STATISTICAL MODELS and VERIFICATION

Mihaela NEACSU & Otilia DIACONU

ANM/LMN/AMASC

COSMO-September 2013

SUMMARY

SUMMARY

AMASC group - Main activities

STATISTICAL POST_PROCESSING - cooperation with Météo France

- **MOS_ECMWF** - 00 and 12 UTC
- **MOS_ARPEGE** - 00,06,12,18 UTC
- **MOS_ALADIN** - 00 and 12 UTC
- **MOS_EPS**

FORECAST VERIFICATION - short and medium range

- final forecasts
- user delivered forecasts
- MOS forecasts
- NWP forecasts

Other duties: **comparative display** of the direct model output: ECMWF, ARPEGE, ALADIN, ALARO, COSMO

RESULTS - dedicated web site http://neptun.meteoromania.ro

AMASC group - Main activities

STATISTICAL POST_PROCESSING - cooperation with Météo France

- **MOS_ECMWF** - 00 and 12 UTC
- **MOS_ARPEGE** - 00,06,12,18 UTC
- **MOS_ALADIN** - 00 and 12 UTC
- **MOS_EPS**

FORECAST VERIFICATION - short and medium range

- final forecasts
- user delivered forecasts
- MOS forecasts
- NWP forecasts

Other duties: **comparative display** of the direct model output: ECMWF, ARPEGE, ALADIN, ALARO, COSMO

RESULTS - dedicated web site http://neptun.meteoromania.ro

AMASC group - Main activities

STATISTICAL POST_PROCESSING - cooperation with Météo France

- **MOS_ECMWF** - 00 and 12 UTC
- **MOS_ARPEGE** - 00,06,12,18 UTC
- **MOS_ALADIN** - 00 and 12 UTC
- **MOS_EPS**

FORECAST VERIFICATION - short and medium range

- final forecasts
- user delivered forecasts
- MOS forecasts
- NWP forecasts

Other duties: **comparative display** of the direct model output: ECMWF, ARPEGE, ALADIN, ALARO, COSMO

RESULTS - dedicated web site http://neptun.meteoromania.ro

SUMMARY

MOS systems

The main features

- mathematically simple, yet powerful
- need historical record of observations at forecast points (Hopefully a long, stable one!)
- equations are applied to future run of similar forecast model

Predictands

- temperatures: spot temperatures(every 6h),extreme temperatures
- wind: direction and speed. MOS equation are developped for U and V components and speed.
- total cloudiness: 3 classes(clear sky, variable sky(scattered) and cloudy sky)
- total precipitation in 6h: 3 classes(No pp, small amounts, moderate/intense amounts)

MOS systems

The main features

- mathematically simple, yet powerful
- need historical record of observations at forecast points (Hopefully a long, stable one!)
- equations are applied to future run of similar forecast model

Predictands

- temperatures: spot temperatures(every 6h),extreme temperatures
- wind: direction and speed. MOS equation are developped for U and V components and speed.
- total cloudiness: 3 classes(clear sky, variable sky(scattered) and cloudy sky)
- total precipitation in 6h: 3 classes(No pp, small amounts, moderate/intense amounts)

MOS systems

Predictors

- A long list..., derived from NWP Forecasts: MSLP,Geopotential, Temperatures...
- there are a total of **54** primary potential predictors and **15** derived predictors.

Statistical Methods used

- Multiple Linear Regression - (Forward Selection) for temperatures and wind
- Discriminant Analysis (Forward Selection) for total cloudiness and total precipitation
- Canonical Analysis to summarize spatial information of **16** grid points around the station
- the **equations** - are developped for each RUN/parameters/stations/forecast time

MOS systems

Predictors

- A long list..., derived from NWP Forecasts: MSLP,Geopotential, Temperatures...
- there are a total of **54** primary potential predictors and **15** derived predictors.

Statistical Methods used

- Multiple Linear Regression - (Forward Selection) for temperatures and wind
- Discriminant Analysis (Forward Selection) for total cloudiness and total precipitation
- Canonical Analysis to summarize spatial information of **16** grid points around the station
- the **equations** - are developped for each RUN/parameters/stations/forecast time

MOS systems

MOS - update and dissemination

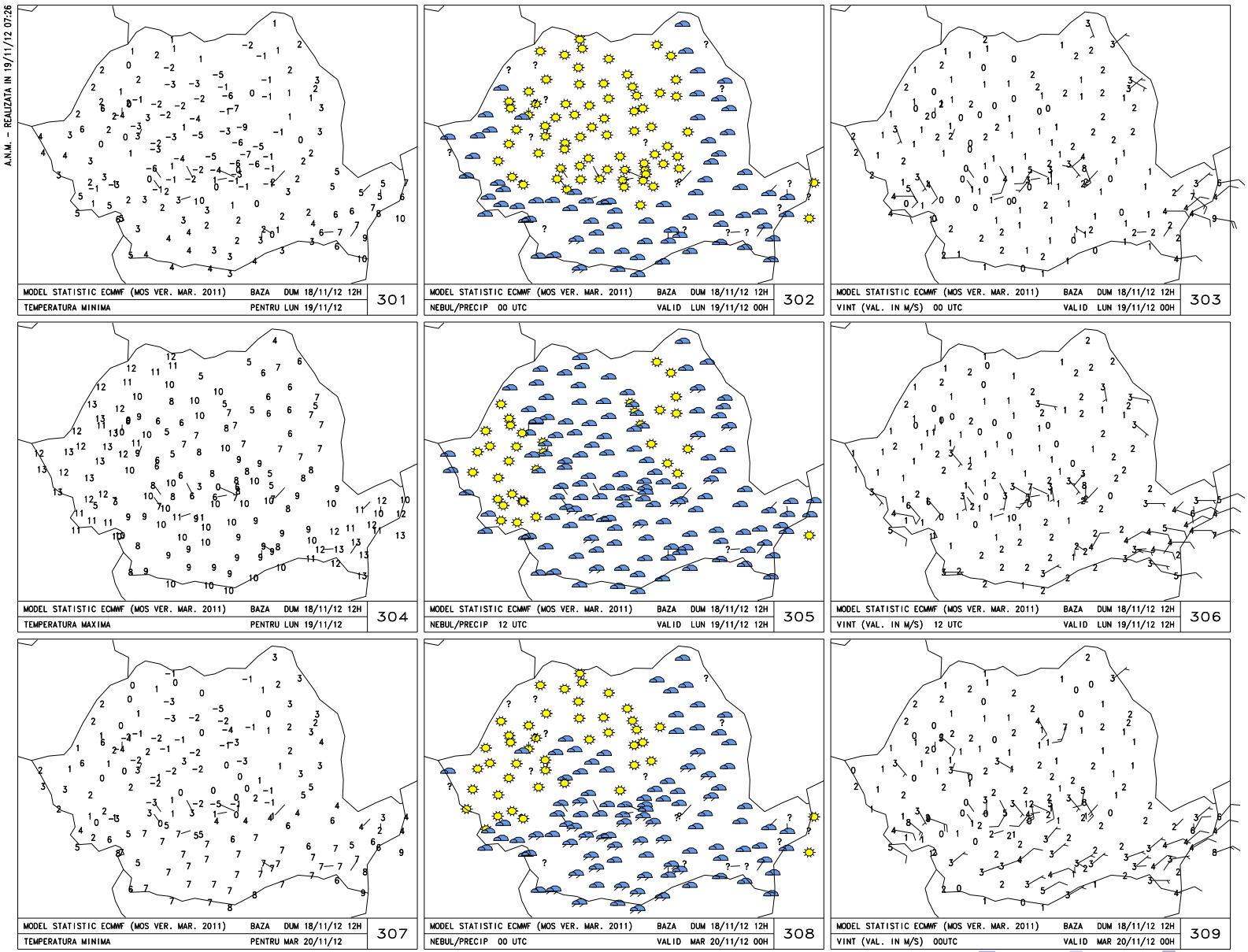
- all MOS systems are updated every two years - *cooperation programme with Météo France*
- MOS results are disseminated on the web-site : maps formats and text in various agregation formats

MOS systems

MOS - update and dissemination

- all MOS systems are updated every two years - *cooperation programme with Météo France*
- MOS results are disseminated on the web-site : maps formats and text in various agregation formats

MOS systems - results - maps examples



MOS systems - results - text examples - by region

Prognoze MOS pe statii grupate pe regiuni

Model MOS: ECMWF **REGIUNEA:** BANAT **RUN:** 00 UTC *Avans regiune:* + -

ECMWF * MODEL STATISTIC MOS ** Ver. 2009 - RUN 00 BAZA: 20121114

LEGENDA Nebulozitate : S=Senin, V=Variabil, N=Noros Precipitatii : F = Fara, S=Slabe, M=Moderate

BANAT

DATA DE VALIDITATE:	20121115				20121116				20121117				20121118				20121119			
ORA:	00	06	12	18	00	06	12	18	00	06	12	18	00	06	12	18	00	06	12	18
SINICOLAU_MARE																				
Tmin/Tmax	3.9		11.5		0.7		11.2		2.4		12.3		2.7		13.3		4.5		13.4	
Nebulozitate	N	N	N	S	N	S	S	S	S	S	S	S	S	S	S	S	S	S	S	N
Precipitatii	F	F	F	F	M	F	F	F	F	F	F	F	F	F	F	F	F	F	F	F
Prob.Precip(%)	79	56	80	78	70	93	88	68	80	90	80	75	90	73	68	69	74	73	67	66
Viteza vant (m/s)	0	0	1	1	0	1	1	1	1	1	2	2	2	2	4	2	2	2	2	1
Directie vant (deg)	70	129	90	80	39	49	80	110	120	110	130	130	130	130	150	140	130	119	140	39
JIMBOLIA																				
Tmin/Tmax	5.0		11.4		1.3		11.7		2.8		12.9		2.9		13.5		4.5		13.8	
Nebulozitate	N	N	N	S	N	S	S	S	S	S	S	S	S	S	S	S	S	S	S	N
Precipitatii	?	?	F	F	?	?	F	F	?	?	F	F	?	?	F	F	?	?	F	F
Prob.Precip(%)	-99	-99	80	69	-99	-99	90	71	-99	-99	84	81	-99	-99	66	71	-99	-99	69	59
Viteza vant (m/s)	1	0	1	0	2	1	1	1	4	1	2	2	3	2	4	2	2	2	3	1
Directie vant (deg)	90	119	129	79	49	80	110	120	59	129	140	140	89	130	159	160	130	130	160	80
TIMISOARA																				
Tmin/Tmax	5.2		11.7		1.1		12.2		1.9		13.1		2.5		14.0		3.1		14.3	
Nebulozitate	N	N	N	V	S	S	S	S	S	S	S	S	S	S	S	S	S	S	S	N
Precipitatii	F	F	S	F	F	F	F	F	F	F	F	F	F	F	F	F	F	F	F	F
Prob.Precip(%)	78	71	86	74	79	82	89	78	90	82	85	84	91	89	72	71	75	76	71	57
Viteza vant (m/s)	1	1	1	1	1	1	1	1	1	1	1	1	1	1	2	1	1	1	2	1
Directie vant (deg)	59	90	49	49	59	80	29	49	70	80	100	90	80	80	140	120	70	80	99	19

MOS systems - results - text examples - by station

Prognoze MOS pe statii

STATIA: Buc_Baneasa *Avans statie:* + - CLICK-AFISAR

ECMWF * MODEL STATISTIC MOS ** Ver. 2009 - RUN 12 BAZA: 20121114

DATA DE VALIDITATE:	20121116				20121117				20121118				20121119				20121120			
ORA:	00	06	12	18	00	06	12	18	00	06	12	18	00	06	12	18	00	06	12	18
BUC_BANEASA																				
Tmin/Tmax	-1.9		11.0		-1.6		11.2		-0.6		10.4		4.2		10.2		4.0		9.3	
Nebulozitate	S	N	S	S	S	S	S	S	N	N	N	N	N	N	N	N	N	N	N	S
Precipitatii	F	F	F	F	F	F	F	F	F	F	F	F	M	M	M	F	F	S	S	F
Prob.Precip(%)	98	98	96	91	92	93	93	89	72	74	68	67	63	59	66	66	61	57	57	79
Viteza vant (m/s)	1	0	1	1	1	1	2	1	1	2	2	1	1	1	2	2	2	2	2	2
Directie vant (deg)	269	289	149	39	39	39	79	49	39	29	59	49	39	49	69	39	39	39	69	49

ARPEGE * MODEL STATISTIC MOS ** Ver. 2009 - RUN 00 BAZA: 20121115

DATA DE VALIDITATE:	20121116				20121117				20121118			
ORA:	00	06	12	18	00	06	12	18	00	06	12	18
BUC_BANEASA												
Tmin/Tmax	-0.1		6.2		-2.2		3.1		-0.9		5.9	
Nebulozitate	S	N	N	S	S	N	N	S	N	N	N	N
Precipitatii	F	F	F	F	F	F	F	F	F	F	F	F
Prob.Precip(%)	93	92	95	91	93	95	95	94	72	67	84	59
Viteza vant (m/s)	2	1	1	1	1	1	2	1	2	1	2	1
Directie vant (deg)	230	250	240	59	39	29	79	49	29	19	79	49

ROMANIA

SUMMARY

Why verify ? - Main purposes

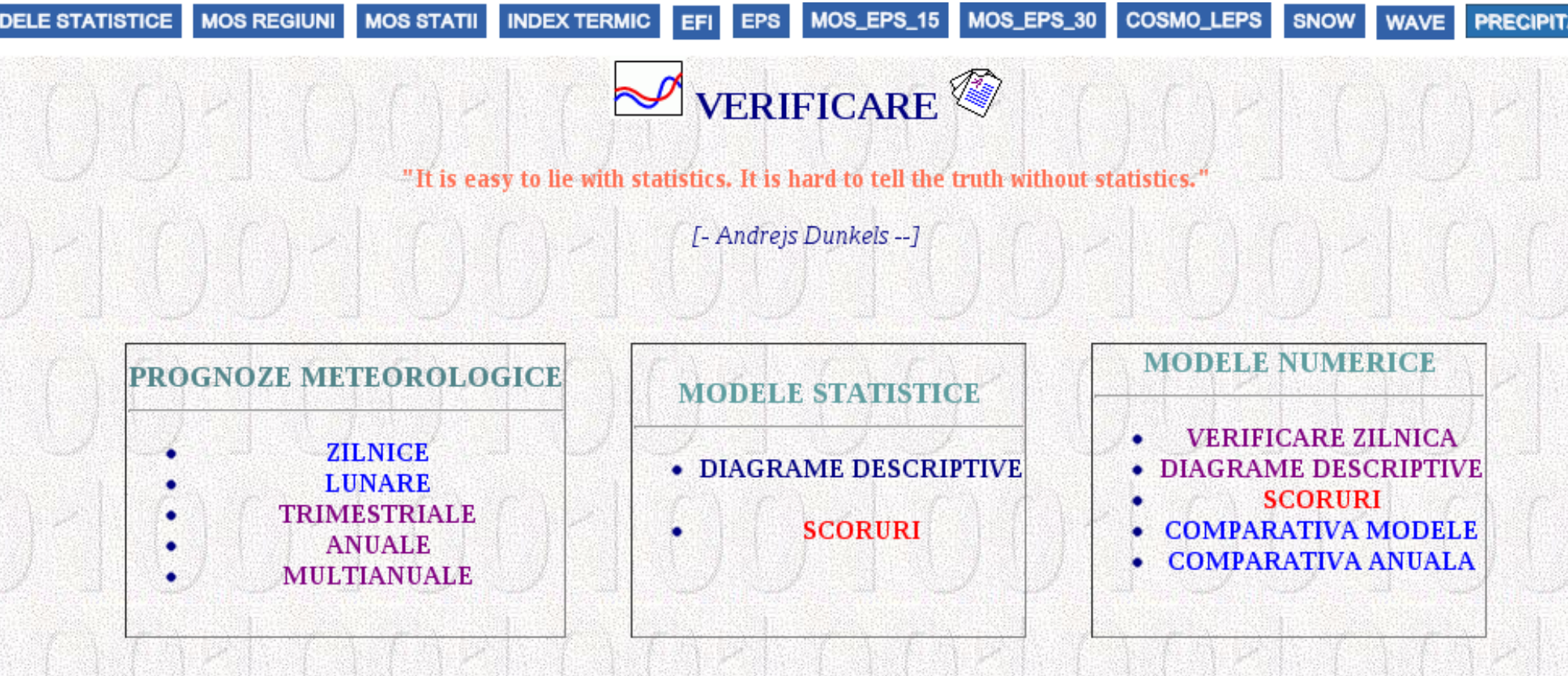
- Administrative purposes
- Scientific purpose
- Economic purposes

"Producing forecasts without verifying them systematically is an implicit admission that the quality of the forecasts is a low priority" - Harold E. Brooks,NOAA/National Severe Storms Laboratory

..reasons to verify meteorological forecasts

- help operational forecasters understand model biases and select models for use in different conditions
- help “users“ interpret forecasts
- identify forecasts weaknesses, strengths differences

Verification in NMA

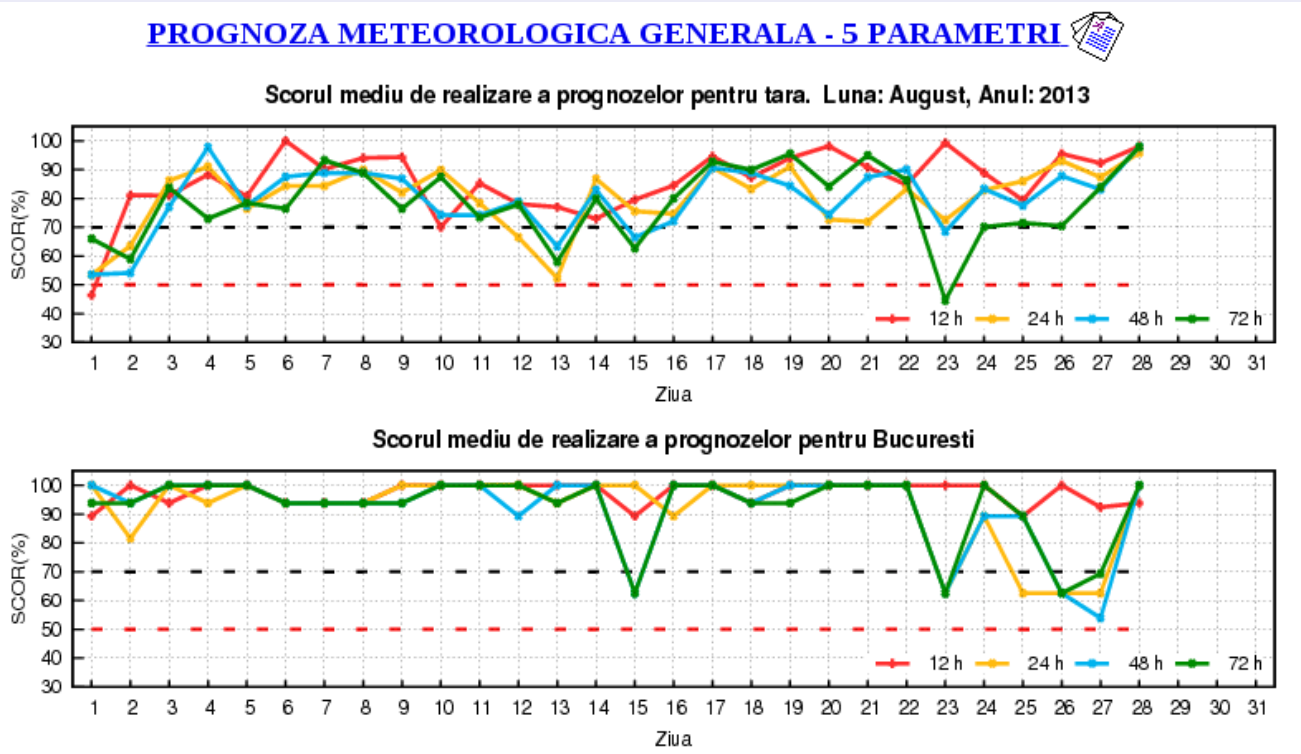


- Forecaster products - Final forecast
- MOS verification
- Numerical models verification

Verification in NMA

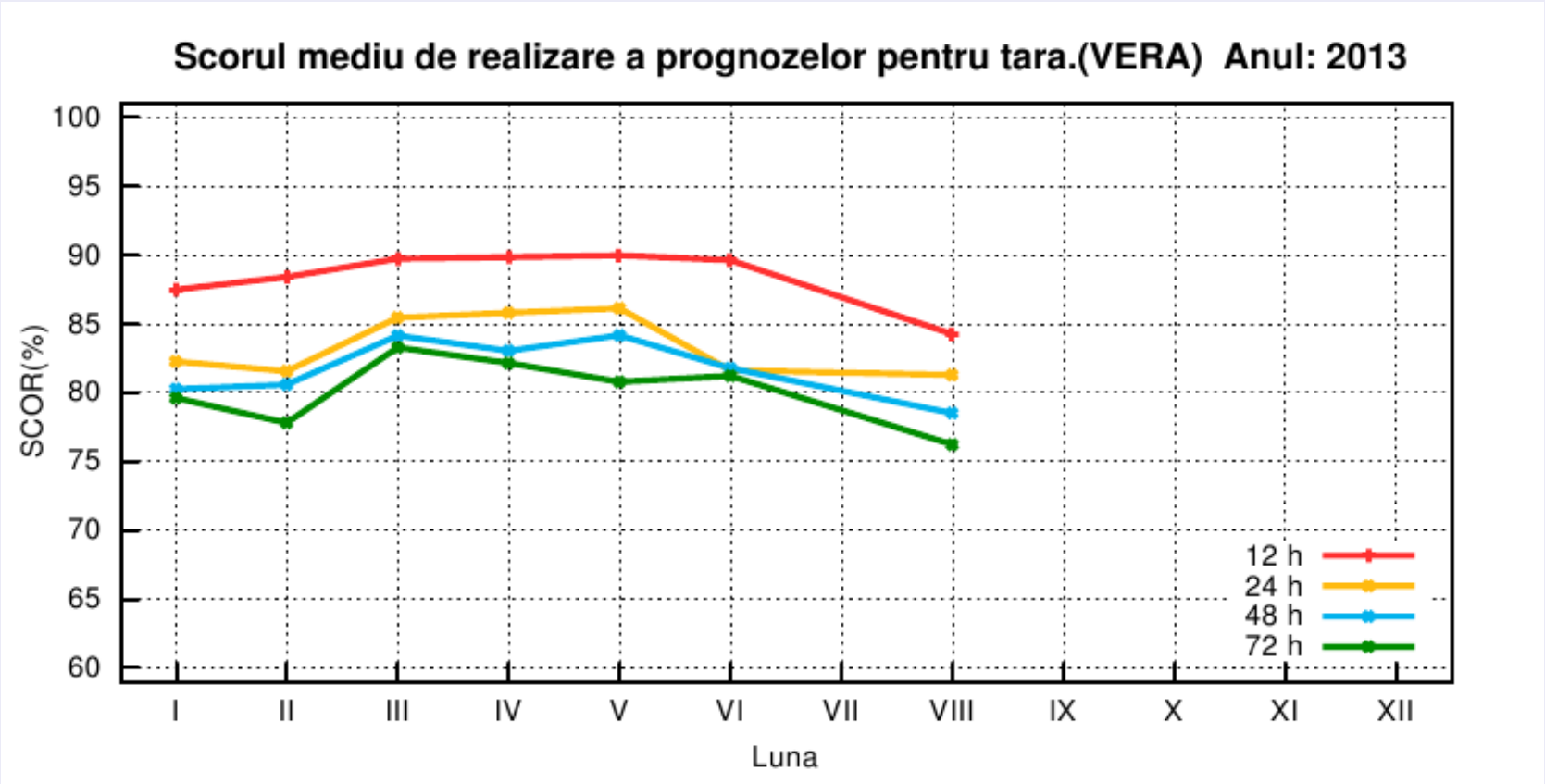
- **Final forecast** - forecaster product
- **Method**: adaptation of the method described in: K.Colls and all,1981: A Forecast verification procedure for public weather forecasts, Australian, *Met.Mag, Vol 29,No.1, 9-25*

Daily verification



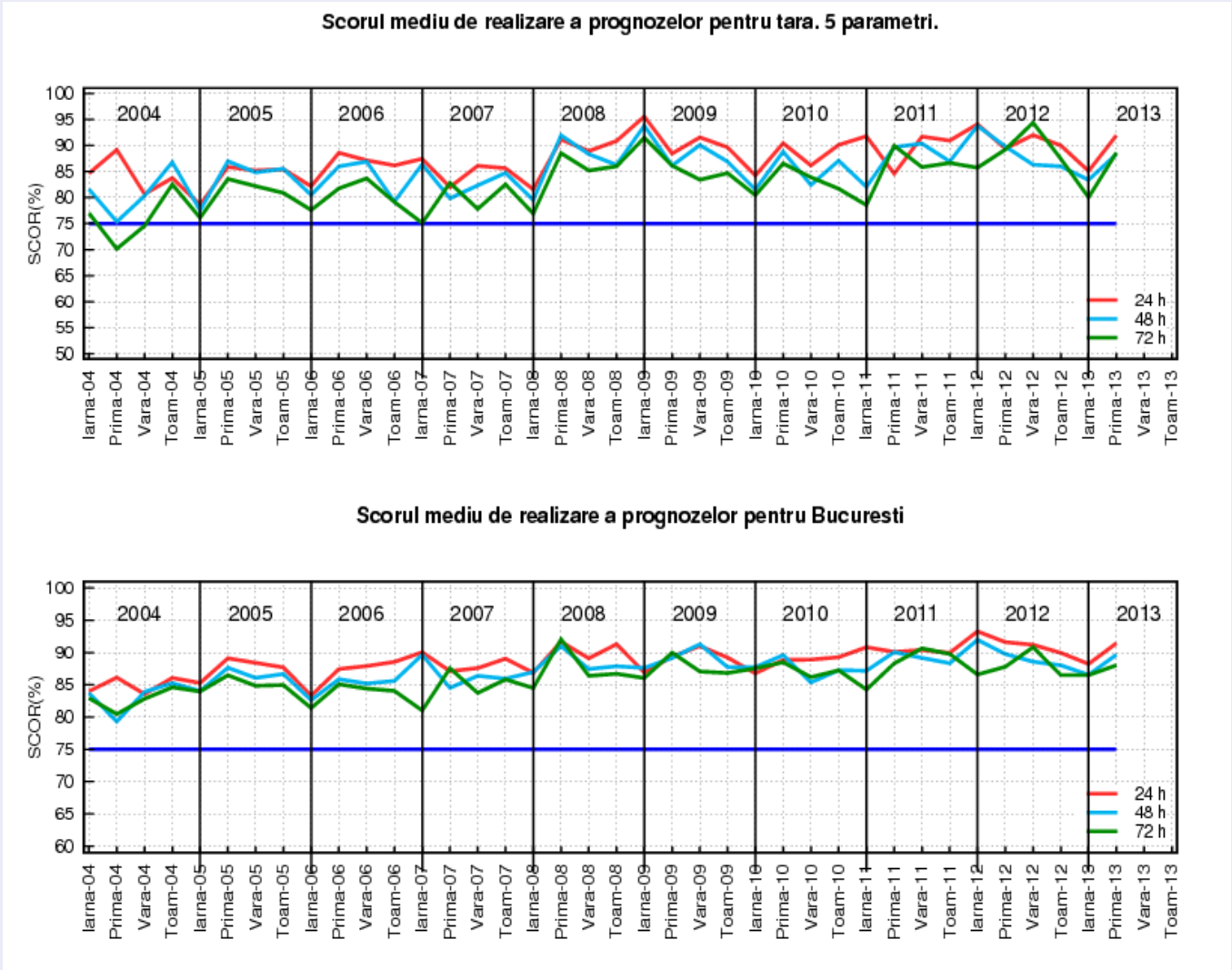
Verification in NMA

Monthly scores



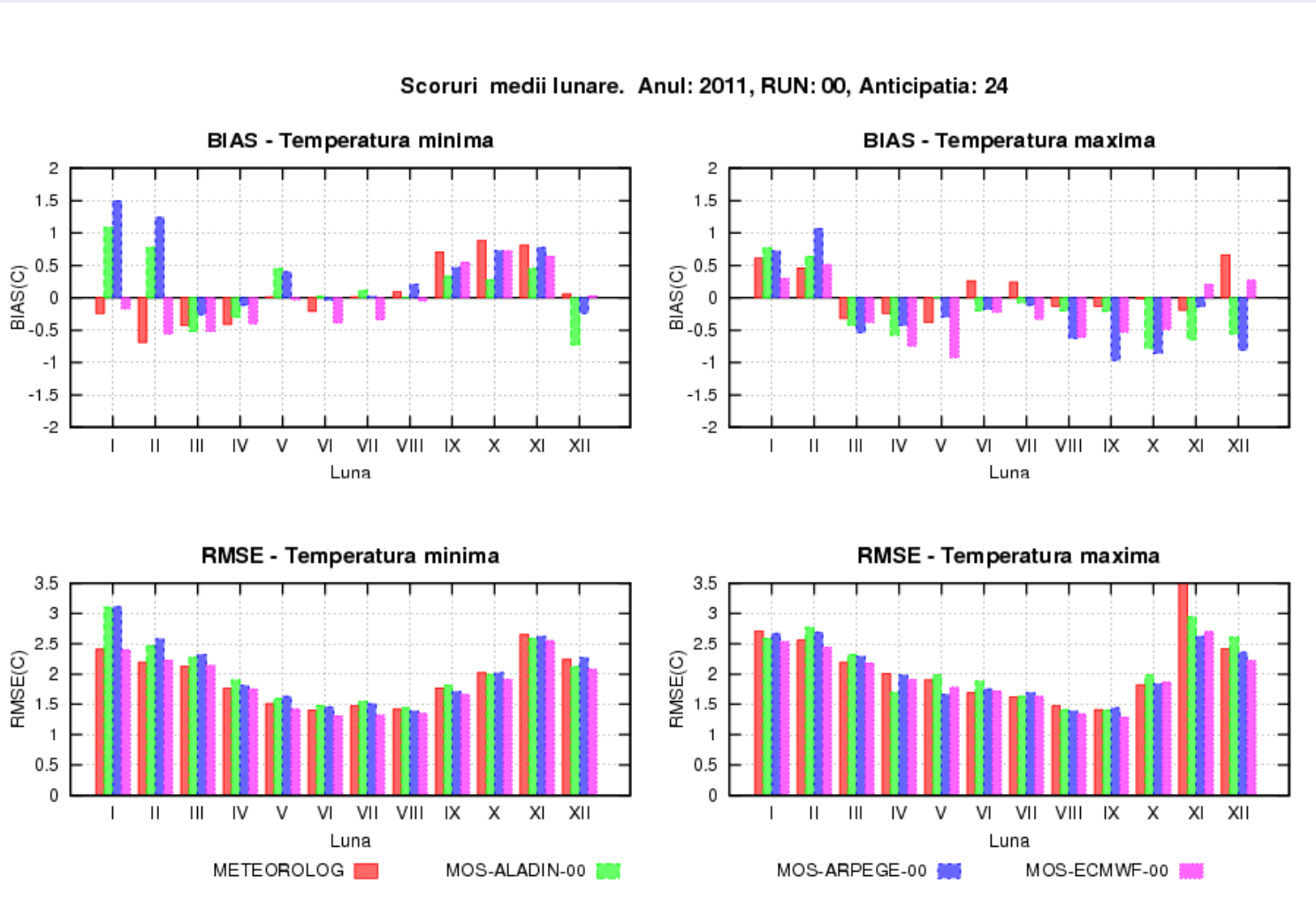
Verification in NMA

Seasonal-multi-annual scores

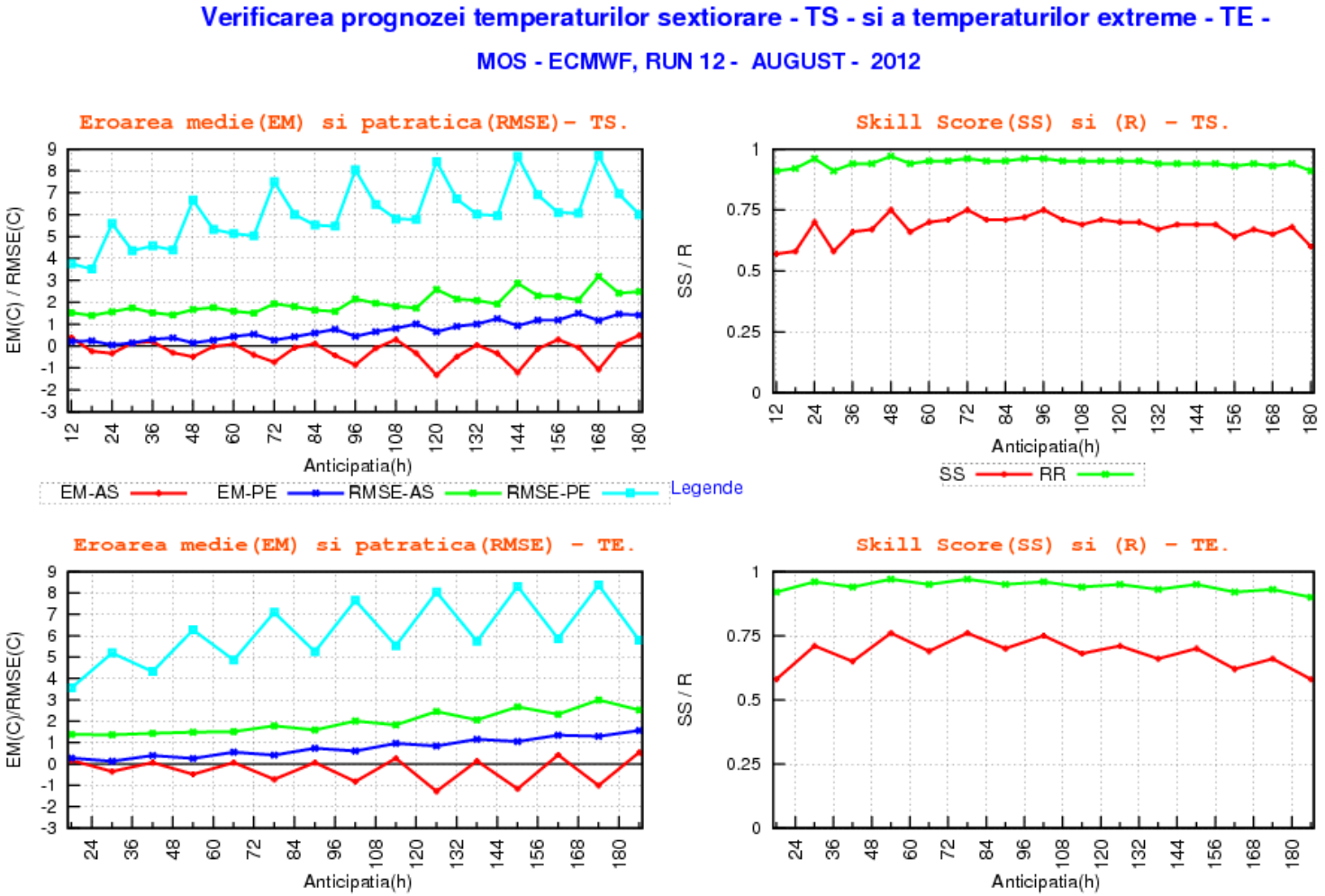


Verification in NMA

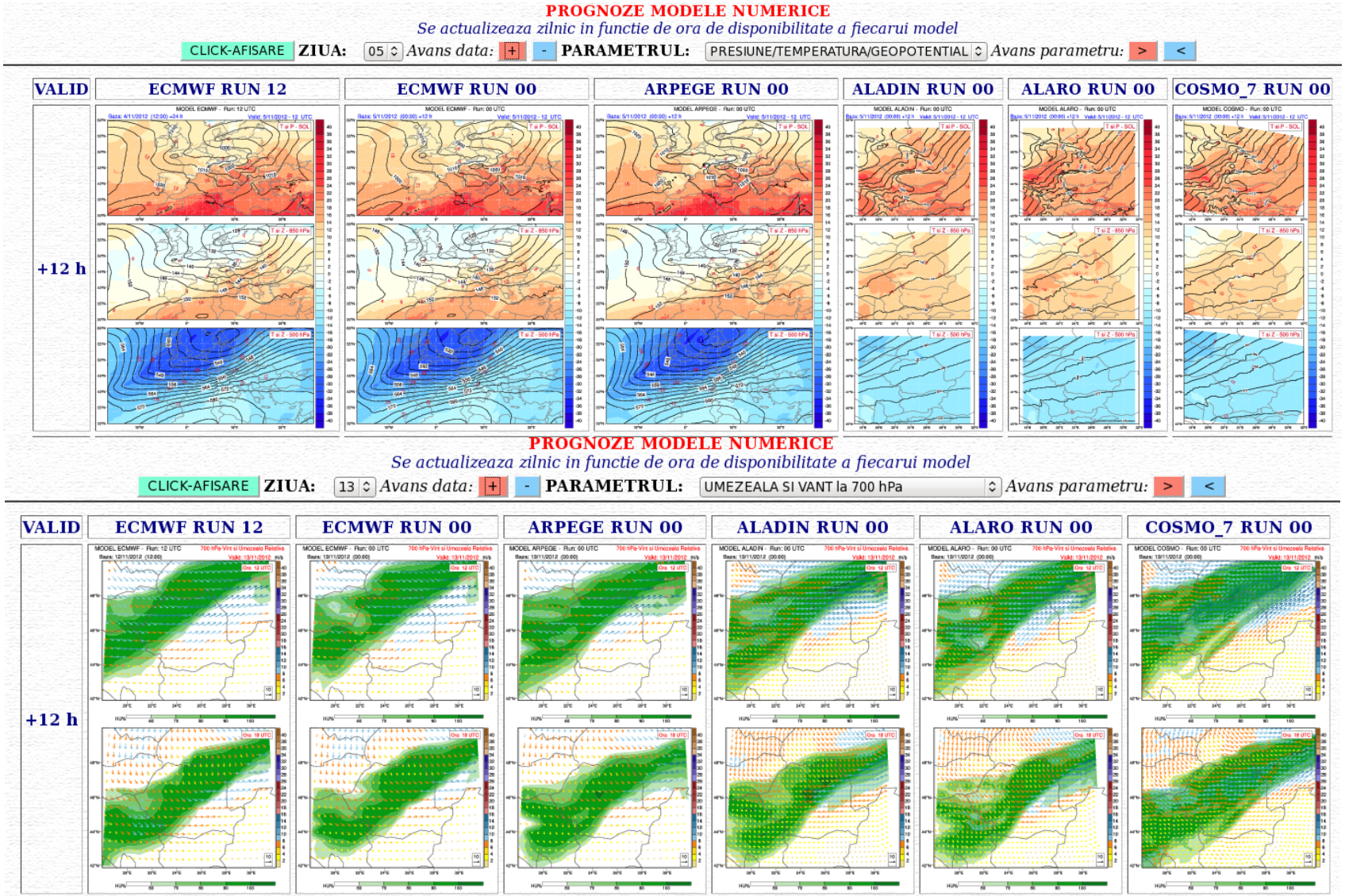
MOS against forecaster - standard scores



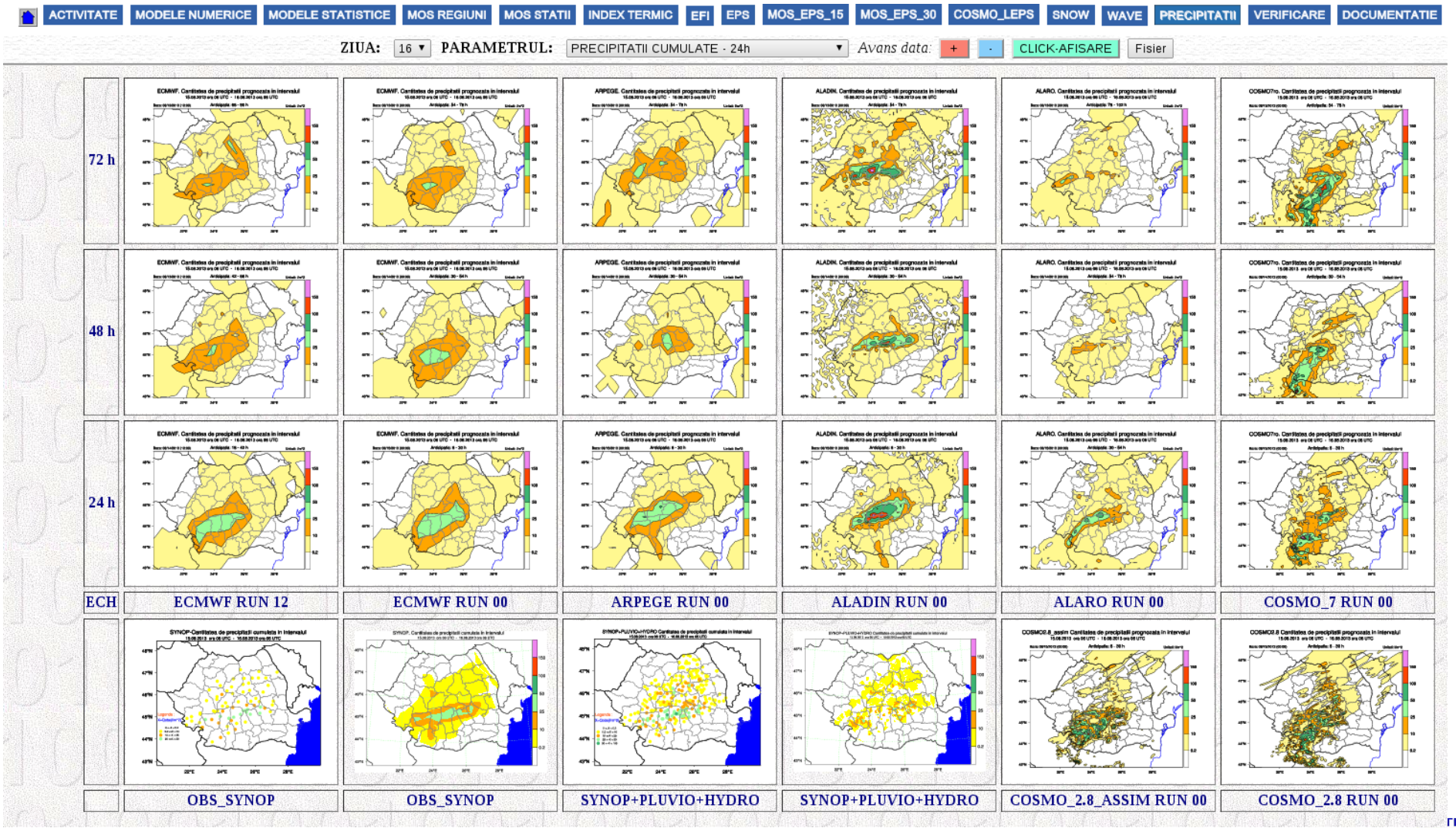
MOS verification - VERMOD-standard scores



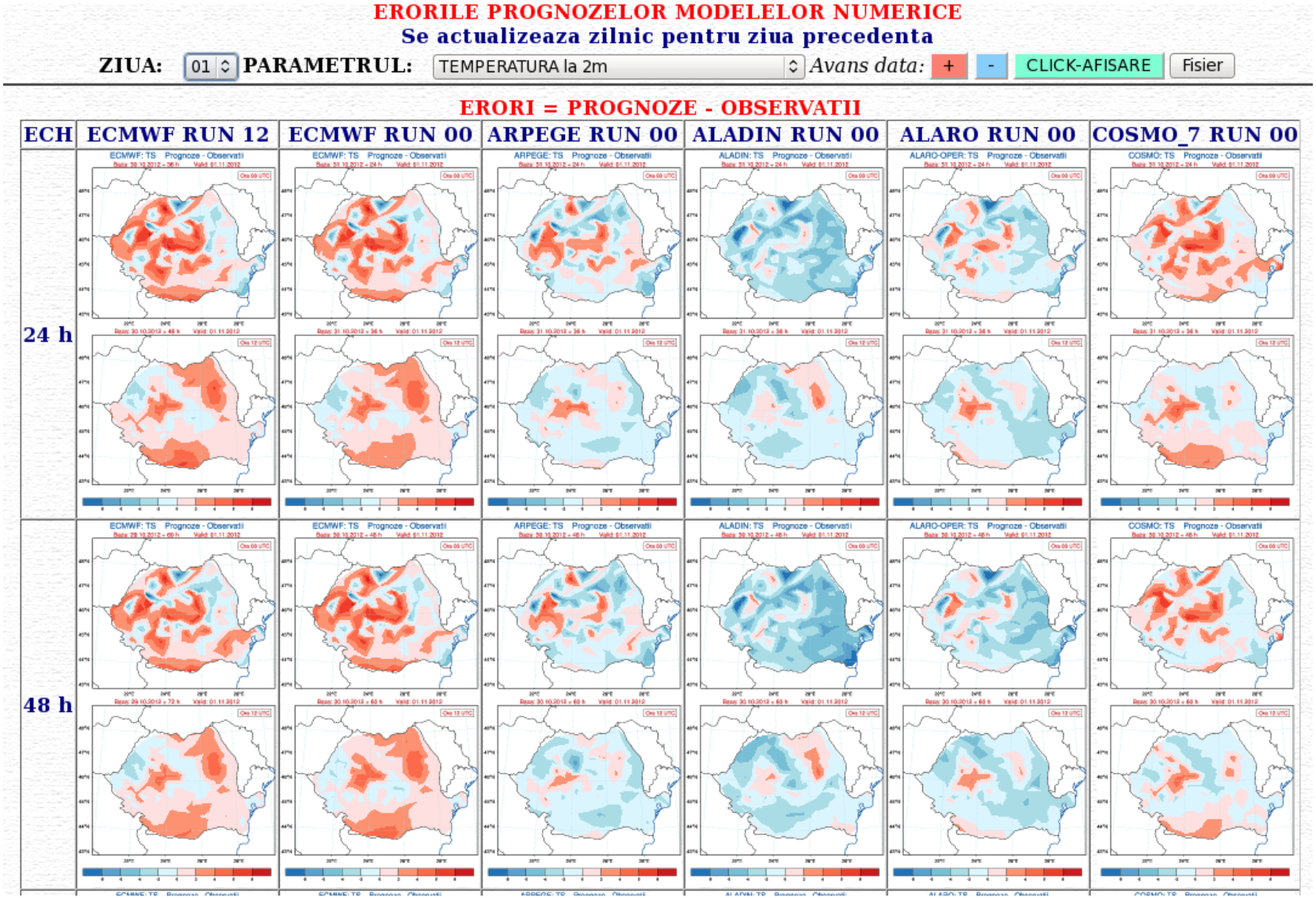
Comparative DMO display - a few parameters



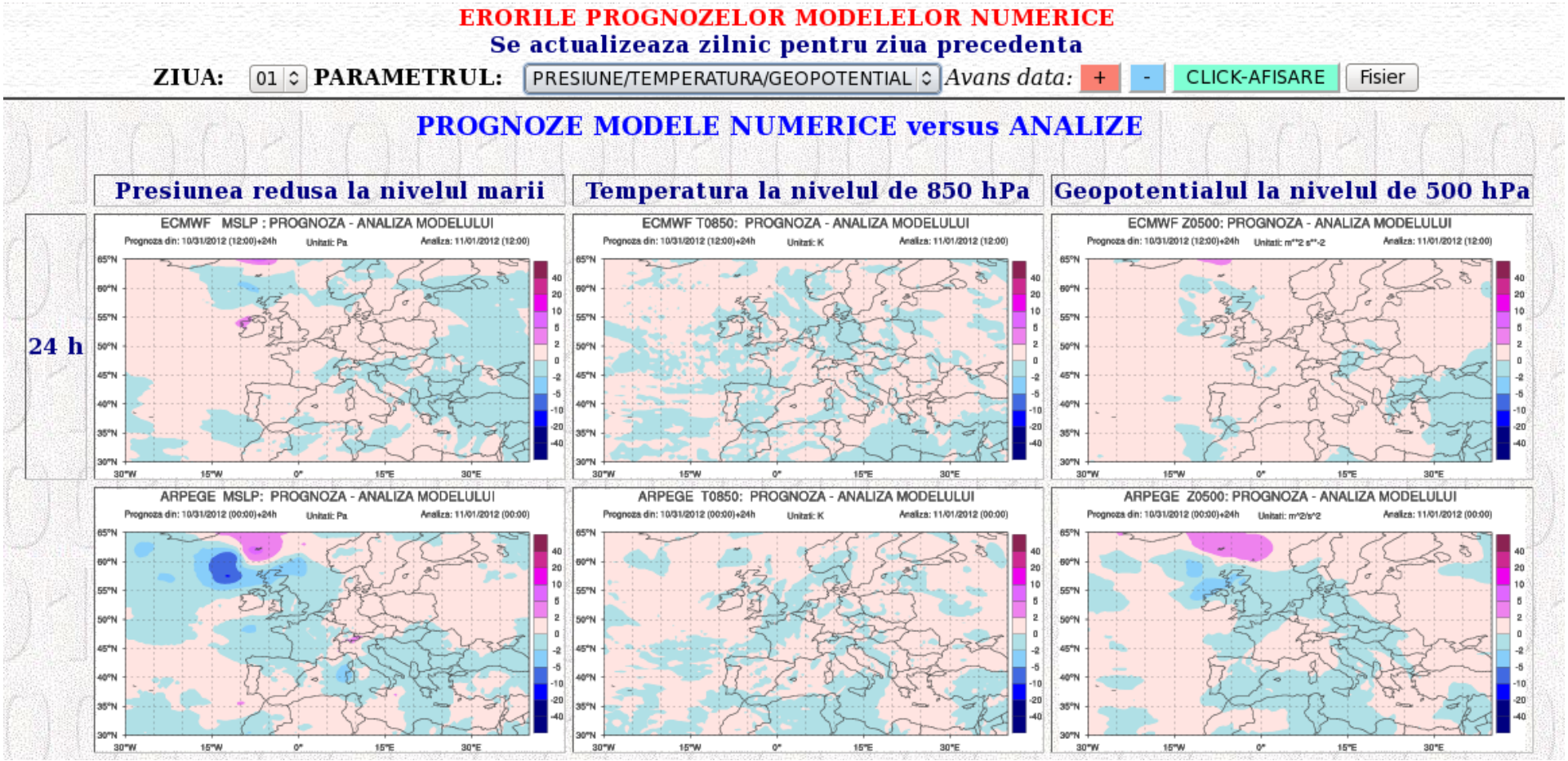
DMO - verification - daily - visual



DMO - daily errors against observations



DMO - daily errors against analysis



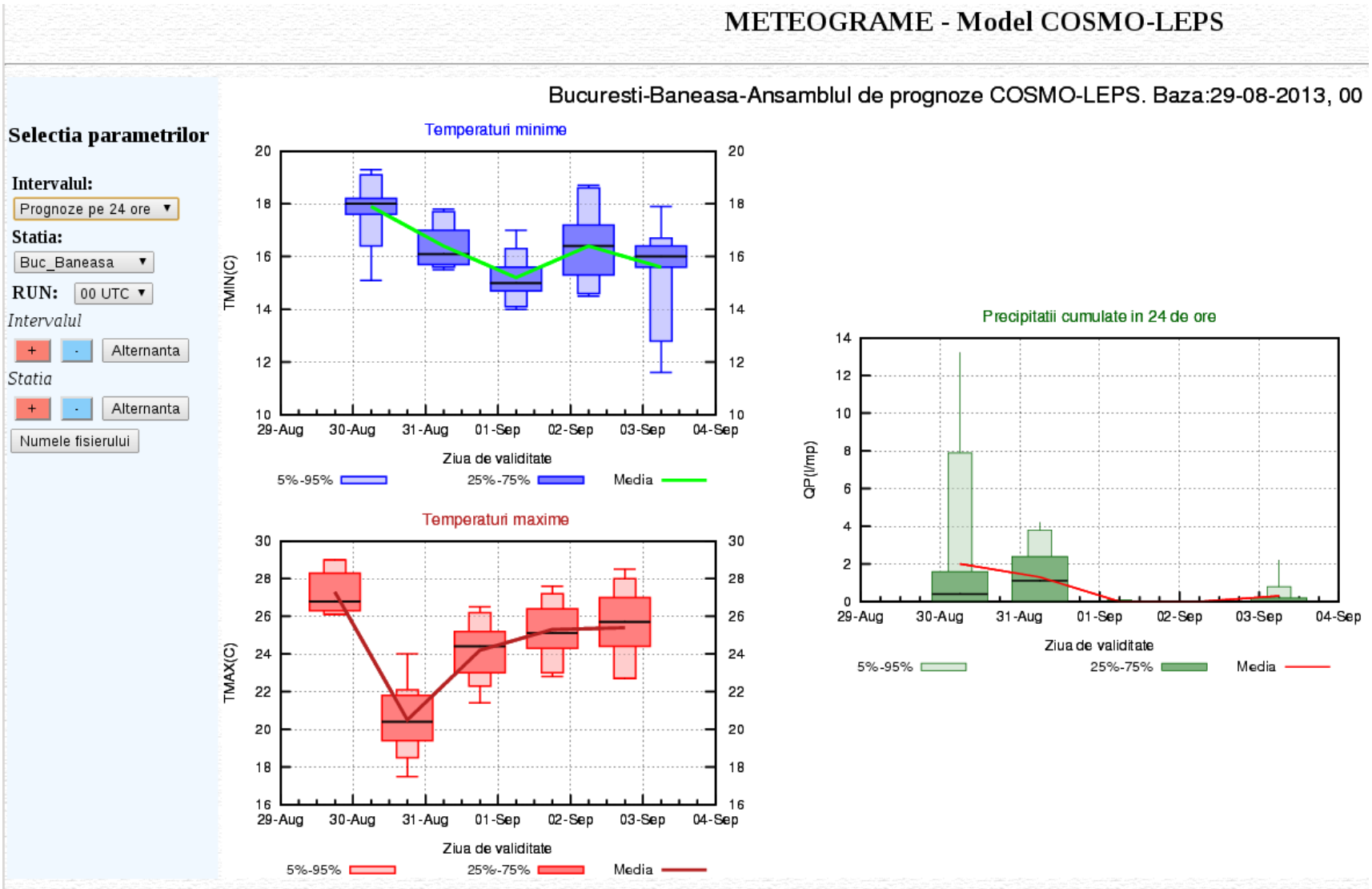
Purpose of showing comparative maps

- help forecaster making decission
- use of the same plotting area for all models forecasts and observations
- use the same threshold and the same color palette
- make faster and easier comparison between models forecasts and/or observations
- all small maps can be zoomed

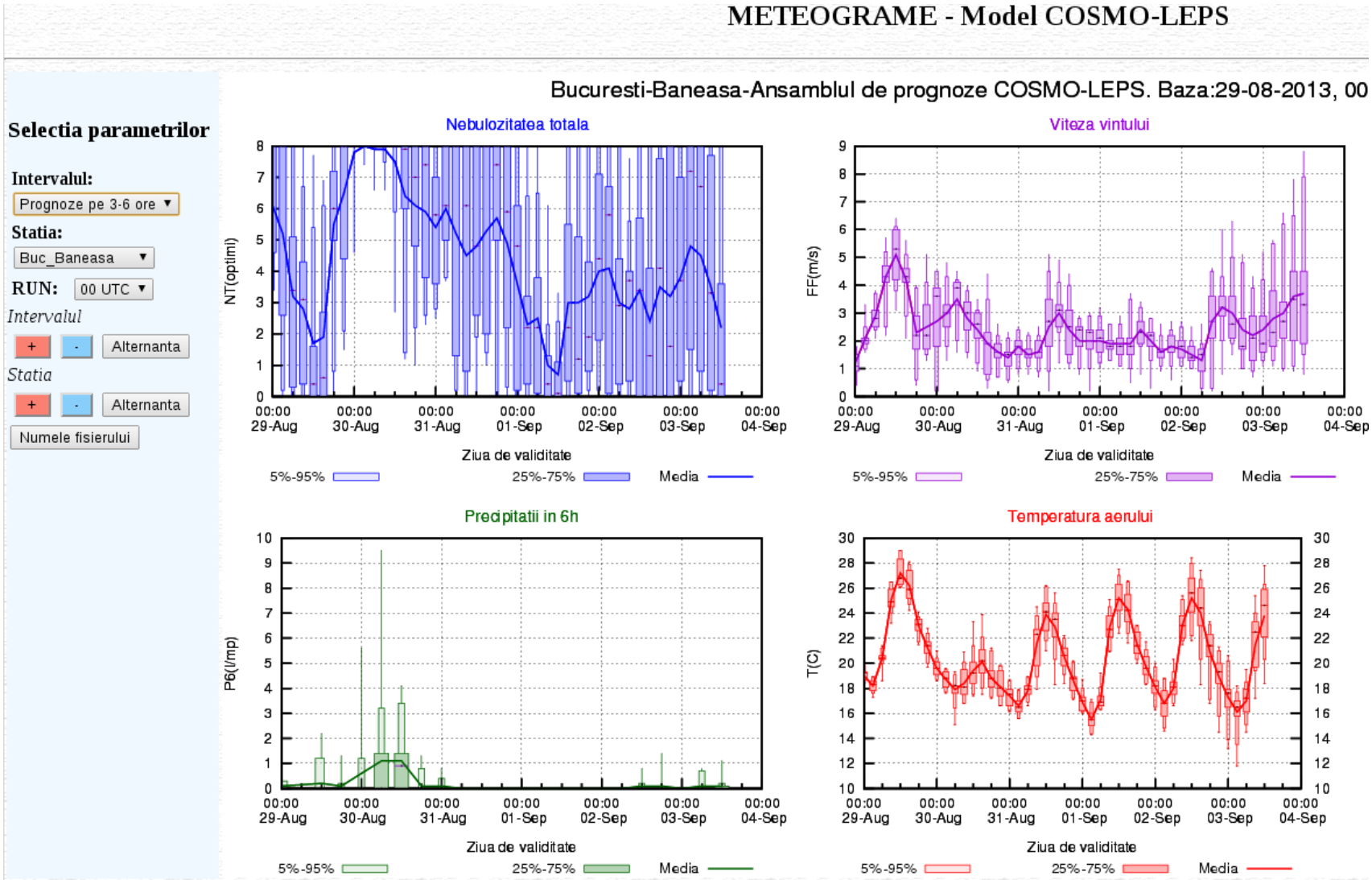
Parameters:

- 24,12,6 and 3 h total amounts of precipitation
- mslp, geopotential and temperature - 850 hPa and 500 hPa
- relative humidity and wind(speed and direction) at 10m, 700 hPa
- CAPE and MU

COSMO-LEPS -meteograms - a few graphical products



COSMO-LEPS -meteograms - a few graphical products



VERMOD 1.0 - Direct Model Verification system

The system uses:

- daily files, monthly archives
- shell scripts, programms and routines in Fortran (developed in AMASC group)
- **R** scripts and Gnuplot for graphics

A. Data

Observations

- the system accepts observations from whatever source provided that they are compatible with forecasts
- in the current version - SYNOP observations for parameters: temperature, wind (direction, speed, components), cloudiness, precipitation - amounts in different intervals: 6, 12, 24 hours
- date are subject to control procedures: climatological limits, minimum and maximum limits of the parameter

Forecasts

- GRIB format - cut out Romania domain for ECMWF and ARPEGE models (the same domain is used forthe other models)
- is made interpolation in station with nearest grid point method: same procedure for all models
- forecasts turns in the same units as the observations (temperature in degrees Celsius, rainfall in l/sqm)

A. Data

Observations

- the system accepts observations from whatever source provided that they are compatible with forecasts
- in the current version - SYNOP observations for parameters: temperature, wind (direction, speed, components), cloudiness, precipitation - amounts in different intervals: 6, 12, 24 hours
- date are subject to control procedures: climatological limits, minimum and maximum limits of the parameter

Forecasts

- GRIB format - cut out Romania domain for ECMWF and ARPEGE models (the same domain is used forthe other models)
- is made interpolation in station with nearest grid point method: same procedure for all models
- forecasts turns in the same units as the observations (temperature in degrees Celsius, rainfall in l/sqm)

A. Data

Data format- Observations

```
TS   SYN00   4  163 38
 F6.2 (I5.5,  4(1X, F6.2))
-99.90
08 09 01 00 00 08 09 01 00 00
08 09 01 06 00 08 09 01 06 00
08 09 01 12 00 08 09 01 12 00
08 09 01 18 00 08 09 01 18 00
15000  12.00  12.60  21.00  18.10
15004   8.80  10.20  23.20  16.60
15007  12.00  12.50  20.90  14.90
15010   8.30  11.40  24.00  18.70
15014  10.10  11.90  23.10  16.90
15015   8.30  12.50  21.80  15.10
```

A. Data format

Forecasts

```
TS   DMCEP  29  163 30
 F8.2 (I5.5,  8(1X, F8.2),  2(/,5X,  8(1X, F8.2)),/,5X,  5(1X, F8.2))
  -99.90
08 09 01 00 00 08 09 01 12 00
08 09 01 00 00 08 09 01 18 00
08 09 01 00 00 08 09 02 00 00
08 09 01 00 00 08 09 02 06 00
.............................
08 09 01 00 00 08 09 08 06 00
08 09 01 00 00 08 09 08 12 00
15000    20.52    14.78    12.06    14.84    24.88    17.56    13.17    16.74
         25.94    19.05    15.25    17.33    26.62    20.45    14.77    18.47
         28.09    21.91    17.35    19.58    29.95    23.28    18.25    20.69
         31.40    23.95    19.55    21.25    29.66
15004    21.40    19.11    12.00    17.18    23.79    16.70    13.52    18.70
         24.83    17.96    15.08    19.66    26.23    19.24    16.80    21.08
         27.20    20.70    17.85    21.40    28.65    22.75    20.80    22.72
         30.34    23.01    20.47    24.04    29.10
```

B. Analysis

Steps

1. sincronizing in time the generation of the sets **Forecasts - Observations**
2. the descriptive diagrams are made for a number of stations
3. the calculation of the specific scores, depending on the number of the parameters, for the three types of stratifications:
   - on each station
   - on each corespondent region of the CMR area of observation
   - on each country stations

B. Analysis

Steps

1. sincronizing in time the generation of the sets **Forecasts - Observations**
2. the descriptive diagrams are made for a number of stations
3. the calculation of the specific scores, depending on the number of the parameters, for the three types of stratifications:
   - on each station
   - on each corespondent region of the CMR area of observation
   - on each country stations

B. Analysis

Steps

1. sincronizing in time the generation of the sets **Forecasts - Observations**
2. the descriptive diagrams are made for a number of stations
3. the calculation of the specific scores, depending on the number of the parameters, for the three types of stratifications:
   - on each station
   - on each corespondent region of the CMR area of observation
   - on each country stations

B. Analysis

Steps

1. sincronizing in time the generation of the sets **Forecasts - Observations**
2. the descriptive diagrams are made for a number of stations
3. the calculation of the specific scores, depending on the number of the parameters, for the three types of stratifications:
   - on each station
   - on each corespondent region of the CMR area of observation
   - on each country stations

B. Analysis

Steps

1. sincronizing in time the generation of the sets **Forecasts - Observations**
2. the descriptive diagrams are made for a number of stations
3. the calculation of the specific scores, depending on the number of the parameters, for the three types of stratifications:
   - on each station
   - on each corespondent region of the CMR area of observation
   - on each country stations

B. Analysis

Steps

1. sincronizing in time the generation of the sets **Forecasts - Observations**
2. the descriptive diagrams are made for a number of stations
3. the calculation of the specific scores, depending on the number of the parameters, for the three types of stratifications:
   - on each station
   - on each corespondent region of the CMR area of observation
   - on each country stations

Analysis - continuous parameters

For a set P_i and O_i the following scores are calculated:

The average of observation	$MO = \frac{1}{N}\sum_{i=1}^{N} O_i$
The average of forecasts	$MP = \frac{1}{N}\sum_{i=1}^{N} P_i$
Average ratios	$RM = \frac{MO}{MP}$
Average differences	$DM = MO - MP$
Variance of observation	$VAROBS = \frac{1}{N}\sum_{i=1}^{N}(O_i - MO)^2$
Variance of forecasts	$VARPREV = \frac{1}{N}\sum_{i=1}^{N}(P_i - MP)^2$
Ratio of variances	$RVAR = \frac{VAROBS}{VARPREV}$

Analysis - continuous parameters

Difference of variances	$DIFVAR = VAROBS - VARPREV$
Covariance error	$COV = \frac{1}{N}\sum_{i=1}^{N}(O_i - MO)(P_i - MP)$
Correlation coefficient	$COR = \frac{COV}{\sqrt{VAROBS * VARPREV}}$
Standard deviation	$STD = \sqrt{\frac{1}{N-1}\sum_{i=1}^{N}(P_i - O_i - BIAS)^2}$

Analysis - continuous parameters

Bias - Average error	$BIAS = \frac{1}{N}\sum_{i=1}^{N}(P_i - O_i)$
MAE - Absolute average error	$MAE = \frac{1}{N}\sum_{i=1}^{N}\lvert P_i - O_i\rvert$
MSE - Mean Squared error	$MSE = \frac{1}{N}\sum_{i=1}^{N}(P_i - O_i)^2$
RMSE - Root mean square error	$RMSE = \sqrt{MSE}$
RV- Variance reduction	$RV = \frac{\sum_{i=1}^{N}(P_i - O_i)}{\sum_{i=1}^{N}(MO - O_i)}$

Analysis - categorical parameters (clasess/categories)

For a set P_i and O_i the tabel of contingency is calculated:

	Observations		
Forecasts	YES	NO	**TOTAL**
YES	a	b	**a+b**
NO	c	d	**c+d**
TOTAL	**a+c**	**b+d**	**a+b+c+d=n**

Analysis- categorical parameters- Associated scores

Score name	Definition
Probability of observation	$s = \frac{a+c}{n}$
Probability of forecast	$r = \frac{a+b}{n}$
Bias	$Bias = \frac{a+b}{a+c}$
Percentage of success	$H = \frac{a}{a+c}$
False alarm ratio	$FAR = \frac{b}{a+b}$
False	$F = \frac{b}{b+d}$
Percentage correct	$PC = \frac{a+d}{n}$
Heidke Scor	$HS = \frac{PC-E}{1-E}$

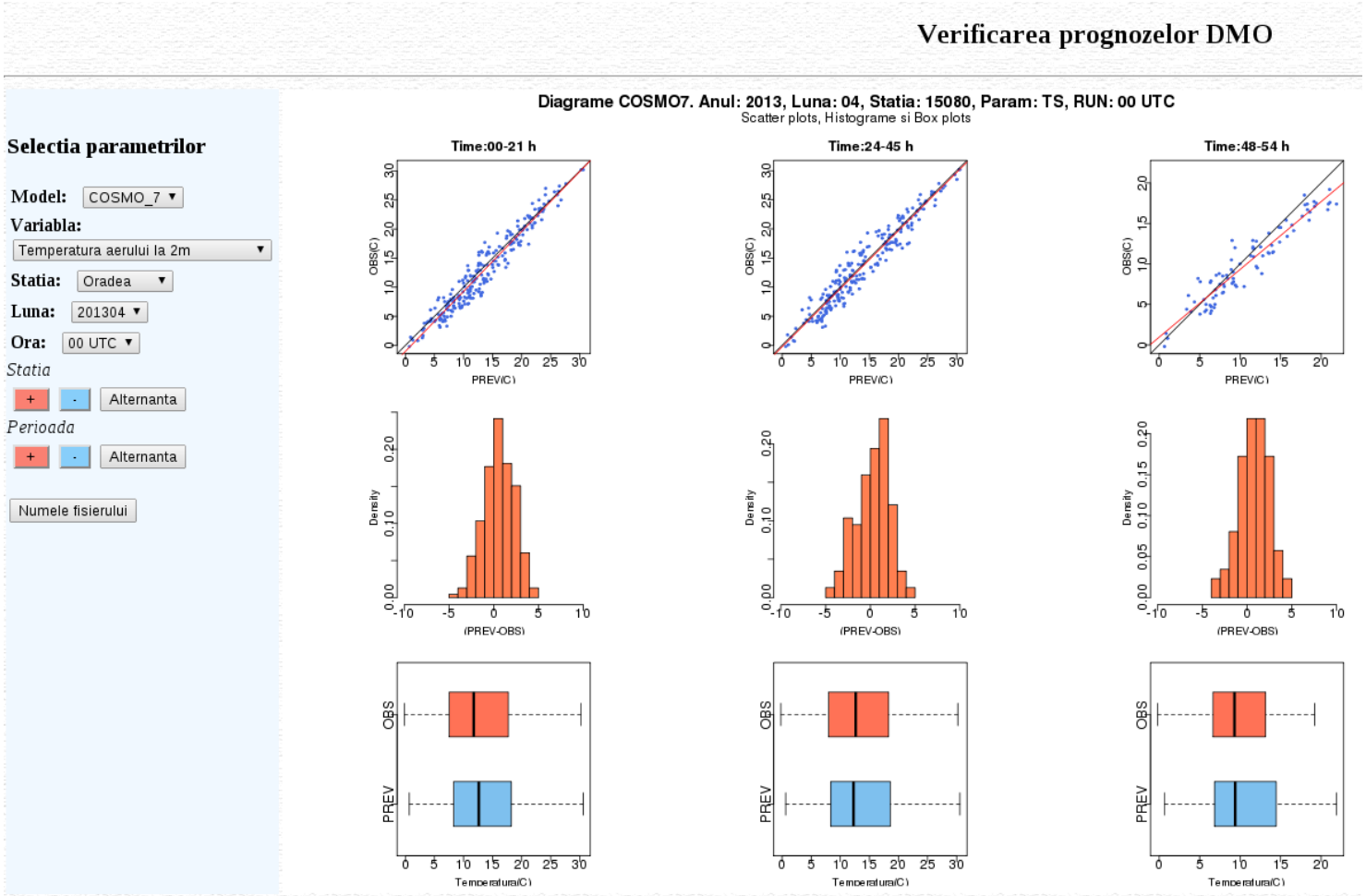
where E is PC hazard forecasting

Analysis- categorical parameters- Associated scores

Score name	Definition
Critical success indicator	$CSI = \frac{a}{a+b+c}$
GSS or ETS	$ETS = \frac{a-ar}{a+b+c-ar}$ where $ar = \frac{(a+b)(a+c)}{n}$
Yule's Q	$Q = \frac{ad-bc}{ad+bc}$
Odds Ratio - OR	$OR = \frac{ad}{bc}$

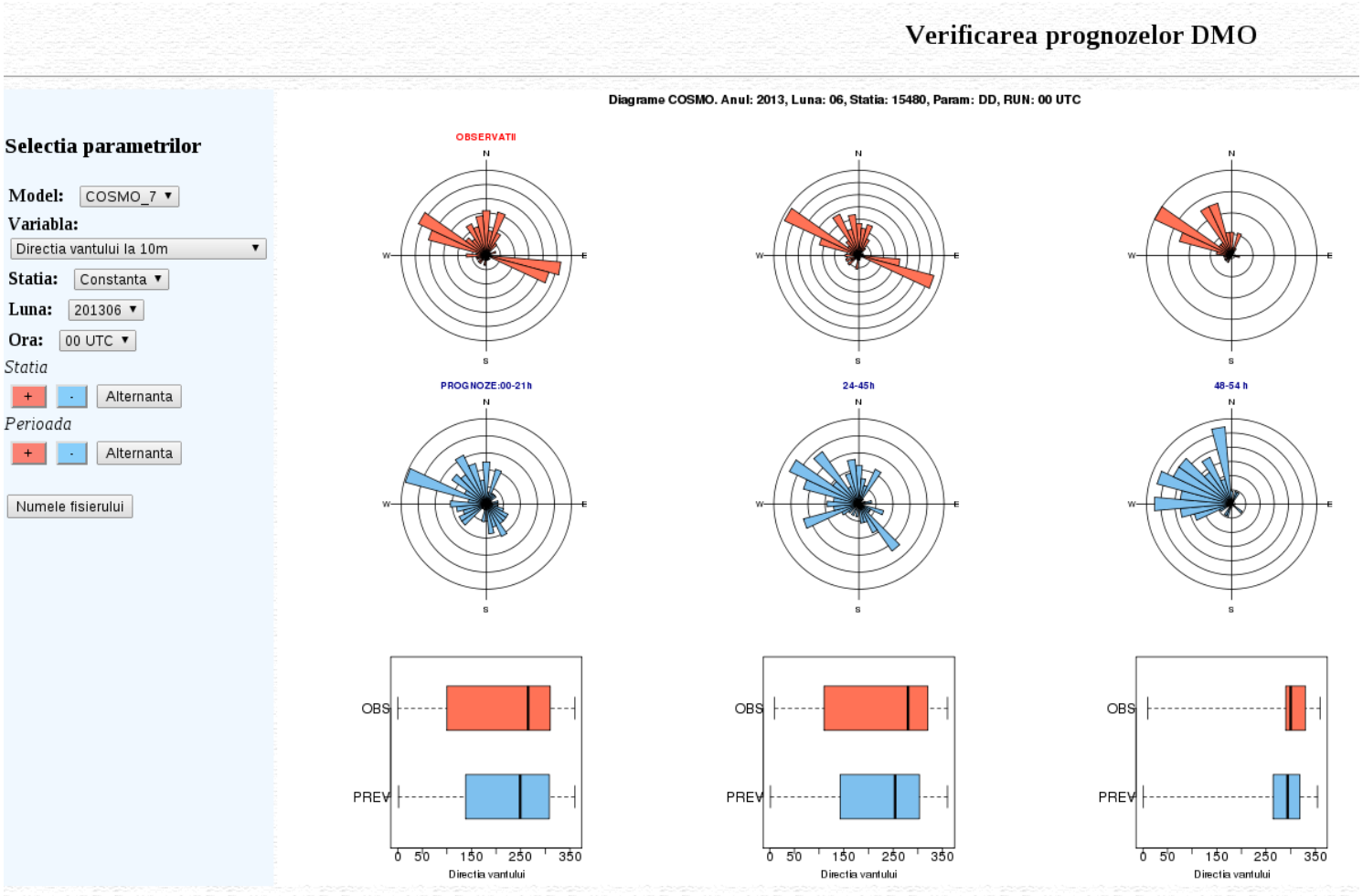
DMO verification- results -

Diagrams-examples



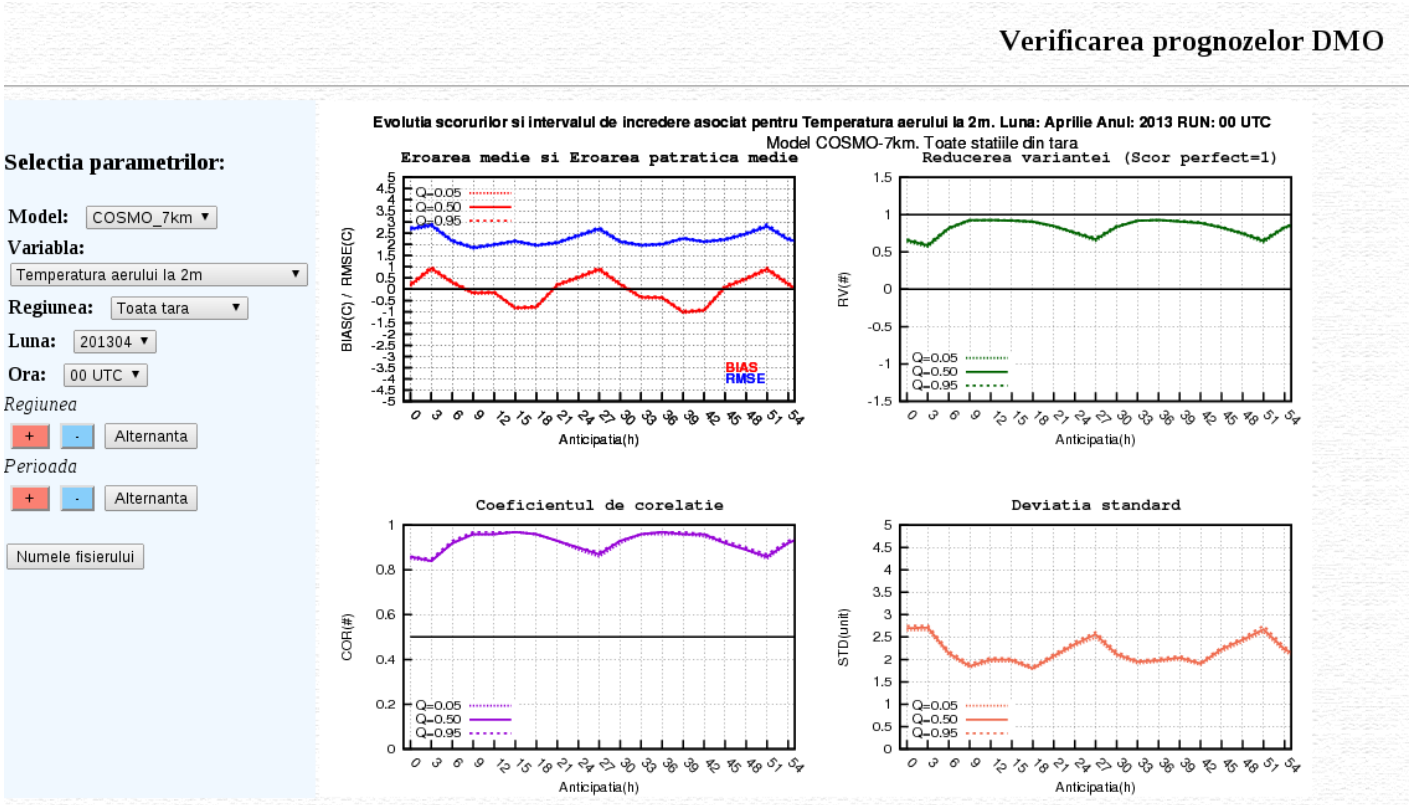
DMO verification- results

Diagrames descriptive-examples



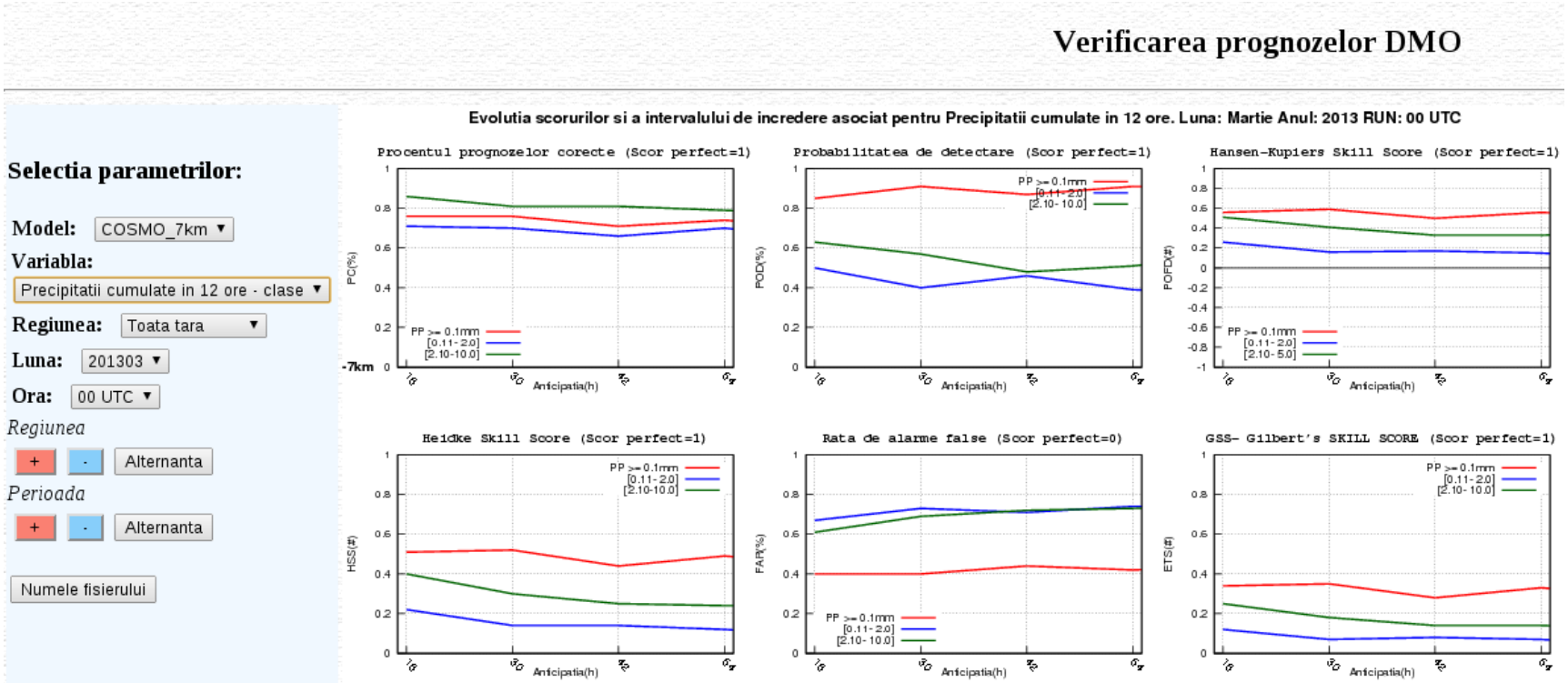
DMO verification- results

Monthly scores-examples



DMO verification- results

Monthly scores-examples



DMO verification- results

Monthly scores -examples

Comparativa DMO

Selectia parametrilor:

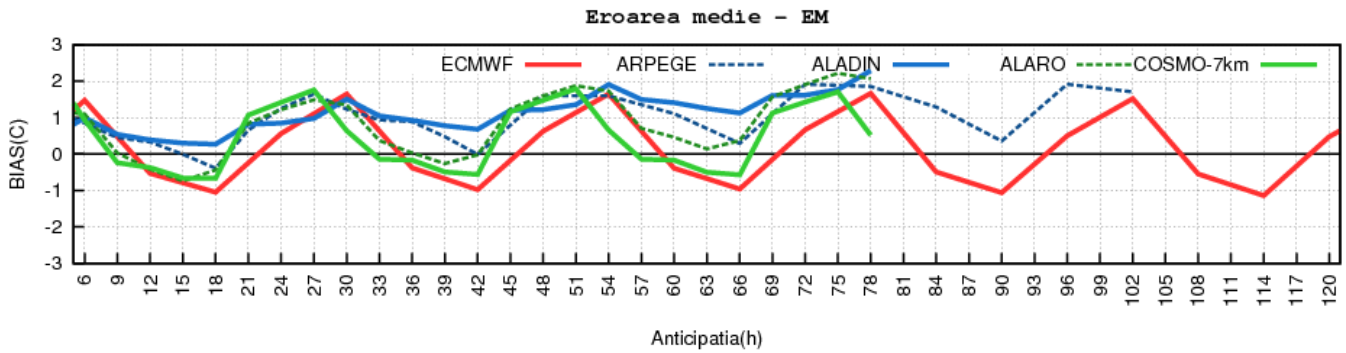
Variabla: Temperatura aerului la 2m

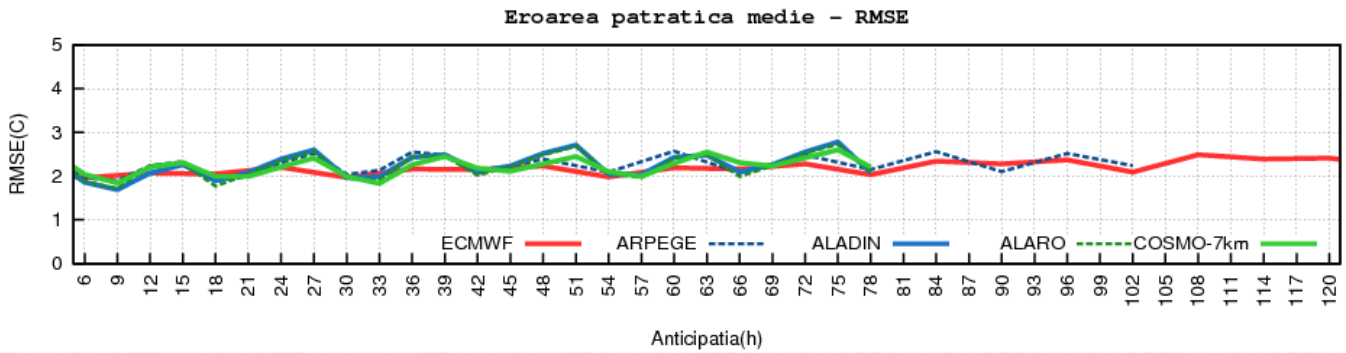
Luna: 201307 **Ora:** 00 UTC

Perioada + - Alternanta

Numele fisierului

Evolutia scorurilor pentru TS. Luna: 07 Anul: 2013 RUN: 00 UTC
Comparativa modele. Toate statiile din tara





DMO verification- results

Monthly scores -examples

Comparativa DMO

Selectia parametrilor:

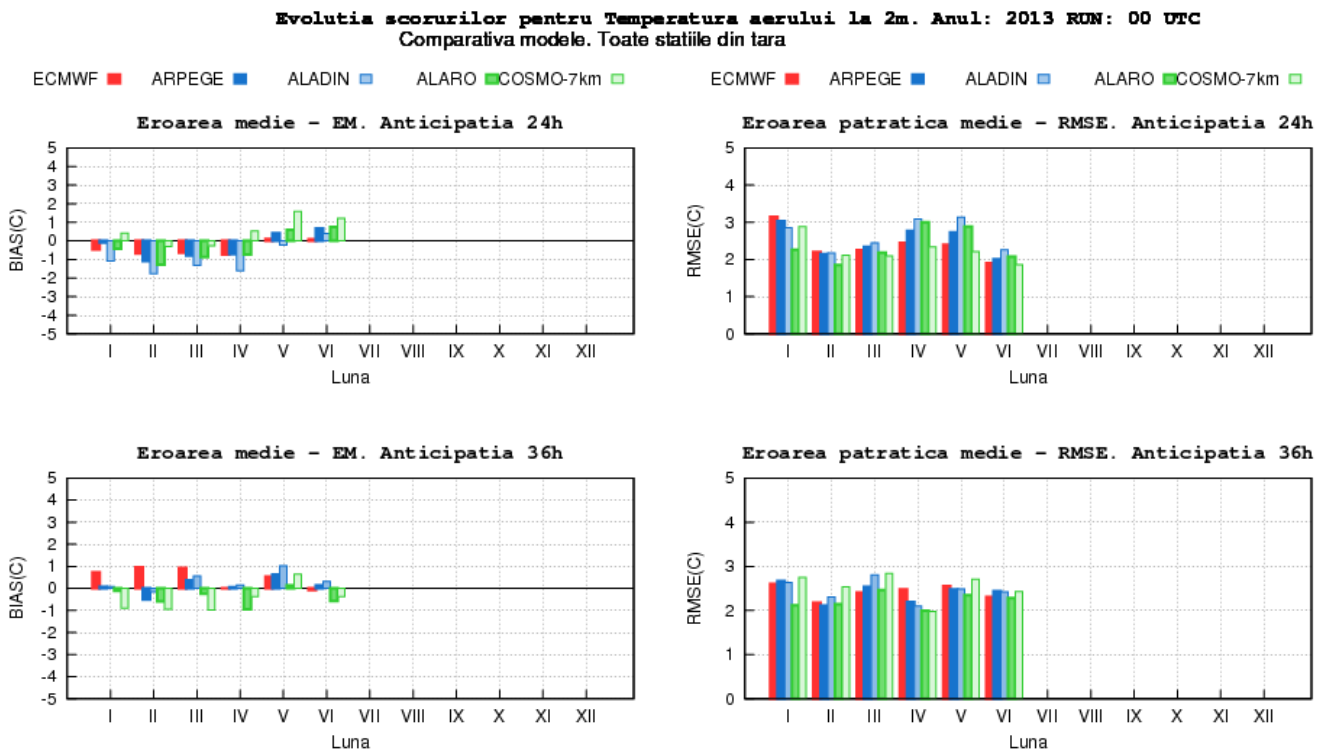
Variabla: Temperatura aerului la 2m

ANUL: 2013 **Ora:** 00 UTC

Perioada

\+ - Alternanta

Numele fisierului



SUMMARY

TO DO...

Statistical postprocessing

- calibration of MOS_EPS - some research using BMA
- use Kalman filter on MOS output (for temperature)
- more MOS_MIXTE - for now (ECMWF+ARPEGE) is in use
- to **improve efficient use of VarEPS**

Verification

- improve our **VERMOD** software
- database for verification
- MOS_EPS verification
- explore other verification methods: **MODE, Wawelets,**
- some research - with **MODE** - METv4.0-ncar, on precipitation, using ALADIN

ROMANIA

TO DO...

Statistical postprocessing

- calibration of MOS_EPS - some research using BMA
- use Kalman filter on MOS output (for temperature)
- more MOS_MIXTE - for now (ECMWF+ARPEGE) is in use
- to **improve efficient use of VarEPS**

Verification

- improve our **VERMOD** software
- database for verification
- MOS_EPS verification
- explore other verification methods: **MODE, Wawelets,**
- some research - with **MODE** - METv4.0-ncar, on precipitation, using ALADIN

ROMANIA

http://neptun.meteoromania.ro

Thank you for your attention !