

Operational verification in Italy

Angela Celozzi Elena Oberto – Naima Vela Maria Stefania Tesini





Overview



- Verification using VERSUS (CNMCA & ARPA-ER)
 - Standard : T 2m, Td 2m, MSLP, WIND 10m, TCC
 - Upper air : T , WIND
 - Conditional:
 T & TCC (obs) & Wind Speed
 Td & WIND speed
- Precipitation using High-Res Rain-gauges (ARPA-P)
 - Comparison of COSMO models over Italian territory







Summary

- Upper Air verification: Temperature, Wind Speed
 - COSMO-I7, COSMO-ME
- Total Cloud Cover:
 - COSMO-I7 , COSMO-I7 / ECMWF, (COSMO-I7 / ECMWF /COSMO-I2 for MAM2012)
 - COSMO-ME , COSMO-ME / ECMWF , COSMO-ME / COSMO-IT
- Mean Sea Level pressure and 10 m Wind:
 - COSMO-I7 , COSMO-I7 / ECMWF
 - COSMO-ME , COSMO-ME / ECMWF , COSMO-ME / COSMO-IT
- 2m Temperature and Dew Point Temperature:
 - COSMO-I7 , COSMO-I7 / ECMWF
 - COSMO-ME , COSMO-ME / ECMWF , COSMO-ME / COSMO-IT
 - COSMO-ME conditional verification
- Precipitation:
 - Long trend
 - Cosmo models inter-comparison







Upper air

TEMPERATURE



Italian radio-sounding stations

only few of them do sounding at 6 and 18 UTC so we focus on verification of mainly 00 and 12 UTC











Upper Air Temperature COSMO-I7 JJA 2011

- underestimation above 250 hPa
- nearly no bias from 250 hPa to 700 hPa with lower MAE & RMSE
- overestimation at 700 hPa increasing with the forecast time
- underestimation under 700 hPa in particular at 00 UTC









Upper Air Temperature COSMO-I7 SON 2011

• similar to JJA but at the lower level the underestimation is bigger at 00 UTC









Upper Air Temperature COSMO-I7 DJF 2011-12

- same underestimation at higher levels as in the other seasons
- the underestimation start at 300 hPa and increases moving to lower layer
- at midnight the errors sees to be mainly due to bias















Upper Air Temperature COSMO-ME JJA 2011

- underestimation above 250 Hpa
- overestimation at 250
 hPa and 700 hPa
 increasing with forecast
 time
- ME increases with forecast time ("the model seems to warm")
- •Compared to COSMO-I7 seems warmer in particular from 700 hPa and 1000 hPa





Lugano - Cosmo General Meeting - 10-13 September 2012 agenzio regionale prevenzione e ambiente dell'amilio-romagno

12 UTC



Upper Air Temperature COSMO-ME SON 2011

- similar behavior above 250 hPa
- little overestimation at 700 hPa increasing with forecast time
- at midnight negative bias under 850 hPa, reducing with forecast time





Lugano - Cosmo General Meeting - 10-13 September 2012 12 UTC





Upper Air Temperature COSMO-ME DJF 2011-12

• very similar to COSMO-I7 even if the errors are a little smaller





Lugano - Cosmo General Meeting - 10-13 September 2012 agenzia regionale prevenzione e ambiente dell'emilia-romagna

■ 12 UTC



Upper Air Temperature COSMO-ME MAM 2012

- underestimation
 under 700 hPa only at
 00 UTC
- overestimation at 700 hPa increasing with forecast time
- nearly no bias at 12
 UTC (except 700 hPa)

 "usual" overestimation at higher levels





Lugano - Cosmo General Meeting - 10-13 September 2012 agenzia regionale prevenzione e ambiente dell'emilio-romogno

12 UTC



Upper air

WIND SPEED



Italian radio-sounding stations

only few of them do sounding at 6 and 18 UTC so we focus on verification of mainly 00 and 12 UTC











Upper Air Wind Speed COSMO-I7 JJA 2011

•MAE near 2 m/s, RMSE a bit bigger (but < 4m/s)

•Largest error at about 250 hPa (but relative to Jet stream speed is small)

•Tendency to increase negative bias in particular in the lower layers and during night with forecast step









Upper Air Wind Speed COSMO-I7 SON 2011

• more or less as JJA









Upper Air Wind Speed COSMO-I7 DJF 2011-12

• bias nearly 0 but little increase with forecast time of MAE and RMSE especially at 400-500 hPa















COSMOME Seasonal Run 00 Wind speed- Italy - 00 Run Stratification : All Italian TEMP Station - Period: JJA 2011

Upper Air Wind Speed COSMO-ME JJA 2011

Negative ME , in particular at "Jet stream height" where errors are bigger and grow with forecast steps









COSMOME Seasonal Run 00 Wind speed- Italy - 00 Run Stratification : All Italian TEMP Station - Period: SON 2011

Upper Air Wind Speed COSMO-ME SON 2011

•Errors at about 250 hPa increase with forecast step

• small negative bias, especially at 12 UTC and below 925 hPa also during night









Upper Air Wind Speed COSMO-ME DJF 2011-12

• Errors at about 400-500 hPa grow with forecast step (seasonal feature as for COSMO-I7)

Scale is different!















TOTAL CLOUD COVER























TCC :COSMO-I7,COSMO-I2,ECMWF

Arpa-ER started to use Versus!































TCC considerations

- COSMO-I7 tends to overestimate cloud cover with a bigger bias during the night
 - Errors in the early hours of integration are bigger then in the following steps
 - For MAM2012 COSMO-I2 reduces the overestimation during night but also underestimate during the day, but no significant differences in MAE and RMSE
 - ECMW has a bias oscillating around zero and errors are a little smaller than COSMO-I7
- COSMO-ME has a positive bias during night and a bit negative during day
 - Night overestimation is more pronounced than ECMWF but RMSE are nearly the same in JJA and DJF slightly better in SON and MAM
 - COSMO-IT has bigger errors in the first 6 hours of integration respect to COSMO-ME







MEAN SEA LEVEL PRESSURE



























- COSMO-I7 and COSMO-ME
 - different bias depending on season
 - MAE and RMSE have similar trend in all the season except in winter when the errors grow with forecast step
 - respect to ECMWF
 - COSMO models have bigger errors in DJF and MAM, but the difference are less than 1 hPa
- COSMO-ME vs COSMO-IT
 - very similar in MAM2012
 - COSMO-IT a little worst in RMSE, especially in DJF
 - difference (except MAM2012) in bias






10 M WIND



Lugano - Cosmo General Meeting - 10-13 September 2012















*

ambiente dell'emilia-romagna





























Wind speed considerations

- COSMO-I7 and COSMO-ME
 - negative bias, a little bigger for COSMO-ME (ECMWF has bigger bias)
 - MAE and RMSE constant during the forecast step
 - MAE around 1.5 m/s (2 m/s in DJF)
 - RMSE around 2.5 m/s (3 m/s in DJF)
 - Non significant differences respect to ECMWF (COSMO-ME slightly better)
 - COSMO-ME vs COSMO-IT
 - COSMO-IT has positive bias in JJA, SON and MAM (DJF has negative value for ME at 15 UTC)
 - RMSE similar except for DJF when COSMO-IT is better







2M TEMPERATURE



Lugano - Cosmo General Meeting 10-13 September 2012































*

















- ME with a "typical" diurnal cycle: overestimation during night and underestimation during day
- This behavior is masked during DJF for both 7 Km-COSMO and MAM for COSMO-ME because model temperature is generally too low
- ECMWF has opposite diurnal cycle of ME respect to COSMO models

MAE varies between 1.5 and 2.5 degrees (COSMO-ME slightly better) and RMSE between 2.5 and 3 degrees. Errors over all Italian stations are smaller than ECMWF

• COSMO-IT is a little better than COSMO-ME, the ME is less negative but the cycle of the error is the same





T2m when observed TCC $\leq 25\%$



T2m when observed TCC $\leq 25\%$ & wind speed ≤ 2 m/s









T2m when observed TCC \geq 75%



T2m when observed TCC \geq 75% & wind speed \leq 2 m/s









T2m CV considerations

COSMO-ME:

- In clear sky conditions (observed) the amplitude of the diurnal cycle of bias increases
 - In very low wind conditions the amplitude of the diurnal cycle of the bias further increase. Bias during the night hours becomes positive, bias in the central hours of the day becomes more negative
 - MAE and RMSE increase in clear sky condition and the error lines show several peaks during nighttime and at 15 UTC when the underestimation is bigger.
 - Windless condition amplify these features
- In cloudy condition the errors has a lesser amplitude, the mean bias is shifted toward higher temperature and the diurnal cycle is a bit different
 - This cause an increase in the error in JJA during daytime and a general decrease of errors in the other season.
 - Windless condition has no clear effects
- Further conditional verification studies have been done at ARPA-ER on COSMO-I7 (shown in the web conference of August 27th) – results are in the same direction!







2M DEW POINT TEMPERATURE



Lugano - Cosmo General Meeting 10-13 September 2012

















































ambiente dell'emilia-romagna

TD 2m when observed wind speed \leq 2 m/s





- COSMO-I7
 - MAE from 2.5 to 3 °C, RMSE from 3.5 to 4 °C bigger errors at 12-15 UTC
 - Different diurnal cycle of bias error depending on season but generally the dew point is under-predicted
- COSMO-ME
 - Smaller RMSE than COSMO-I7 but very similar behavior
 - Negative bias error except MAM at 15 and 18 UTC (up to 0.5 °C overestimation for the 3 day of forecast)
 - Conditional verification in windless situations does not show significant differences
- Respect to ECMWF
 - COSMO-I7 better in JJA and SON, equal in the other seasons
 - COSMO-ME better in all seasons
- COSMO-IT
 - Equal to COSMO-I7 in JJA and SON
 - Better RMSE in DJF and MAM, but the bias error in these seasons differs







PRECIPITATION



Lugano - Cosmo General Meeting 10-13 September 2012





Stephenson et al. Introduce the extreme dependency score (EDS) as a good alternative to standard scores for verification of rare events.

	Event observed yes	Event observed no	Total
Forecast yes	А	b	a + b
Forecast no	С	d	c + d
Total	a + c	b+d	n= a + b + c + d





frequency bias	FBI = (a + b)/(a + c)	[0,∞] best 1	The frequency bias index indicates whether the forecasting system under or over-	
			forecasts the number of events.	
hit rate (POD)	H = a/(a + c)	[0,1] best 1	The hit rate represents the probability that the event is forecast when it occurs	
false alarm rate (POFD)	F = b/(b + d)	[0,1] best 0	The false alarm rate represents the probability of forecasting the event when it did not occur.	
			% not events obs. Not correctly forecasted.	
			fraction of the observed "no" events were	
			incorrectly forecast as "yes".	
true skill score	TSS = H - F	[-1,1] best 1	The true skill score gives information on how the forecasting system distinguishes between occurrences and not occurrences.	
base rate	BR = (a + c)/n	[0,1]	The base rate represents the probability that the event occurs. By definition, 1-BR plotted versus increasing thresholds represents the probability that precipitation amount does not exceed a certain threshold.	
extreme dependency score	EDS= 2[ln((a+c)/n)/ln(a/n)] -1	[-1,1] best 1	What is the association between forecast and observed rare events? Converges to 2η -1 as event frequency approaches 0, where η is a parameter describing how fast the hit rate converges to zero for rarer events. EDS is independent of bias, so should be presented together with the frequency bias	



- To get clear information about how the forecasting system detects the extreme events, it would be fair if the EDS is compared for events having the same base rate. One has to investigate if better value of the EDS are related to an improvement in the quality of the forecasting system or if they are due to the event variability over the years.
- The equation defining the EDS uses the left hand side of a contingency table and the total number of cases (sample size). This results in an increased freedom for false alarms and correct negatives, which can freely vary with the only restriction that their sum has to be constant. Therefore, it is paramount to use the EDS in combination with other scores that include the right hand side of the contingency table, as the F and/or the FBI to show that improvements are not due to an increase of false alarms. (Ghelli&Primo,2009)





The affect of the base rate on the extreme dependency score (Ghelli&Primo,2009)

The Extreme Dependency Score (EDS) has been introduced as an alternative measure to verify the performance of numerical weather prediction models for rare events, taking advantage of the non-vanishing property of the score when the event probability tends to zero.

This score varies from 1 (best value) to -1 (worst value).

The EDS is written as a function of BR:

EDS = [ln(BR) - ln(HR)] / [ln(BR) + ln(HR)]

Equation presents the EDS as a function of the base rate and the hit rate. when HR = 1, the EDS = 1 and when BR = 1, the EDS = -1. On the other hand, when the base rate is equal to one, the event happens all the time and so the EDS is not an appropriate score since it is focused on verification of extreme events (low probability of occurrence). Therefore, if different data samples need to be compared, it is imperative to have similar base rate.







•Thus, even if there are no misses and the EDS value is maximum, the forecasting system might have a high number of false alarms. Therefore, an EDS = 1 does not imply a skilful system. If values of the EDS for different periods need to be compared, then the base rate must be constant in time to avoid changes in the EDS to be just a reflection of changes in the BR.

• If the base rate is constant, an increase of the EDS implies a better probability of detection (hit rate), i.e. a more skilful system. If only the hit rate is constant, then an increase of the EDS is only due to a higher event probability. If neither the base rate nor the hit rate is constant, then the improvement of the EDS could be due to any of the previous reasons.





The extreme dependency score: a non-vanishing measure for forecasts of rare events (Stephenson et al.)

EDS takes the value of 1 for perfect forecasts and 0 for random forecasts, and is greater than zero for forecasts that have hit rates that converge slower than those of random forecasts

EDS has demonstrated here that there is dependency between the forecasts and the observations for more rare events, which is masked by the traditional skill scores that converge to zero as the base rate vanishes. EDS does not explicitly depend on the bias in the system for vanishing base rate and so is less prone to improvement by hedging the forecasts. EDS has the disadvantage that it is based only on the numbers of hits and misses, and so ignores information about false alarms and correct rejections. Therefore, EDS is non-informative about forecast bias, and a forecasting system with a good EDS could be very biased. Therefore, one should present EDS together with the frequency bias as a function of threshold in order to provide a complete summary of forecast performance.




















ETS 201012-201205 10mm/24h



















-100 - -60 ·60 · ·20

·20 · 20

20 - 60 60 - 100

🞽 🔀 100 - 🕐











• Very difficult to summarize the results



Thanks for your

attention!





Lugano - Cosmo General Meeting - 10-13 September 2012