Assimilation of 10 m wind from SYNOP observations over land

(Report)

Marián Jurašek
marian.jurasek@shmu.sk

Toulouse, France

14.05.-30.06.2003

1 Introduction

The effort of the objective and accurate forecast led to a big development of numerical weather models. These models solve a very complicated system of differential equations, but for successful forecast, they need a very accurate initial condition. The accuracy of the initial condition is improved by assimilation of various real atmosphere observations (e.g. SYNOP, TEMP, PILOT, SATEM, TOVS...). The development of assimilation started from optimal interpolation and continued through 3D-VAR to the 4D-VAR, and future will show what will be the next. The 4D-VAR incremental assimilation tries to find initial condition as near to the observations in the space and time as possible, taking into account the restrictions originating from the numerical solving of the differential equations mentioned above.

The current 4D-VAR assimilation at Météo-France processes only the 10 m wind over sea from SHIP observations. The main aim of my stay was to find good (simple and objective) parameters to select the appropriate SYNOP 10 m wind observations for the assimilation of 10 m wind over land and consequently find the impact of this assimilation to the forecast.

2 Evaluated parameters

We used RMSE for (observation - guess) for all non blacklisted SYNOP stations and both U and V components of 10 m wind over the whole period from 1^{st} of January, 2003 to 23^{rd} of May, 2003. Departures (observation - guess) are stored in field **fg_depar** in **update** table of **ODB**. We used **MANDAODB** with the following sql select to gather data from **ODB** for all 4 (00, 06, 12, 18) terms and for each day from the processed period:

```
CREATE VIEW mandalay AS
SELECT varno, statid, date, time, fg_depar
FROM hdr, body, update
WHERE ((obstype=1) && ((varno=41) || (varno=42)) && ((obschar.codetype=11)
|| (obschar.codetype=14)) && !(status.blacklisted@body))
```

```
SORTBY statid
```

where **obstype** is observation type: 1 for SYNOP; **varno** is subobservation type: 41 for U wind component and 42 for V wind component; **statid** is the indicative of the station; **obschar.codetype** is kind of observation: 11 for SYNOP surface and 14 for SYNOP automatic. RMSE was calculated for each stations with different **statid** and one RMSE for all processed period. It was done by script **/users/obs/jurasek/odb/pracuj.ksh** on machine **obs2** and by my own c program **/users/obs/jurasek/odb/stat_synop.c** on the same machine.

Statistic relation was evaluated for three parameters: interpolated standard orography deviation to the observation point, interpolated model orography to the observation point and SYNOP station altitude.

2.1 Model orography

The value of this parameter for observation point is stored in the field **modoro** in **hdr** table of **ODB**. Following sql select was used:

```
CREATE VIEW mandalay AS
SELECT statid, modoro
FROM hdr, body, update
WHERE ((obstype=1) && ((varno=41) || (varno=42)) && ((obschar.codetype=11)
|| (obschar.codetype=14)) && !(status.blacklisted@body))
SORTBY statid
```

This data was joined with the RMSE data and the results are on the graph in Fig.1.

As Fig. 1. shows, RMSE values have the same structure for every value of the model orography. We can't choose any interval of model orography, where RMSE values are significantly lower than in the others parts. Model orography is not suitable objective parameter for our purposes.

2.2 Station altitude

In other words: real orography. The value of this parameter for the observation point is stored in the field **stalt** in **hdr** table of **ODB**. Following sql select was used:

```
CREATE VIEW mandalay AS
SELECT statid, stalt
FROM hdr, body, update
WHERE ((obstype=1) && ((varno=41) || (varno=42)) && ((obschar.codetype=11)
|| (obschar.codetype=14)) && !(status.blacklisted@body))
SORTBY statid
```

This data was joined with the RMSE data and the results are at Fig.2.

As Fig.2 shows, RMSE values have the same structure for every value of the station altitude. We can't choose any interval of station altitude, where RMSE values are significantly lower than in the others parts. Station altitude is not suitable objective parameter for our purposes.

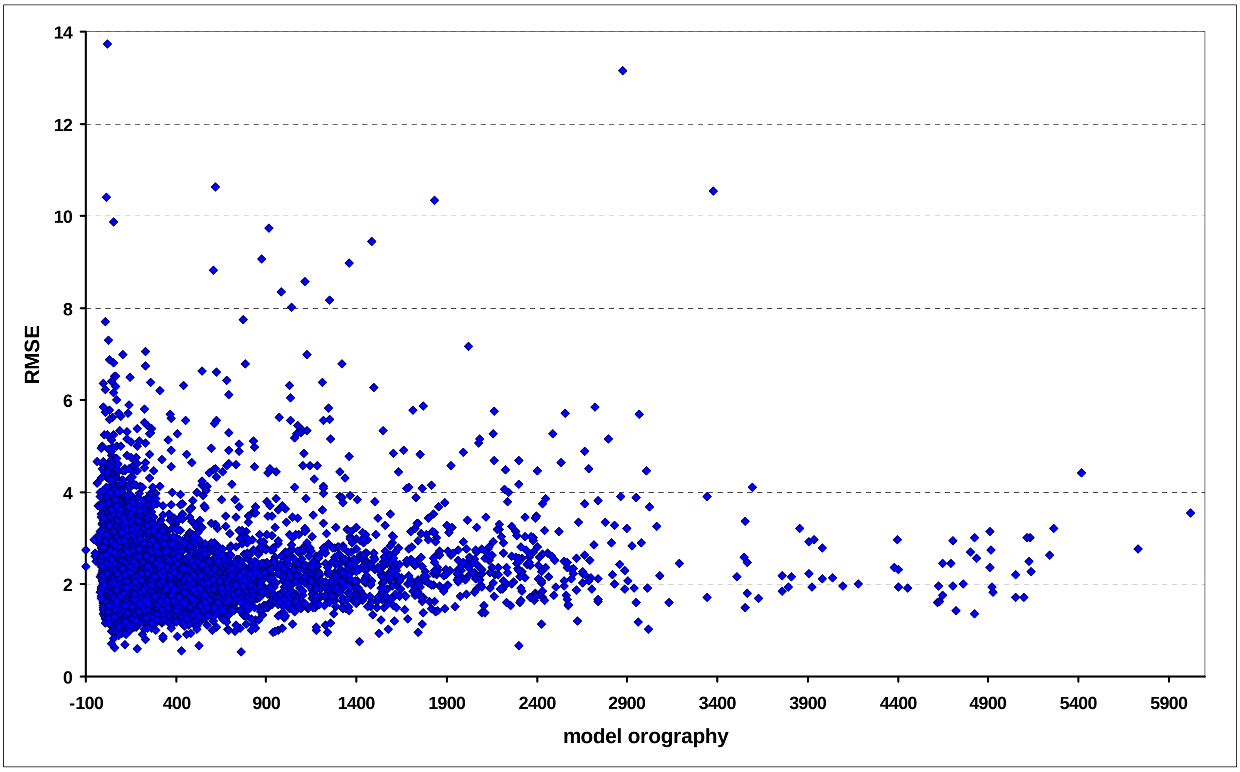

Figure 1: RMSE for 10m wind SYNOP observation at different model orography.

2.3 Standard orography deviation

This parameter is a little bit different from previous two, because normally it is neither stored in **ODB**, nor interpolated to the observation point. The values of this parameter are calculated in the model only for the grid points. On the Fig.3 we can see standard orography deviation in the Europe part of the ARPÈGE domain. The values are multiplied by **gz**. We had to modified the source code of the model to obtain this parameter. It was done on way of the smallest resistance. Our changes leaded to store the standard orography deviation in the field **modoro** in the **hdr** table of the **ODB**. We used already existing field for the ozone observations to pass our parameter from part where it is calculated to the part of the source code, where it is interpolated to the observation point and saved to the **ODB**. Following changes were made in the ARPÈGE source code based on **cy25t1_op3.03**:

- subroutine **arp/var/cobs.F90**
  This subroutine calls the computation of the standard orography deviation which is consequently stored in the **ZGP** field on the position **MVFGETRL** from **PTRGPD**. This value is copied to the ozone part of the buffer **PB1(JROF,ILEVO3)**. The output from **diff** from original shows changes:

```
71c71
<        &MVXPWP ,MVXORO ,MVXTPC ,MVFGETRL
---
>        &MVXPWP ,MVXORO ,MVXTPC
260,262c260
<!          PB1(JROF,ILEVO3) = PO3T(JROF,JLEV)
```

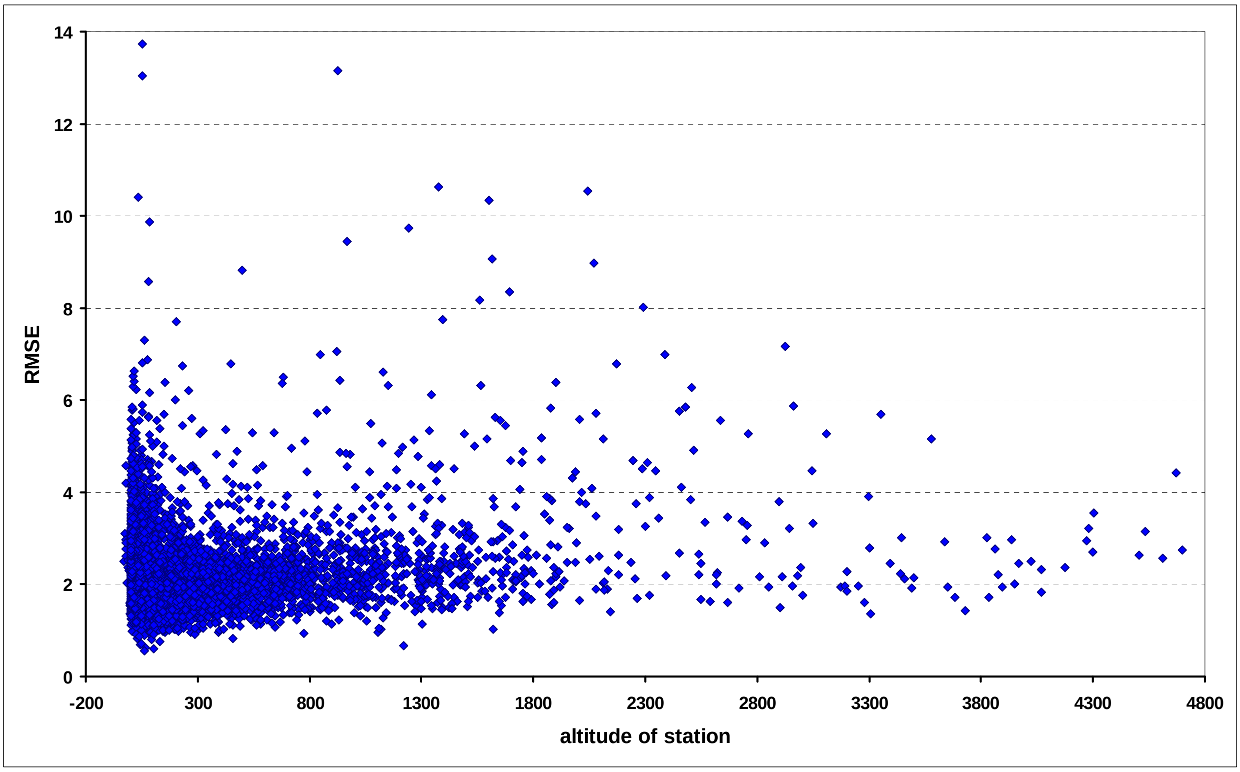

Figure 2: RMSE for 10m wind SYNOP observation at different altitude of SYNOP station.

```
<          PB1(JROF,ILEVO3) = _ZERO_
<          PB1(JROF,ILEVO3) = ZGP(JROF,MVFGETRL)
---
>          PB1(JROF,ILEVO3) = PO3T(JROF,JLEV)
264d261
<!          PB1(JROF,ILEVO3) = _ZERO_
266,267c263
<          PB1(JROF,ILEVO3) = ZGP(JROF,MVFGETRL)
<        ENDIF
---
>        ENDIF
```

- subroutine **arp/pp_obs/mpobseq.F90**
  "New ozone part" of the buffer is written to the global ozone array **GSMO3** from **YOMMVO**:

```
499a500,504
<!       jurasek for orogvar
<        DO J=1,NFLEVG
<              GSMO3 (J,INUM) = PSLO3 (IOBSL,J)
<        ENDDO
<!       jurasek orogvar
```

- subroutine **arp/obs_preproc/sualobs.F90**

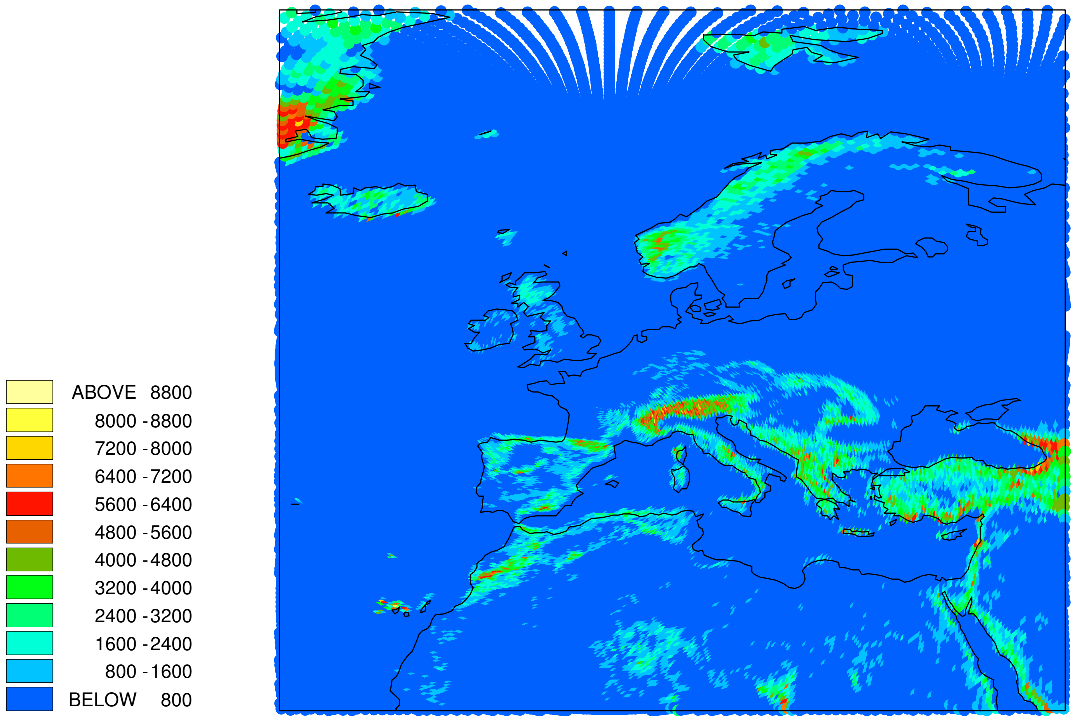

Figure 3: Standard orography deviation in the Europe part of the ARPÈGE domain. The values are multiplied by **gz**.

And, of course, **GSMO3** should be well-allocated for all of the observations:

```
191,192c191,193
<!        IF(LGSMO3) THEN
<!          ALLOCATE(GSMO3 (NFLEVG,NOBTOV))
<         ALLOCATE(GSMO3 (NFLEVG,NOBTOT))
- - -
>        IF(LGSMO3) THEN
>         ALLOCATE(GSMO3 (NFLEVG,NOBTOV))
194c195
<!        ENDIF
- - -
>        ENDIF
```

- subroutine **arp/pp_obs/preint.F90**
  The "ozone" field is prepared for the interpolation to the observation point:

```
259a260,263
<!        jurasek orogvar
<        PO3F5(JOBS,1:NFLEVG) = GSMO3(1:NFLEVG,IPOSM)
<!        jurasek orogvar
```

- subroutine **arp/pp_obs/slint.F90**
  The interpolation of the "ozone" field should be done for all observations:

```
447c447
<!       IF(LGOMO3.OR. LGSMO3) THEN
- - -
>       IF(LGOMO3.OR. LGSMO3) THEN
459c459
<!       ENDIF
- - -
>       ENDIF
```

- subroutine **odb/odbport/ctxinitdb.F90**
  We allow to update **modoro** field from **ODB** in the subroutine **HOP.F90**:

```
1167c1167
<       allocate(ctx(idctx,it)%view(1)%update_col(2))
- - -
>       allocate(ctx(idctx,it)%view(1)1168a1169
<       ctx(idctx,it)%view(1)%update_col(2) = 'modoro@hdr'
```

- subroutine **arp/pp_obs/hop.F90**
  The "ozone" field is written to the **modoro** part of **ODB** buffer:

```
352a353,358
<!       jurasek orogvar
<       ROBHDR(JOBS,MDBMOR)=ZO3F5(JOBS,NFLEVG)
<!       jurasek orogvar
```

- subroutine **arp/obs_preproc/prech.F90**
  Writing original model orography to the **modoro** part of **ODB** buffer should be disabled:

```
111,112c111,112
<!       jurasek
<!       ROBHDR(JOBS,MDBMOR) = ECOROG(IABNOB)/RG
- - -
>
>       ROBHDR(JOBS,MDBMOR) = ECOROG(IABNOB)/RG
```

Modified source code is on **kami** in the folder **PACK_orogdev**. Only one **SCREENING** with this modified source was run. The obtained **ODB** files were processed as in the case of model orography. The results are shown in Fig.4. The part with the biggest density is shown in the detail in Fig.5. As both figures show, RMSE values have the same structure for every value of the standard orography deviation. We can't choose any interval of standard orography deviation, where RMSE values are significantly lower than in the others parts. Neither standard orography deviation is suitable objective parameter for our purposes.

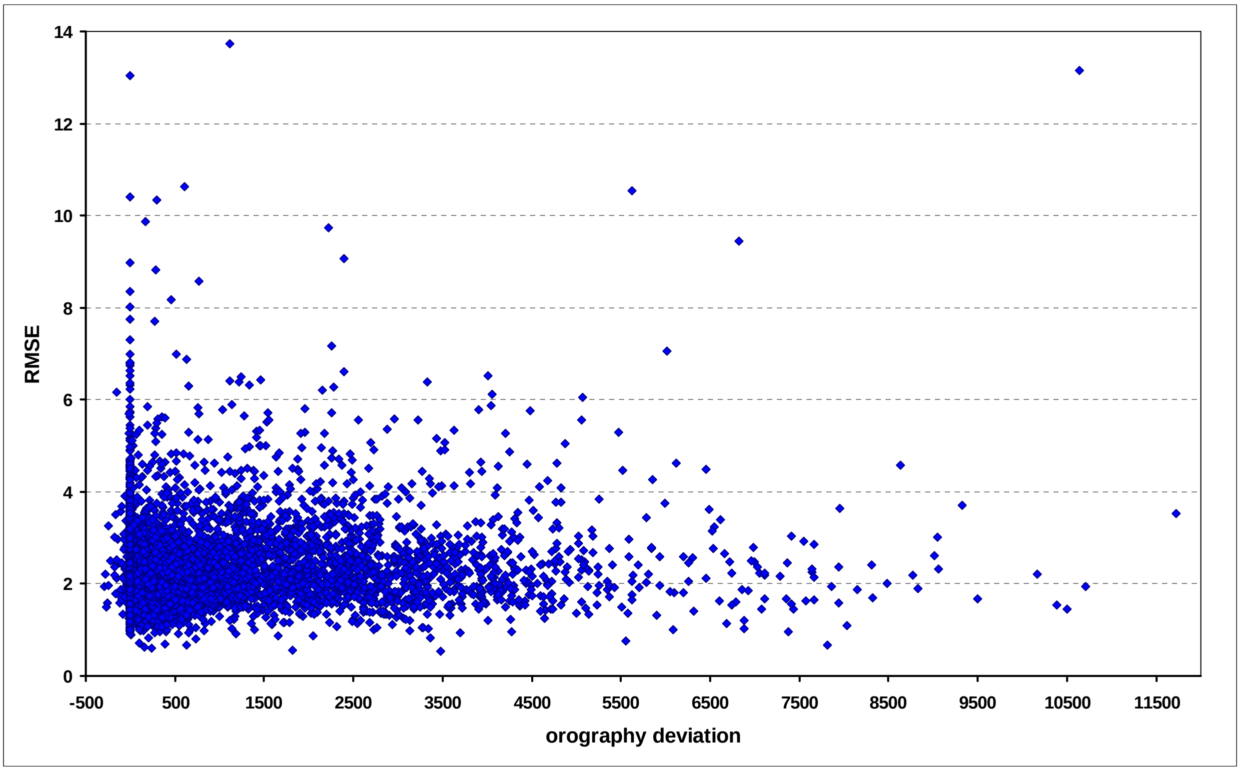

Figure 4: RMSE for 10m wind SYNOP observation at different standard orogprahy deviation. The vlues are multiplied by **gz**.

2.4 Blacklisting

We showed that none of our parameters is applicable and we had to change the strategy. We have processed RMSE values for the 143 days period and we have sufficiently big statistical data set. For some SYNOP stations we have more than 1700 observations of 10 m wind. We decided to prepare a **blacklist** of the SYNOP stations containing all stations, which RMSE was over the selected level or had less than 200 processed observations of the 10 m wind. The first try for determining the level of the RMSE was to take the value from the SHIP 10 m wid observations from the same period. Following sql select was used:

```
CREATE VIEW mandalay AS
SELECT varno, statid, date, time, fg_depar
FROM hdr, body, update
WHERE ((obstype=1) && ((varno=41) || (varno=42)) && ((obschar.codetype=22)
|| (obschar.codetype=23) || (obschar.codetype=24)) && (status.active@body))
SORTBY statid
```

where obschar.codetype=22, 23, 24 represent SHIP observations and **status.active@body** selects only active observations after screening. Obtained data was consequently processed

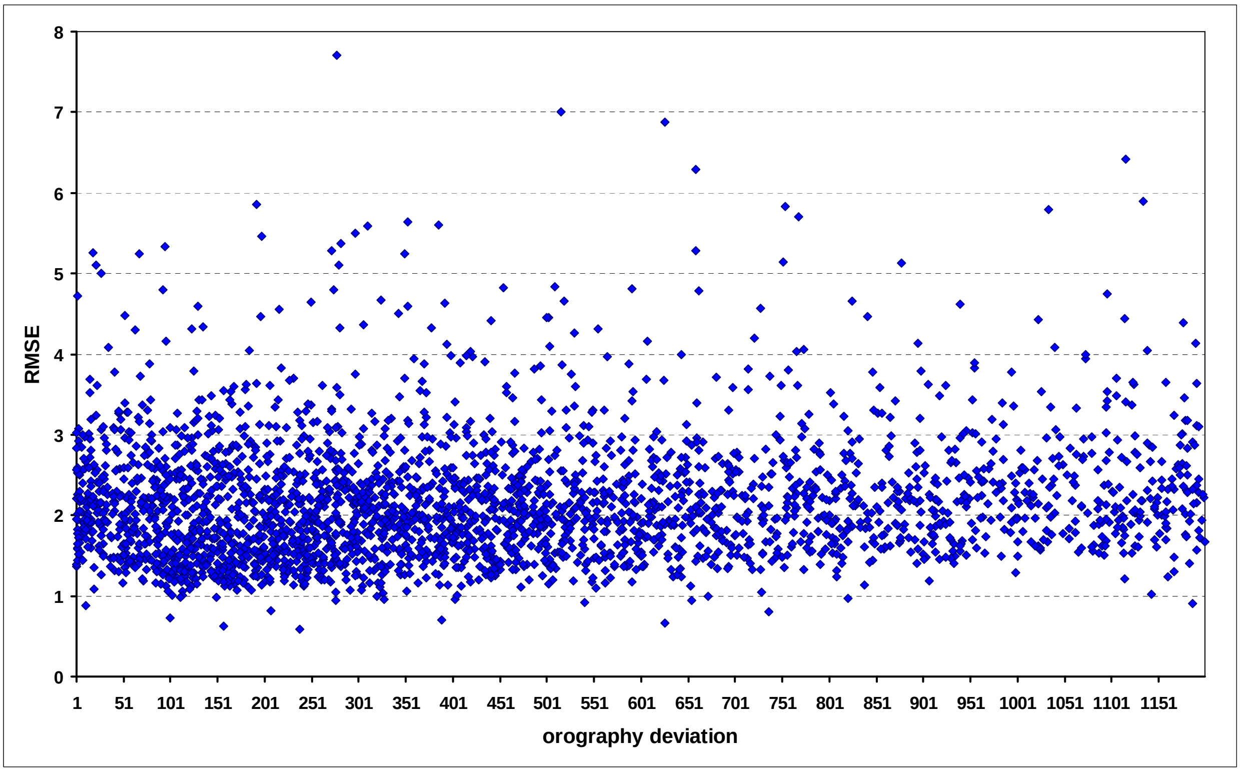

Figure 5: Detailed view of the most dense part (from 1 to 2000) of the previous graph of RMSE for 10m wind SYNOP observation at different standard orography deviation. The values are multiplied by **gz**.

like SYNOP data. Results are shown in Fig.6. We see that RMSE structure of SHIP observations of 10 m wind is very similar to the SYNOP RMSE structure. We cannot determine the limit value of RMSE. We set the limiting level of the RMSE in order to create a blacklist for SYNOP 10 m wind observations to value 4. Both mentioned conditions rejected 15

1. observation type code (1 for SYNOP)
2. observation type name (SYNOP)
3. observation kind code (11 for manual SYNOP, 14 for automatic SYNOP)
4. observation subtype code (41 for U component of 10 m wind, 42 for V component of 10 m wind)
5. SYNOP station indicative (e.g. 11518)
6. date of blacklisting (e.g. 19012003)

A small sample of the file is for example:

```
                        ⋮
1 SYNOP      14  41  80063   19012003
1 SYNOP      14  41  84143   19012003
1 SYNOP      14  41  13492   19012003
1 SYNOP      11  42  15469   19012003
```

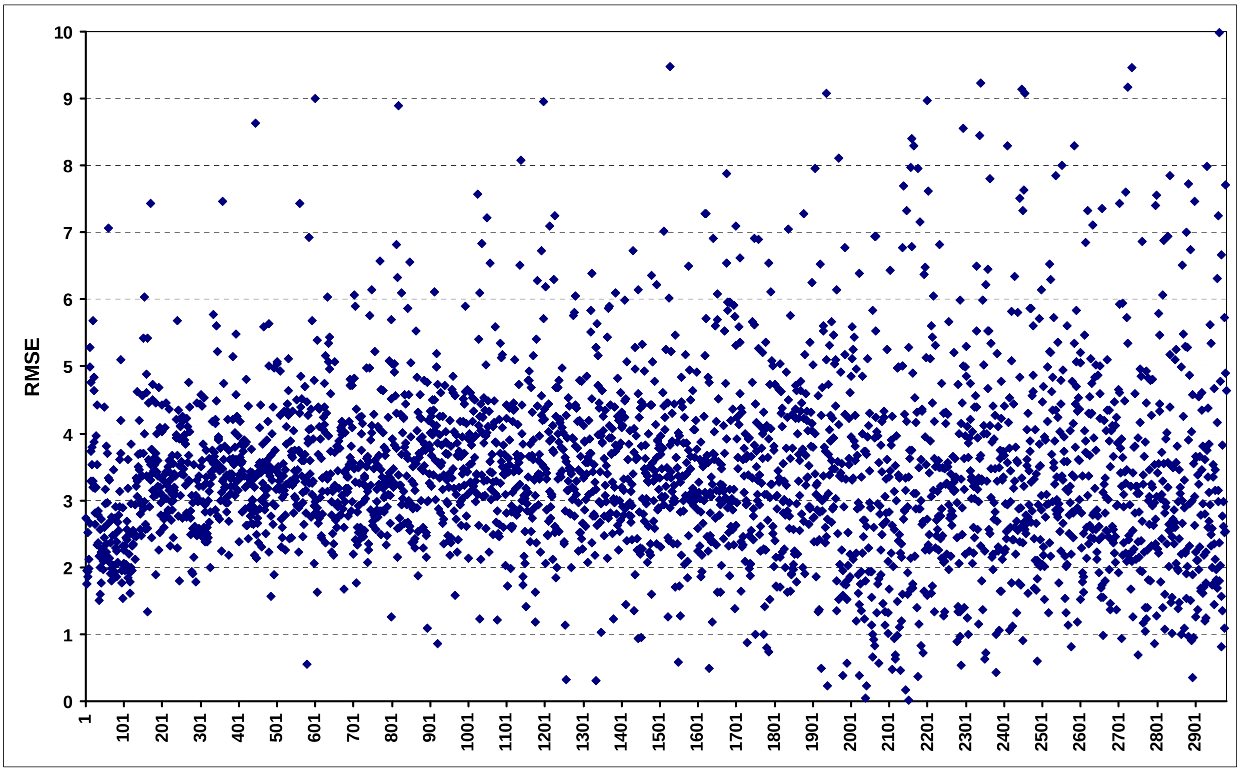

Figure 6: RMSE for 10m wind SHIP observations.

1 SYNOP	11	42	15434	19012003
1 SYNOP	11	42	67895	19012003
		⋮		

The whole file is on **kami:/u/gp/mrpa/mrpa682/blacklist_10m_wind** .

3 Experiment

3.1 Period selection

During our experiment we tested the impact of using our blacklist file to the scores of forecast. For this purpose, we decided to investigate a 15-day period. Daily RMSE for French SYNOP observations of 10 m wind was calculated from RMSE data for SYNOP stations. Why French SYNOP stations? We used these stations because there is the biggest density of gridpoints in France region. Results are in the Fig.7. For our experiment, the period with the biggest average RMSE, from 19^{th} January, 2003 00.00 UTC to 5^{th} February, 2003 00.00 UTC was selected.

3.2 Experiment description

Our experiment was based on ARPÈGE **cy25t1_op1.21** with 4D-VAR assimilation. We created it under **OLIVE** and its number is **0217**. Its short description is the following:

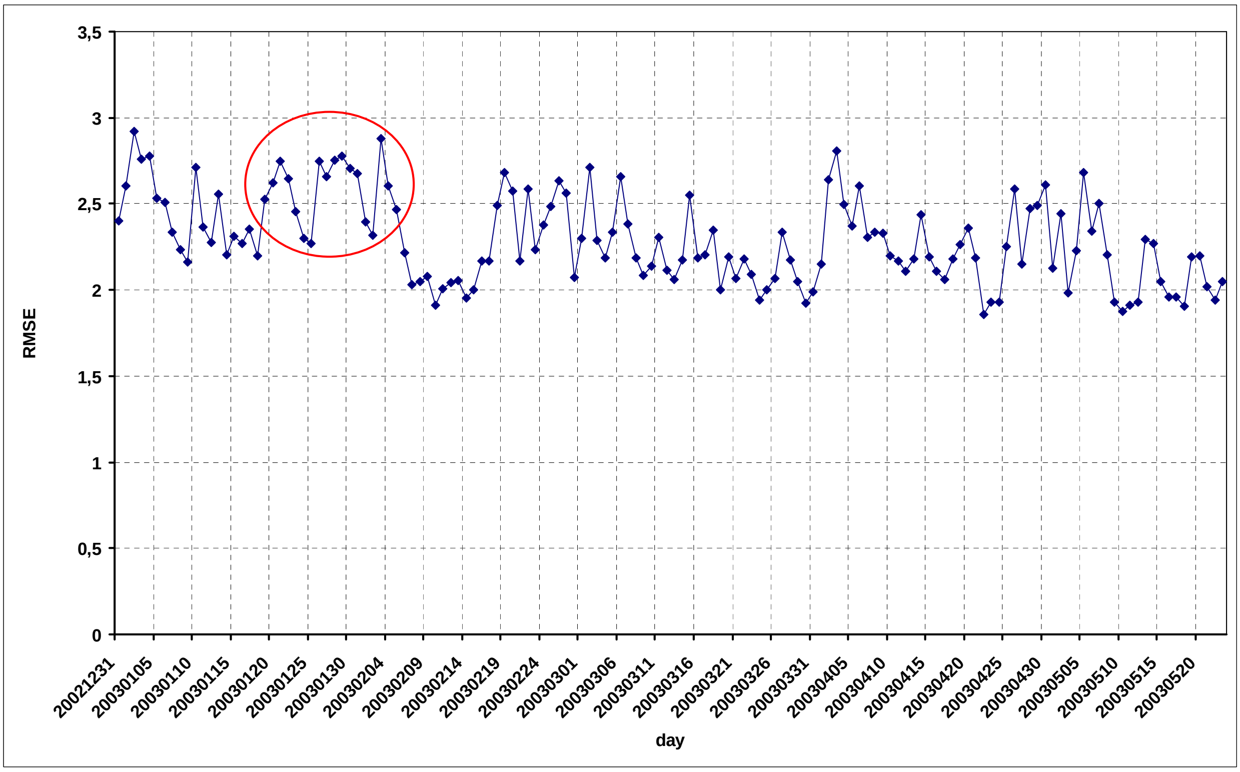

Figure 7: Daily RMSE for 10m wind SYNOP observation for French SYNOP stations. 15-day period with the highest RMSE (in red ellipse) chosen for next processing.

Activate production:	1
Production base hours:	00,12
Activate assimilation cycle:	1
Assimilation base hours:	00,06,12,18
Experiment begin hour:	00
Experiment end hour:	00
Taken from suite:	oper cy25t1_op1.21

The production was made for 96 hours and we added **batodb** to the observation preprocessing part. **Batodb** creates an **ECMASCR ODB** file from **OBSOUL**. Blacklist file is read during this creation and "blacklisting" information is put to the **ECMASCR ODB** as well.

3.3 Namelists modifications

We had to switch off the rejection of SYNOP 10 m wind observations over land made by default value of **LSLRW10=.TRUE.** in **NAMOBS** namelist. We had to change this value to **LSLRW10=.FALSE.** in namelists files for screening (**namelistscreen**) , minimization (**namelistmin1311**, **namelistmin1312**) and first update of trajectory (**namelisttraj1**, **namelisttraj2**) in **NAMOBS** section. The namelists we applied are stored in folder **/u/gp/mrpa/mrpa682/namelists** on **kami**.

3.4 Impact of 10 m wind assimilation

To see impact of 10 m wind assimilation to the accuracy of forecast the standard scores were calculated. And we compared them with scores of the operational forecast.

In Figures 8. - 12., we can compare which forecast was closer to the analyse. Green colour means that the forecast from our experiment was closer to our experiment analyse. For GEOPOTENTIAL (Fig. 8) we slightly improved long term prediction only for Australia region. For "rest" of the world, impact was zero or very little negative. Fig. 9 shows scores for TEMPERATURE. We can see only very small improvement in Asia region for long-term forecast. Score for HUMIDITY (Fig. 10) was absolutely not affected. Best improvement we obtained for wind (as we expected). In Fig. 11, we have score for WIND as vector, in this part direction of WIND is included. In Fig. 12, we have score for WIND SPEED. We improved forecast mainly for Australia and Asia region.

Different approach how to evaluate accuracy of forecast is compare forecast with observations. Fig. 13. - 17. show scores for TEMP observations. Green is again good. As we can see, impact on scores is slightly negative or neutral except for HUMIDITY. The worst impact to the forecast was recognised for GEOPOTENTIAL over Europe. Figs. 18. - 24. show scores for SYNOP observations. Values closest to the zero are the best. Green lines are scores for our experiment, violet lines are scores for operational forecast. Lines with "x" are for rout mean square error **RMSE**, lines without "x" are for **BIAS**. As figures show, scores for SYNOP were mainly not affected. In some region we little bit improved short term forecast for wind (up to 6 hours), Fig. 21. and 22.

4 Closure

Assimilation of 10 m wind over land from SYNOP observations didn't dramatically improve quality of forecast and fortunately didn't make our forecast worse. We had relatively better forecast only for Australia and Asia region. It can have various reason:

1. We used too strong criteria for creation our blacklist. It means, that all our observation are very close to guess and their assimilation not very affected analyse.
2. Test period was too short and probably without strong convection, which affects wind over Europe.

Taking into account all mentioned things I suggest next steps:

1. Prepare new experiment for testing with longer period and with convective cases.
2. Build new blacklist with more relaxes criteria.
3. Try to assimilate 10 m wind from SYNOP observations only over regions with positive impact - Australia and Asia.

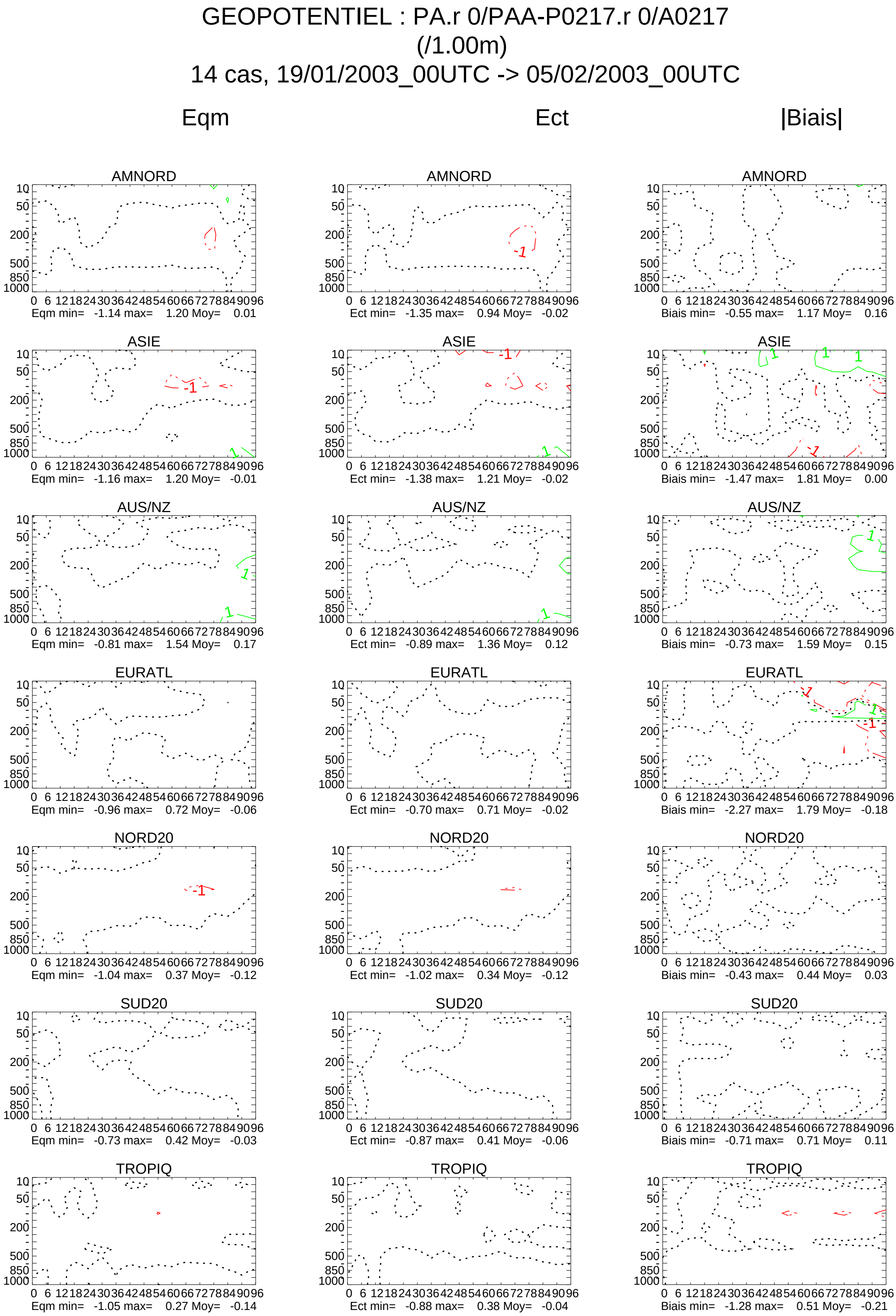

Figure 8: Scores for GEOPOTENTIEL - analyse vs. forecast, experiment vs. operational run.

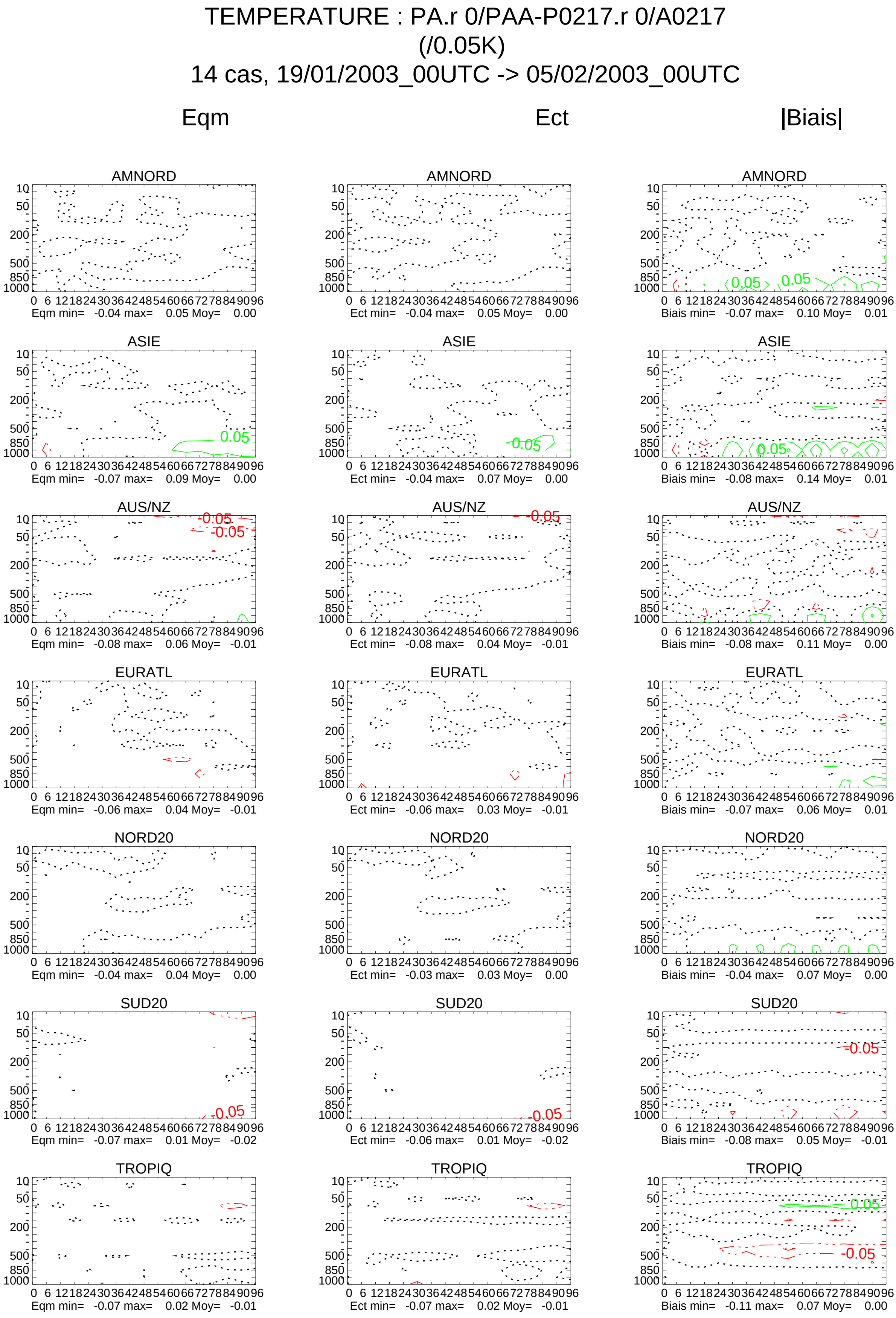

Figure 9: Scores for TEMPERATURE - analyse vs. forecast, experiment vs. operational run.

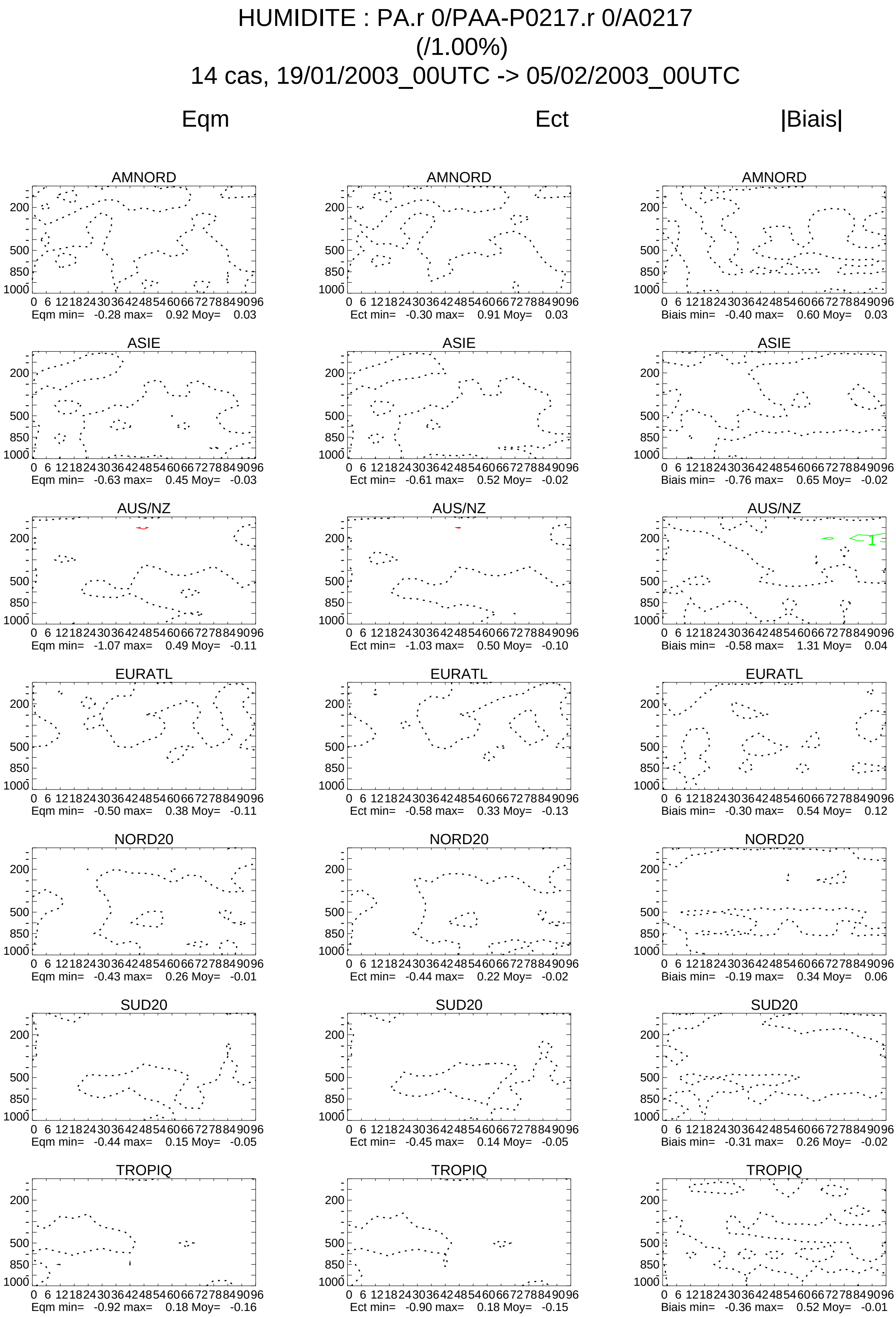

Figure 10: Scores for HUMIDITY - analyse vs. forecast, experiment vs. operational run.

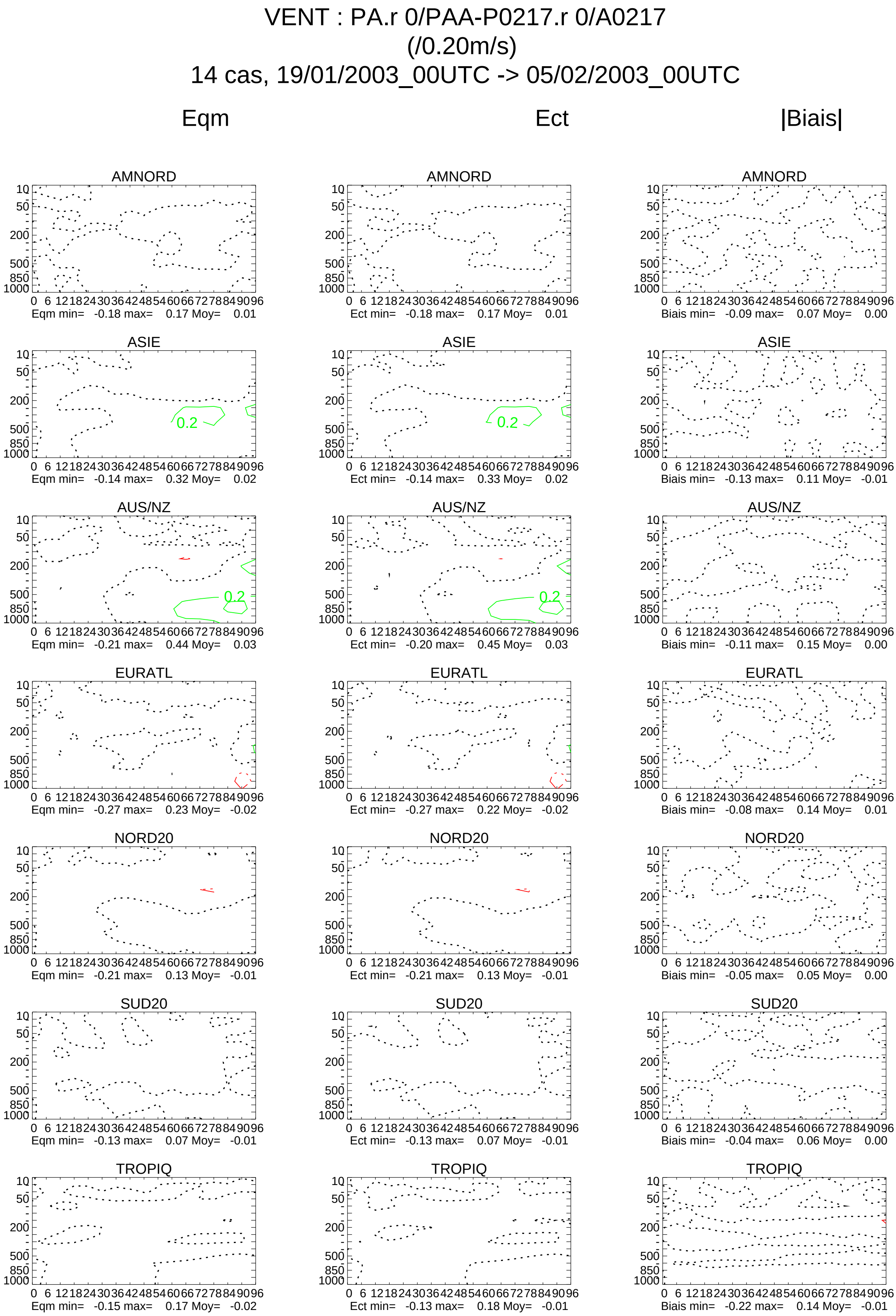

Figure 11: Scores for WIND (vector) - analyse vs. forecast, experiment vs. operational run.

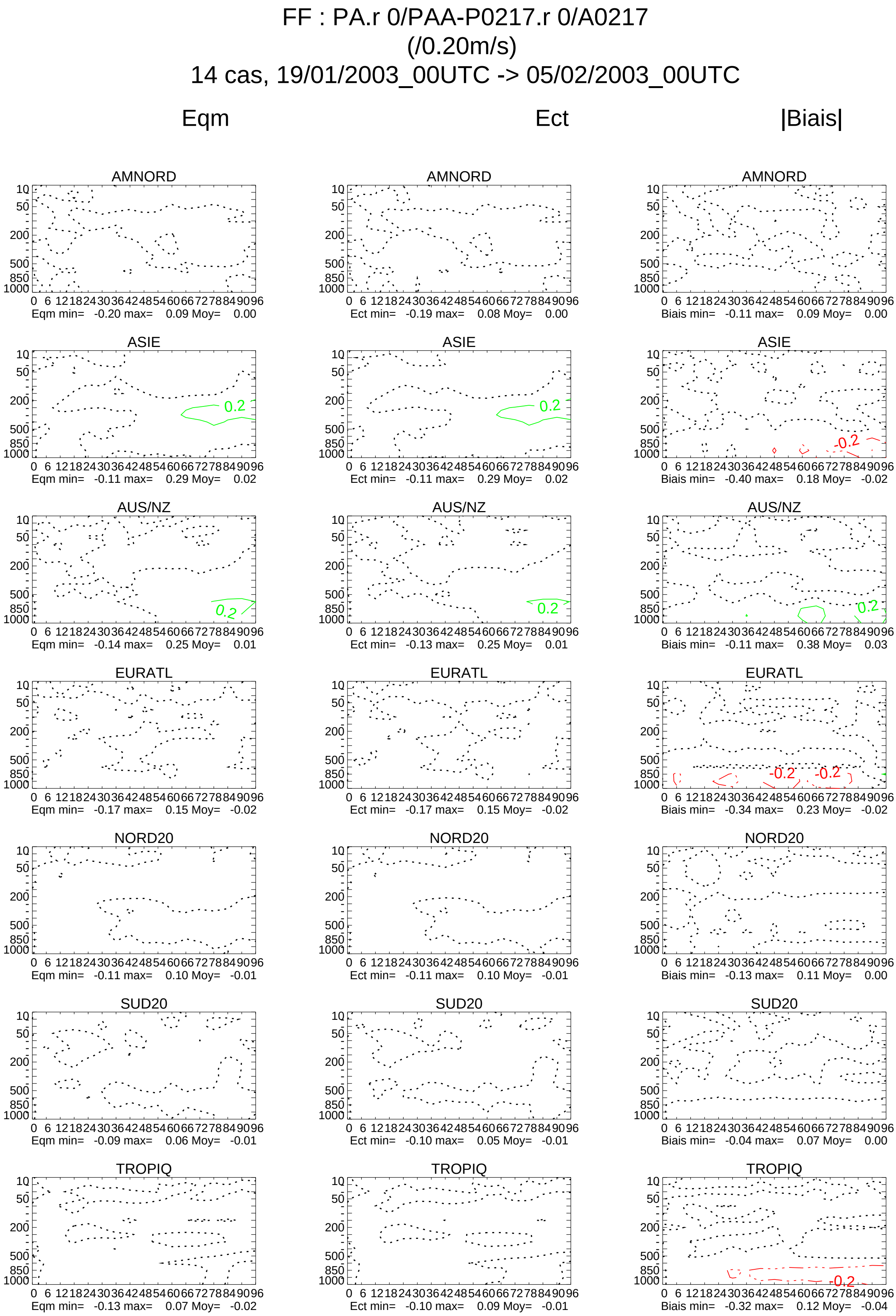

Figure 12: Scores for WIND (speed) - analyse vs. forecast, experiment vs. operational run.

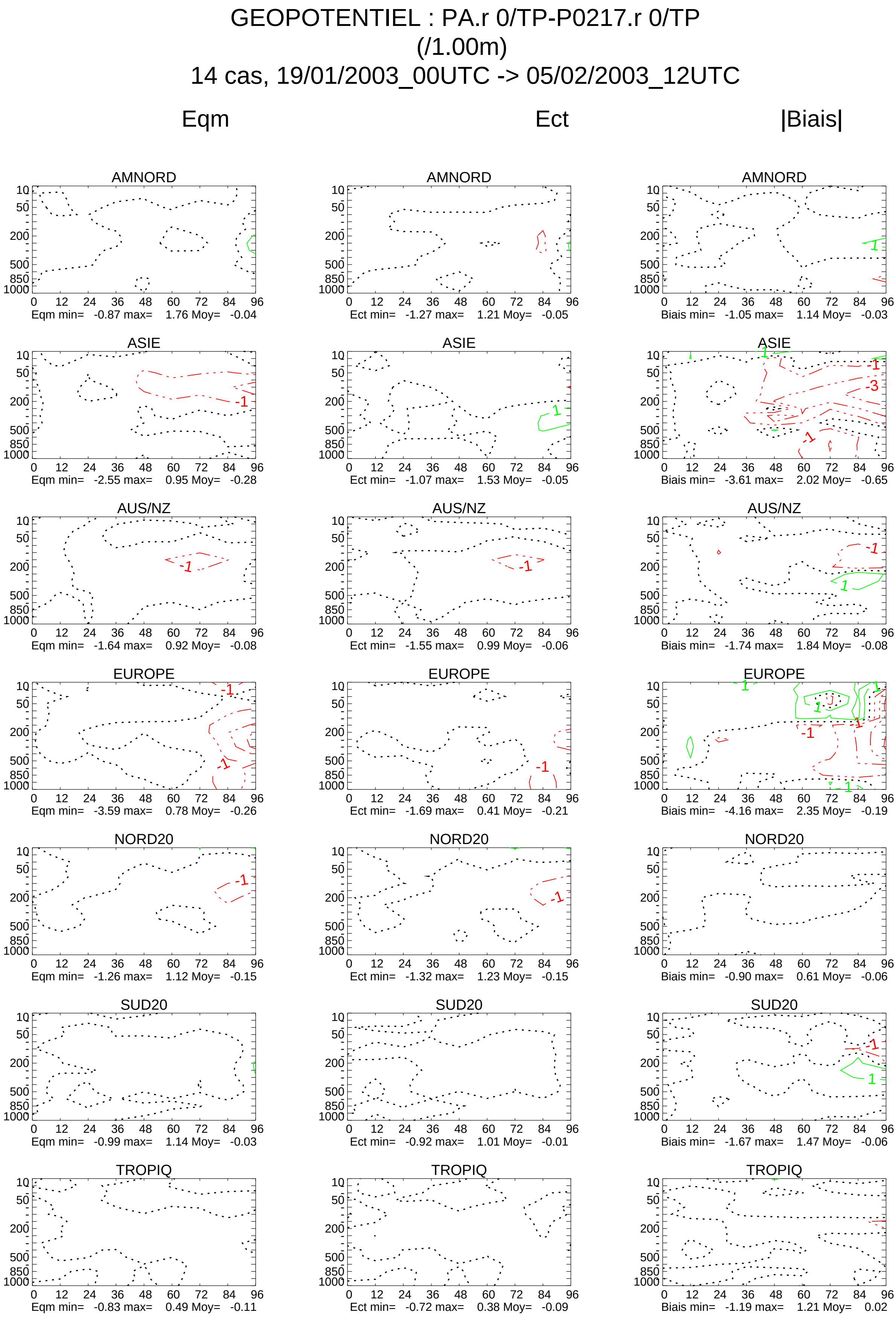

Figure 13: Scores for GEOPOTENTIEL - TEMP observation vs. forecast, experiment vs. operational run.

Figure 14: Scores for TEMPERATURE - TEMP observation vs. forecast, experiment vs. operational run.

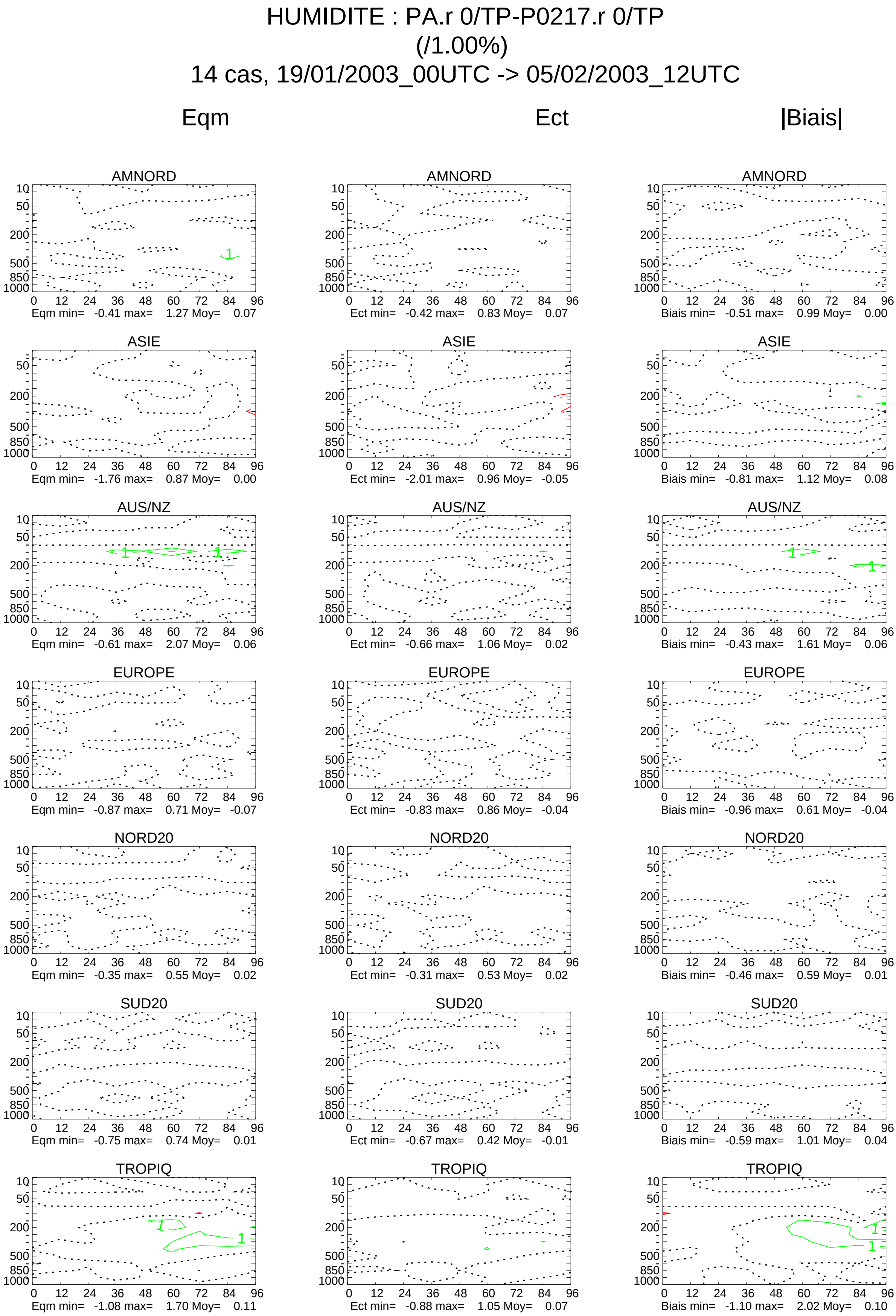

Figure 15: Scores for HUMIDITY - TEMP observation vs. forecast, experiment vs. operational run.

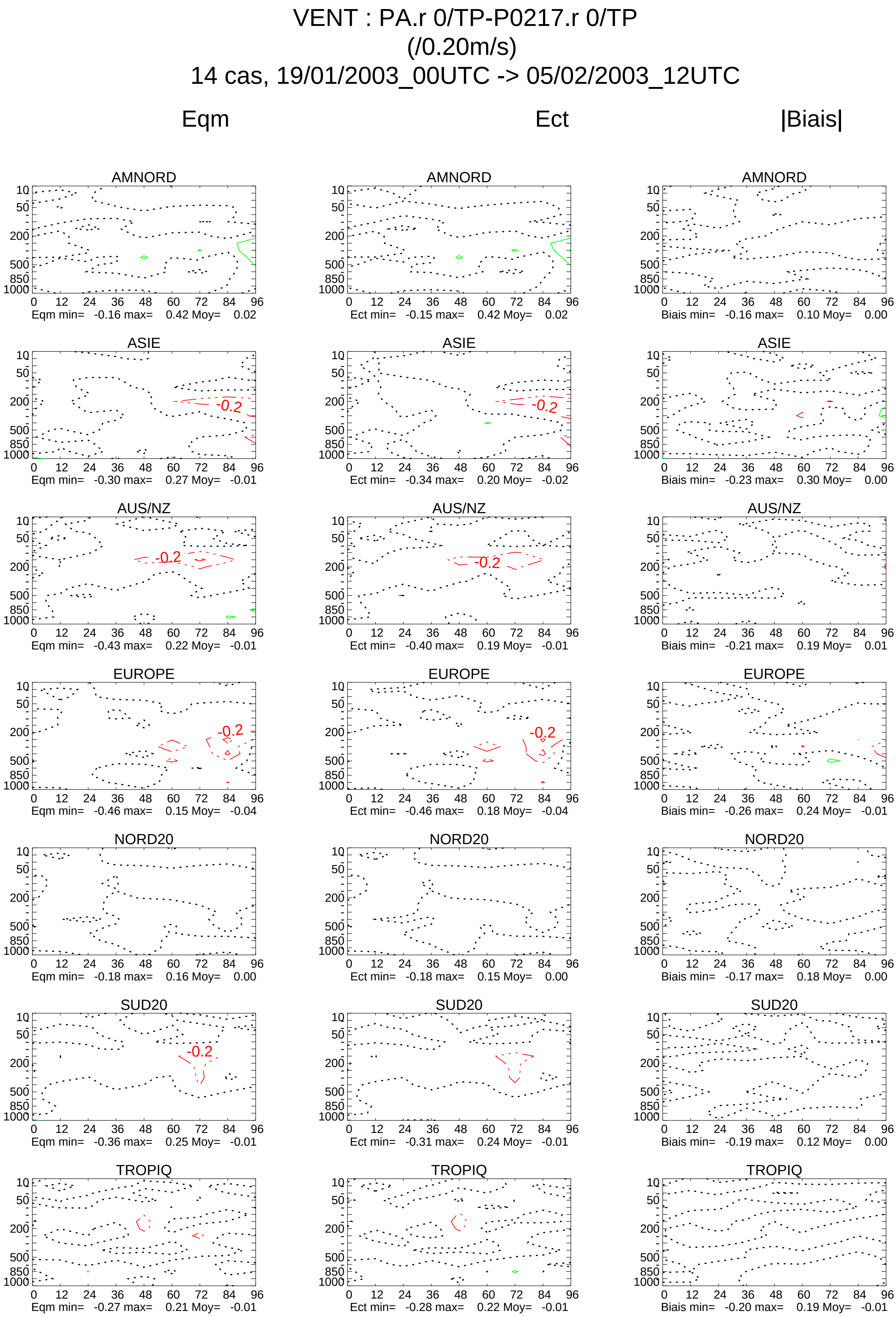

Figure 16: Scores for WIND (vector) - TEMP observation vs. forecast, experiment vs. operational run.

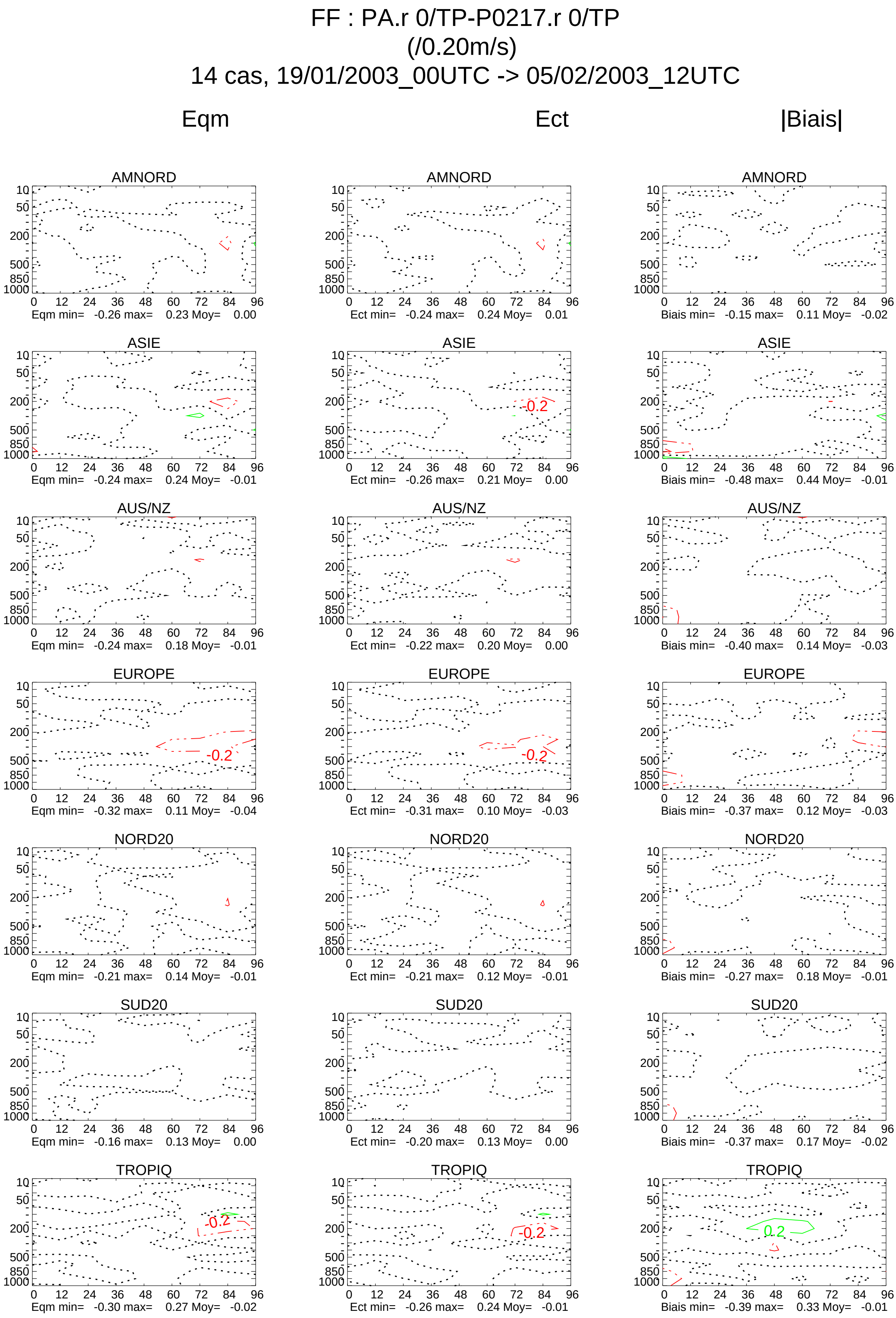

Figure 17: Scores for WIND (speed) - TEMP observation vs. forecast, experiment vs. operational run.

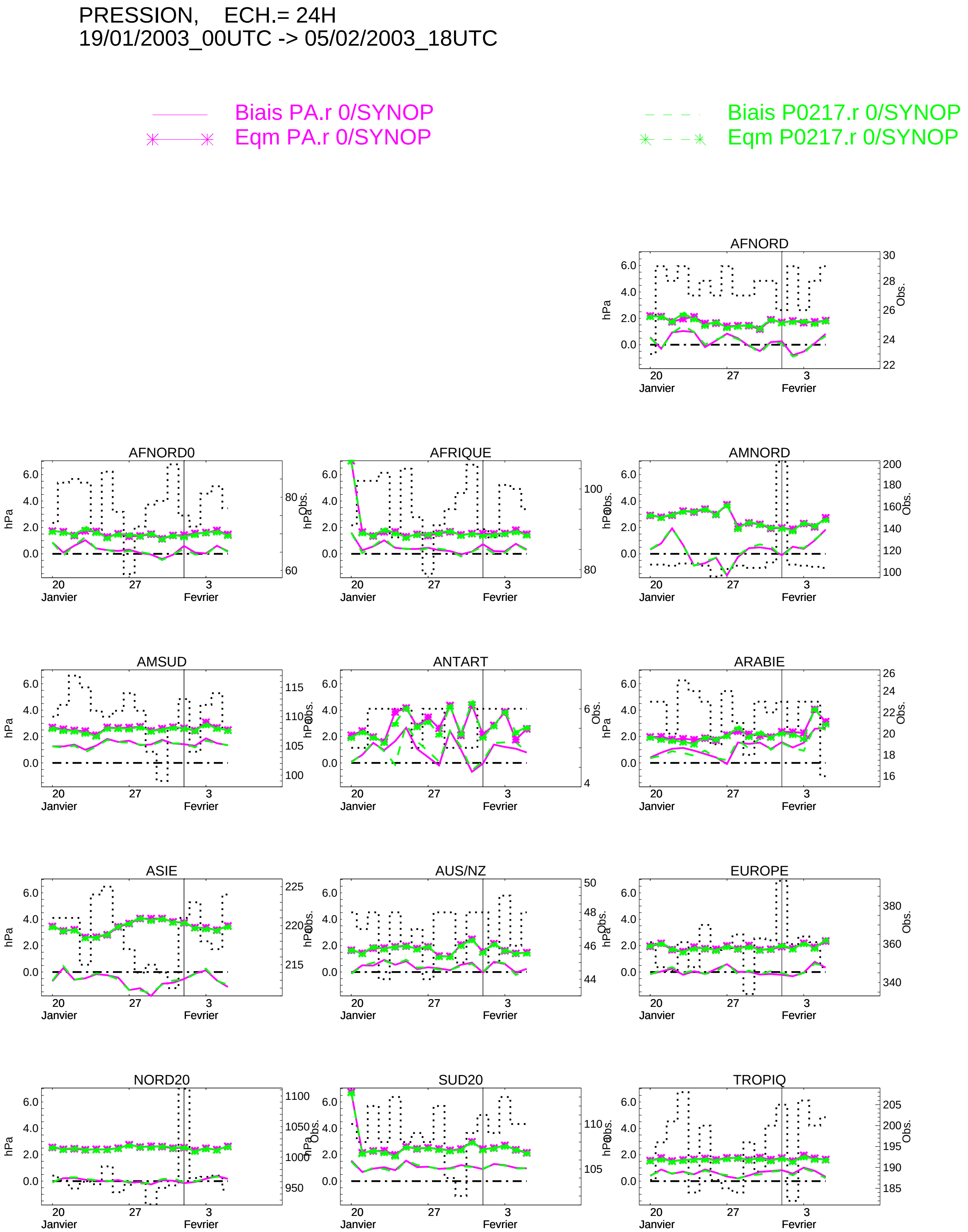

Figure 18: Scores for PRESSURE - SYNOP observation vs. forecast, experiment vs. operational run.

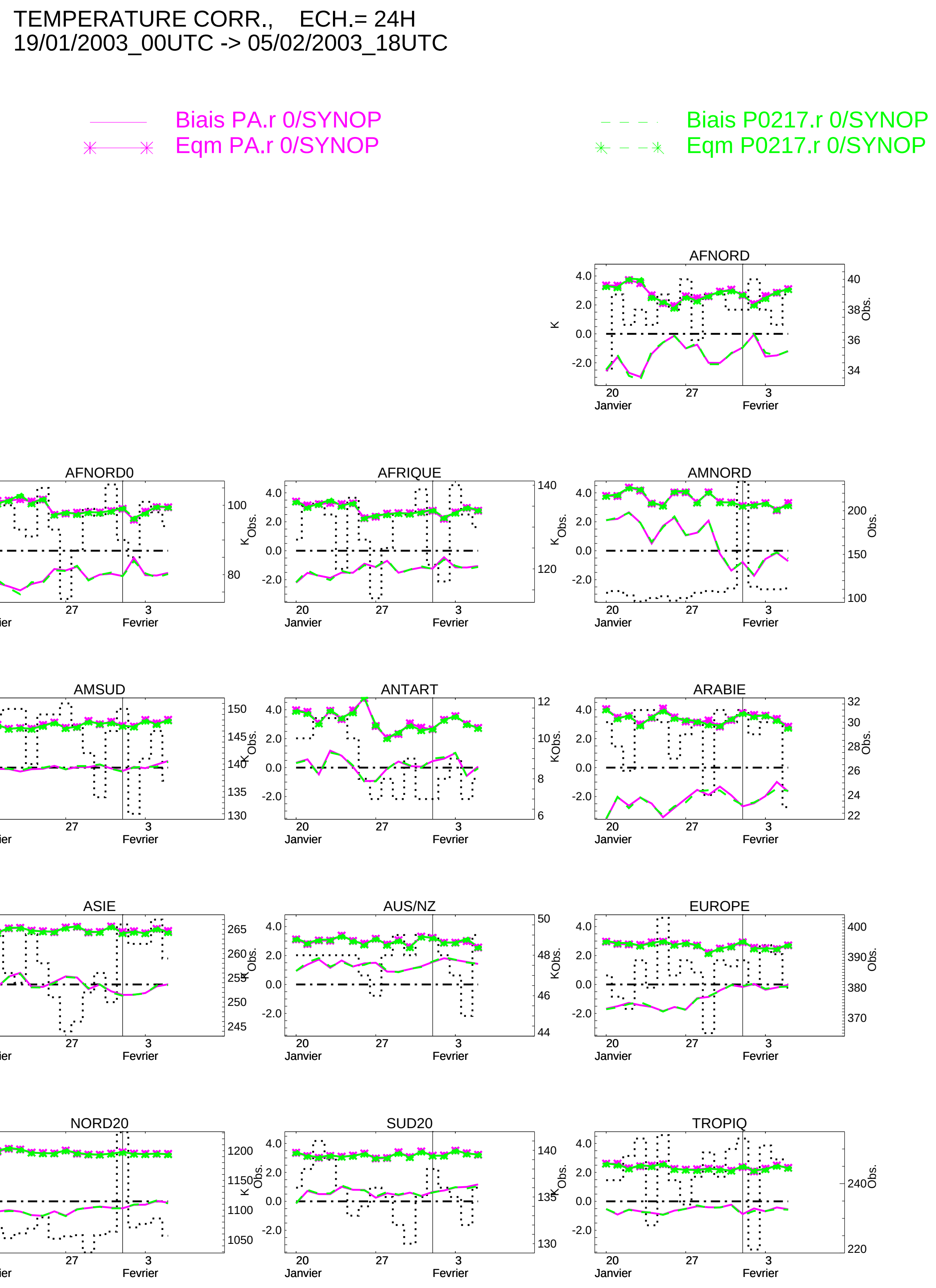

Figure 19: Scores for TEMPERATURE - SYNOP observation vs. forecast, experiment vs. operational run.

Figure 20: Scores for HUMIDITY - SYNOP observation vs. forecast, experiment vs. operational run.

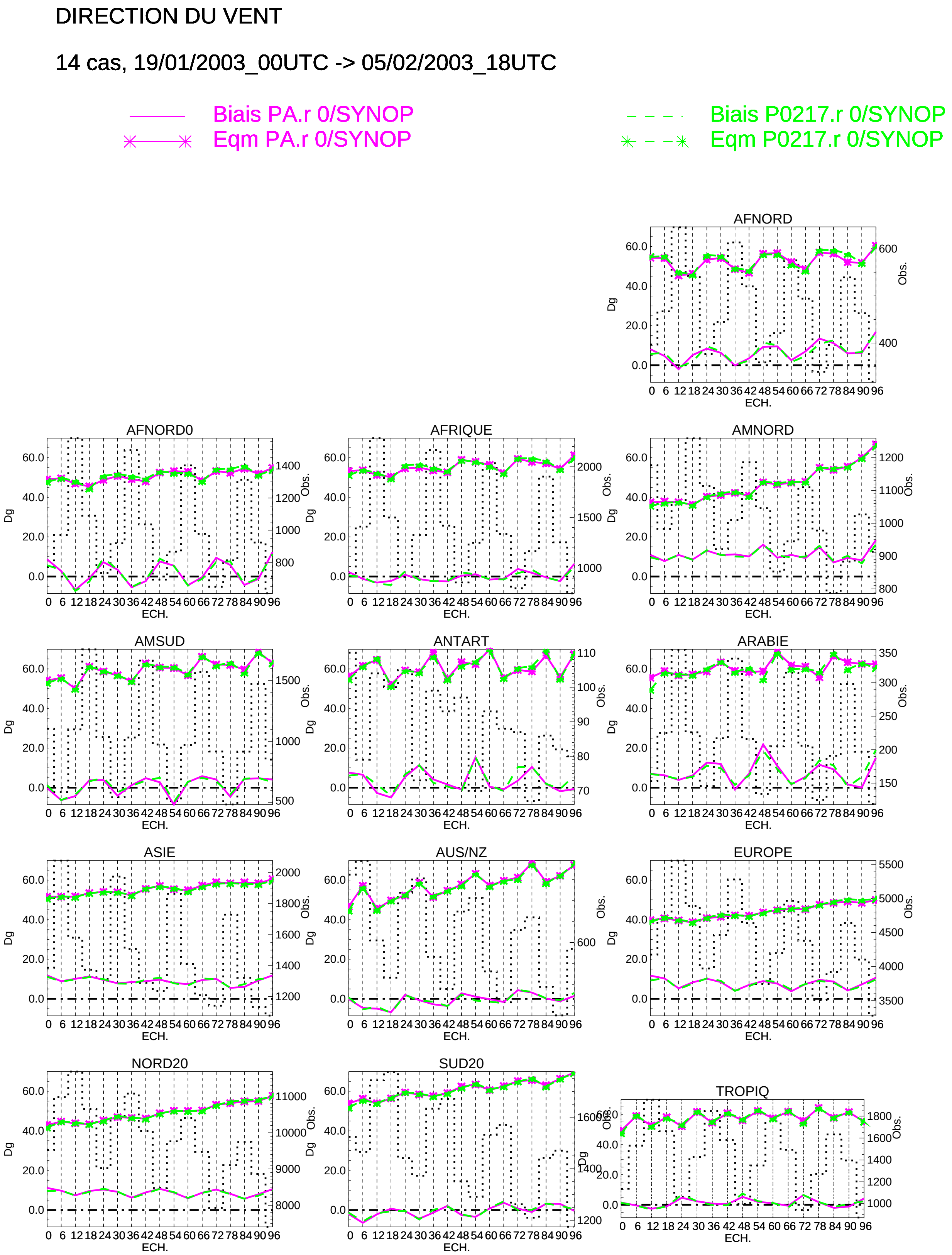

Figure 21: Scores for WIND DIRECTION - SYNOP observation vs. forecast, experiment vs. operational run.

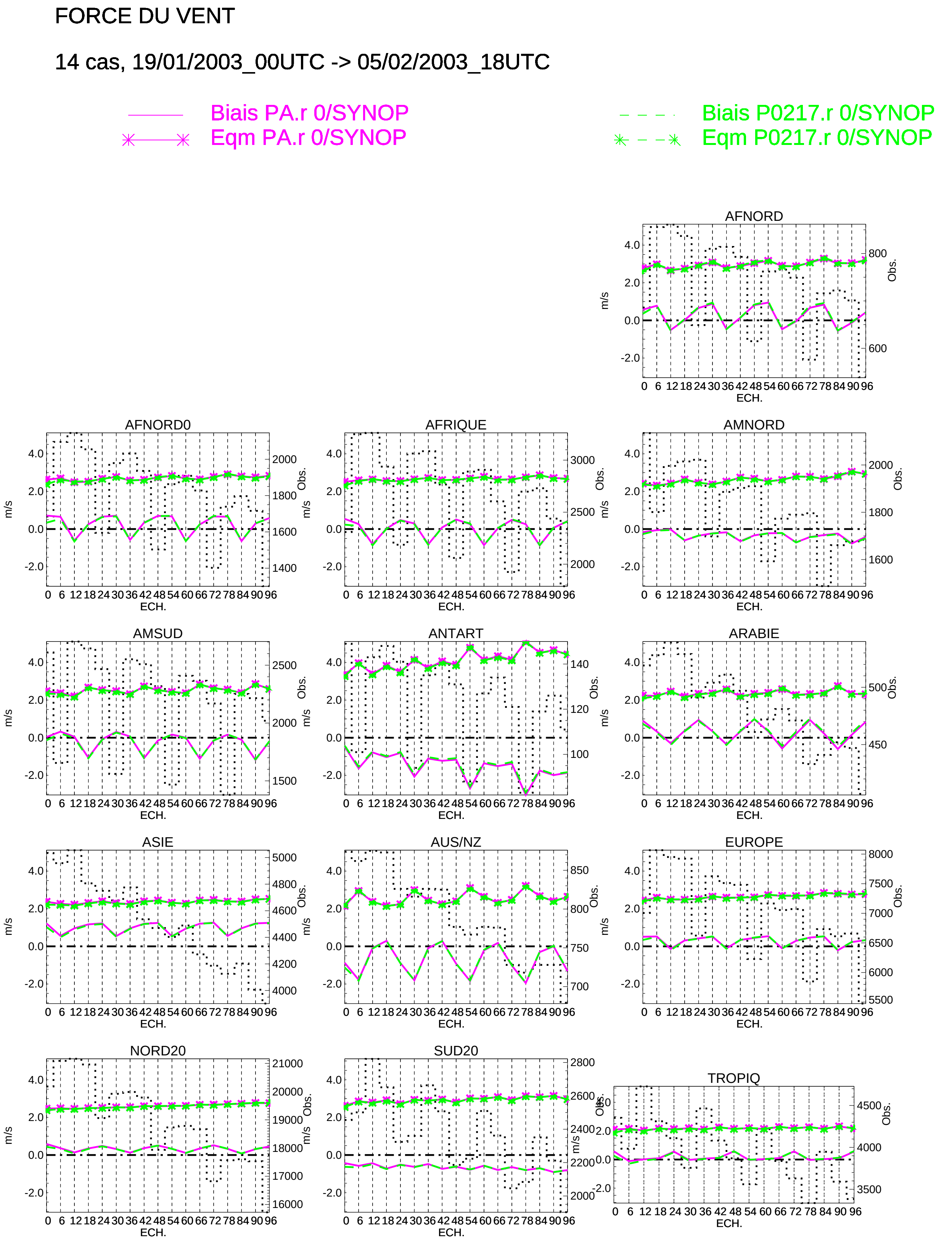

Figure 22: Scores for WIND SPEED - SYNOP observation vs. forecast, experiment vs. operational run.

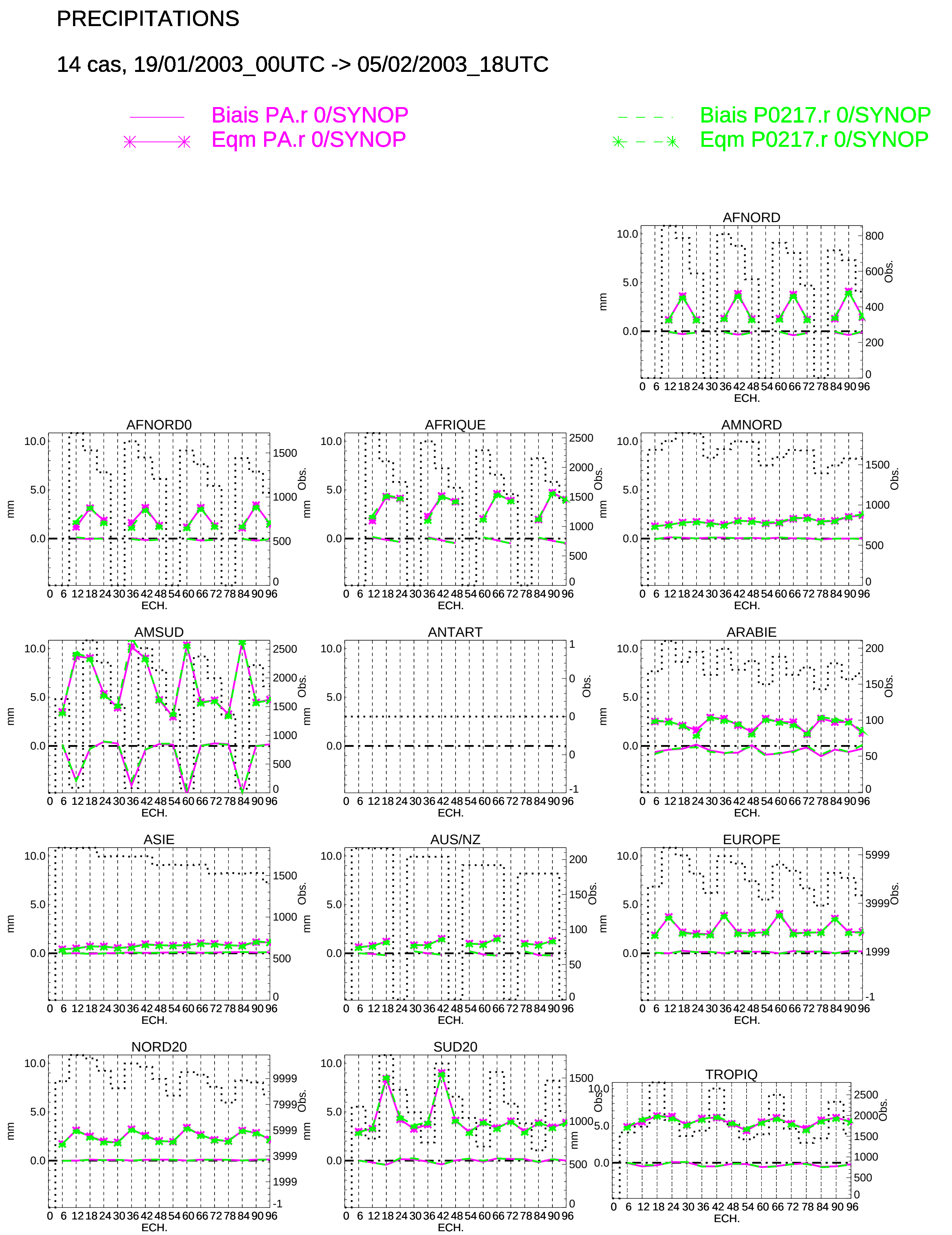

Figure 23: Scores for PRECIPITATION - SYNOP observation vs. forecast, experiment vs. operational run.

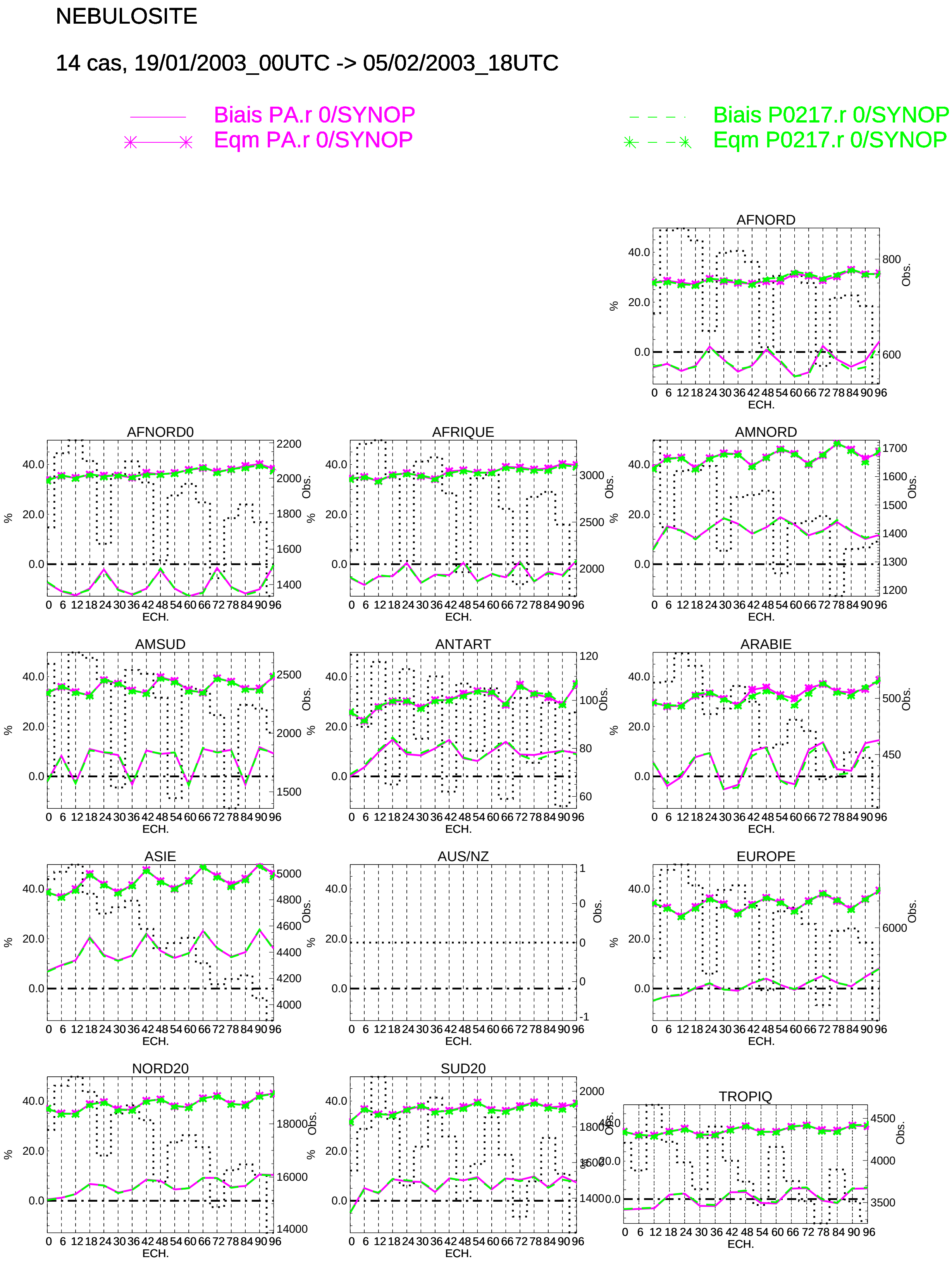

Figure 24: Scores for CLOUDINESS - SYNOP observation vs. forecast, experiment vs. operational run.