New operational COSMO model setup in Italy

Chiara Marsigli, Virginia Poli, Andrea Montani, Thomas Gastaldo, Davide Cesari, Tiziana Paccagnella

Arpae Emilia-Romagna
Servizio Idro-Meteo-Clima
Bologna, Italy

CONSORTIUM FOR SMALL SCALE MODELING
COSMO

September 2017, Jerusalem

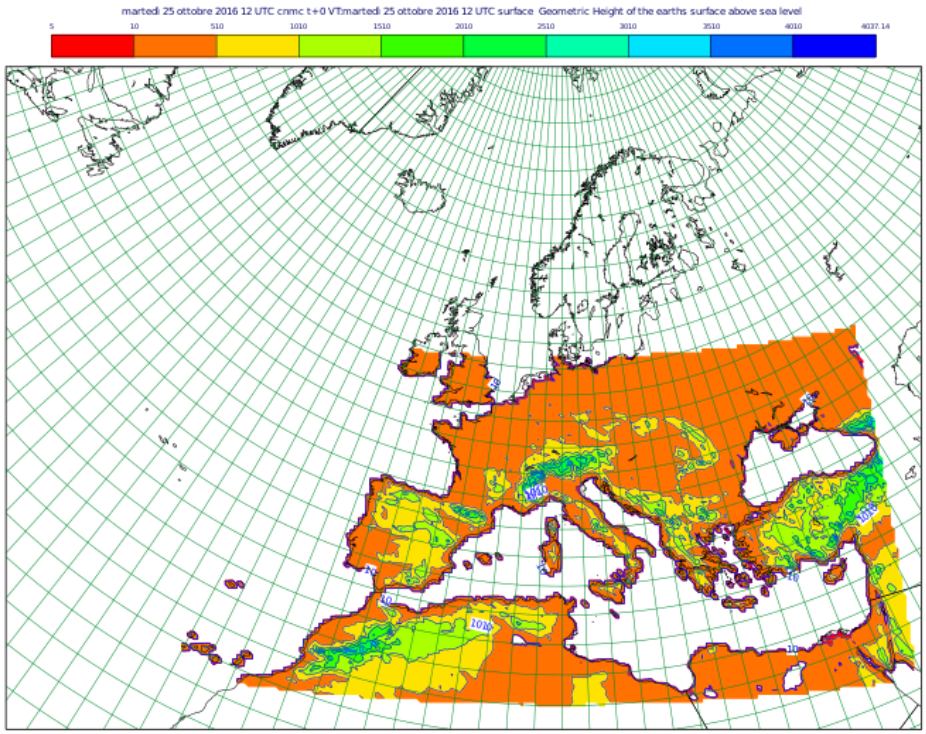
Outline

Update of the Italian COSMO model national operational setup

Starting from 2016 (but preparation started much earlier), the operational setup of COSMO model run under the framework of LAMI agreement (ARPA Piemonte, Arpae Emilia-Romagna, COMET) is undergoing a major revision:

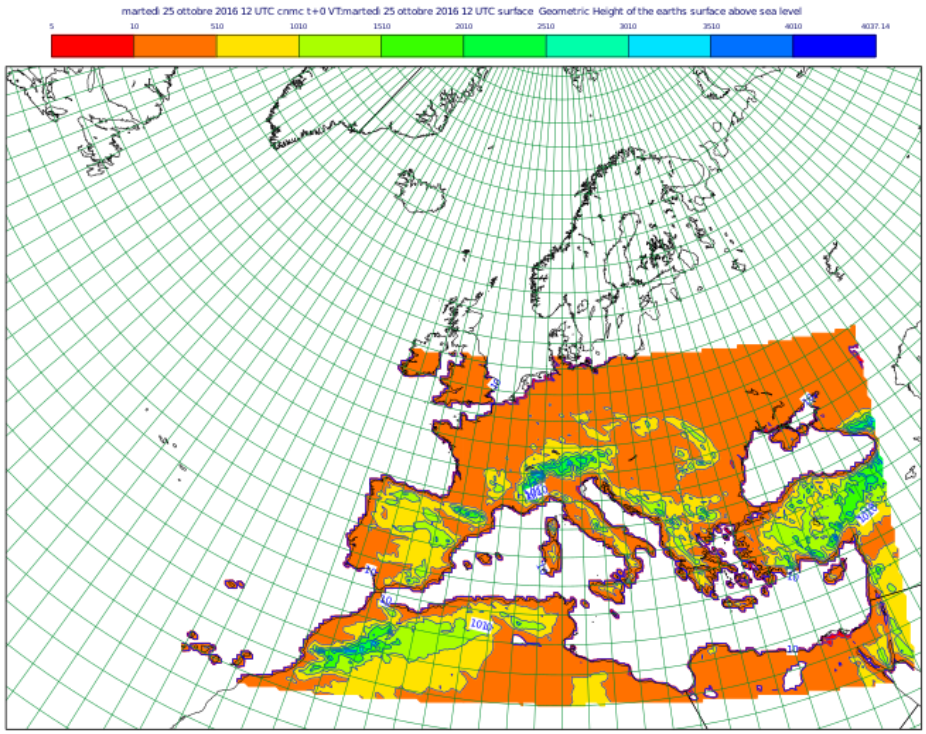
- Lower resolution setup upgraded to a grid step of 5 km (0.045°), 45 vertical levels, and domain extended to cover all the Mediterranean area
- Higher resolution setup ugraded to a grid step of 2.2 km (0.02°) and 65 vertical levels, on a domain covering Italian territory.

Lower resolution setup - 5 km



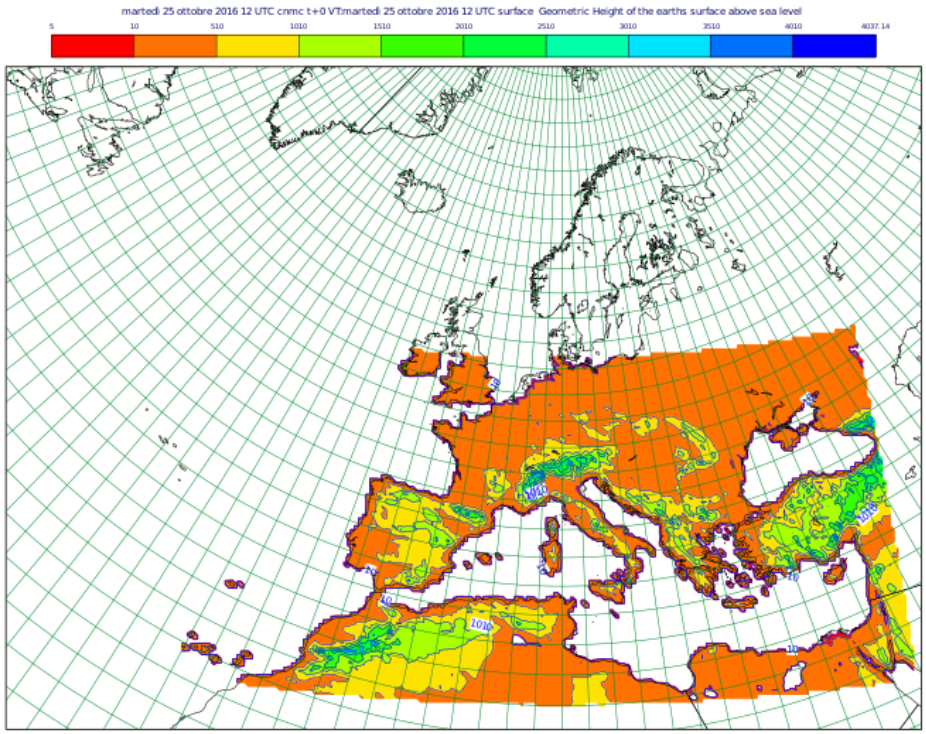
- 1083x559x45 gridpoints
- 2 runs per day up to 72 hours
- Initial conditions provided by ETKF analysis system run by COMET (Italian National Meteorological Service)
- The configuration is going to be aligned with COMET operational setup so that the output data can be used as a reciprocal backup.

Lower resolution setup - 5 km



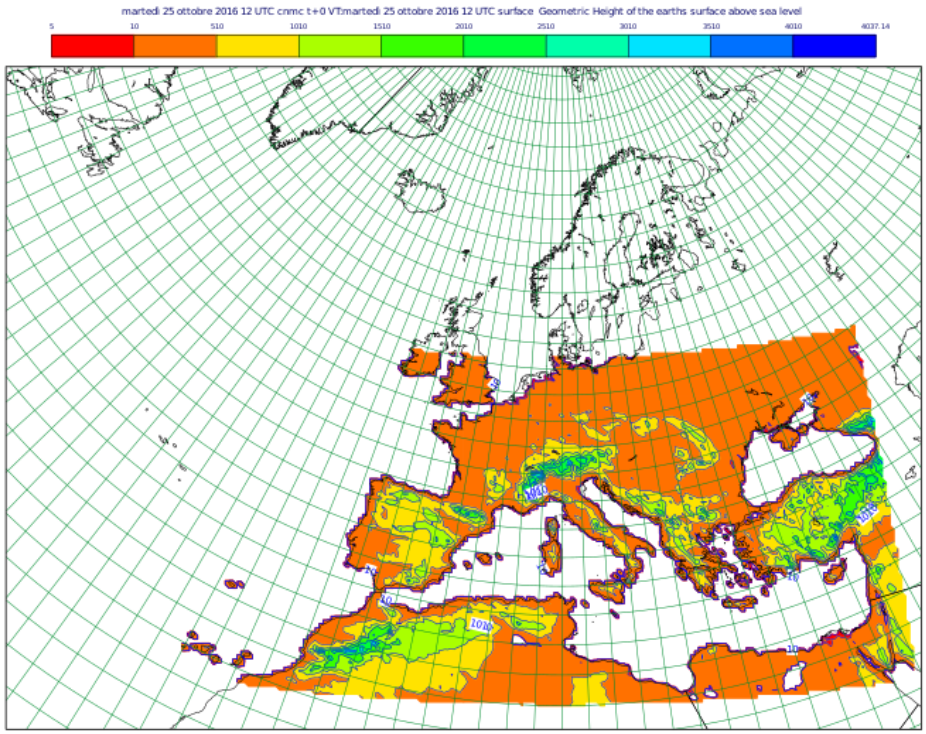
- 1083x559x45 gridpoints
- 2 runs per day up to 72 hours
- Initial conditions provided by ETKF analysis system run by COMET (Italian National Meteorological Service)
- The configuration is going to be aligned with COMET operational setup so that the output data can be used as a reciprocal backup.

Lower resolution setup - 5 km



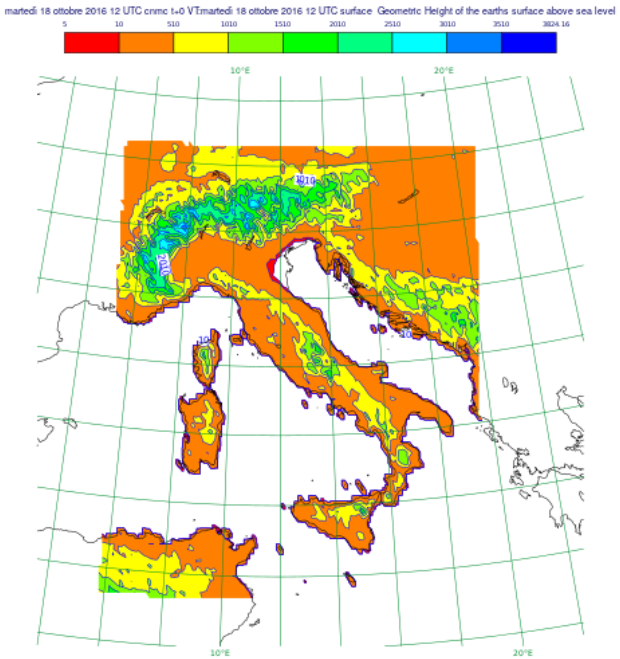
- 1083x559x45 gridpoints
- 2 runs per day up to 72 hours
- Initial conditions provided by ETKF analysis system run by COMET (Italian National Meteorological Service)
- The configuration is going to be aligned with COMET operational setup so that the output data can be used as a reciprocal backup.

Lower resolution setup - 5 km



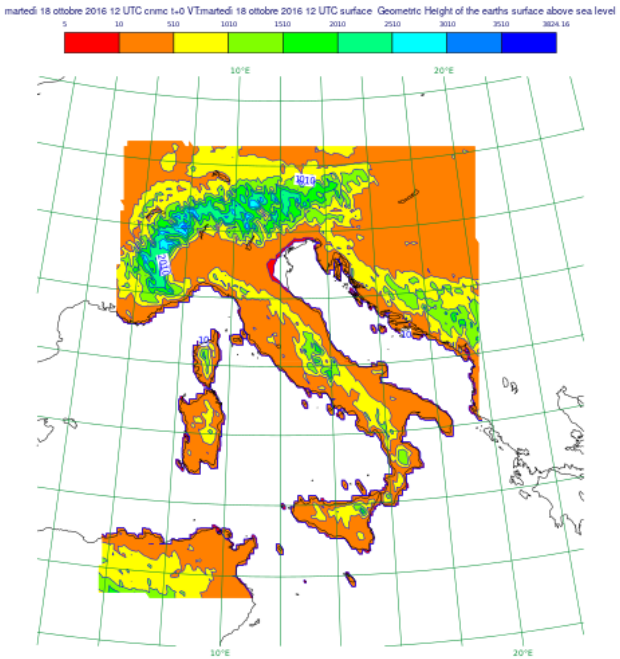
- 1083x559x45 gridpoints
- 2 runs per day up to 72 hours
- Initial conditions provided by ETKF analysis system run by COMET (Italian National Meteorological Service)
- The configuration is going to be aligned with COMET operational setup so that the output data can be used as a reciprocal backup.

Higher resolution setup - 2.2 km



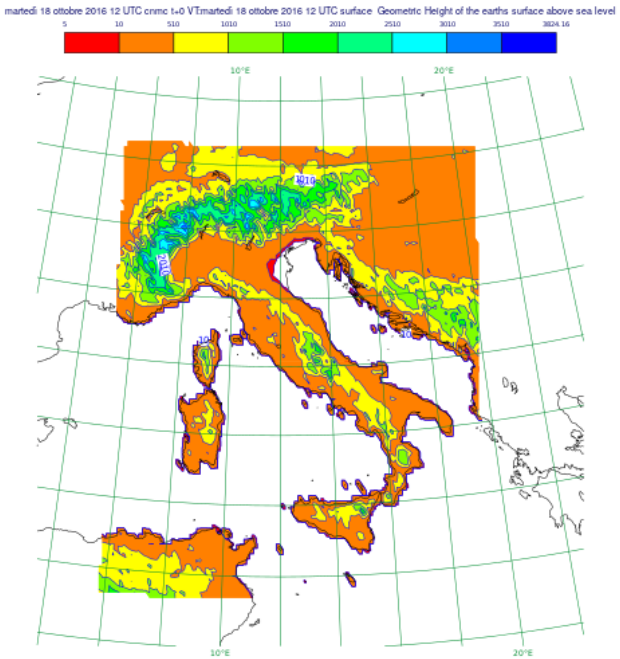
- 576x701x65 gridpoints
- Quasi-continuous ensemble data assimilation cycle with KENDA + Latent Heat Nudging on Italian national radar composite
- 3-hour cycles with 20 ensemble members, plans to reduce the first and increase the second
- boundary conditions from the 5km run on Mediterranean (“deterministic” member) and from COMET 10 km COSMO ensemble data assimilation runs (“perturbed” members)
- already operational

Higher resolution setup - 2.2 km



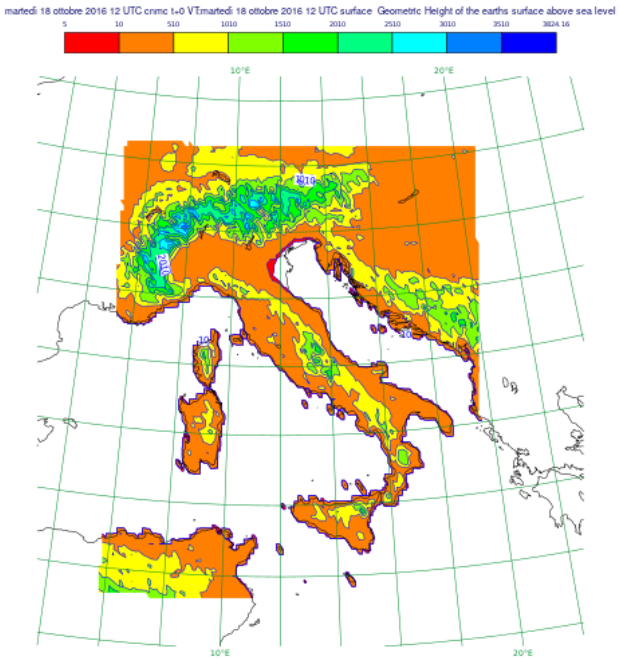
- 576x701x65 gridpoints
- Quasi-continuous ensemble data assimilation cycle with KENDA + Latent Heat Nudging on Italian national radar composite
- 3-hour cycles with 20 ensemble members, plans to reduce the first and increase the second
- boundary conditions from the 5km run on Mediterranean (“deterministic” member) and from COMET 10 km COSMO ensemble data assimilation runs (“perturbed” members)
- already operational

Higher resolution setup - 2.2 km



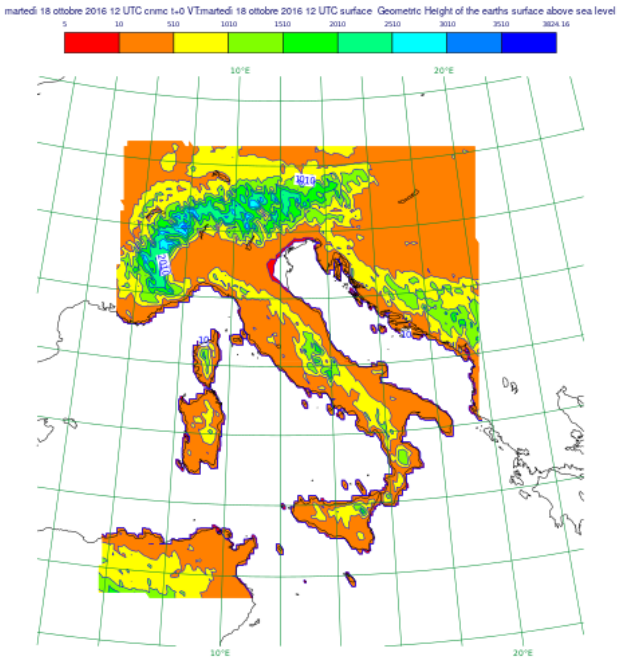
- 576x701x65 gridpoints
- Quasi-continuous ensemble data assimilation cycle with KENDA + Latent Heat Nudging on Italian national radar composite
- 3-hour cycles with 20 ensemble members, plans to reduce the first and increase the second
- boundary conditions from the 5km run on Mediterranean (“deterministic” member) and from COMET 10 km COSMO ensemble data assimilation runs (“perturbed” members)
- already operational

Higher resolution setup - 2.2 km



- 576x701x65 gridpoints
- Quasi-continuous ensemble data assimilation cycle with KENDA + Latent Heat Nudging on Italian national radar composite
- 3-hour cycles with 20 ensemble members, plans to reduce the first and increase the second
- boundary conditions from the 5km run on Mediterranean (“deterministic” member) and from COMET 10 km COSMO ensemble data assimilation runs (“perturbed” members)
- already operational

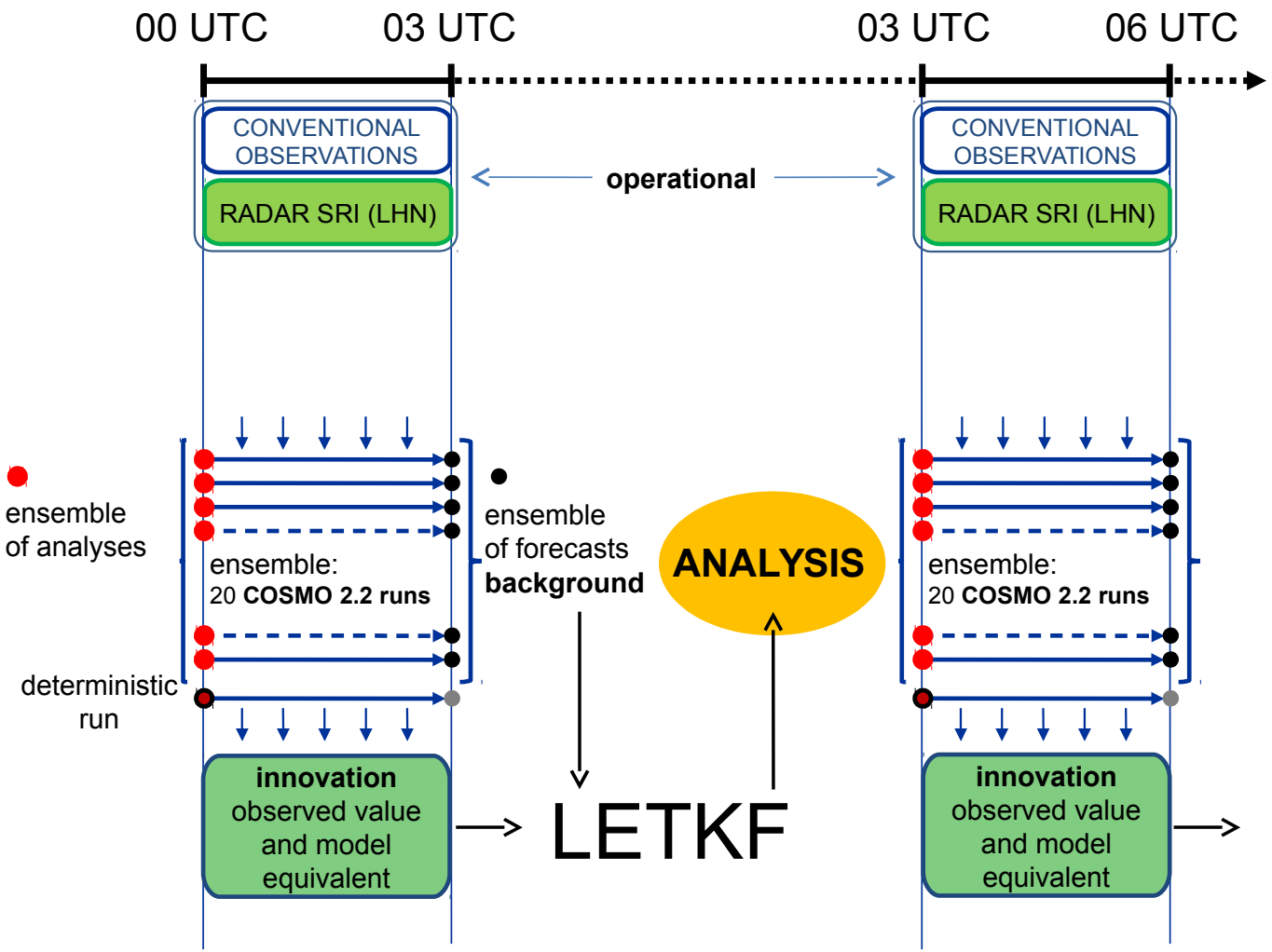
Higher resolution setup - 2.2 km



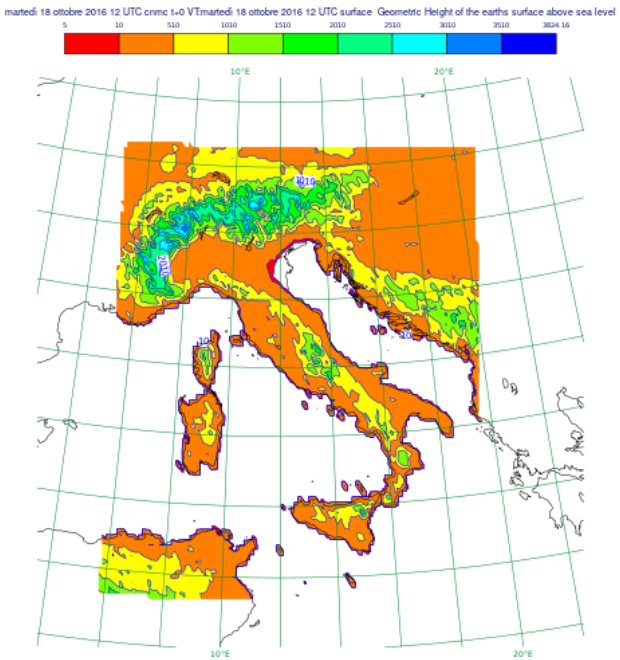
- 576x701x65 gridpoints
- Quasi-continuous ensemble data assimilation cycle with KENDA + Latent Heat Nudging on Italian national radar composite
- 3-hour cycles with 20 ensemble members, plans to reduce the first and increase the second
- boundary conditions from the 5km run on Mediterranean (“deterministic” member) and from COMET 10 km COSMO ensemble data assimilation runs (“perturbed” members)
- already operational

Higher resolution setup - 2.2 km

KENDA (Kilometer-scale ENsemble Data Assimilation)



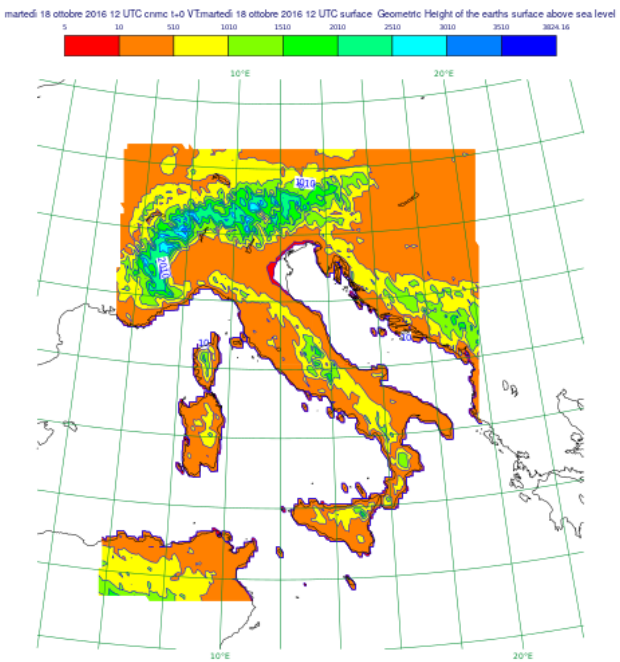
Higher resolution setup - 2.2 km



The deterministic analysis is used for initialising:

- 8 very-short-range (+18h) deterministic forecasts per day with a short cutoff (RUC - already operational)
- 2 longer range deterministic forecasts per day (+48h) with a longer cutoff and more recent boundary conditions (under test in these days)
- all using boundary conditions from the 5km run on Mediterranean

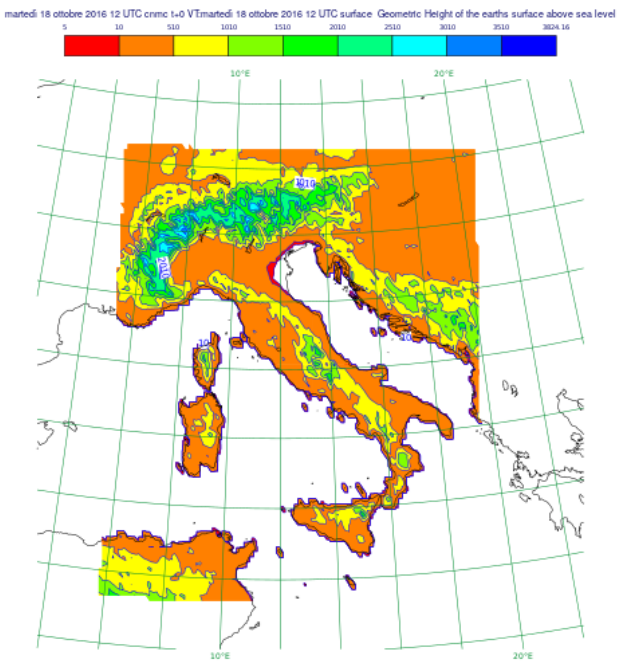
Higher resolution setup - 2.2 km



The deterministic analysis is used for initialising:

- 8 very-short-range (+18h) deterministic forecasts per day with a short cutoff (RUC - already operational)
- 2 longer range deterministic forecasts per day (+48h) with a longer cutoff and more recent boundary conditions (under test in these days)
- all using boundary conditions from the 5km run on Mediterranean

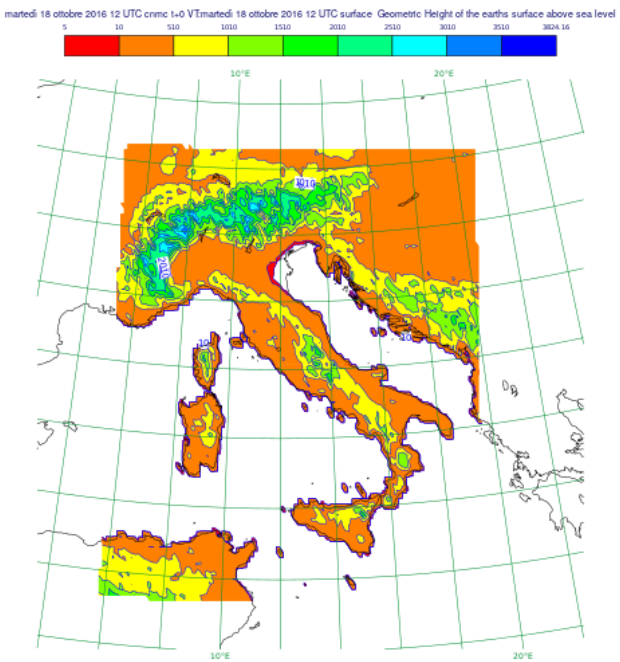
Higher resolution setup - 2.2 km



The deterministic analysis is used for initialising:

- 8 very-short-range (+18h) deterministic forecasts per day with a short cutoff (RUC - already operational)
- 2 longer range deterministic forecasts per day (+48h) with a longer cutoff and more recent boundary conditions (under test in these days)
- all using boundary conditions from the 5km run on Mediterranean

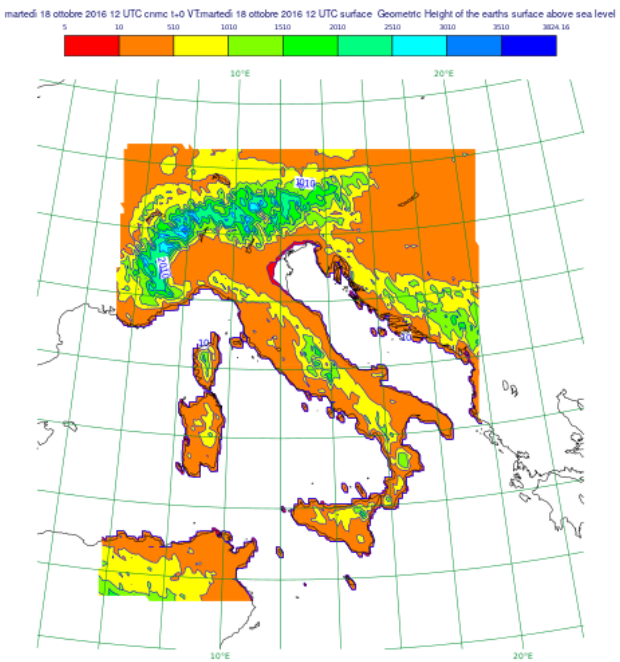
Higher resolution setup - 2.2 km



The perturbed analyses are used for initialising:

- 1 ensemble forecast per day, +48h, 20 members, (under implementation in these days)
- perturbed boundary conditions from the COMET 10 km COSMO ensemble forecast runs.

Higher resolution setup - 2.2 km



The perturbed analyses are used for initialising:

- 1 ensemble forecast per day, +48h, 20 members, (under implementation in these days)
- perturbed boundary conditions from the COMET 10 km COSMO ensemble forecast runs.

Computing centre

The operational procedures are hosted in the CINECA computing centre, chosen by means of a public tender, and the financial resources are provided by the National Civil Protection Department.

The most critical procedures are managed directly by the CINECA operators while the more innovative ones (e.g. high resolution ensemble analysis and forecast) are directly managed by Arpae personnel.

We will not bore you with the number of nodes/processors/cores, the speed of the interconnections, etc. involved in the procedures. They are simply high enough to have the forecast done in a resonable time (2.5 hours for completing the 5km and the 2.2km forecast phases).

Computing centre

The operational procedures are hosted in the CINECA computing centre, chosen by means of a public tender, and the financial resources are provided by the National Civil Protection Department.

The most critical procedures are managed directly by the CINECA operators while the more innovative ones (e.g. high resolution ensemble analysis and forecast) are directly managed by Arpae personnel.

We will not bore you with the number of nodes/processors/cores, the speed of the interconnections, etc. involved in the procedures. They are simply high enough to have the forecast done in a resonable time (2.5 hours for completing the 5km and the 2.2km forecast phases).

Computing centre

The operational procedures are hosted in the CINECA computing centre, chosen by means of a public tender, and the financial resources are provided by the National Civil Protection Department.

The most critical procedures are managed directly by the CINECA operators while the more innovative ones (e.g. high resolution ensemble analysis and forecast) are directly managed by Arpae personnel.

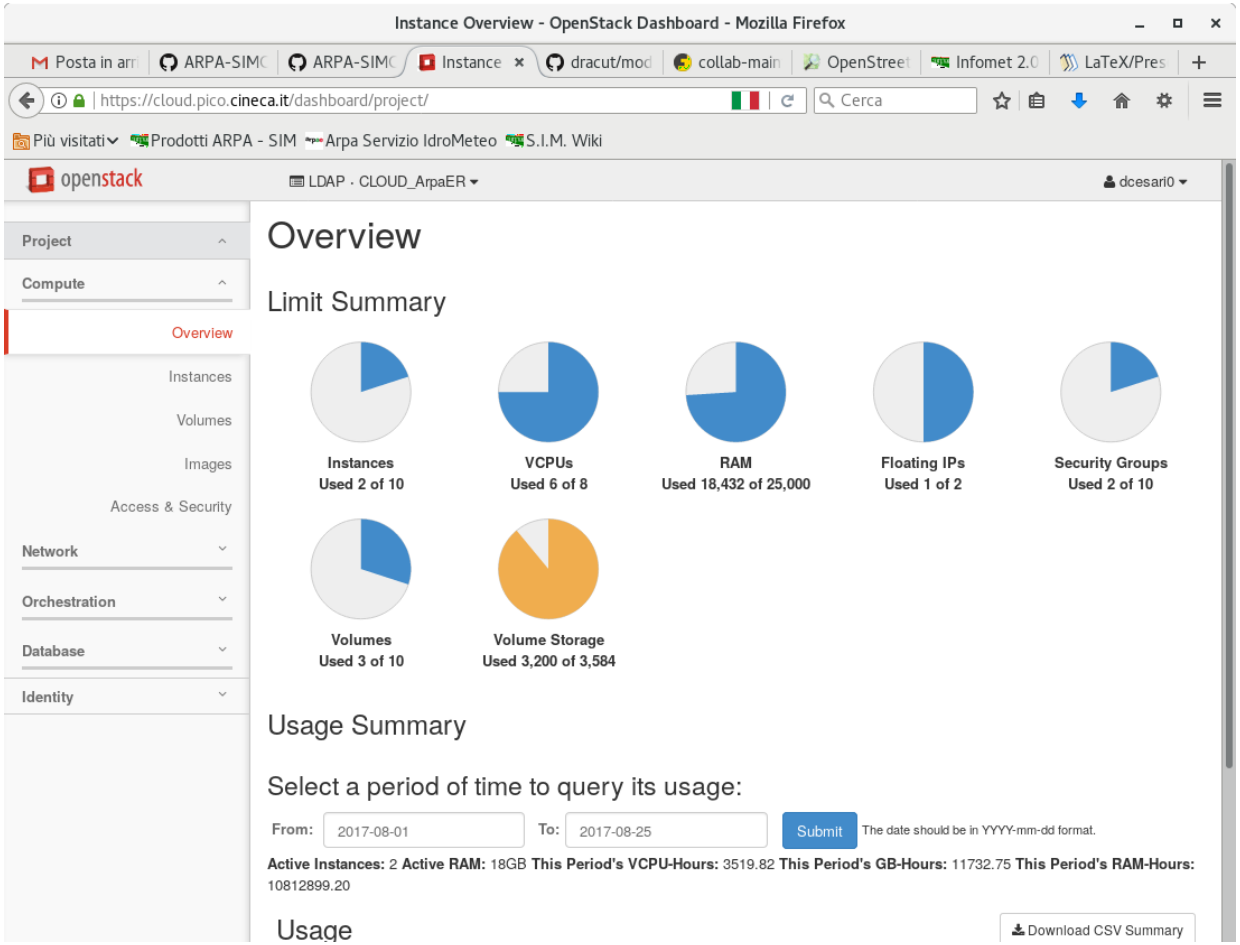
We will not bore you with the number of nodes/processors/cores, the speed of the interconnections, etc. involved in the procedures. They are simply high enough to have the forecast done in a resonable time (2.5 hours for completing the 5km and the 2.2km forecast phases).

Cloud infrastructure

In addition to HPC resources, within the Computing Centre hosting the operational procedures, Arpae has access to a small cloud infrastructure located in the data centre itself.

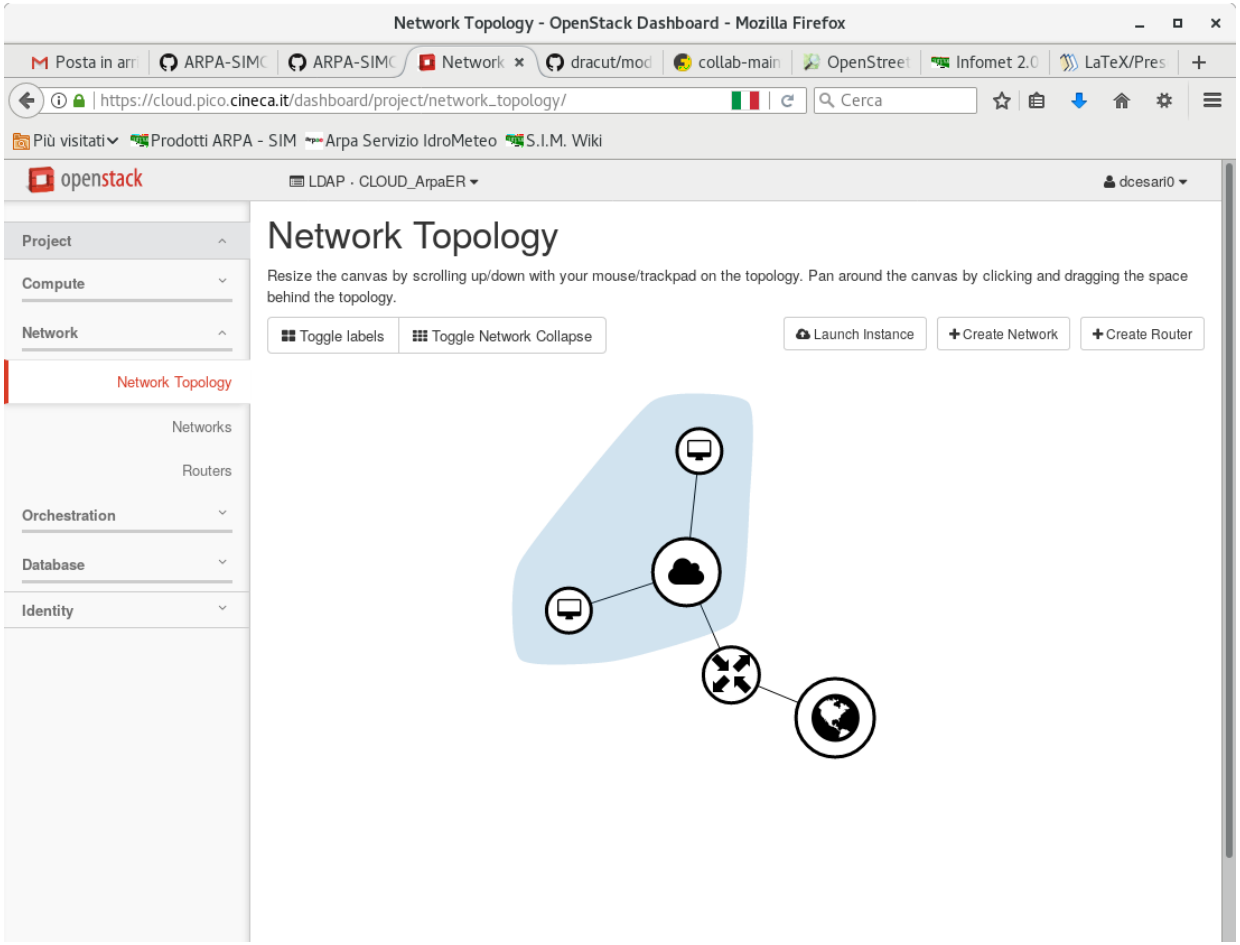
Cloud infrastructure

In CINECA this infrastructure is based on Openstack/kvm environment, and it gives Arpae access to a private network of virtual machines with arbitrary operating system and root access, open to the internet and tightly connected to the HPC systems where COSMO model is run.



Cloud infrastructure

In CINECA this infrastructure is based on Openstack/kvm environment, and it gives Arpae access to a private network of virtual machines with arbitrary operating system and root access, open to the internet and tightly connected to the HPC systems where COSMO model is run.



Cloud infrastructure

This setup is very suitable for performing tasks which do not well fit into a traditional shared-resources HPC system:

- Download and prepare input data for the operational procedures
- Perform serial data processing on model output
- Keep a rolling data archive of the full model input and output, remotely accessible by authorised users through command line and web interface
- Host a public web server with documentation about the operational procedures and data access.

Cloud infrastructure

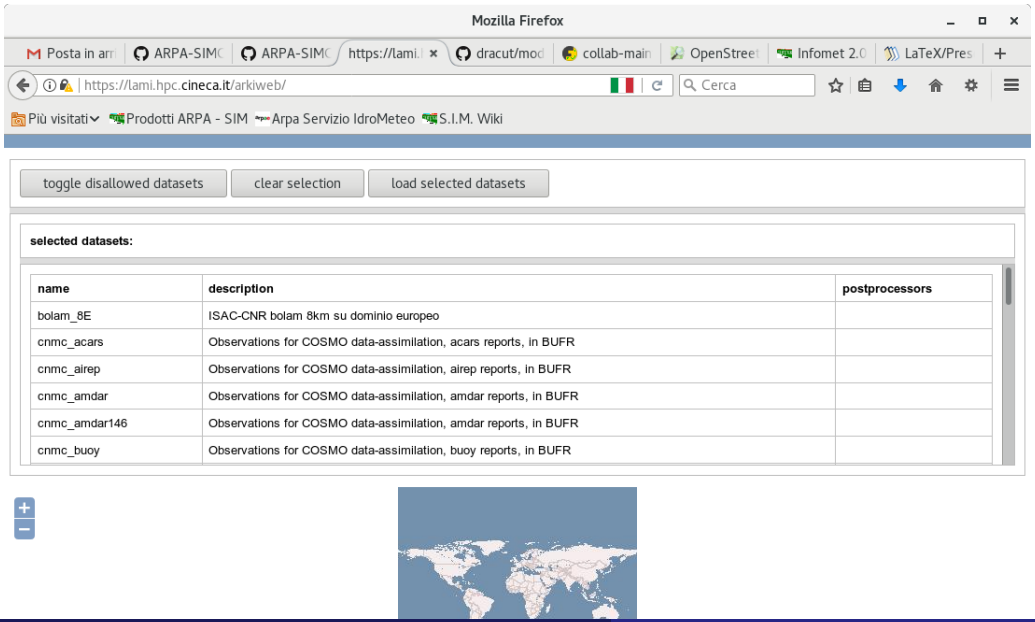
This setup is very suitable for performing tasks which do not well fit into a traditional shared-resources HPC system:

- Download and prepare input data for the operational procedures
- Perform serial data processing on model output
- Keep a rolling data archive of the full model input and output, remotely accessible by authorised users through command line and web interface
- Host a public web server with documentation about the operational procedures and data access.

Cloud infrastructure

This setup is very suitable for performing tasks which do not well fit into a traditional shared-resources HPC system:

- Download and prepare input data for the operational procedures
- Perform serial data processing on model output
- Keep a rolling data archive of the full model input and output, remotely accessible by authorised users through command line and web interface
- Host a public web server with documentation about the operational procedures and data access.



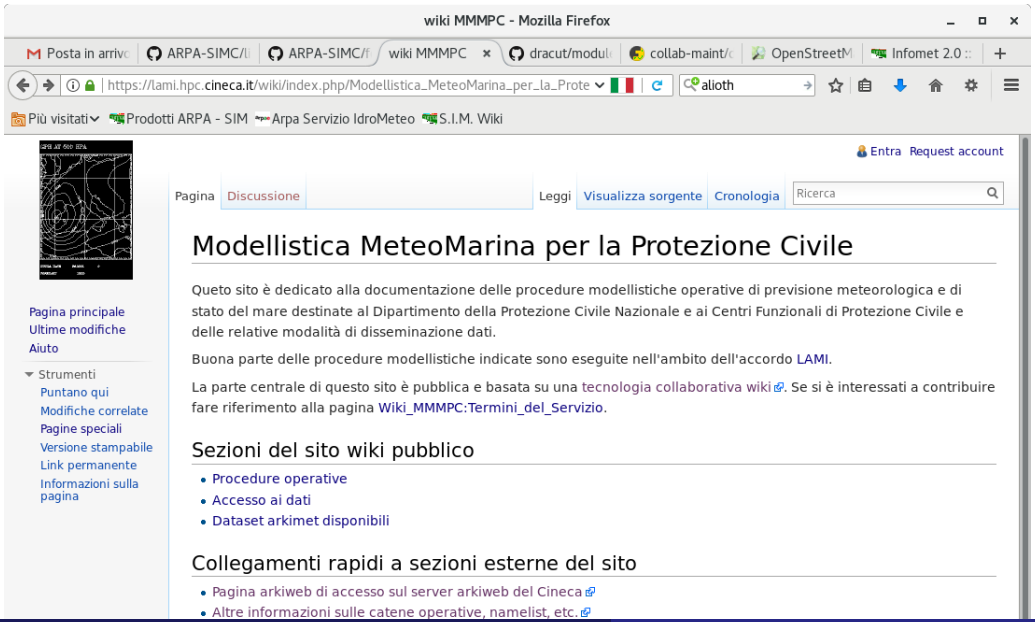
Server address:
`https://lami.hpc.cineca.it/arkiweb`

Archiving software repository:
`https://github.com/ARPA-SIMC/arkimet`

Cloud infrastructure

This setup is very suitable for performing tasks which do not well fit into a traditional shared-resources HPC system:

- Download and prepare input data for the operational procedures
- Perform serial data processing on model output
- Keep a rolling data archive of the full model input and output, remotely accessible by authorised users through command line and web interface
- Host a public web server with documentation about the operational procedures and data access.



Server address:
`https://lami.hpc.cineca.it/`

Cloud infrastructure

For the small organizations that cannot afford an own HPC data center this has some advantages:

- Allow to process and disseminate big volumes of model data without the need to transfer all the raw model output to the own site (big-data approach)
- Simplify some tasks related to the operational model procedures
- Bypass or customize the typical corporate network blockings (e.g. proxies, firewalls, virus filters) that sometimes make NWP life hard.

And disadvantages:

- Need to replicate part of the internal operational infrastructure on a different site
- Need to take care of the network security
- Possible confusion for the external users about where services are hosted.

Cloud infrastructure

For the small organizations that cannot afford an own HPC data center this has some advantages:

- Allow to process and disseminate big volumes of model data without the need to transfer all the raw model output to the own site (big-data approach)
- Simplify some tasks related to the operational model procedures
- Bypass or customize the typical corporate network blockings (e.g. proxies, firewalls, virus filters) that sometimes make NWP life hard.

And disadvantages:

- Need to replicate part of the internal operational infrastructure on a different site
- Need to take care of the network security
- Possible confusion for the external users about where services are hosted.

Cloud infrastructure

For the small organizations that cannot afford an own HPC data center this has some advantages:

- Allow to process and disseminate big volumes of model data without the need to transfer all the raw model output to the own site (big-data approach)
- Simplify some tasks related to the operational model procedures
- Bypass or customize the typical corporate network blockings (e.g. proxies, firewalls, virus filters) that sometimes make NWP life hard.

And disadvantages:

- Need to replicate part of the internal operational infrastructure on a different site
- Need to take care of the network security
- Possible confusion for the external users about where services are hosted.

Cloud infrastructure

For the small organizations that cannot afford an own HPC data center this has some advantages:

- Allow to process and disseminate big volumes of model data without the need to transfer all the raw model output to the own site (big-data approach)
- Simplify some tasks related to the operational model procedures
- Bypass or customize the typical corporate network blockings (e.g. proxies, firewalls, virus filters) that sometimes make NWP life hard.

And disadvantages:

- Need to replicate part of the internal operational infrastructure on a different site
- Need to take care of the network security
- Possible confusion for the external users about where services are hosted.

Cloud infrastructure

For the small organizations that cannot afford an own HPC data center this has some advantages:

- Allow to process and disseminate big volumes of model data without the need to transfer all the raw model output to the own site (big-data approach)
- Simplify some tasks related to the operational model procedures
- Bypass or customize the typical corporate network blockings (e.g. proxies, firewalls, virus filters) that sometimes make NWP life hard.

And disadvantages:

- Need to replicate part of the internal operational infrastructure on a different site
- Need to take care of the network security
- Possible confusion for the external users about where services are hosted.

Cloud infrastructure

For the small organizations that cannot afford an own HPC data center this has some advantages:

- Allow to process and disseminate big volumes of model data without the need to transfer all the raw model output to the own site (big-data approach)
- Simplify some tasks related to the operational model procedures
- Bypass or customize the typical corporate network blockings (e.g. proxies, firewalls, virus filters) that sometimes make NWP life hard.

And disadvantages:

- Need to replicate part of the internal operational infrastructure on a different site
- Need to take care of the network security
- Possible confusion for the external users about where services are hosted.

Cloud infrastructure

For the small organizations that cannot afford an own HPC data center this has some advantages:

- Allow to process and disseminate big volumes of model data without the need to transfer all the raw model output to the own site (big-data approach)
- Simplify some tasks related to the operational model procedures
- Bypass or customize the typical corporate network blockings (e.g. proxies, firewalls, virus filters) that sometimes make NWP life hard.

And disadvantages:

- Need to replicate part of the internal operational infrastructure on a different site
- Need to take care of the network security
- Possible confusion for the external users about where services are hosted.

Conclusions

- This new setup is an important milestone in a long process of renewal of the Italian operational setup of COSMO-based weather forecast
- This process involved a many-body interaction among different players
- Hopefully this is not an end point but the start of a long-lasting infrastructure ensuring stability and ability to innovate at the same time.