The NWP Test Suite, a COSMO tool for quality assurance

Massimo Milelli
and the
NWP TS team:
Montani, Gofa, Iriza, Bogdan

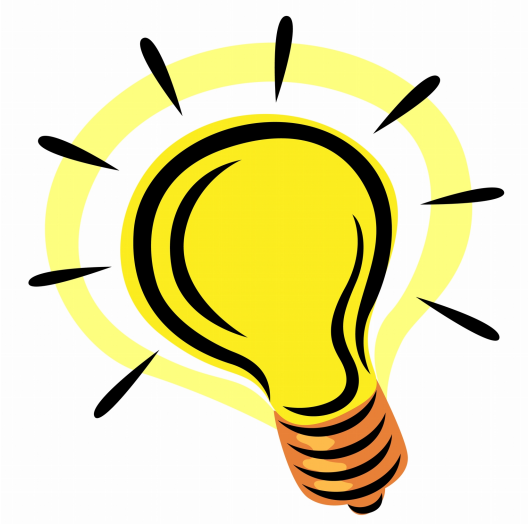
Outlook

- Quality check
- Goal of the meteorological suite
- Suite setup
- Verification setup
- Verification examples
- Next steps

Outlook

- ***Quality check***
- Goal of the meteorological suite
- Suite setup
- Verification setup
- Verification examples
- Next steps

A few words about the standard process of software development, from the initial idea to the final release of a new version



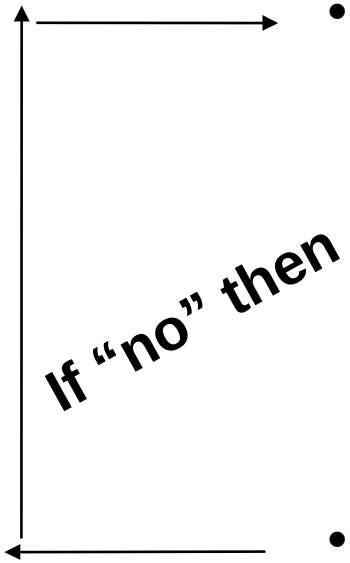
- idea → developer(s)

- *idea → developer(s)*
- first decision → SMC (Scientific Management Committee)

- *idea → developer(s)*
- *first decision → SMC (Scientific Management Committee)*
- development → according to proper coding standards
  - ✓ the source code provided conforms to the coding rules
  - ✓ all source code modifications are documented
  - ✓ all changes have been tested. The results are published appropriately
  - ✓ a Code Responsible Person is available in the future

- *idea → developer(s)*
- *first decision → SMC (Scientific Management Committee)*
- *development → according to proper coding standards*
  - ✓ *the source code provided conforms to the coding rules*
  - ✓ *all source code modifications are documented*
  - ✓ *all changes have been tested. The results are published appropriately*
  - ✓ *a Code Responsible Person is available in the future*
- second decision → SMC

- *idea → developer(s)*
- *first decision → SMC (Scientific Management Committee)*
- *development → according to proper coding standards*
  - ✓ *the source code provided conforms to the coding rules*
  - ✓ *all source code modifications are documented*
  - ✓ *all changes have been tested. The results are published appropriately*
  - ✓ *a Code Responsible Person is available in the future*
- second decision → SMC



- *idea → developer(s)*
- *first decision → SMC (Scientific Management Committee)*
- *development → according to proper coding standards*
  - ✓ *the source code provided conforms to the coding rules*
  - ✓ *all source code modifications are documented*
  - ✓ *all changes have been tested. The results are published appropriately*
  - ✓ *a Code Responsible Person is available in the future*
- *second decision → SMC*
- official implementation → SCA (Source Code Administrator)

- *idea → developer(s)*
- *first decision → SMC (Scientific Management Committee)*
- *development → according to proper coding standards*
  - ✓ *the source code provided conforms to the coding rules*
  - ✓ *all source code modifications are documented*
  - ✓ *all changes have been tested. The results are published appropriately*
  - ✓ *a Code Responsible Person is available in the future*
- *second decision → SMC*
- *official implementation → SCA (Source Code Administrator)*
- standard Technical Test Suite ***

***Technical Test Suite

- independence of processor configurations (MPI and OpenMP - for parallel code)
- reproducibility of results with older versions (if applicable)
- restart functionality
- I/O with Grib/NetCDF
- tests with array bound checking
- possibility to run with input data from different models (ICON, IFS, ERA, etc.)
- timings / efficiency / scalability
- portability

- *idea → developer(s)*
- *first decision → SMC (Scientific Management Committee)*
- *development → according to proper coding standards*
  - ✓ *the source code provided conforms to the coding rules*
  - ✓ *all source code modifications are documented*
  - ✓ *all changes have been tested. The results are published appropriately*
  - ✓ *a Code Responsible Person is available in the future*
- *second decision → SMC*
- *official implementation → SCA (Source Code Administrator)*
- standard Technical Test Suite

If "no" then

- *idea → developer(s)*
- *first decision → SMC (Scientific Management Committee)*
- *development → according to proper coding standards*
  - ✓ *the source code provided conforms to the coding rules*
  - ✓ *all source code modifications are documented*
  - ✓ *all changes have been tested. The results are published appropriately*
  - ✓ *a Code Responsible Person is available in the future*
- *second decision → SMC*
- *official implementation → SCA (Source Code Administrator)*
- *standard Technical Test Suite*
- **Meteorological (NWP) Test Suite**

- *idea → developer(s)*
- *first decision → SMC (Scientific Management Committee)*
- *development → according to proper coding standards*
  - ✓ *the source code provided conforms to the coding rules*
  - ✓ *all source code modifications are documented*
  - ✓ *all changes have been tested. The results are published appropriately*
  - ✓ *a Code Responsible Person is available in the future*

If "no" then

- *second decision → SMC*
- *official implementation → SCA (Source Code Administrator)*
- *standard Technical Test Suite*
- **Meteorological (NWP) Test Suite**

- *idea → developer(s)*
- *first decision → SMC (Scientific Management Committee)*
- *development → according to proper coding standards*
  - ✓ *the source code provided conforms to the coding rules*
  - ✓ *all source code modifications are documented*
  - ✓ *all changes have been tested. The results are published appropriately*
  - ✓ *a Code Responsible Person is available in the future*
- *second decision → SMC*
- *official implementation → SCA (Source Code Administrator)*
- *standard Technical Test Suite*
- *Meteorological (NWP) Test Suite*
- STC (Steering Committee) approval

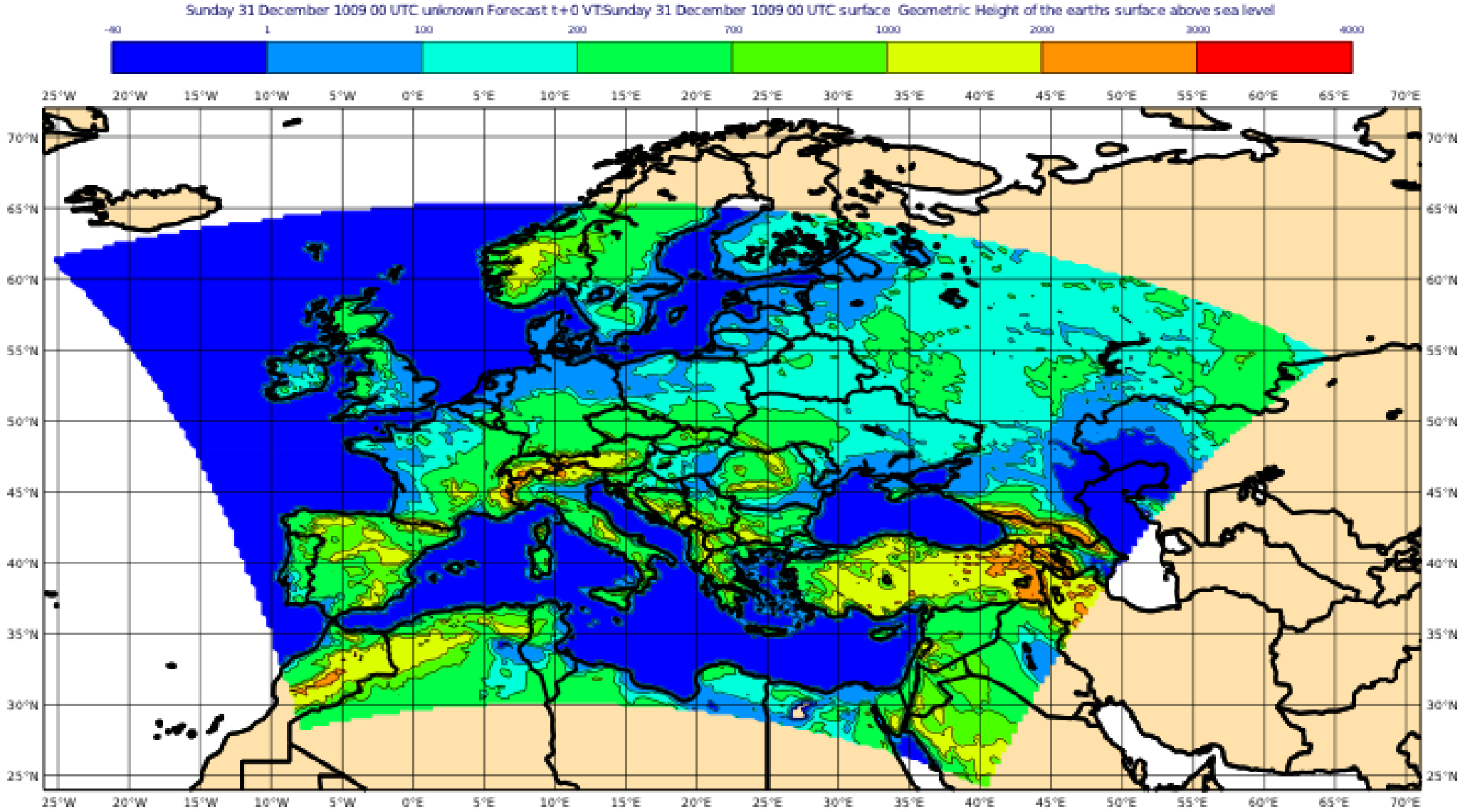
Outlook

- Quality check
- ***Goal of the meteorological suite***
- Suite setup
- Verification setup
- Verification examples
- Next steps

- to perform carefully-controlled and **rigorous testing**, including the calculation of **verification statistics**, for any COSMO model test-version
- to offer necessary information on the **model forecasting performance**
- to facilitate the decision about the **upgrade** of a model test version to a new official release
- to evaluate the **impact** that all implemented numerical or physical processes have on the model
- to provide the COSMO community with **standards** against which the impacts of new developments in the model should be evaluated

Outlook

- Quality check
- Goal of the meteorological suite
- ***Suite setup***
- Verification setup
- Verification examples
- Next steps


Sunday 31 December 1009 00 UTC unknown Forecast t+0 VT:Sunday 31 December 1009 00 UTC surface Geometric Height of the earths surface above sea level
-40
1
100
200
700
1000
2000
3000
4000
25°W 20°W 15°W 10°W 5°W 0°E 5°E 10°E 15°E 20°E 25°E 30°E 35°E 40°E 45°E 50°E 55°E 60°E 65°E 70°E
70°N
65°N
60°N
55°N
50°N
45°N
40°N
35°N
30°N
25°N

COSMO@7p0: ie_tot = 745 ; je_tot = 569; 40 ML; dlon = dlat = 0.0625 **(7 km)**; fc+72h

COSMO@2p8: ie_tot = 1799 ; je_tot = 1369; 50 ML; dlon = dlat = 0.025 **(2.8 km)**; fc+48h

- both initial and boundary (forecast) conditions are provided by IFS HRES
- as for observations, synop reports from a domain covering most of Europe and the Middle East are used (about 3600 stations x day)
- output fields are stored and provided to the verification software (also installed at ECMWF) for the comparison of the 2 model versions by the computation of scores and plots at both resolutions
- verification period: January and July 2013
- special project (Germany, Italy and Greece) for BU (2013-2015, 2016-2018)
- final report published on the COSMO web page

	Resources allocated	Resources used (up to 2 releases per year)
High Performance Computing Facility	5000000 BU	~4800000 BU
Data storage capacity	1 Tb	~0.7 Tb
BU average usage per day		
INT2LM for IFS to COSMO-7km, **~40 BU** per day up to +72h		
COSMO-7km, **~4000 BU** per day up to +72h		
INT2LM for COSMO-7km to COSMO-2.8km, **~300 BU** per day up to +48h		
COSMO-2.8km, **~35000 BU** per day up to +48h		

Outlook

- Quality check
- Goal of the meteorological suite
- Suite setup
- ***Verification setup***
- Verification examples
- Next steps

surface continuous parameters (2mT, Dew Point T, WindSp, TCC, MSLP):
BIAS, RMSE – up to +72h for COSMO-7km, up to +48h for COSMO-2.8km

precipitation (6h, 12h, 24h) for selected thresholds (greater than 0.2, 0.4, 0.6, 0.8, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 12, 14, 16, 18, 20, 25, 30 mm):
ETS, FBI, Performance diagrams – up to +72h for COSMO-7km, up to +48h for COSMO-2.8km

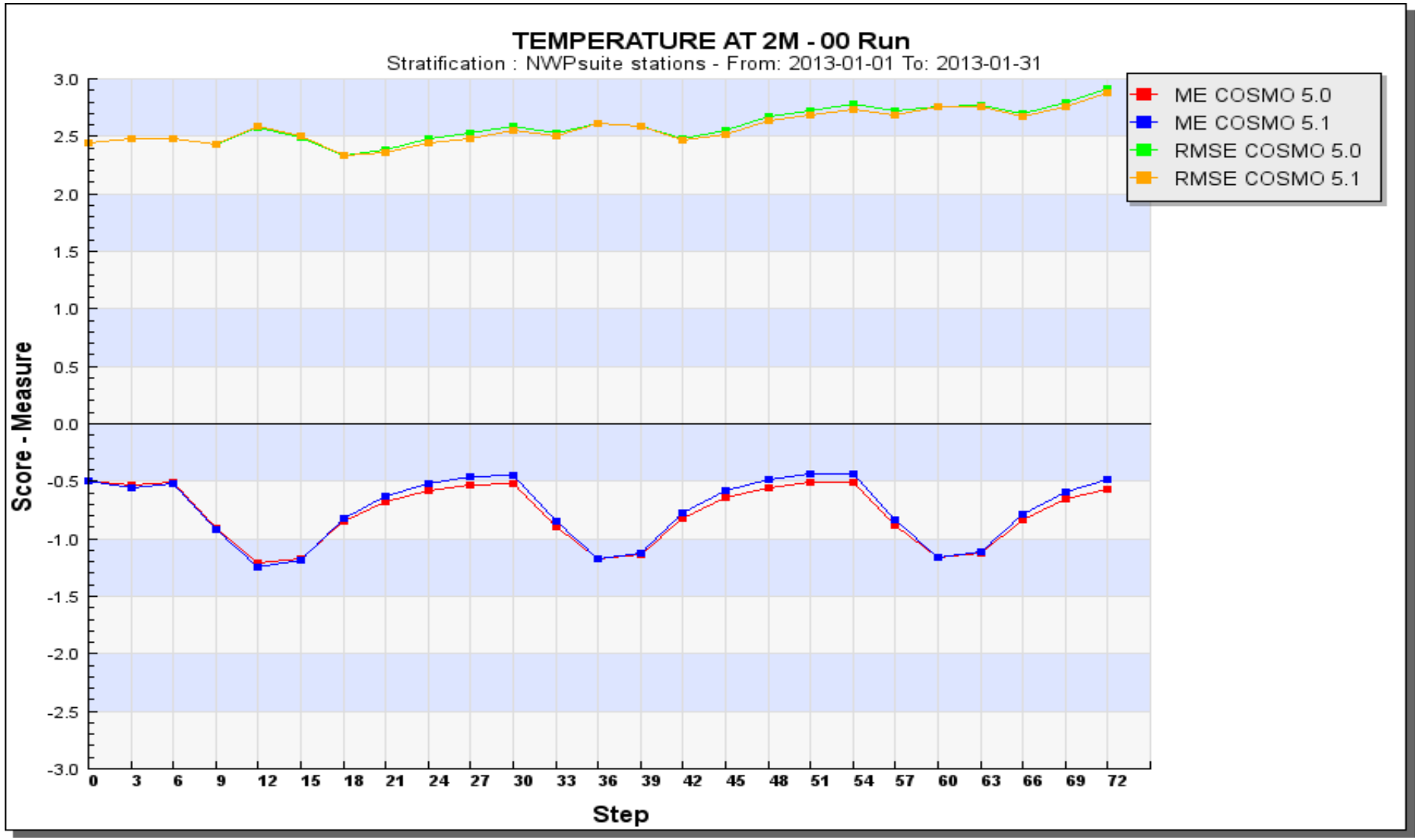
upper air parameters (T, RH, WindSp for selected pressure levels, i.e. 250., 500., 700., 850., 925., 1000 hPa):
BIAS, MAE, RMSE – up to +72h for COSMO-7km, up to +48h for COSMO-2.8km

Outlook

- Quality check
- Goal of the meteorological suite
- Suite setup
- Verification setup
- ***Verification examples***
- Next steps

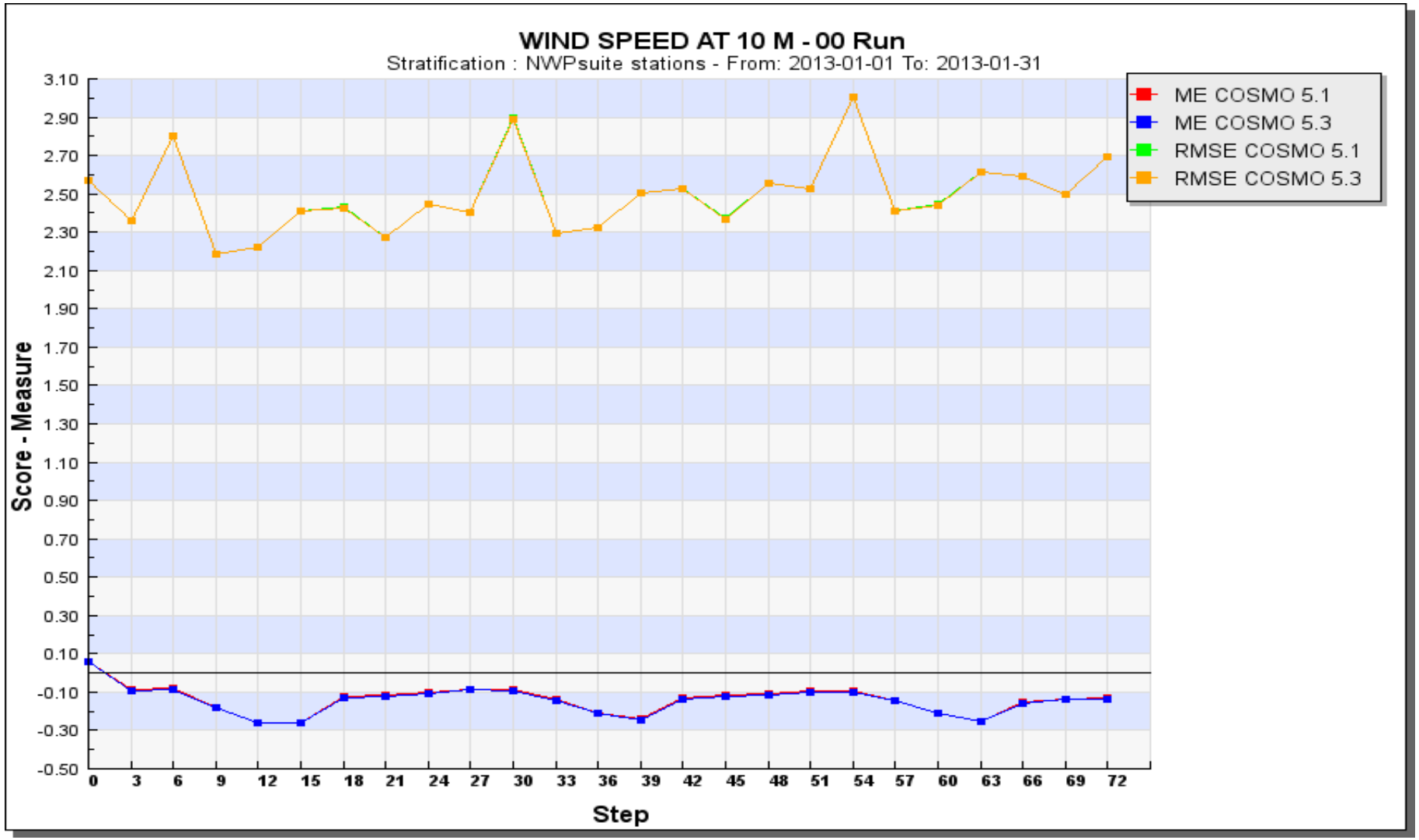
COSMO v5.1 vs v5.0

COSMO@7km



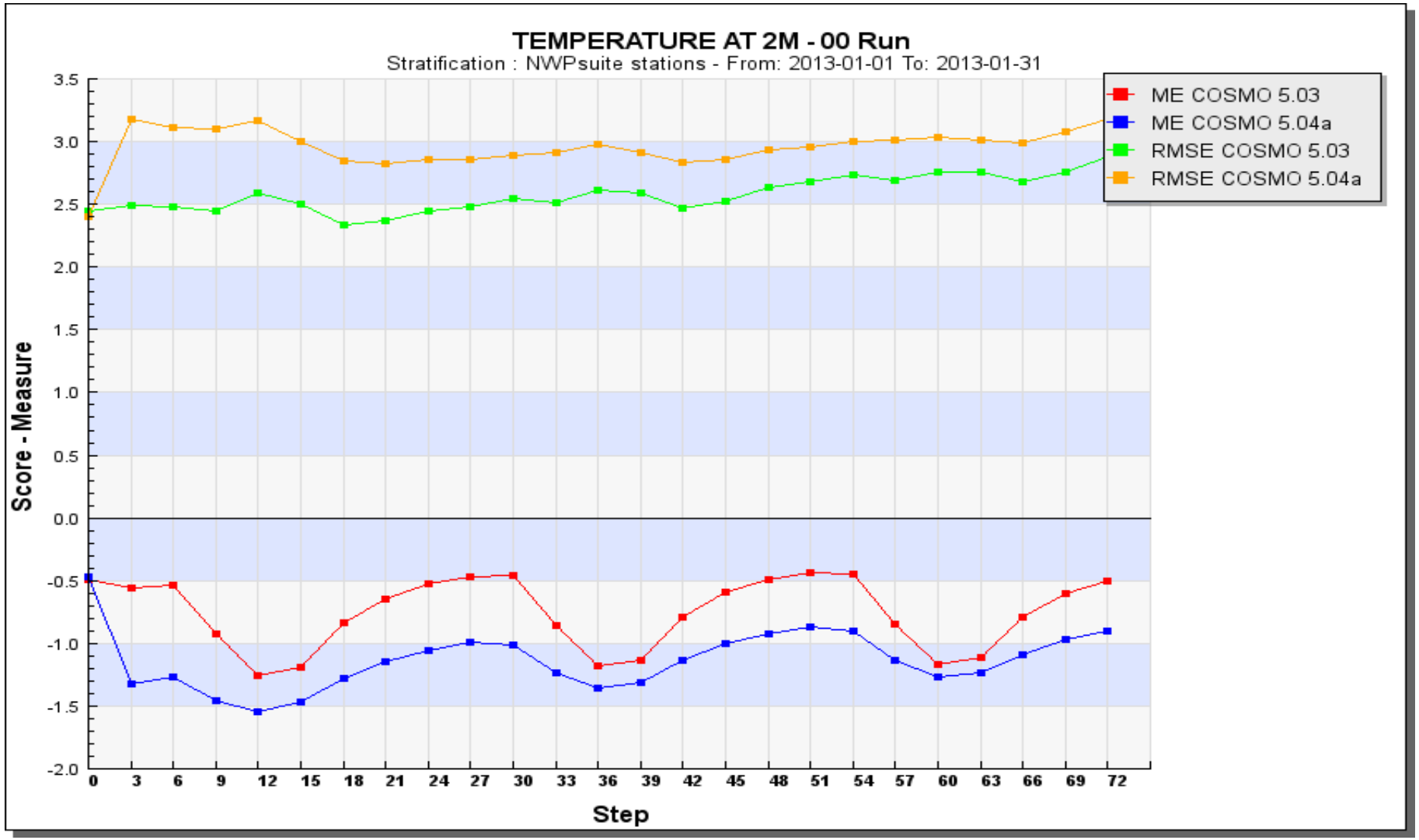
COSMO v5.3 vs v5.1

COSMO@7km



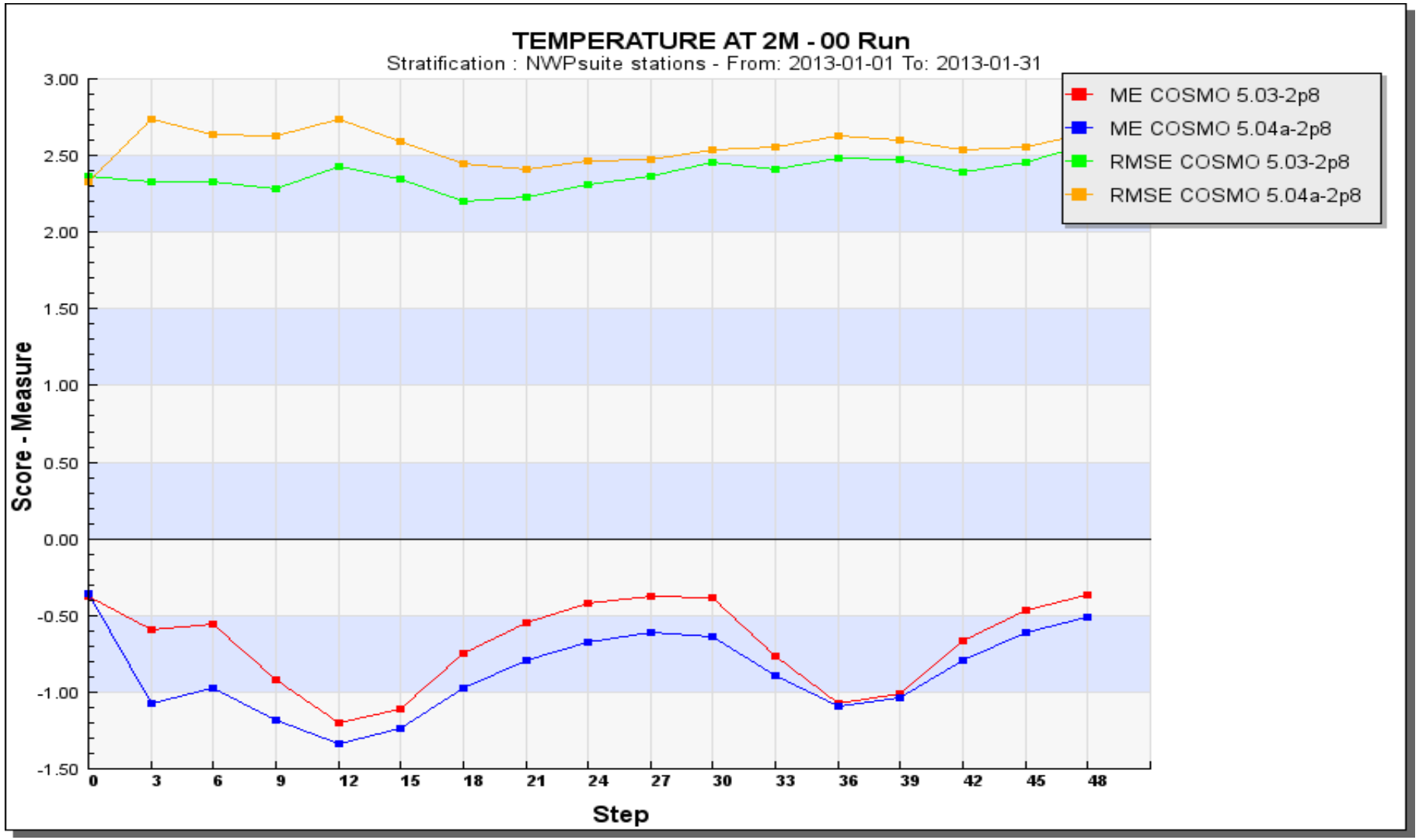
COSMO v5.4 vs v5.3

COSMO@7km



COSMO v5.4 vs v5.3

COSMO@2.8km



Outlook

- Quality check
- Goal of the meteorological suite
- Suite setup
- Verification setup
- Verification examples
- ***Next steps***

- **trying to understand why v5.4 performances are worse compared to v5.3**
- runs in single precision to save BU
- introduction of statistical significance (bootstrap) as differences are often marginal
- possibility to add a unified score (combining the performance of various parameters)
- introduction of a Score Card

Domain	Parameter	Level	Anomaly correlation										RMS error									
			Forecast day										Forecast day									
			1	2	3	4	5	6	7	8	9	10	1	2	3	4	5	6	7	8	9	10
Europe	Relative humidity	300hPa	▲			■			■		■			■	■		■	■	▼		■		
		700 hPa	▲	■									▲		■	■	■		■				
	Temperature	100 hPa	▲	▲	■				■	■	■	■	▲	▲	▲	▲	▲	▲	▲	▲	▲	■
		500 hPa	▲	▲		■			■				▲	■		▲						
		850 hPa	▲	■						■			▲		■	■	■		■	■		■
		1000 hPa	■									■	▼					■				
	Wind	200 hPa	▲	▲		■			■			■	▲	▲	■	▲	■					
		850 hPa	▲	▲	■	■	■						▲	▲	▲	■	■					■
	Geopotential	100 hPa	▲	▲	▲								▲	▲	▲	▲	▲	▲	▲	▲	▲	▲
		500 hPa	▲	■		■	■					■	▲	▲	■	▲	■		■			■
		850 hPa	▲				■	■			■	■	▲			▲	■			■	■	■
		1000 hPa	▲					■			■	■	▲				■	■		■	■	■
Extratropical Northern Hemisphere	10 m wind		▲	▲	■	■		■					▲	▲	▲	▲	▲	▲	▲	■	■	
	Relative humidity	300hPa	▲	■	■				■					▼	▼	▼	▼	▼	▼	■	■	■
		700 hPa	▲	▲	■		■					■	▲		■	■		■	▼	■		▼
	Waves	swh			■		■	■	▲	▲			▲	▲	▲	▲	▲	▲	▲	▲	■	
		mwp	▲	▲	▲	▲	▲	▲	▲	■	■		▲	▲	▲	▲	▲	▲	▲	▲	▲	
	Temperature	100 hPa	▲		▼	▼	▼	■		■	■	■	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲
		500 hPa	▲	▲	■	■							▲	▲	■	▲						
		850 hPa	▲	▲	▲	■			■			■	▲	▲	▲	■			■	■		
		1000 hPa	▼	▼							■	■	▼	■		■	■				■	■
	Wind	200 hPa	▲	▲	■	■						■	▲	▲	▲	▲	▲	▲	■			
		850 hPa	▲	▲	▲	■							▲	▲	▲	■						
	Geopotential	100 hPa	▲	▲	▲	■							▲	▲	▲	▲	▲	▲	▲	▲	▲	▲
		500 hPa	▲	▲	▲	■							▲	▲	▲	▲	■					
		850 hPa	▲	▲			■						▲	▲	■	■						
		1000 hPa	▲	■			■						▲	■		■						

Symbol legend: for a given forecast step...
(d: score difference, s: confidence interval width)

- ▲ Cy38r2 **better** than Cy38r1 – **statistically highly significant**
- ▲ Cy38r2 **better** than Cy38r1 – **statistically significant**
- Cy38r2 better than Cy38r1 – not statistically significant
- Little difference between Cy38r2 and Cy38r1
- Cy38r2 worse than Cy38r1 – not statistically significant
- ▼ Cy38r2 **worse** than Cy38r1 – **statistically significant**
- ▼ Cy38r2 **worse** than Cy38r1 – **statistically highly significant**

- **new verification software ?**

  This possibility will be explored to understand if there is a clear advantage.

- **from forecast to hindcast** (using IFS or ICON analyses) ?

  As the hindcast run is long, the soil variables have time to adjust to the atmospheric forcing. Moreover from a technical point of view the system is cheaper (less BU, less time). On the other hand the suite should be reshuffled completely.

Thank you for your attention !

(...and any suggestions are welcome...)