4.2 Study area

The setup of the experiments is applied to the Tiber river catchment, one of the most important basins in central Italy. The Tiber river is the third longest Italian river (405 km) and the second widest in terms of basin area extension. The total area of the catchment is 17.375 km$^2$ and represents the 5% of the total Italian territory (Multirisk plan, risk flood analysis, 2014).

The 90% of the basin area belongs to the Lazio and Umbria regions, while the remaining 10% falls in the Emilia-Romagna, Toscana, Marche and Abruzzo regional administrations. The territory is characterized by terrigenous sediments and flysch mainly composed by clay schists and silty-clay deposits. Most of the basin (54%) has an agricultural soil use, while the rest is covered by forests (39%) and urban areas (5%) (Multirisk plan, risk flood analysis, 2014).

This catchment is characterized by complex topography and typical Apennine precipitation regime (annual precipitation characterized more by liquid precipitation, than by snow accumulation) that differs from the Alpine regime. Snowfall is really rare above 500m a.s.l.. Most of the weather patterns come from the Tirreno sea, less often from the Adriatic sea. The climate is typically Mediterranean with average annual rainfall of 950 mm, and average temperature of 13.5°C with maximum temperature values in June and minimum values in January (Tarpanelli et al., 2012). In the last centuries the Tiber river flooding events are increased in frequency and intensity. Regional climate studies (Camici et al. (2010), Viterbo et al. (2011) and Multirisk plan, risk flood analysis, (2014)) confirm the global tendency of increased average temperatures, decreased total rainfall over the area and increased occurrence of extreme precipitation events. In the last 10 years, the Tiber river catchment has been interested by six extreme rainfall events defined "extraordinary" (2005, 2008, two events in 2010, 2012 and 2013) and by three intense drought periods (2003, 2007, 2012). This kind of extremes events have caused almost one billion euro of total damages to houses, infrastructures and public and private heritage in 10 years. (Multirisk plan, risk flood analysis, 2014) On average, every year the Tiber river catchment experiences a major flood during the fall/winter season. Moreover, also the secondary network and minor channels are interested by flooding events, inundations, flash floods and debris flows. The majority of these phenomena happen in the fall-winter period, in the months in between November and January of the following year.

The upper-middle part of the basin belongs mostly to the Umbria region and covers an area of 12700 km$^2$. This region represent the most interesting part of the basin for high resolution hydrometeorological modelling, since it is characterized by hill-mountain topography (from 50 m to 2500 m above sea level).

I choose to simulate the year 2012 in my experiment, since in that year the Tiber river basin experienced an intense drought period (JJA 2012) and an extreme flood event (11-14 November 2012). In both of these intense events, the regional administration asked for special recovery procedures such as the natural catastrophe status and the national state of emergency. In summer 2012, the Tiber river basin was interested by a sensible decreasing of the precipitation contributions during summer season, leading to
a reduction of the fresh water storage in the aquifer and a decrease of the water level in the principal water bodies (such as the Trasimeno Lake). Moreover, in November 2012 it experienced an extraordinary flood, with an estimated return period of roughly 100 years. In this event the precipitation amounts reached in some areas the cumulated value of 350 mm in 72 hours, equivalent to one third of the average annual rainfall. During this flood event most of the watershed and sub-basins experienced diffused inundations, debris flows and damages to the infrastructure, for a total amount of 230 million euros (Costantini et al., 2012).

4.2.1 Domain setting

The WRF Model model version 3.7.1 is used to perform both coupled and uncoupled simulations with the WRF-Hydro architecture over the study area. The two domains of my WRF and WRF/Hydro simulations are centered in the Tiber river basin and extended over central Italy in the range 35.34°N - 48.47°N and 4.61°E - 20.25°E in the external domain, resolved at 12 km (98 x 121 grid points) and and in the range 40.59°N - 45.06°N and 7.73°E - 16.67°E in the internal domain, resolved at 4 km grid spacing (177 x 123 grid points) (Fig. 4.2)

Figure 4.2: The two nested domains used for the simulations: external domain d01 resolved at 12 km resolution and inner domain d02 resolved at 4 km.

The disaggregation factor between the atmospheric simulations and the hydrological routing grid is 20, meaning that the hydrological processing is performed at 200 m grid
resolution, only in the most interior domain (d02). The time step of the WRF model is 15 seconds on the large d01 domain and 5 seconds in the smallest d02 domain.

4.2.2 Observational datasets

Due to the historical importance of the basin, the area is characterized by a pretty dense hydro-meteorological network (on average one station every 150 km$^2$) and a long time series of 30 minutes real time data measurements over more 25 years of data collection (Tarpanelli et al., 2012). The total number of ground measurement stations (raingauges and thermometers) inside the basin is 192. In the whole d02 domain, the number of stations is 3991 (Fig. 4.3). Station observations are provided by the Italian civil protection network every 30 minutes.

![Figure 4.3: Meteorological stations over the d02 domain. In blue the meteorological stations inside the basin.](image)

The station data are used after a simple filtering operation to control the quality of the observations: only stations with more than 80% of valid measurements over the year (not "not a number") are left. Moreover, the raingauge measurements that are negative or over 200 mm/h are set to "not a number" in the timeseries. The 200 mm/h threshold is decided considering a corresponding return period of more than 500 years for a rainfall duration of one hour, according to the most updated height duration frequency curves Morbidelli et al. (2016). Even if there are more sophisticated methods to eliminate spurious measurements from databases, this one that I used is just intended to provide a preliminary quality check for my measurements.

Inside the basin there are also 43 hydrometers over different sections of the main channel and in the secondary channel network. The Monte Molino river section (5269
km$^2$ of catchment area underlied) is selected to compare the observed hydrograph with
the model hourly time steps outputs in terms of streamflow. I choose this section for the
WRF and WRF-Hydro model comparison, since it is one of the most complete in the
data time-series and results to be located in the medium part of the basin, far enough
from the main reservoir of Montedoglio ($168x10^6 m^3$ of maximum water volume stored)
(Brocca et al., 2011).

In addition to that the model hydrological surface processes has been verified using
the in-situ soil moisture measurements (Fig. 4.4) in the basin and the available station
flux observations in the region of interest (Fig. 4.5). Eleven soil moisture stations
are available in the basin, mostly concentrated in the upper part of the basin for the
years 2012 (evaluation and calibration period) and 2013 (for the validation period). The
soil moisture stations data has been collected from the Umbria Region CFD (Centro
Funzionale Decentrato) database.

Figure 4.4: Soil moisture stations inside the basin.
The available flux stations for the time period of my simulations are five, from the FLUXNET database (DAAC), 2016). Even the flux stations are located inside the d02 domain, the resulting informations from the station has to be taken with cautions since their location is spatially concentrated in two areas (western Lazio region in Viterbo province for the stations named "ITCA1, ITCA2, ITCA3" and "ITRO4" and western Abruzzo region in Aquila province for the station named "ITCOL") that results to be influenced by different large scale rainfall patterns (the first area is more influenced by precipitating systems coming from the Tirrenian sea, while the second area is more Adriatically-influenced) (Fig. 4.5. For all the aforementioned reasons, the flux station data observations has to be taken with care when compared to the models simulations.

Figure 4.5: Flux stations inside the domain.

4.2.3 Computational resources

Running WRF and WRF-Hydro models require the use of high performance computing (HPC) resources. Even if these aspects are less related to the specific scientific question, the computational resources are fundamental pre-requisite to be addressed in order to perform high resolution model simulations at the event and, even more, at seasonal scale.
For this reason I strongly believe that is important to dedicate one section of the present chapter to discuss more in deep the computational resources required for the seasonal scale experiment.

The WRF and WRF-Hydro codes are written in FORTRAN90 and are parallelized for the execution of the code on HPC architectures including LINUX systems community clusters, multi-processor desktops and IBM 'blue gene' supercomputers. The parallelization of the code utilizes a geographic decomposition and 'halo' array passing structures. Even if the parallelization, together with a responsible administration of the available resources permit to the code to be more efficient, the high resolution hydrometeorological simulations are, in general, very demanding in terms of computational costs. An important part of the experiment is related to access to the computational resources, considering a good preliminary evaluation of the computational cost by means of scaling tests. The major part of the Tiber river experiment was run over the Tier-1 system Galileo at Consorzio Interuniversitario per il Calcolo Automatico dell'Italia Nord Orientale (CINECA) in Italy, even if some preliminary studies have been carried using the Supermuc cluster at LRZ/Germany cluster (more detailed information about the two cluster can be found at http://www.hpc.cineca.it/hardware/galileo and at https://www.lrz.de/services/compute/supermuc/systemdescription/). Once that the computational resources are obtained, an important part of the work consisted in choosing a good compromise between the number of core necessary for the simulations and the wall-clock time required for the simulations. Assuming a good code scalability, the higher is the number of cores used and the faster is the code to run, but meaning that it is also more computationally costly. After some preliminary studies, I decided to run my simulations using 64 cores on Galileo HPC machine.

The total amount of computational resources used in this project is around the 220000 core hours, spanned over a total of three ISCRA-C project on the Galileo machine. The WRF stand alone model setting required about 25000 core hours for each run, for a total of 50000 hours; the WRF-Hydro stand alone calibration required 50000 hours to perform a total of 32 runs; the WRF/WRF-Hydro comparison required a total of 81000 hours to perform the WRF and WRF/WRF-Hydro comparison at the seasonal scale. Finally an additional 60000 hours were required for debug, test cases, preliminary investigations at the event scale and scaling test analysis. As evident from these numbers, the WRF-Hydro stand alone results to be less computationally costly than the WRF and WRF/WRF-Hydro that have to compute also the atmospheric processes. While for a WRF-Hydro stand alone run are needed 45 hours of wall-clock time to simulate one year, the WRF and WRF/WRF-Hydro runs require 32 and 42 days of wall-clock time, respectively. In general, the WRF/WRF-Hydro result to be more computationally costly of 25-30% of the total core hours respect to the WRF stand alone configuration. In the light of future possible operational application of the model, the present configuration will need 2,1 hours and 2,7 hours in WRF and WRF/WRF-Hydro mode respectively to run a forecast for the next 24 hours.
4.3 WRF stand-alone model experiment

The WRF stand-alone model is run over the study area, according to the domain setting illustrated in section 4.2.1, in a two-way nesting mode. A set of different model configurations in terms of convective closures and microphysics schemes is explored to find a model configuration that can properly reproduce the main atmospheric processes over the area. The different model runs are illustrated in tab 4.1.

<table>
<thead>
<tr>
<th>Configuration</th>
<th>Convective closure d01</th>
<th>Convective closure d02</th>
<th>Microphysics</th>
</tr>
</thead>
<tbody>
<tr>
<td>Exp-WSM6</td>
<td>Kain-Fritsch</td>
<td>Explicit</td>
<td>WSM6</td>
</tr>
<tr>
<td>Exp-Thompson</td>
<td>Kain-Fritsch</td>
<td>Explicit</td>
<td>Thompson</td>
</tr>
</tbody>
</table>

The outermost domain at 12 km adopts Kain-Fritsch parameterization, while the innermost domain at 4 km has been resolved explicitly. These choices in terms of convective closure are motivated both by the experience gained in the previous study on Pakistan flood 2010 (see chapter 3) and according to the study of Pieri et al. (2015). In the Pakistan flood 2012 experiment, it resulted to be more effective to run the WRF model in explicit mode at 3.5 km, instead that using a convective closure at "grey zone" resolution (corresponding 1-5 km) (see section 3.6.1 for more detailed explanation). In addition to that, Pieri et al. (2015) performed WRF sensitivity studies over the EURO-CORDEX domain with a special focus over complex topography areas at 4 km resolution (the same of my simulations) at the climatological time scale (1979-1998). The Pieri et al. (2015) final conclusions indicate that the best results over mid-latitude european complex topography areas are gained resolving convection explicitly at 4 km of grid-spacing and running the simulations with Thompson microphysics.

In this analysis, the author decided to run the model explicitly and to perform the comparison between the WSM6 single moment (WSM6) versus Thompson double-moment microphysics. The main features of these schemes are already introduced in section 3.4.1 and further informations can be found in Hong and Lim (2006) and Thompson et al. (2008).

The vertical dimension is discretized in 60 levels, since a good vertical resolution is needed to study the complex orography area. The turbulent parameterization for the planetary boundary layer is the Yonsei University Scheme (Hong et al., 2006). The radiation scheme used is the RRTM scheme for longwave radiation (Mlawer et al., 1997) and the Goddard scheme (Chou and Suarez, 1999) for the short-wave parameterization. The land use dataset is given by the USGS 24-category data available from the WRF Preprocessing System (WPS) and the land surface model is the Noah-MP. The topography for WRF model simulation is given at the finest available resolution (30 arc-seconds) from the WPS database.

The initial and boundary conditions are provided by ERA-Interim re-analysis (Dee
et al., 2011b) every 6 hours at the 0.75° native resolution, performing the experiment in hindcast mode. Sea Surface Temperature updates (SST) are given by ERA-Interim reanalysis, as well. More informations about ERA-Interim reanalysis product are given in section 2.2.2.

4.3.1 WRF-stand alone model setting

The different WRF model configurations are run for the year 2012 over the study area. Model output are compared to observations mainly in terms of rainfall variables, looking both at precipitation patterns and intensity. The observed precipitation fields are obtained interpolating the raingauges station measurements at 4 km using Kriging technique (using a radius and a spherical kernel) (Krige (1966), Matheron (1967), Wackernagel (1995) and Minasny and McBratney (2005)). Since raingauge observations are only available over land, the precipitation maps obtained both from model runs and observations are masked over the ocean (results are compared only over land). The model runs are evaluated at different time scales in terms of accumulated precipitation maps, scatterplots, CDFs, average daily cycle of precipitation and rainfall accumulation values and RMSE at the catchments scale.

In the first place the monthly accumulated rainfall over the area is compared, as simulated by the model using the two selected microphysics and the raingauges based observed rainfall depth maps. Because of the huge amount of figures, only the most explicative figures are shown in the text. For the complete set of figures, the reader is referred to the appendix B.1. An example of the monthly accumulated map is given both for a summer month (August) (Fig. 4.6) and a winter month (December) (Fig. 4.7). In both of the figures, the model shows a good accordance with the observed precipitation (first lines of Fig. 4.6 and Fig. 4.7). In order to understand if the two model runs systematically reproduce some model induced patterns due to interaction with the orography or other specific model induced factors, I also had a look at the accumulated rainfall over 15 days within the month, randomly chosen. I consider not only the accumulated rainfall over the whole month, but I also check if the different set-up of WRF model can reproduce the precipitation variability properly inside the month. If we look the model configuration columns ((a), (d) and (g) and (b), (e), (h)) of figure 4.6 and 4.7 it is possible to check if there are some systematic patterns of precipitation that repeats over the same model configurations, or if the randomly chosen daily rainfall is varying enough in the month. The second and third lines of the figures 4.6 and 4.7 does not show a systematic behavior in reproducing specific patterns over the orography in both configurations.
Figure 4.6: Accumulated monthly rainfall map analysis for the month of August. Total accumulated rainfall values over the month ((a)-(c)), accumulated rainfall over a first sample of 15 random days ((d)-(f)), accumulated rainfall over a second sample of 15 random days (for (g)-(i) for (from left to right) Exp-Thom, Exp-WSM6 and observations.

Figure 4.7: Accumulated monthly rainfall map analysis for the month of December. Total accumulated rainfall values over the month ((a)-(c)), accumulated rainfall over a first sample of 15 random days ((d)-(f)), accumulated rainfall over a second sample of 15 random days (for (g)-(i) for (from left to right) Exp-Thom, Exp-WSM6 and observations.
Even if the monthly accumulated maps helps on a first instance to understand the precipitation patterns and intensity over the area at the monthly scale, the Cumulative Distribution Functions (CDFs) gives an additional quantitative representation of the precipitation distribution over the domain of the hourly values over the each month (Fig. 4.8).

In general the Thompson scheme tends to better reproduce low values of the distributions, especially for the months of January, February and March. It is more difficult to individuate a clear tendency to overestimate or underestimate for the highest values of the distributions for both configurations (WSM6 and Thompson), compared to observations. In general the two microphysics schemes tend to overestimate/underestimate the observations accordingly for the same months. Nevertheless, the WSM6 configuration seems to produce more often precipitation distributions with higher values than Thompson simulation. In the months of July, August, September and October the two configurations exhibit few differences in terms of distribution.

The monthly averaged diurnal cycle of the precipitation is obtained by averaging the rainfall intensity over each month for the same hour of the day. The results provide the monthly average rainfall variability over the 24 hours. The graphs of the two simulations shows an overall good agreement with the average observed diurnal precipitation cycle. During summer season (months of June, July, August) and in September, both model configurations tends not to well depict the observed double diurnal peak over the day, even if the Thompson manage only partially to reproduce a less intense second peak in August and September and a better correlation. The example of this behavior is given in Fig. 4.9 for August, the rest of the graphs are reported in the appendix B.2.
Figure 4.8: Cumulative Distribution functions (CDFs) for the different model configurations: Explicit Thompson (blue line), Explicit WSM6 (green line) and raingauge observations (red line). CDFs from (a) to (n) refer to the different months of the year 2012 (from January (a) to December 2012 (n) from left to right and from the top to the bottom).
On February all the model runs strongly deviate from the observed behavior not only in terms monthly rainfall maps (Fig. 4.10), but also if we look at the diurnal cycle (totally uncorrelated) (Fig. 4.11) and at the CDF (Fig 4.8 panel (b)). For both configurations, model results are quite different from the observations.
Figure 4.10: Accumulated monthly rainfall map analysis for the month of February. Total accumulated rainfall values over the month ((a)-(c)), accumulated rainfall over a first sample of 15 random days ((d)-(f)), accumulated rainfall over a second sample of 15 random days (for (g)-(i) for (from left to right) Exp-Thom, Exp-WSM6 and observations.

Figure 4.11: Average precipitation diurnal cycle over the month of February.
A possible explanation to this behavior is given by an exceptional snowfall that interested most of the Italian peninsula in the first half of February 2012 (from February 1st to 12th). WRF model seems not able to properly reproduce the characteristic of precipitation, producing different rainfall dynamics for liquid precipitation instead of solid precipitation. In order to prove this, the hourly averaged rainfall inside the d02 domain is analyzed for different configurations and observations. As evident from figure 4.12, the two configurations perform similar simulations, different from the observations. In the first half of the month, when actually the main snowfall event occurred, the WRF model tends to overestimate liquid precipitation and the observations seem to be strongly influenced by melting processes. In the second half of the month, when the snowfall is over, the model manage to reproduce better the observations than the first half of February.

![Figure 4.12: Average hourly rainfall over the d02 domain for the month of February 2012: comparison between WRF model simulations and interpolated raingauge observations.](image)

In addition to that, the results from scatterplots are also analyzed. Fig. 4.13 represents scatterplot of the daily rainfall averaged over the d02 domain, for the whole year 2012. The graph also shows the associated statistical values in terms of regression coefficient, correlation coefficient $R^2$ and RMSE for every configuration.

The points are not too dispersed along the "perfect simulation" line (in green) for both configurations and, in general the two simulations results to be quite similar. While the regression coefficient result to be closer to one in the Thompson configuration (in blue), the $R^2$ and RMSE scores result to be better for the WSM6 configuration.

In order to move from the pure meteorological evaluation of WRF to a more hydrologically based investigation, the last part of these analyses focuses more to the basin scale, Tab. 4.2 shows the calculated monthly averaged values over the catchment area.
The evaluation of the precipitation amounts over the catchment confirms the fact that most of the time both configurations are in accordance in terms of overestimation/underestimations over the single months, compared to the observed fields. Nevertheless, in some months, the two model settings behave differently and one microphysics scheme overestimates rainfall, while the other underestimates. This happens mostly during summer months (June, July) and the months of transition to summer (May and September). These different behavior between two model simulations suggests an higher model uncertainty during summer season in predicting local convective events, as stated in numerous works in literature (e.g. Olson et al. (1995), Moser et al. (2013)). Biggest differences in precipitation amounts between WRF simulations and observations are detected for the months of February, October and November. In February the WSM6 and Thompson microphysics tends to overestimate the liquid contribution of precipitation without solid precipitation due to the snowfall event. The months of October and November falls in the most rainy season (1-14 November the catchment experienced a major flood event - see for more details section 4.2) and the WRF stand alone model tends to produce less rain then observed over the basin in both the chosen configurations. Overall, in some cases the Thompson microphysics scheme is closer to the observations than the WSM6, whereas in other cases is also true the contrary.

From all the analyses shown before it is not trivial to choose an outstanding model
settings, since the two model configurations behave similarly both at basin scale and at domain scale. The two model configurations can reproduce the main physical rainfall processes, except the snow storm in February. In some cases the WSM6 seems to produce better rainfall estimates, while in other months Thompson microphysics provides better performances. Since the evaluation is not giving a clear signal in choosing one microphysics instead of another, and considering that for a more aware parameterization choice a climatological study over several years can be considered more reliable (especially if we consider the analysis of the diurnal cycle of precipitation), I decided to follow the recommendations of the paper of Pieri et al. (2015). That study investigated sensitivity of precipitation statistics to resolution, microphysics, and convective parameterization over EURO-CORDEX domain, using WRF model operated at 4 km of grid resolution over 30 years with a special focus over complex topography areas. Pieri et al. (2015) conclude that the best configuration for this area in terms of precipitation statistics result to be Kain-Fritsch convective closure over d01, explicit configuration at 4 km innermost domain and Thompson microphysics scheme.

Since the WRF stand-alone configuration is chosen, it is now possible to proceed to the calibration of the WRF-Hydro stand alone over the area in order to set the land surface model and hydrological routing parameters to reproduce the main characteristic of the water cycle over the catchment.

### 4.4 WRF-Hydro calibration experiment

In the second part of this thesis, the hydrological model needs to be calibrated at seasonal scale, before performing the WRF-Hydro fully-coupled run. As introduced in section 2.4,
the WRF-Hydro suite needs meteorological input variables in order to perform flux calculations and hydrological routing. In order to calibrate the WRF-Hydro stand alone model, the model is fed with meteorological forcing derived from the available observations (instead than provided by WRF), at the same grid spacing and extent of the WRF d02 domain (Fig. 4.2).

The observational forcing data have been provided to WRF-Hydro at hourly time resolution and the time-step of the simulation over the high resolution grid is 5 seconds. This time step value respects the CFL condition at 200m hydrological grid resolution (see section 2.4).

In these WRF-Hydro stand alone simulations the overland flow routing, channel routing and base flow module are activated for the routing calculations.

In addition to meteorological variables, also terrain data information has to be provided to the model suite, in order to build the high-resolution routing grid on which the hydrological processes are resolved. Terrestrial information of stream channel network and water bodies (such as lake, reservoir, and ocean), groundwater basins and routing are derived by DEM information, using GIS tools. Moreover, static data information about land use, soil type and vegetation are provided to the model, using WRF WPS pre-processor. While for the Pakistan flood experiment (section 3.4) and for the WRF stand alone experiment of section 4.3 the standard WRF initialization dataset is used, for the WRF-Hydro calibration experiment land use has been provided using the most updated CORINE land use dataset Büttner (2014). More details about terrestrial data preparation are given in section 4.4.2.

The model hourly time steps output in terms of channel routing are compared against the observed stream flow measurements at Monte Molino river section. In addition to that the routing processes and the infiltration and exfiltration processes are verified using the available station flux observations in the region of interest and the in-situ soil moisture measurements (section 4.2.2, Fig. 4.3, Fig. 4.4 and Fig. 4.5). In the process of hydrological calibration, since the model is fed by inputs that are as closer as possible to reality, the target is to adjust the model parameters in order to reproduce model results that are as close as possible to the observations. The main WRF-Hydro model parameters concerns LSM calibration (vegetation parameters, soil parameters etc.) and hydrological routing processes (channel routing, surface runoff, ground water routing etc.). Even if the parameters are contained in different parameters tables, thus reflecting in principle different aspects of the routing, infiltration and exfiltration processes, there is a strong intercorrelations phenomena among them and it is complicated to separate each single parameter contribution. For a more detailed description of hydrological calibration method the reader is referred to the 1.7 section.

Finally, after that the model is calibrated over the year 2012, it needs to be validated over a different period in order to finally consider the model parameter able to be representative of the hydrological processes of the area, for slightly different case studies. In the experiment, the validation is carried over the following year (2013).
4.4.1 Meteorological forcing

The meteorological forcing is prepared in order to provide the model with meteorological variables that are as close as possible to the observations. The required variables are incoming shortwave and longwave radiation (W/m²), specific humidity (kg/kg), air temperature (K), surface pressure (Pa), near surface u- and v-components for wind (m/s) and liquid water precipitation rate (mm/s).

Rainfall and temperature hourly observations are provided by the Italian Civil protection agency ground station network (3991 stations) and interpolated at the meteorological model grid spacing (4 km). The input forcing data in terms of shortwave and longwave radiation, near surface wind speed (u and v components), specific humidity and surface pressure are given by IFS analysis data at 0.125° resolution re-interpolated at 4 km resolution by mean of linear interpolation. IFS analysis are an ECMWF product and further informations can be found in White (2002).

In order to interpolate the ground-station derived fields of temperature and rainfall, I initially try different interpolation techniques (linear, Inverse Distance Weighting, Kriging) on shorter time periods (2-3 months) to find a method that can be a good compromise between the necessity to have good precision in the interpolated fields and the computational time required for the interpolation. After this first comparison (not shown), ordinary Kriging is applied to rainfall measurements (radius of 30 pixels on the 4 km WRF d02 grid and a spherical kernel) and a linear interpolation is used for temperature stations. In addition to that, the obtained temperature fields are corrected considering the topography, applying a linear regression relationship for every pixel function of height (equation 4.1). Δh is calculated as the difference between the WRF topography aggregated at 12 km (resolution comparable to the IFS 0.125°) and the original 4 km WRF DEM. The coarse resolution DEM is calculated aggregating the HGT_M WRF variable from 4 km to 12 km using using a moving average filter. The resulting air temperature \( T_a \) is derived from from the original value \( T_{a0} \) according to the environmental lapse rate \( L_a = 6.49 \text{ K/km} \) as defined in Standard ICAO atmosphere (equation 4.1) (ICAO, 1993).

\[
T_a = T_{a0} + L_a \Delta h \quad \text{where} \quad L_a = -0.00649(\text{K/m}) \tag{4.1}
\]

Specific humidity is calculated in a prognostic way (equation 4.2), function of dew point temperature \( T_d \) (equation 4.3) surface pressure \( (P_s) \) (equation 4.4) and vapor pressure \( (e) \) (equation 4.5).

\[
rho_o = \frac{R_L}{R_W} \times \frac{e}{P_s + e \left(1 - \frac{R_L}{R_W}\right)} \tag{4.2}
\]

The interpolated meteorological variables from IFS are corrected as a function of height, according to the following relationships (equations 4.3, 4.4).

\[
T_d = T_{d0} + L_d \Delta h \quad \text{where} \quad L_d = -0.004(\text{K/m}) \tag{4.3}
\]
\[ P_s = P_{s0} + L_a \left( \frac{\Delta T_a}{T_{a0}} \right) \left( \frac{g R}{M} \right) \] (4.4)

Dew point temperature \((T_d)\) is calculated from the original IFS-derived fields \((T_{d0})\), function of the height gradient \((\Delta h)\) and wet adiabatic lapse rate \((L_d = -4 \text{ K/km})\) (equation 4.3). Surface pressure is corrected function of topography, considering the thermodynamic relation between pressure and temperature and the air temperature corrections \((\Delta T_a)\) calculated in equation 4.1.

In equation 4.4 \(g\) is equal to the earth-surface gravitational acceleration \((g = 9.80665 \text{ m/s}^2)\), \(R\) is the universal gas constant \((R = 8.31447 \text{ J/(molK)})\) and \(M\) is the molar mass of dry air \((M = 0.0289644 \text{ kg/mol})\).

Vapor pressure \(e\) (equation 4.5) is derived from dew point temperature \((T_d)\) in Celsius and calculated in equation 4.3.

\[ e = 6.1078 \times 10^{\left(\frac{(T_d-A)}{(T_d+B)}\right)} \] where \(A = 7.5\) and \(B = 273.3\) (4.5)

All the meteorological input has been calculated using this method at hourly timescale for the whole 2012 year in order to provide the forcing for the WRF-Hydro calibration. The same forcing has been prepared also for the year 2013 with the same procedure, for the model evaluation phase (see section 4.4.6).

4.4.2 Terrestrial data

In addition to meteorological forcing, static terrain informations are needed by the WRF-Hydro model. Some of the gridded informations required by the LSM can be provided by the WRF geogrid file prepared using WPS or can be manually built by the user using netcdf files. The LSM data required are topographic elevation, geographical information about latitude and longitude in the domain, land use fraction, soil texture class, monthly mean vegetation greenness fraction and monthly mean albedo. In all my experiments, the LSM static initializations are given by geogrid files prepared with WPS.

In order to provide the model with the most state of the art land use, the soil category information are updated using the CORINE land use dataset (Büttner (2014)), instead of the original USGS 24 category land use dataset provided by the default options in WPS and used in the initial WRF stand-alone analysis of section 4.3. CORINE land cover dataset is ingested in the WPS pre-processing system using the method described in Arnold et al. (2010). The CORINE dataset is reclassified in 24 USGS categories according to the Table 1 of the paper of Pineda et al. (2004). In this way it is possible to obtain a remapped land use based on the most updated CORINE land cover dataset but in a format that is compliant to the WPS preprocessing system.

In addition to that, high resolution terrain data that describes topography, channel network, stream order, the presence of lakes or reservoirs, stream gauge points in the channel networks, ground water river basin, geographical information of latitude and longitude are needed as static input for the hydrological part to route water across the
landscape. Moreover the routing grid include eventual distributed specifications for over-
land flow roughness and surface retention depth scaling parameters. The high resolution
routing grid is prepared using an ArcGIS tool developed by NCAR-RAL for WRF-Hydro
initialization, which requires as input WRF topography from geogrid file and an high res-
olution DEM. The DEM used in this study is the 3 arc-second (approximately 90 m at
the equator) void-filled HYDROSHED DEM (Jarvis et al., 2004). In this experiment I
do not add any lake or reservoir specification to the grid and the scaling parameters for
overland flow roughness and surface retention depth are initially set to zero over all the
domain.

Once that all the static informations for LSM and hydrological routing are given, the
WRF-hydro can be run for the manual calibration.

4.4.3 Manual calibration

WRF-Hydro calibration over the area is performed with the aim of reproducing the ob-
served hourly hydrograph of the Monte Molino gauging station, by using a multi-step
manual calibration approach similar to the study of Yucel et al. (2015) and Hogue et al.
(2000). At the beginning the most sensible parameters for calibration are identified, and
then the obtained results are refined testing the sensitivity of the model to other parame-
ters. All the calibration is performed manually in order to better understand the physical
processes connected to the single variation of each parameter (Yilmaz et al., 2008). As
many physic-based modelling approaches, there is a strong parameters interaction in
some of the modelled processes and the risk of equifinality may occur. Nevertheless,
a manual calibration requires less computational time than an automatic calibration
technique (Yucel et al., 2015).

According to the previous calibration studies (Senatore et al. (2015); Yucel et al.
(2015); and Givati et al. (2016)) the calibration runs appear to be most sensible to
REFKDT, REFDK, OVROUGHRTFAC values, adjusting infiltration/runoff generation,
hydraulic conductivity and overland routing roughness scaling coefficient. In particular
Yucel et al. (2015) find that the model is more sensible to volume adjustments of the hy-
drograph in terms of infiltration factor (REFKDT) and surface retention depth (RETDE-
PRTFAC), while it is more sensible to timing adjustments with the roughness coefficients
of Manning for channel routing (MANN) and surface roughness (OVROUGHRTFAC).
The adopted step-wise calibration strategy in a first phase permits to adjust the model
to correctly reproduce the water balance in terms of volumes and then, when the amount
of right amount of water in the system is calibrated, the time is adjusted. All the cali-
bration parameters investigated in the study are varied from the original default values
of the parameters table.

I started the calibration procedure varying the REFKDT parameter that regulated
the water that infiltrate to the soil between surface and subsurface runoff. The parameter
is varied in a range between 0.1 and 5 to adjust the hydrograph volume, for a total of 12
calibration runs (Tab. 4.3). According to the resulting runs, the higher is the REFKDT
and the higher is the volume of the hydrograph and the more pronounced are the flow
peaks (Tab. 4.3).
Table 4.3: Summary statistics for the different calibration runs analyzed.

<table>
<thead>
<tr>
<th>REFKDT</th>
<th>RETDEPRTFAC</th>
<th>SATKDT</th>
<th>Q Peak</th>
<th>Q Obs</th>
<th>T peak</th>
<th>T Obs</th>
</tr>
</thead>
<tbody>
<tr>
<td>5</td>
<td>default</td>
<td>default</td>
<td>253.98</td>
<td>1337.9</td>
<td>01/12/12 15:00</td>
<td>13/11/12 05:00</td>
</tr>
<tr>
<td>2.5</td>
<td>default</td>
<td>default</td>
<td>241.43</td>
<td>1337.9</td>
<td>01/12/12 16:00</td>
<td>13/11/12 05:00</td>
</tr>
<tr>
<td>1</td>
<td>default</td>
<td>default</td>
<td>479.52</td>
<td>1337.9</td>
<td>01/12/12 11:00</td>
<td>13/11/12 05:00</td>
</tr>
<tr>
<td>0.5</td>
<td>default</td>
<td>default</td>
<td>1023.3</td>
<td>1337.9</td>
<td>12/11/12 15:00</td>
<td>13/11/12 05:00</td>
</tr>
<tr>
<td>0.45</td>
<td>default</td>
<td>default</td>
<td>1786.1</td>
<td>1337.9</td>
<td>13/11/12 02:00</td>
<td>13/11/12 05:00</td>
</tr>
<tr>
<td>0.4</td>
<td>default</td>
<td>default</td>
<td>1289.7</td>
<td>1337.9</td>
<td>12/11/12 02:00</td>
<td>13/11/12 05:00</td>
</tr>
<tr>
<td>0.37</td>
<td>default</td>
<td>default</td>
<td>1532.8</td>
<td>1337.9</td>
<td>13/11/12 01:00</td>
<td>13/11/12 05:00</td>
</tr>
<tr>
<td>0.35</td>
<td>default</td>
<td>default</td>
<td>1175.2</td>
<td>1337.9</td>
<td>13/11/12 02:00</td>
<td>13/11/12 05:00</td>
</tr>
<tr>
<td>0.3</td>
<td>default</td>
<td>default</td>
<td>3467.4</td>
<td>1337.9</td>
<td>01/12/12 15:00</td>
<td>13/11/12 05:00</td>
</tr>
<tr>
<td>0.4</td>
<td>default</td>
<td>default</td>
<td>1354.3</td>
<td>1337.9</td>
<td>12/11/12 02:00</td>
<td>13/11/12 05:00</td>
</tr>
<tr>
<td>0.3</td>
<td>default</td>
<td>default</td>
<td>165.98</td>
<td>0.9 x 10^-6</td>
<td>13/11/12 02:00</td>
<td>13/11/12 15:00</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>REFKDT</th>
<th>RETDEPRTFAC</th>
<th>SATKDT</th>
<th>RMSE</th>
<th>RHO</th>
<th>Nash-Sutcliffe</th>
</tr>
</thead>
<tbody>
<tr>
<td>5</td>
<td>default</td>
<td>default</td>
<td>99.27</td>
<td>0.53</td>
<td>0.26</td>
</tr>
<tr>
<td>2.5</td>
<td>default</td>
<td>default</td>
<td>94.97</td>
<td>0.60</td>
<td>0.32</td>
</tr>
<tr>
<td>1</td>
<td>default</td>
<td>default</td>
<td>79.16</td>
<td>0.81</td>
<td>0.53</td>
</tr>
<tr>
<td>0.5</td>
<td>default</td>
<td>default</td>
<td>58.46</td>
<td>0.89</td>
<td>0.74</td>
</tr>
<tr>
<td>0.45</td>
<td>default</td>
<td>default</td>
<td>56.88</td>
<td>0.89</td>
<td>0.76</td>
</tr>
<tr>
<td>0.42</td>
<td>default</td>
<td>default</td>
<td>66.09</td>
<td>0.84</td>
<td>0.67</td>
</tr>
<tr>
<td>0.4</td>
<td>default</td>
<td>default</td>
<td>56.71</td>
<td>0.88</td>
<td>0.76</td>
</tr>
<tr>
<td>0.35</td>
<td>default</td>
<td>default</td>
<td>57.70</td>
<td>0.87</td>
<td>0.75</td>
</tr>
<tr>
<td>0.3</td>
<td>default</td>
<td>default</td>
<td>65.89</td>
<td>0.83</td>
<td>0.68</td>
</tr>
<tr>
<td>0.1</td>
<td>default</td>
<td>default</td>
<td>64.75</td>
<td>0.85</td>
<td>0.69</td>
</tr>
<tr>
<td>0.4</td>
<td>default</td>
<td>default</td>
<td>165.98</td>
<td>0.68</td>
<td>-1.06</td>
</tr>
<tr>
<td>0.3</td>
<td>default</td>
<td>default+0.9 x 10^-6</td>
<td>55.88</td>
<td>0.88</td>
<td>0.77</td>
</tr>
</tbody>
</table>
The model performance is evaluated in terms of simulated streamflow \( (m^3/s) \) with the observed streamflow values at Monte Molino. The WRF-Hydro evaluation is done both considering the results at the event scale and the seasonal scale of the calibration. The main flood event in the time series is evaluated considering the timing and the maximum flow of the hydrograph. Moreover, a quantitative evaluation based on RMSE, correlation coefficient (RHO) and Nash-Sutcliffe efficiency coefficient values is calculated over all the year (Table 4.3). Using these strategies, the quantitative analysis considers both the main flow peak at the event scale and the main target of reproducing correctly the seasonal water cycle all over the year.

According to Tab. 4.3, the highest values of Nash Sutcliffe value has been reached by the model run with REFKDT=0.4. Also the other values of RHO and RMSE have very positive evaluation within this run. The REFKDT=0.4 calibration run reproduce very well the timing of the hydrograph over the whole year and the main peak of the flooding event that occurred in November (1346.6 m\(^3\)/s maximum flow peak from the WRF-Hydro simulation Vs. 1337.9 m\(^3\)/s observed). Nevertheless the streamflow simulated values seem as to drain too fast in the descending branch of the hydrograph for the highest flow peaks. In addition to that the model overestimates low peak flows occurring in the driest season (from January to August) (Fig. 4.14).

![Figure 4.14: Comparison between the calibration run with REFKDT=0.4 (black line) and the streamflow observations (red line) at the Monte Molino river section.](image)

Since the next comparison between WRF and WRF/WRF-Hydro simulation will be focused on the evaluation over all the year 2012 with a specific focus on summer convective season and not only on a specific flood event, it necessary to investigate further and to
try to adjust the model calibration to fix these specific problems.

The RETDEPRTFAC scaling factor for surface retention depth has been varied from the initial value of 0 to 500 (Fig. 4.15 and Tab. 4.3). Even if the range of variation is way wider than the one recommended in Yucel et al. (2015), (between 0 and 10), the model seems unexpectedly to have little sensitivity to this parameter. The hydrograph results to be affected more in terms of timing of the hydrograph, instead that in terms of volume as stated by Yucel et al. (2015). This result is quite unexpected and can be explained by the fact that, in a distributed model, a single parameter can affect different aspects of the physics of the model and produce different results depending on the soil type, vegetation and the local characteristics of the area. In this basin, the variation of the scaling factor for the surface retention depth affects the streamflow, creating a delay of the flux. Probably the increased surface retention depth can slow down the velocity of formation and contribution of the runoff to the channel network. Since this hypothesis has not been explored yet, it needs further future investigations.

Since there is already a good timing agreement between the observed stream flow and model hydrograph, it is considered not necessary to be varied in the calibration procedure. At the same way, it is not necessary to vary the roughness coefficients (MANN and OVROUGHRT) as described in Yucel et al. (2015) and these coefficients are left with their model default values.

Figure 4.15: Comparison between the calibration run with REFKDT=0.4 and RETDEPFRAC=0 (blu line) and the other run with REFKDT=0.4 and RETDEPFRAC=500 (black line) at the Monte Molino river section.

Since the model runs present an overestimation of the low discharge peaks during dry
season, the sensitivity to the SATKDT parameter is investigated in order to identify if the overestimation is depending on a possible model oversaturation in the dry season. SATKDT parameter represents the saturation soil conductivity and regulates the contribution of the ground water to the total. Even if this parameter can be varied both in the horizontal and in the vertical conductivity; in this study it is assumed to vary isotropically. In order to introduce anisotropy in the subsurface flow calculation, the users can vary the parameter LATKSATFAC, that tends to have a significant influence only in very wet regimes. Since the Tiber river it’s not considered dominated by wet regimes, SATKDT is varied only, using the same saturation both in vertical and horizontal direction. SATKDT is calibrated for the two predominant soil type inside the Tiber basin (category 6 loam and category 9 clay loam). The default value of SATKDT for clay is $2.45 \times 10^{-6}$ and for clay-loam is $3.38 \times 10^{-6}$. This parameter is varied in the range of two order of magnitude difference for a total of 11 runs. An increasing value of the the saturation soil conductivity has the effect of lowering the peaks and increasing the baseflow. In the experiment the strategy to lower the REFKDT (from 0.4 to 0.3) and increasing the SATKDT is adopted. In this case an overall slight improvement of the statistics is shown for SATKDT = default + $0.9 \times 10^{-6}$, even if the draining from the peak is still too fast and the overestimation of the low peak is still present (Fig. 4.16). The model predicts very well the event of November, but at the same time provides some lower floods that are not observed, such as the ones simulated from September until to October. In these cases, the hydrograph seems to respond too much to low peak rainfall, giving the possibility to a basin over-saturation in that season. In addition, the recession limbs of floods are always more steep than the observed ones. This is expected because the model is able to close the water balance, for which if fictitious floods are generated in summer period (runoff greater than the actual one), lower runoff is expected during the recession. The resulting statistics for this run are shown in Tab. 4.3.

Finally, a more extended parameter sensitivity analysis of the model is performed for the GWEXP, SLOPE, BB parameters that regulates the spilling of the ground water into the surface runoff, the drainage out of the last soil layer, the soil retention curve parameter, respectively, for a total of additional 14 runs. In these cases the results are not shown, since the model results not to be sensitive for this specific case to the above mentioned parameters, even if they are explored on a wide range of variability.

From the flow duration curves it is evident that, even if my best calibration run exhibit an high value of correlation and Nash Sutcliffe index for all over the year, the low flow values significantly detach from the observed values (Fig. 4.17).