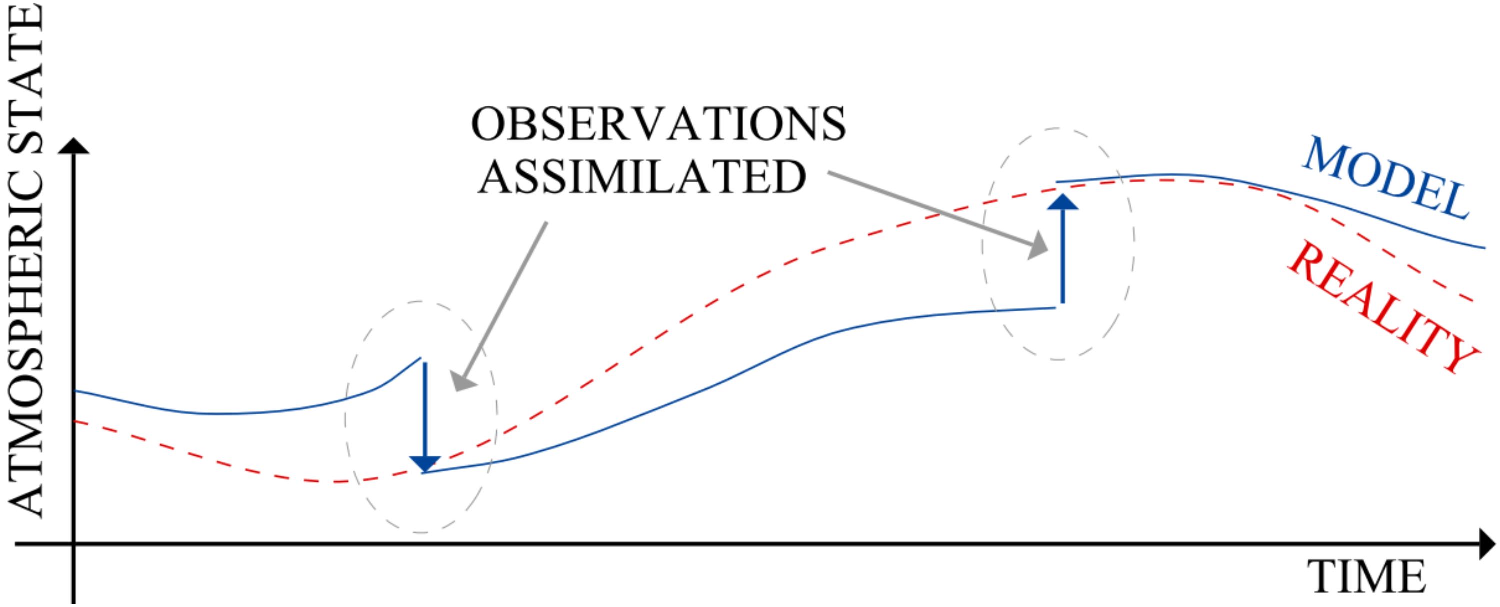
- NWP is an initial/boundary value problem
- Given an estimate of the present state of the atmosphere (initial conditions) appropriate surface and lateral boundary conditions the model simulates or forecasts the evolution of the atmosphere
- The more accurate the estimate of the initial conditions, the better the quality of the forecasts
- This approach is called data assimilation



The DA concept

- ...[the atmosphere] "is a chaotic system in which errors introduced into the system can grow with time... As a consequence, data assimilation is a struggle between chaotic destruction of knowledge and its restoration by new observations"

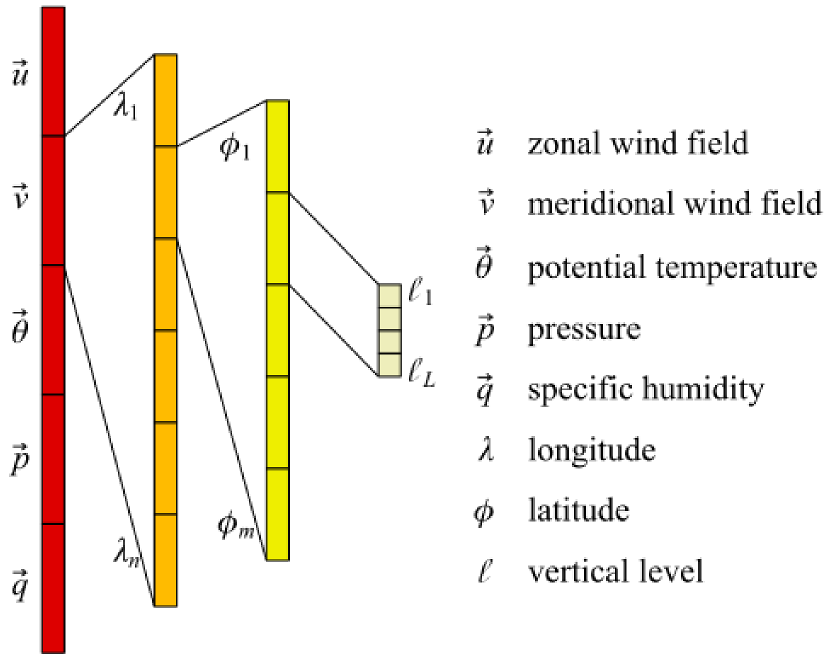
Leith (1993)

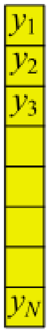


The DA problem

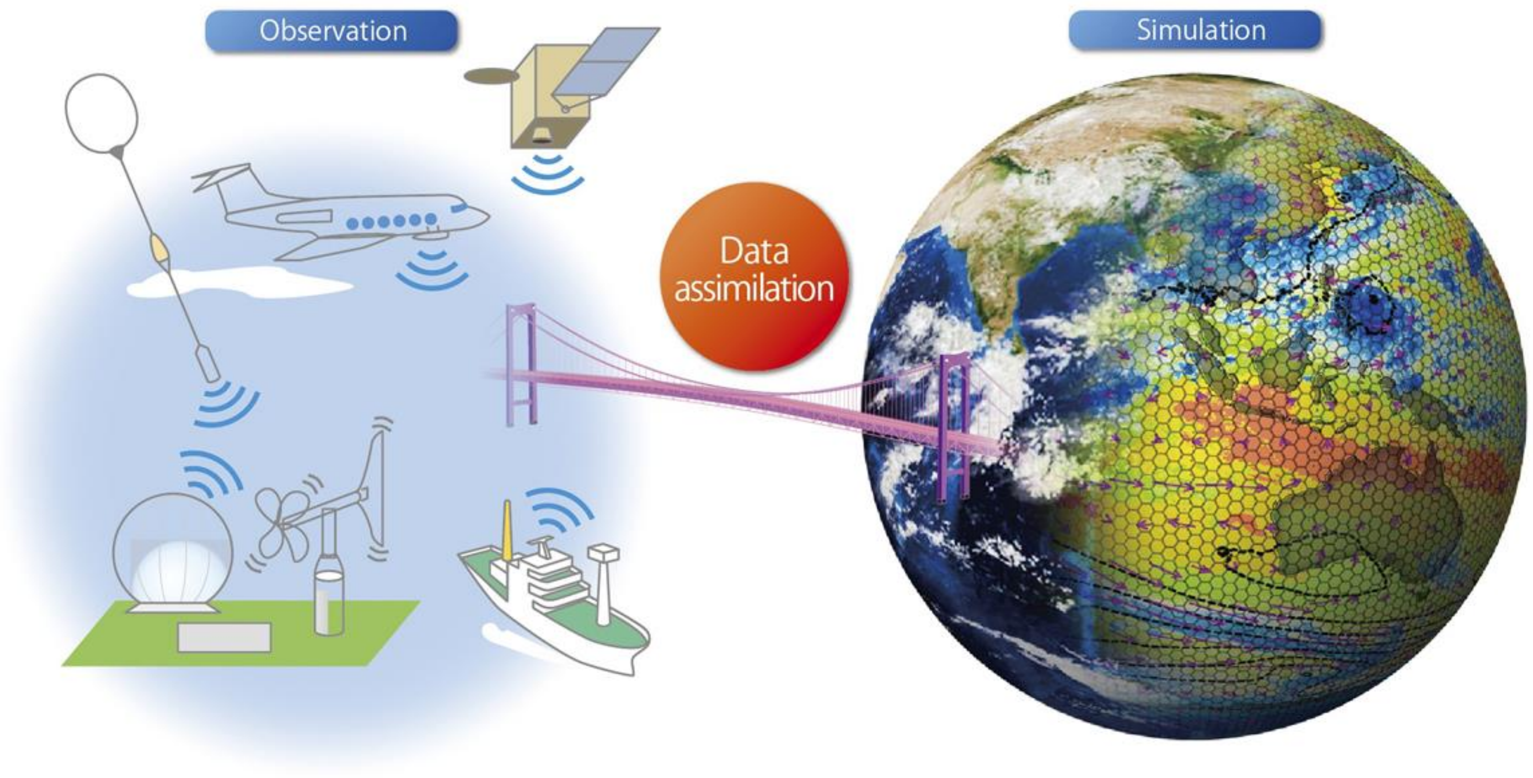
- NWP models: elements > 10^7 (5 × n × m × L)
- No. of obs. << No. of (unknown) elements in $\vec{x}$
- Need to fill-in the missing information with prior knowledge

The state vector $\vec{x}$ The observation vector $\vec{y}$





Variety data combination



Data assimilation serves as a bridge between real-world observations and computer simulations. The accuracy of weather forecasting is greatly improved when a variety of observation data is used to help produce more realistic simulations by means of numerical prediction models.

DA methods

☞Variational analysis. It is based on optimal control theory and the aim is to minimize a given cost function that measures the model-to-data misfit

- Least Squared method
- Optimal Interpolation (OI)
- Three-Dimensional Variational assimilation (3D-Var)
- Four-Dimensional Variational assimilation methods (4D-Var)

☞Sequential methods. The observations are assimilated as soon as they become available

- Kalman Filter (Kalman Filter), Extended Kalman Filter and Ensemble Kalman Filter

How it works…

- We want to measure the temperature in a room, and we have two thermometers that measure with errors

$$T_1 = T_t + \varepsilon_1;\ T_2 = T_t + \varepsilon_2$$

- We assume that the errors are unbiased $\overline{\varepsilon_1} = \overline{\varepsilon_2} = 0$
- Variances $\overline{\varepsilon_1^2} = \sigma_1^2;\ \overline{\varepsilon_2^2}=\sigma_2^2$
- The errors of the two thermometers are uncorrelated $\overline{\varepsilon_1 \varepsilon_2} = 0$
- The question is: how can we estimate the true temperature (T_t) optimally? We call this optimal estimate the "analysis of the temperature (T_a)"

How it works…

- We try to estimate the analysis from a linear combination of the observations $T_a = a_1 T_1 + a_2 T_2$
- Assume that the analysis errors are unbiased $\overline{T_a} = \overline{T_t}$
- This implies that $a_1 + a_2 = 1$
- T_α will be the best estimate of T_t if the coefficients α_1, α_2 are chosen to minimize the MSE of T_α
- $\sigma_\alpha^2 = \overline{(T_a - T_t)^2} = \overline{[a_1(T_1 - T_t) + (1 - a_1)(T_2 - T_t)]^2}$
- Minimizing a function $\frac{\partial \sigma_a^2}{\partial a_1} = 0$ (Fermat)

How it works…

- Minimizing $\frac{\partial \sigma_a^2}{\partial a_1} = 0 \Rightarrow a_1 = \frac{\sigma_2^2}{\sigma_1^2+\sigma_2^2}; a_2 = \frac{\sigma_1^2}{\sigma_1^2+\sigma_2^2}$
- $a_1 = \frac{\frac{1}{\sigma_1^2}}{\frac{1}{\sigma_1^2}+\frac{1}{\sigma_2^2}}; a_2 = \frac{\frac{1}{\sigma_2^2}}{\frac{1}{\sigma_1^2}+\frac{1}{\sigma_2^2}}$
- In the first formula the weight of obs 1 is given by the variance of obs 2 divided by the total error
- In the second formula the weights of the observations are proportional to the "precision" or accuracy of the measurements (defined as the inverse of the variances of the observational errors)

How it works…

- Two measurements and an optimal linear combination (analysis)

$$T_a = a_1 T_1 + a_2 T_2$$

- Since $a_1 + a_2 = 1$

$$T_a = T_1 + a_2 (T_2 - T_1)$$

- Optimal coefficients (min σ_a^2) $a_1 = \frac{\sigma_2^2}{\sigma_1^2+\sigma_2^2}$; $a_2 = \frac{\sigma_1^2}{\sigma_1^2+\sigma_2^2}$

$$a_1 = \frac{\frac{1}{\sigma_1^2}}{\frac{1}{\sigma_1^2} + \frac{1}{\sigma_2^2}}; a_2 = \frac{\frac{1}{\sigma_2^2}}{\frac{1}{\sigma_1^2} + \frac{1}{\sigma_2^2}}$$

- Replacing, we get

$$\sigma_a^2 = \frac{\sigma_1^2 \sigma_2^2}{\sigma_1^2+\sigma_2^2}, \text{or } \frac{1}{\sigma_a^2} = \frac{1}{\sigma_1^2} + \frac{1}{\sigma_2^2}$$

How it works…

- Assume that T_1=T_b (forecast) and T_2=T_0 (observation)

$$T_a = a_1 T_b + a_2 T_0 = T_b + a_2(T_0 - T_b)$$

- A forecast and an observation optimally combined (analysis)

$$T_a = T_b + \frac{\sigma_b^2}{\sigma_b^2 + \sigma_0^2}(T_0 - T_b) = T_b + w\,(T_0 - T_b)$$

$$\frac{1}{\sigma_\alpha^2} = \frac{1}{\sigma_1^2} + \frac{1}{\sigma_2^2}$$

- If the statistics of the errors are exact, and if the coefficients are optimal, then the "precision" of the analysis (defined as the inverse of the variance) is the sum of the precisions of the measurements
- *Least squared method*

Summarizing the theoretical background

$$T_a = T_b + \frac{\sigma_b^2}{\sigma_b^2 + \sigma_0^2}(T_0 - T_b) = T_b + w\,[T_0 - h(T_b)]$$

- Analysis = background + optimal weight x (observational increment)
- Observation operator: model variables => observed variables
- Optimal weight = background error variance/total error variance
- Precision = 1/error variance
- Analysis precision=background precision + obs precision
- We assume no bias, no error correlation

The variational approach

- We have a forecast T_b (prior) and a radiance obs $y_0 = h(T_t) + \varepsilon_0$
- The new information (or innovation) is the observational increment $y_0 - h(T_b)$
- From a 3D-Var point of view, we want to find a T_a that minimizes the cost function J: $J(T_a) = \frac{(T_a - T_b)^2}{2\sigma_b^2} + \frac{(h(T_a) - y_0)^2}{2\sigma_0^2}$
- This analysis temperature T_a is closest to both the forecast T_b and the observation y_o and maximizes the likelihood of $T_a \sim T_t$ given the information we have
- It is easier to find the analysis increment $T_a - T_b$ that minimizes the cost function J

3DVar/4DVar

- 3D-Var based on the 3D cost function minimization

$$J_{3DVar} = \min \frac{1}{2}\left[\left(x^a - x^b\right)^T B^{-1}\left(x^a - x^b\right) + (Hx^a - y)^T R^{-1}(Hx^a - y)\right]$$

Distance to background

Distance to observations

  - *B* is the model error covariance matrix; *R* is the observational error matrix (diagonal);
  - x^a is the control variable; x^b is the background field;
  - *H* is the observation operator (nonlinear); *y* is the observation

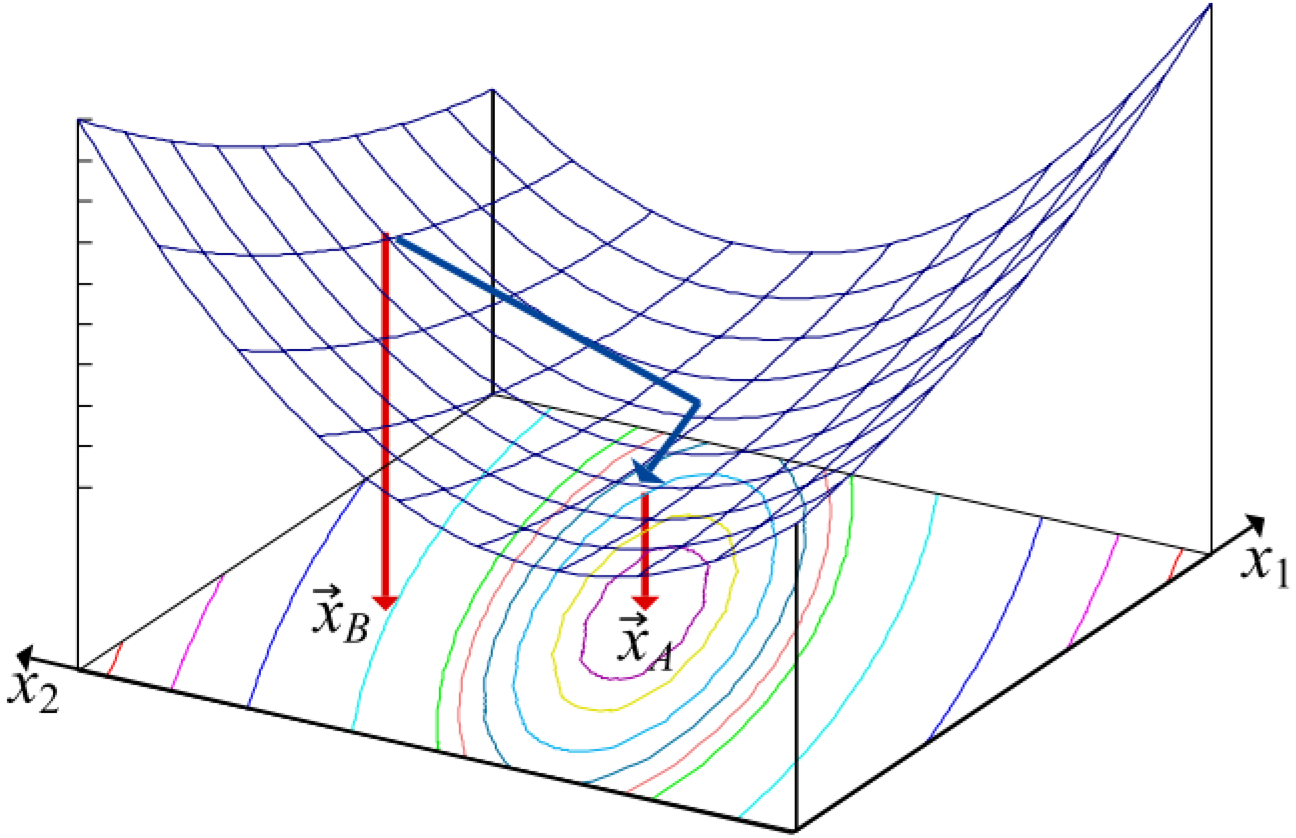
$$J_{4DVar} = \min \frac{1}{2}\left[(x_0 - x_b)^T B^{-1}(x_0 - x_b) + \sum_{i=1}^{S}(Hx_i - y_i)^T R_i^{-1}(Hx_i - y_i)\right]$$

Distance to background at the initial time

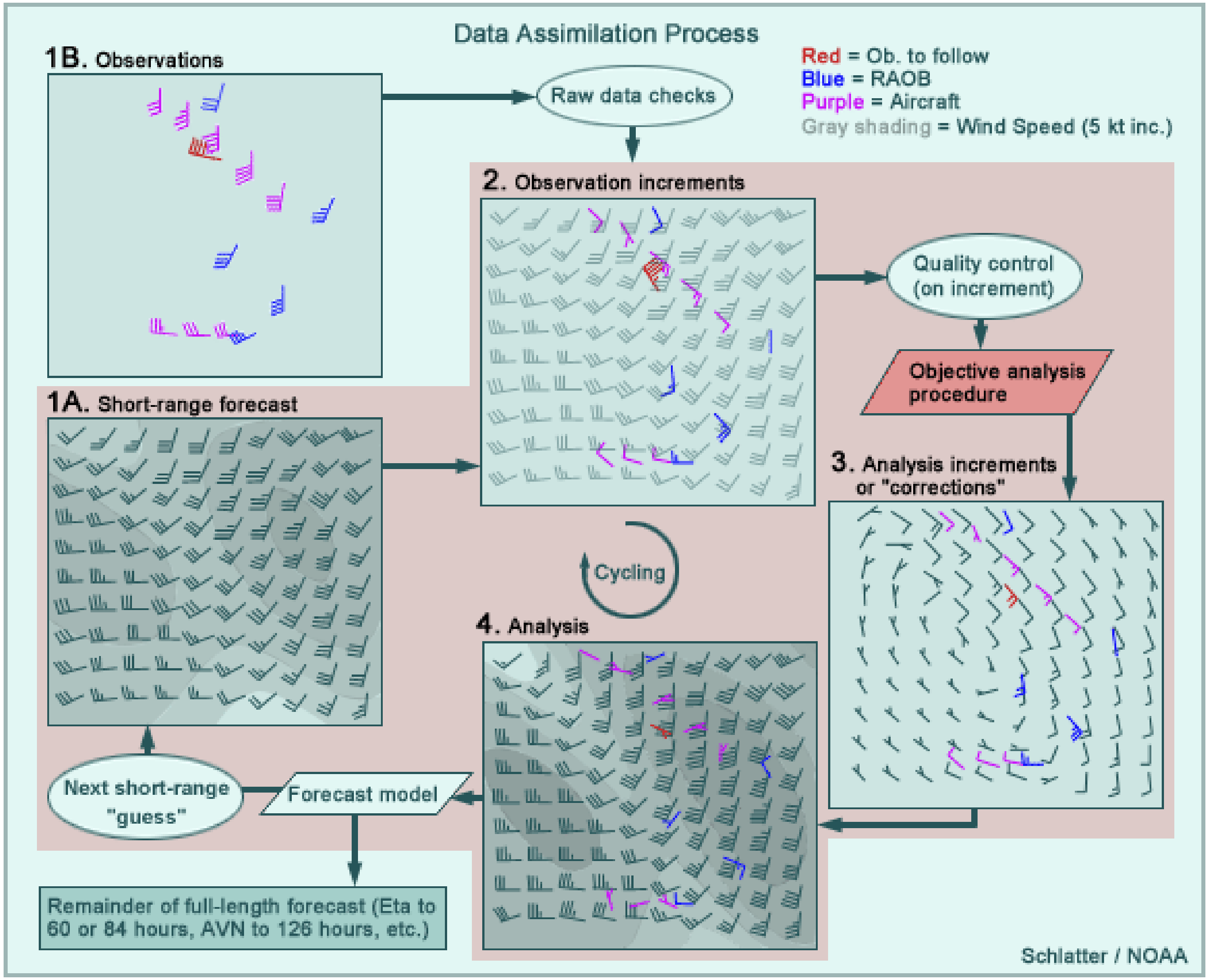
Distance to observations in a time window interval t_0-t_1

Minimizing the cost function

- The problem reduces to an optimization problem in 10^7 dimensional phase space

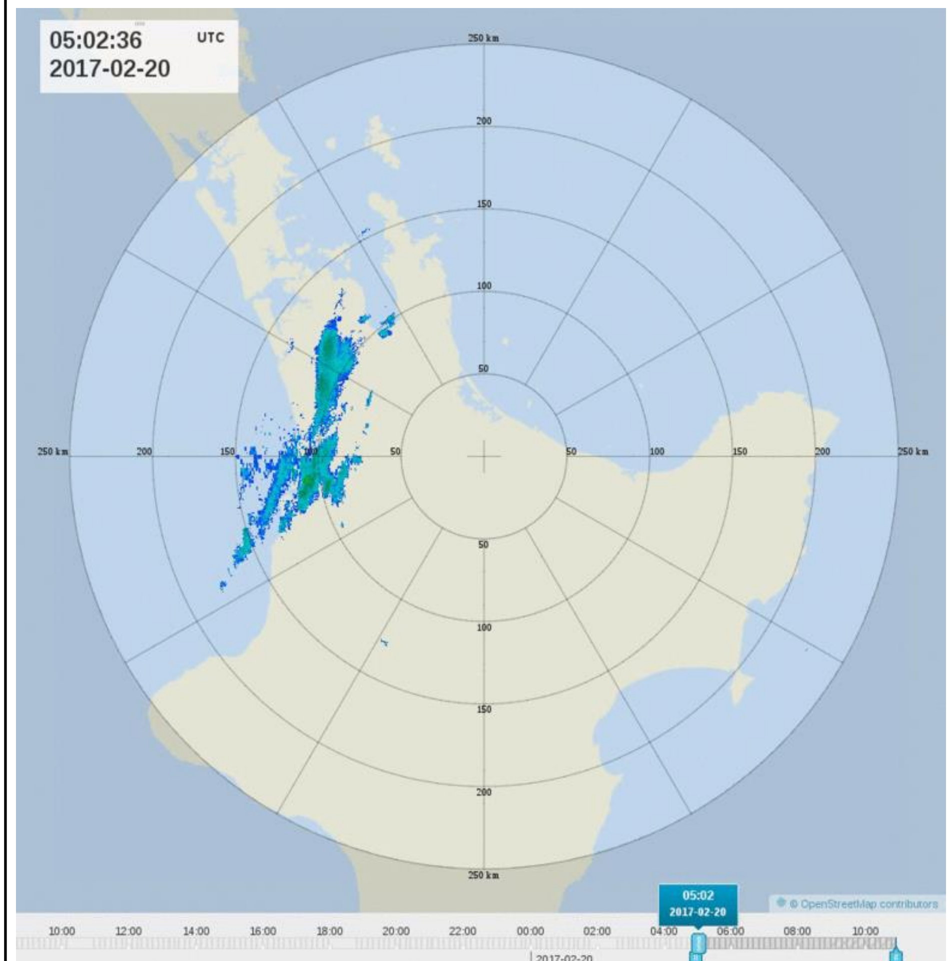


Schematic representation of gridded analyses

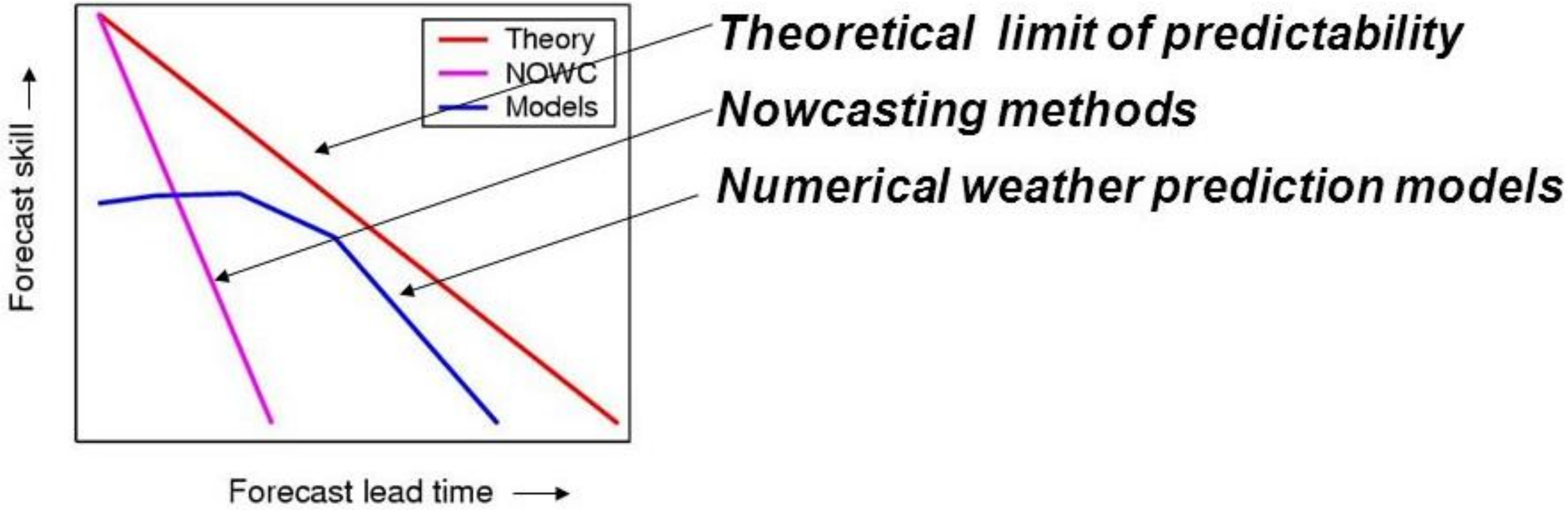


Nowcasting

- Nowcasting is a form of very short-range weather forecasting (usually up to 6 hours) based on very fined observational data (Kalnay, 2003)
- It usually runs every 1 hour (analysis cycle) or even shorter and provides 3hr forecasts on very high resolution (eg. 1km)



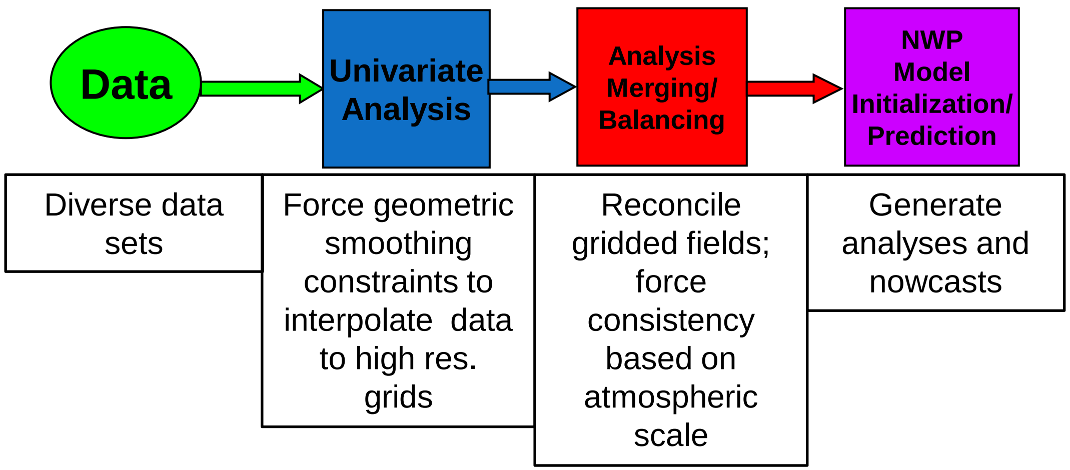
Nowcasting vs NWP forecast skill



Golding (1998), Austin et al. (1987), Wilson et al. (1998)

Basic components

- Data Acquisition and Quality Control



- Univariate Analysis of the
  - Temperature
  - Winds
  - Water Vapor
  - Clouds
    - Microphysical variables
    - Vertical motions

Local Analysis and Prediction System (LAPS)

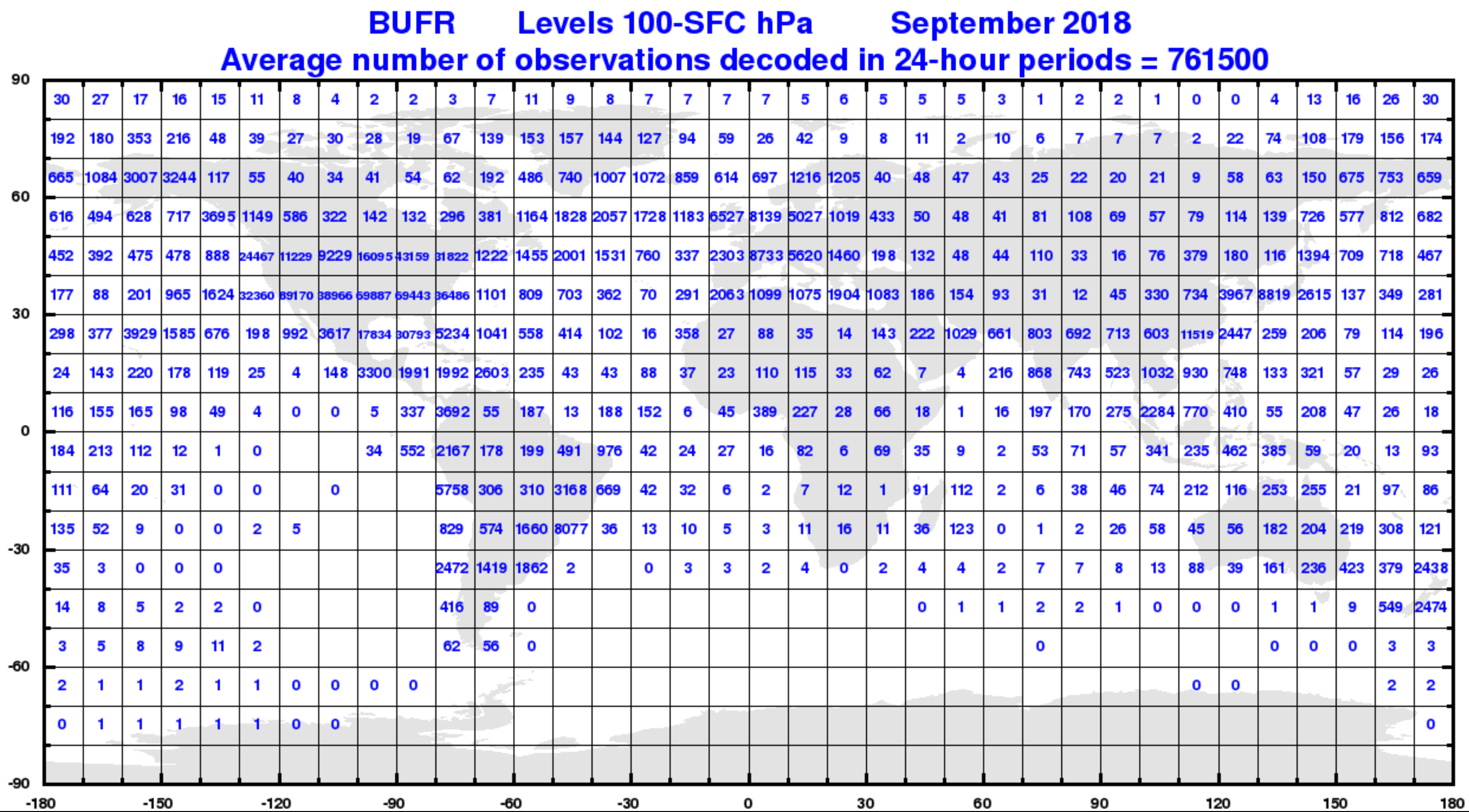
- ✓ NOAA’s Local Analysis and Prediction System (LAPS) is an advanced mesoscale meteorological data assimilation tool designed to exploit all available data sources (local and global) and produce analyzed and guessed grids (*Albers*, 1995)
- ✓ LAPS incorporates a number of surface and upper air observations (METAR/SYNOP, satellite, soundings, radar, etc) to produce high spatial and temporal resolution analysis fields
- ✓ It is based on the traditional objective analysis scheme using the cost function approach in order to minimize differences between the analysis and the observational fields
- ✓ Currently, LAPS has been ported in various European institutes and universities to support operational and research activities

Common data formats

Formats	Description
NetCDF	Network Common Data Form is a set of software libraries and machine-independent data formats that support the creation, access, and sharing of array-oriented scientific data.
GRIB1/2	GRIdded Binary, the standard format for the storage and interchange of meteorological data maintained by Maintainer: World Meteorological Organization (WMO)
HDF	Hierarchical Data Format, a library and multi-object file format for the transfer of graphical and numerical data between machines.

The Binary Universal Form for the Representation of meteorological data (BUFR)

- ✓ BUFR was created in 1988 with the goal of replacing the WMO's dozens of character-based, position-driven meteorological codes, such as SYNOP (surface observations), TEMP (upper air soundings) and CLIMAT (monthly climatological data)



30	27	17	16	15	11	8	4	2	2	3	7	11	9	8	7	7	7	7	5	6	5	5	5	3	1	2	2	1	0	0	4	13	16	26	30
192	180	353	216	48	39	27	30	28	19	67	139	153	157	144	127	94	59	26	42	9	8	11	2	10	6	7	7	7	2	22	74	108	179	156	174
665	1084	3007	3244	117	55	40	34	41	54	62	192	486	740	1007	1072	859	614	697	1216	1205	40	48	47	43	25	22	20	21	9	58	63	150	675	753	659
616	494	628	717	3695	1149	586	322	142	132	296	381	1164	1828	2057	1728	1183	6527	8139	5027	1019	433	50	48	41	81	108	69	57	79	114	139	726	577	812	682
452	392	475	478	888	24467	11229	9229	16095	43159	31822	1222	1455	2001	1531	760	337	2303	8733	5620	1460	198	132	48	44	110	33	16	76	379	180	116	1394	709	718	467
177	88	201	965	1624	32360	89170	38966	69887	69443	36486	1101	809	703	362	70	291	2063	1099	1075	1904	1083	186	154	93	31	12	45	330	734	3967	8819	2615	137	349	281
298	377	3929	1585	676	198	992	3617	17834	30793	5234	1041	558	414	102	16	358	27	88	35	14	143	222	1029	661	803	692	713	603	11519	2447	259	206	79	114	196
24	143	220	178	119	25	4	148	3300	1991	1992	2603	235	43	43	88	37	23	110	115	33	62	7	4	216	868	743	523	1032	930	748	133	321	57	29	26
116	155	165	98	49	4	0	0	5	337	3692	55	187	13	188	152	6	45	389	227	28	66	18	1	16	197	170	275	2284	770	410	55	208	47	26	18
184	213	112	12	1	0			34	552	2167	178	199	491	976	42	24	27	16	82	6	69	35	9	2	53	71	57	341	235	462	385	59	20	13	93
111	64	20	31	0	0		0			5758	306	310	3168	669	42	32	6	2	7	12	1	91	112	2	6	38	46	74	212	116	253	255	21	97	86
135	52	9	0	0	2	5				829	574	1660	8077	36	13	10	5	3	11	16	11	36	123	0	1	2	26	58	45	56	182	204	219	308	121
35	3	0	0	0						2472	1419	1862	2		0	3	3	2	4	0	2	4	4	2	7	7	8	13	88	39	161	236	423	379	2438
14	8	5	2	2	0					416	89	0										0	1	1	2	2	1	0	0	0	1	1	9	549	2474
3	5	8	9	11	2					62	56	0													0						0	0	0	3	3
2	1	1	2	1	1	0	0	0	0																				0	0				2	2
0	1	1	1	1	1	0	0																												0

LAPS nowcasting development and evaluation

- A simplified forward advection scheme has been embedded in LAPS in order to overcome the 'spin up' period which routinely appears in the conventional numerical weather predictions (NWP)
- This way LAPS is able to advance the meteorological parameters in time and provide an estimation of future conditions 3 hours ahead from the analysis hour
- In order to examine the capabilities of the system against conventional NWP, the WRF model with the NMM core was run in operational mode for the same time period
- Using the model output for the same forecast window (1-3 hrs), which is part of the model 'spin-up' time, WRF-NMM was evaluated having as reference the same station data

LAPS advection scheme

- Simple first order advection scheme in 2 dimensions (x,y)
- Solving the advection equation for various meteorological parameters (Temperature, Precipitation, Humidity)
- $\frac{\partial u}{\partial t} + CU\frac{\partial u}{\partial x} = 0$ and $\frac{\partial v}{\partial t} + CV\frac{\partial v}{\partial y} = 0$ (1)
- C is the Courant number defined as $CU = u\frac{\Delta t}{\Delta x}$ and $CV = v\frac{\Delta t}{\Delta y}$
- Equations (1) can be written as $\frac{{u_i}^{t+1}-u^t}{\Delta t}+\frac{{u_{i+1}}^{t+1}-u^t}{\Delta x} = 0$ (2)
- First the Courant number is calculated in order to see if the time-step selected assures the stability of the formulation
- By solving equation (2) we can find the value of the parameter at time t+1 using information at time t

LAPS nowcasting evaluation

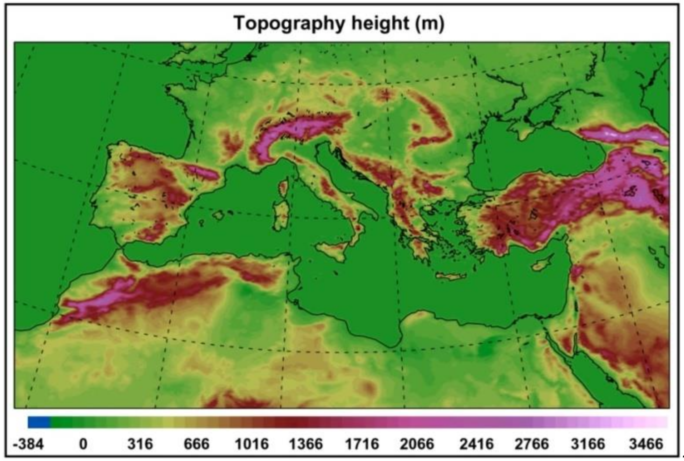
- The performance assessment of the system has as reference the surface measurements available from the WMO network
- The available surface observations from more than 500 conventional stations were used to verify and compare categorical near to analysis forecasts for one month period
- Furthermore in order to examine the capabilities of the system against conventional NWP, the WRF model with the NMM core was run in operational mode for the same time period
- Using the model output for the same forecast window (1-3 hrs), which is part of the model 'spin-up' time, WRF-NMM was evaluated having as reference the same station data

- Spyrou C. V.M. Nomikou, A. Papadopoulos and P. Katsafados, 2017: “LAPS nowcasting – Development and evaluation”. 2nd European Nowcasting Conference (ENC), 03-05 April 2017, DWD, Offenbach, Germany

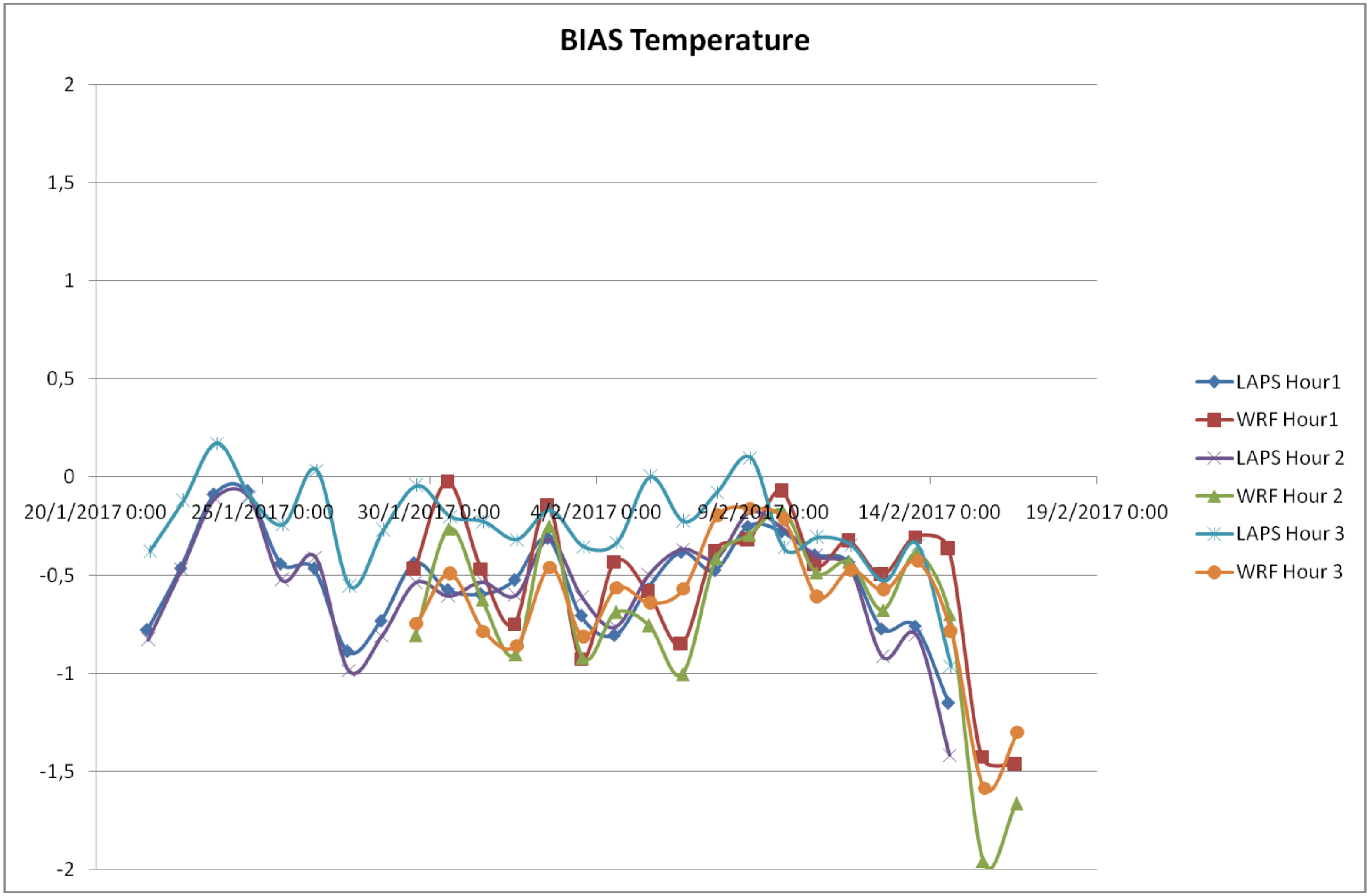
LAPS-WRF configuration

- **LAPS-046-13**
- Horizontal resolution: 15x15km (550x425 grid points), 43 vertical levels
- Background fields: GFS 0.25°x0.25°, 26 isobaric levels up to 50 mb, assimilated cycles 00, 06, 12, 18UTC
  (*NCEP is acknowledged for the data provision*)
- Assimilated data: METAR/SYNOP and RAOB, EUMETSAT Multi-Sensor Precipitation Estimate (MPE), NOAA's QMORPH product
- **WRF-NMM (ver3.0)**
- Horizontal resolution: 0.09°x0.09° (305x273 grid points), 38 vertical levels
- IC & BC: GFS 0.5°x0.5°, 26 isobaric levels up to 50 mb, cycle 12UTC

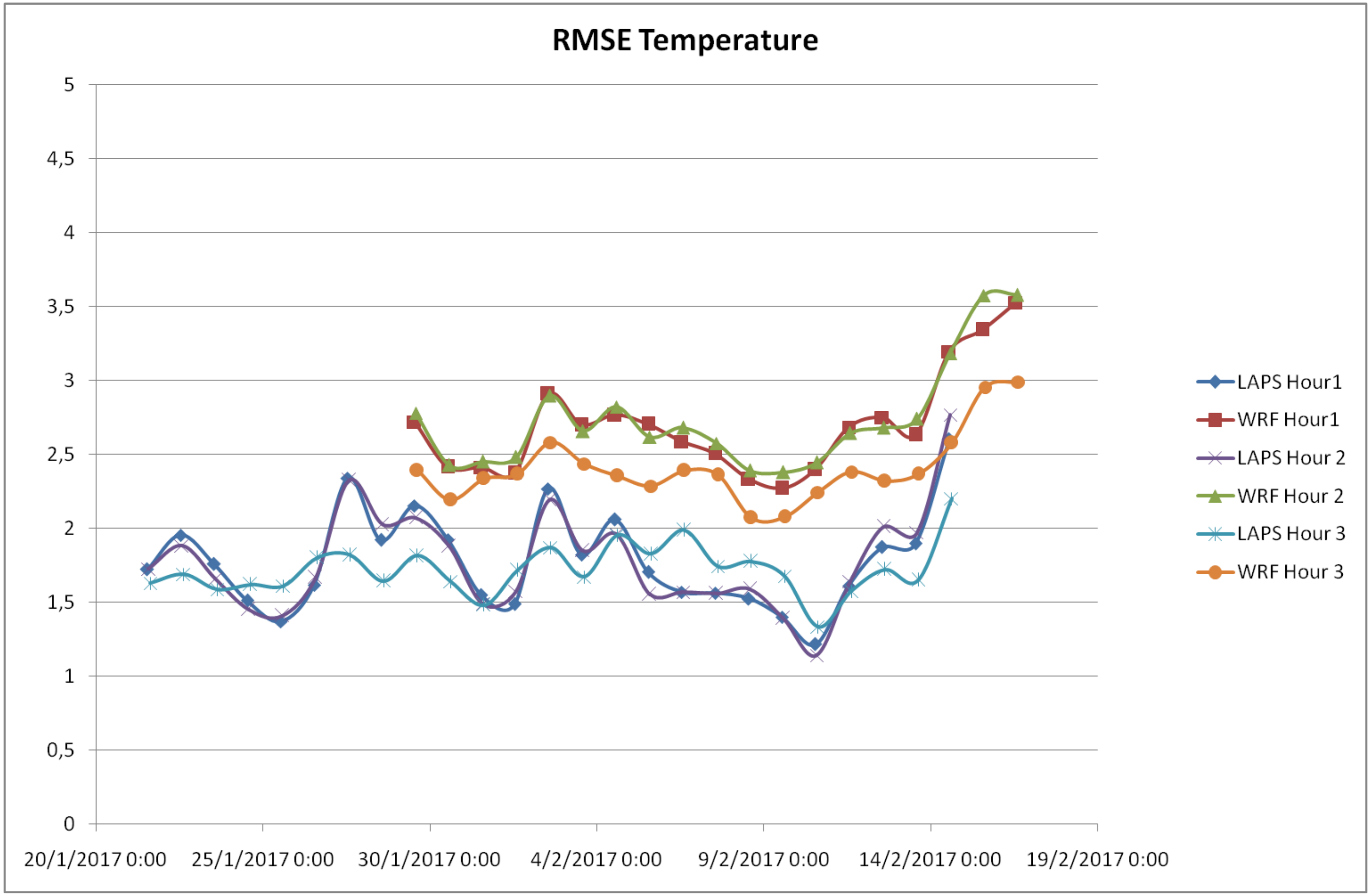
Evaluation period 29/1-16/2/2017



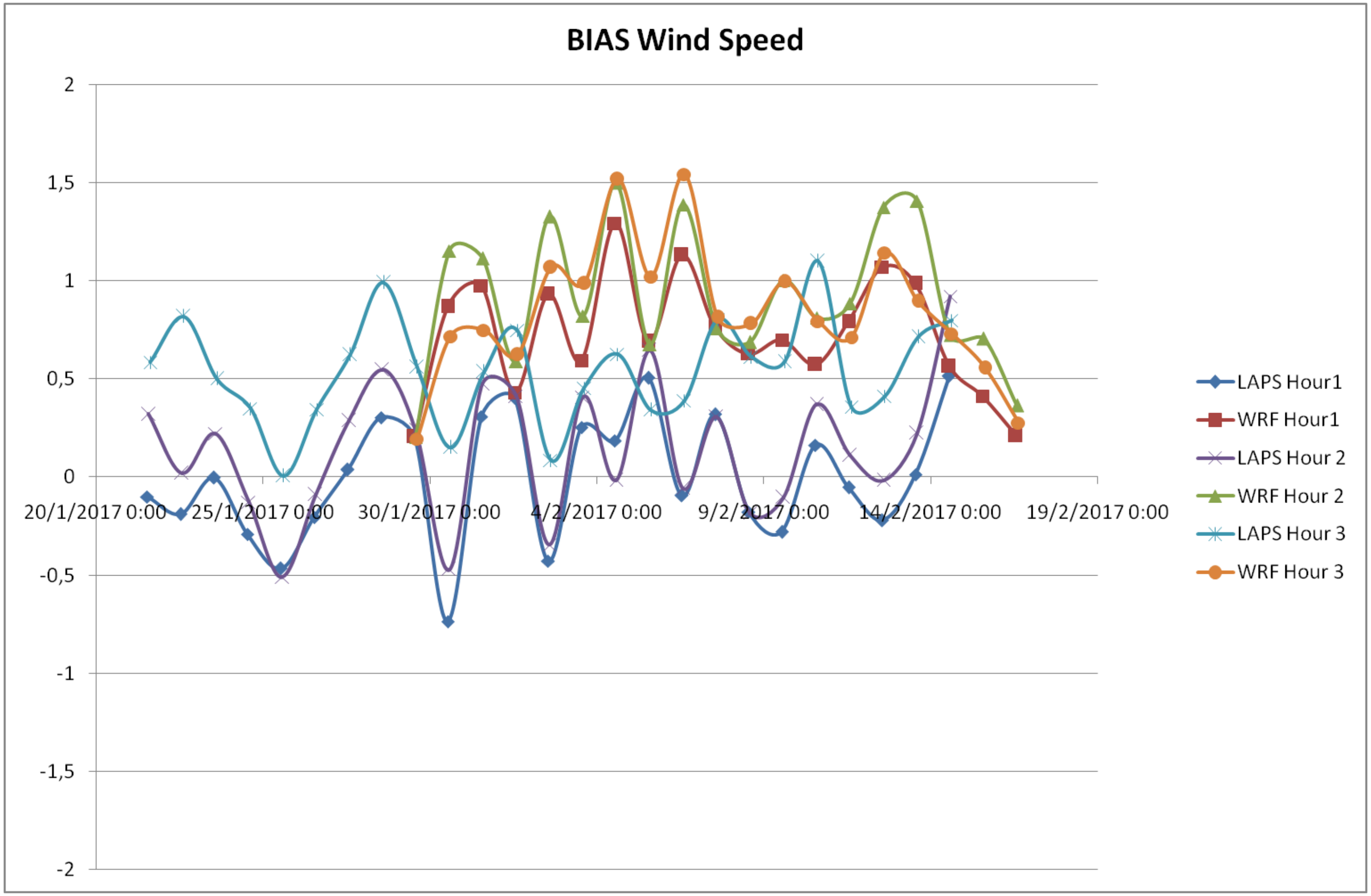
LAPS-WRF temperature evaluation



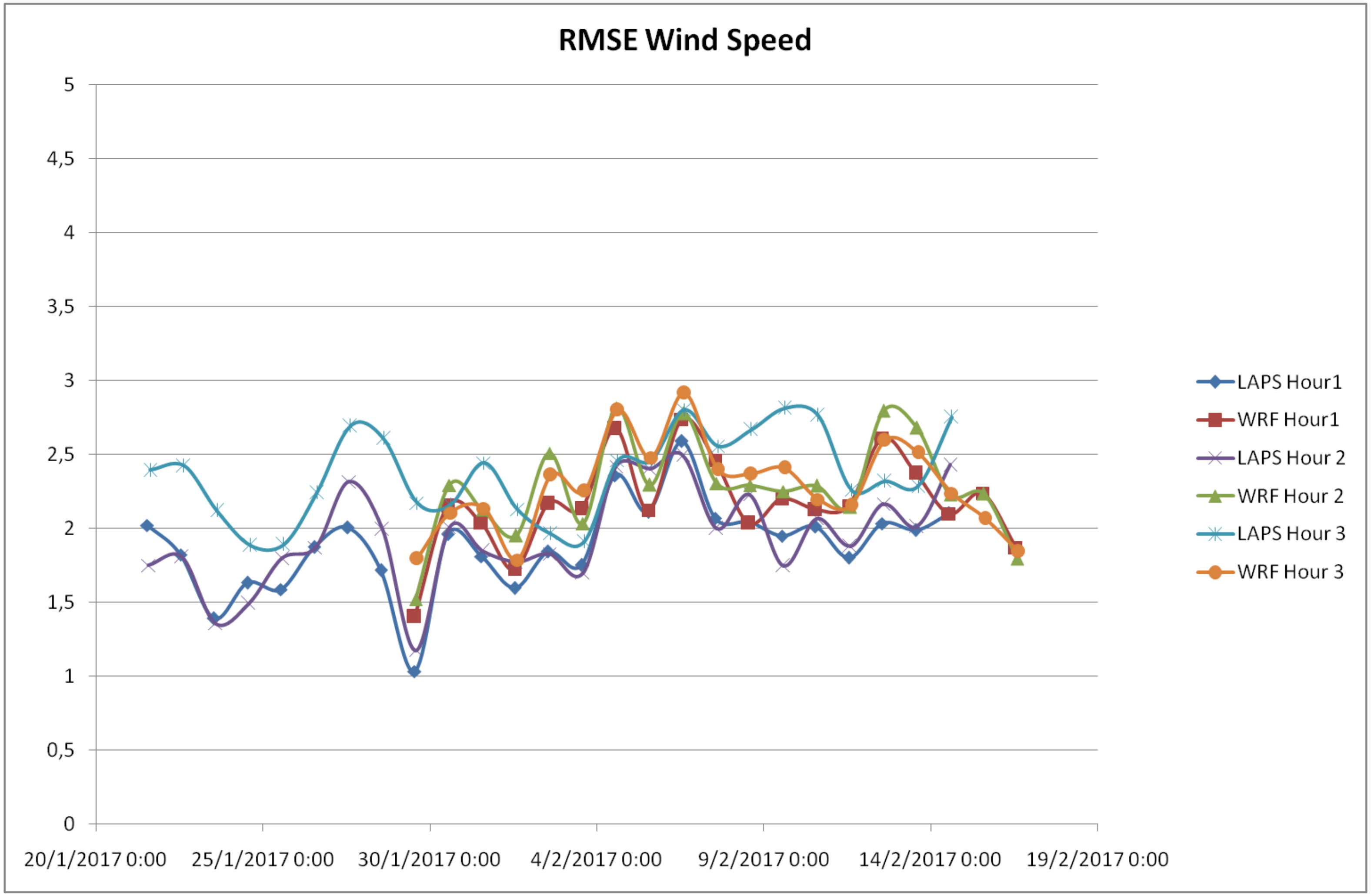
LAPS-WRF temperature evaluation



LAPS-WRF wind speed evaluation



LAPS-WRF wind speed evaluation

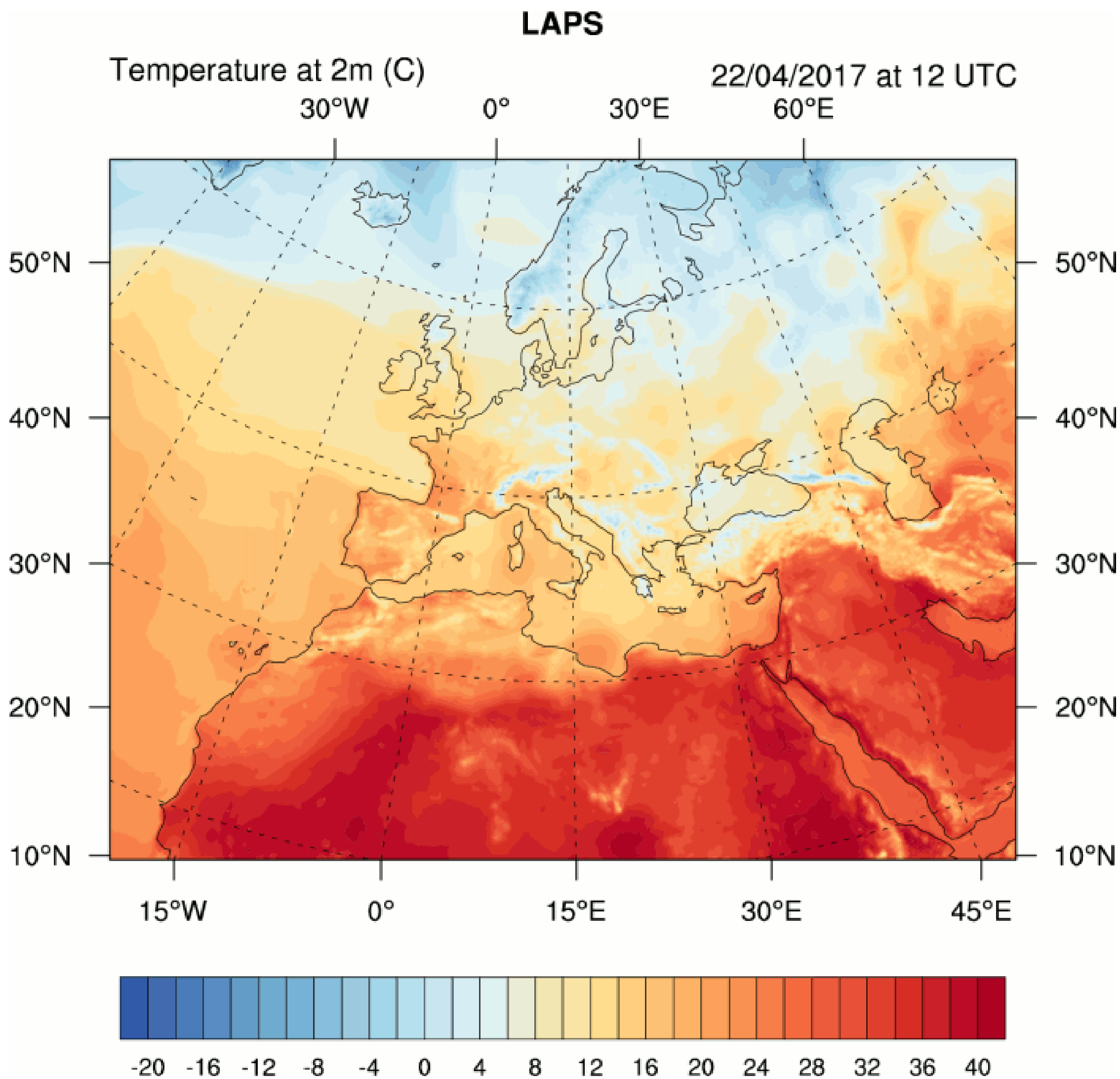


LAPS-WRF aggregated statistics

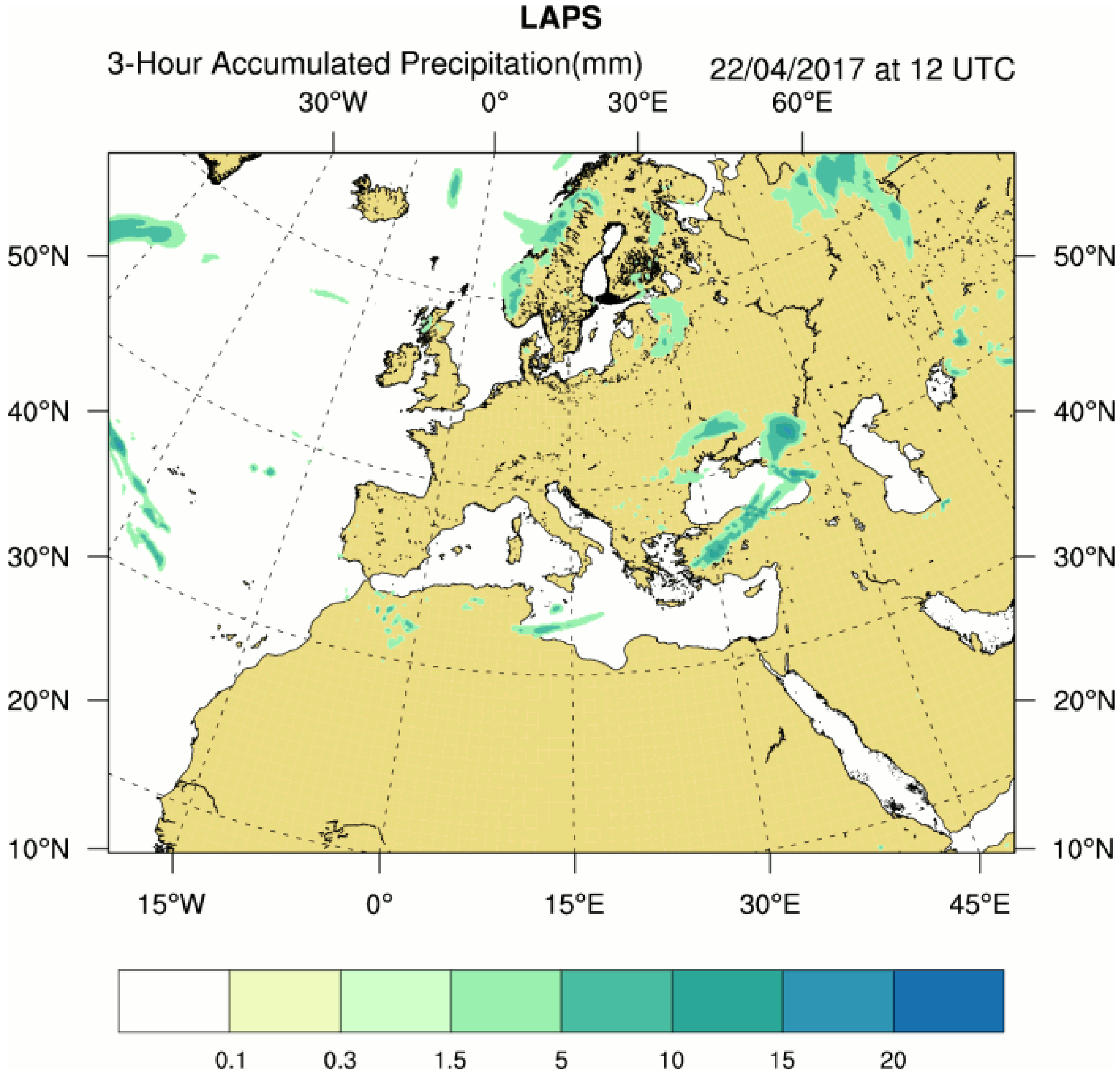
LAPS	Temperature		Wind speed	
	Bias	RMSE	Bias	RMSE
+1	-0,534	1,984	0,212	2,031
+2	-0,557	2,166	0,347	2,142
+3	-0,246	2,398	0,544	2,303

WRF	Temperature		Wind speed	
	Bias	RMSE	Bias	RMSE
+1	-0,541	2,715	0,728	2,2
+2	-0,706	2,758	0,921	2,305
+3	-0,646	2,416	0,848	2,306

LAPS temperature evolution



LAPS precipitation evolution



Relevant textbooks. THANK YOU