

## Increasing the accuracy of the WRF-ARW numerical weather prediction model using Corine Land Cover and JAXA data

Tomáš FEDOR, Jaroslav HOFIERKA

**Abstract:** *The geographic data play an important role in atmospheric and weather forecasting. They influence input physical parameters defining the energy balance and heat and moisture fluxes in the planetary boundary layer. The goal of this paper is to increase the accuracy of the widely used Weather Research and Forecasting (WRF) numerical prediction model using a more detailed, 100 m resolution Corine Land Cover 2018 (CLC) and Japanese Space Agency (JAXA) elevation data in comparison to global data from the United States Geological Survey (USGS) containing the MODIS LULC land cover and elevation data. The WRF-ARW variant of the model is applied to a study area of 50x50 km in the Košice region at a microscale level with a spatial resolution of 250 m for June 24, 2020. The GIS4WRF module in QGIS and other GIS tools are used to perform modelling and input and output data preparation and result analysis. The results for both datasets are compared to data measured in local weather stations and statistically evaluated. The use of higher-resolution geographical data improves the overall accuracy of the model for this microscale level.*

**Keywords:** *Weather Research and Forecasting, Corine Land Cover, Geographic Information System*

### Introduction

The numerical weather prediction (NWP) models are frequently used as important tools for the prediction and research of the state of the atmosphere. The weather forecasting is done by carefully selected models operated by meteorological services around the world. The geographic data representing land surface properties such as elevation and land cover are an integral part of these models and have the highest importance for maintaining the prediction accuracy across different landscapes. The complexity of models and the amount of data require a vast computer resources. With an increasingly available computational power, new possibilities for simulating atmospheric conditions at the microscale level are gradually opening up which, at the same time, require new and more detailed input geographic data. The elevation data are necessary mainly for a correct modelling of the dynamics of the air mass flow while the land cover data are needed for thermodynamic processes taking place in the planetary boundary layer. Several key physical input parameters such as surface roughness, heat capacity, available soil moisture, albedo, or thermal emissivity of the surface can be derived from the land cover data (Jiménez-Estève et al. 2017).

Over the last decade, several studies have shown the positive impact of using newer and more detailed geographic data for high-resolution modelling of atmospheric conditions using NWP (e.g. Beezley et al. 2011, Jiménez-Estève et al. 2017, Vladimirov et al. 2018, Siewert and Kroszczyński 2020). The accuracy and effectiveness of using new data, however, varied considerably because of the diverse character of the monitored sites. For example, in a study focused on modelling of conditions in the rugged terrain of the Pyrenees (Jiménez-Estève et al. 2017), a better topographic data generally proved to be a decisive factor in improving the accuracy of the model output while newer and more detailed land cover data showed a slight

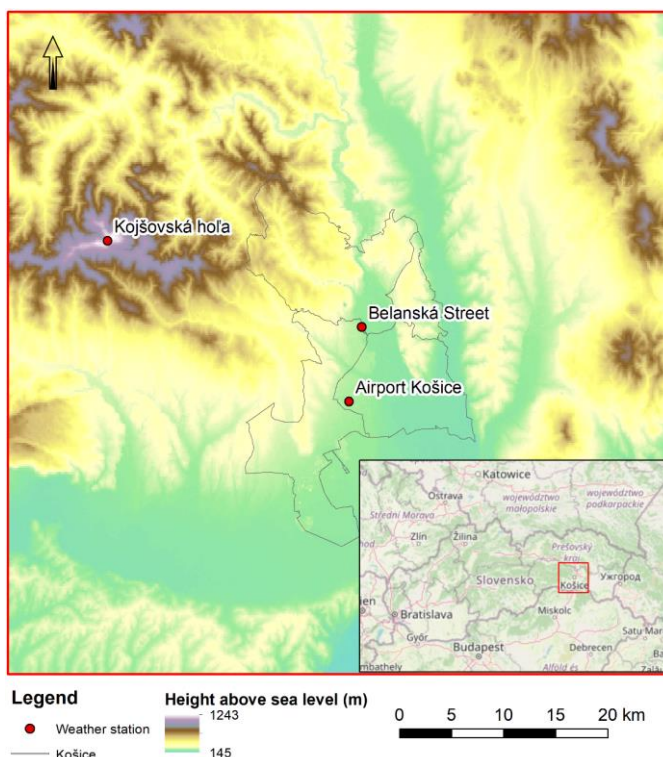
deterioration in accuracy of humidity prediction in comparison to older data. Studies examining the use of newer data for more detailed modelling of conditions for urban areas have generally shown an improvement in accuracy for all monitored parameters. However, the monitored areas mostly included a less rugged terrain (Case et al. 2008, Vučković et al. 2018, Siewert and Kroszczyński 2020) and the only study well documenting the impact of elevation data and at the same time various types of land cover with a significant share of urban built-up areas in temperate climate zone is the study by Vladimirov et al. (2018) for the Sofia region in Bulgaria. The Košice region examined in this paper also exhibits rugged terrain with rapidly changing types of landscapes. This makes this region a very convenient study area for testing the impact of geographic data at a different spatial resolution. Also, this study is the first WRF microscale modelling and evaluation in the Košice region. The aim of this paper is to evaluate the impacts of the use of more detailed geographic data on the accuracy of modelling meteorological variables in the study area using the WRF-ARW model and tools implemented in geographic information system (GIS).

### **Study area and atmospheric conditions**

The study area covers 50x50 km of diverse landscapes in the Košice region (Fig. 1). The eastern part of the study area includes the Slanské vrchy mountains, which act as a natural air flow barrier in the lower levels of the troposphere and makes them an important topographic element influencing microclimatic conditions. Slanské vrchy are mostly covered with dense forest vegetation. The Slovenské Rudohorie mountains are in the western and north-western part of the study area. The highest peak in the Volovské vrchy mountains is Kojšovská hoľa (1,225 m above sea level), which is also an important site for meteorological observations as there is a meteorological radar and an automatic weather station of the Slovak Hydrometeorological Institute (SHMÚ 2020) located. The data obtained from this station were used in this paper to evaluate the accuracy of the model in mountainous areas. The Hornád river flows through the central part of the study area. It is mostly an open flat area characterised by agricultural land with fragmented settlements. The city of Košice with large urban landscapes is in the central part of the region. Significant built-up, airport and industrial areas are also located in the south of the city. The Košice city and the airport are also locations with available data from ground meteorological observations, which were used similarly as in the case of data from Kojšovská hoľa as reference values in the evaluation of the model accuracy. The data from the airport are part of the official SHMÚ measurements and the data from Košice represents measurements from a personal semi-professional weather station Davis Vantage Pro 2, located near the hospital on the Belanská Street. This weather station is also part of the MADIS meteorological observation database system operated by the National Oceanic and Atmospheric Administration (NOAA 2020). Because of the open character of the location the data measured at the airport correspond to rural area (Vučković et al. 2018) and the data from the Belanská Street correspond to the urban built-up area (Fig. 1).

The WRF-ARW modelling and evaluation were done for June 24, 2020, due to the availability of data from ground meteorological observations and a weather situation suitable to compare the modelled conditions with observations. The day before, June 23, 2020, the region was behind a cold front, in a relatively strong northern flow. Such dynamic conditions were suitable for running the simulation since the calculation must be started a few hours in advance of the specified time interval, so that the air can “move” from the initial conditions. On the next day June 24, 2020, the dry and cold postfrontal air mass gradually began to heat up and transform during the day. At the same time, a strong upper-level low advanced towards the territory of Slovakia from the north which pulled drier air above our region at higher levels on its eastern side. This stable layer located during the day at a level of approximately 600 hPa was blocking the development of deep convection, which could complicate the evaluation of model accuracy due to more difficult localisation of convective cells and their significant

and localised impact on meteorological variables. In the surface layer, however, there was still enough moisture after the frontal precipitations from the previous day. During the day, the moisture gradually began to evaporate back into the atmosphere and the humidity increased. The increase of humidity was also supported by the advection of moist air in the lower levels from the north. According to observations, dew points reached around 12 °C in the morning and gradually rose to 15 °C during the day. At the same time, with the diurnal heating, the lower levels gradually became labilised and at noon there was a shallow convection ongoing, delimited from the top by a stable layer. The convection was occasionally producing weak precipitation, but mostly outside the model domain.



*Fig. 1. Košice basin with marked positions of weather stations*

## Methods and data

### The WRF-ARW model and the workflow

WRF is an advanced numerical weather prediction model used for research purposes but also for operational weather forecasting. It is an open-source software developed by the National Center for Atmospheric Research (NCAR). In this paper, the WRF-ARW (Advanced Research WRF) variant of the WRF model was used (version 4.0). This model has been developed by the Mesoscale and Microscale Meteorology Laboratory at the National Center for Atmospheric Research in Boulder, USA. WRF-ARW is also implemented in the QGIS software as the GIS4WRF module (Meyer and Riechter 2018), which was used in this study as a platform for running the model. This module contains both the data pre-processing system (WPS) as well as the numerical weather prediction model for running simulations (WRF).

The simulation of the model works on the principle of applying complex physical calculations following each other in the form of short iterations, simulating the dynamics of air mass.

The longest possible time between iterations is determined in seconds as the product of the size of the grid in the horizontal direction within the domain in kilometres and the value 6. To achieve a higher calculation speed, the largest possible value was used – 6 seconds for the parent domain with a horizontal resolution of 1,000 m and 1.5 seconds for the main domain with a horizontal resolution of 250 m (Tab. 1). To speed up the calculation, the number of vertical levels was set to 34 as in the case of the global model, which is still sufficient for testing new geographical data, though it is usually suggested to use much more vertical levels for such high-resolution simulations. The physical parameterisation of the model was set for both domains using the basic pre-set physical suite "CONUS". This set contains combinations of physical parameterisation options that have been tested and in general achieve the best results (NCAR 2021). Another reason to use the CONUS physical suite is that this set of model physics (Tab. 1) seems to be widely used by the WRF community, which makes it a very suitable option for evaluating the effects of new geographical data. However, it still should be noted that a properly set parameterisation scheme may provide better results, especially in case of microscale simulations.

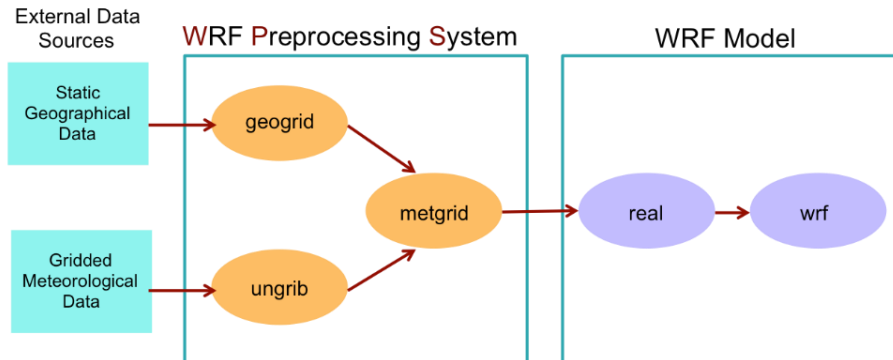
**Tab. 1. The WRF model configuration**

Configuration	Parent domain	Child domain
$\Delta x, \Delta y$	1,000 m	250 m
Time step	6 seconds	1.5 seconds
East-West grid points	200	300
South-North grid points	200	300
Vertical levels	34	34
Microphysics	Thompson graupel scheme	Thompson graupel scheme
Cumulus physics	Tiedtke scheme	Tiedtke scheme
Longwave radiation	RRTMG scheme	RRTMG scheme
Shortwave radiation	RRTMG scheme	RRTMG scheme
Boundary layer	Mellor-Yamada-Janjic (Eta) TKE scheme	Mellor-Yamada-Janjic (Eta) TKE scheme
Surface layer	Monin-Obukhov (Janjic Eta) scheme	Monin-Obukhov (Janjic Eta) scheme
Land-surface	unified Noah land-surface model	unified Noah land-surface model
Surface urban physics	OFF	OFF
Radiation physics calls	30 minutes	30 minutes
Boundary-layer physics calls	- call every time step	- call every time step
Cumulus physics calls	5 minutes	5 minutes

WRF as a regional model requires input meteorological data from a global model. In this paper, input global data from the National Center for Environmental Prediction were used, containing global analyses of atmospheric conditions calculated by the GFS model based on data from the Global Data Assimilation System (GDAS) based on meteorological measurements from around the world. The spatial resolution of the global dataset in a horizontal direction is  $0.25^\circ$ , which roughly corresponds to 25 km, and the vertical resolution is represented by 34 levels (NCEP 2020). Data are generated and archived in 6-hour intervals and can be obtained indirectly on request from the NCAR website. However, data from other global models that the WRF model can incorporate into the calculation can also be used, such as ECMWF or ARPEGE (NCAR 2021).

Data preparation and calculation takes place in several steps (Fig. 2). The first step (geogrid) includes a definition of domains as well as resampling and converting the input geographic data to the multidimensional format. Simultaneously, the next step is to unpack

the input meteorological data (ungrib) from the global forecasting model, which are contained in grib files. In the next step (metgrid), these data are assimilated into one file, while the conditions from the global model are horizontally interpolated according to input geographical and meteorological data. In the next step (real), the conditions are interpolated vertically and the input data for the model run (wrf) are prepared. The data created in the real process are also referred to as initial and boundary conditions, as they serve first to initialise the model and later also as intermediate steps to correct the course of the simulation.



**Fig. 2.** The workflow of the WRF model; Source: NCAR (2021)

In our study area two computational regions – domains were defined. The first, parent domain, with a size of 200x200 km, was used to simulate conditions for a wider area and the second one, the main or child domain, was used with a size of 75x75 km (Fig. 3). The dimensions of the main domain were chosen to be slightly larger than the study area of interest to achieve a smoother simulation of conditions. The spatial resolution of grids was 1,000 m in the first domain and 250 m in the second domain.



**Fig. 3.** Defined domains and their sizes (parent domain – blue, child domain – red)

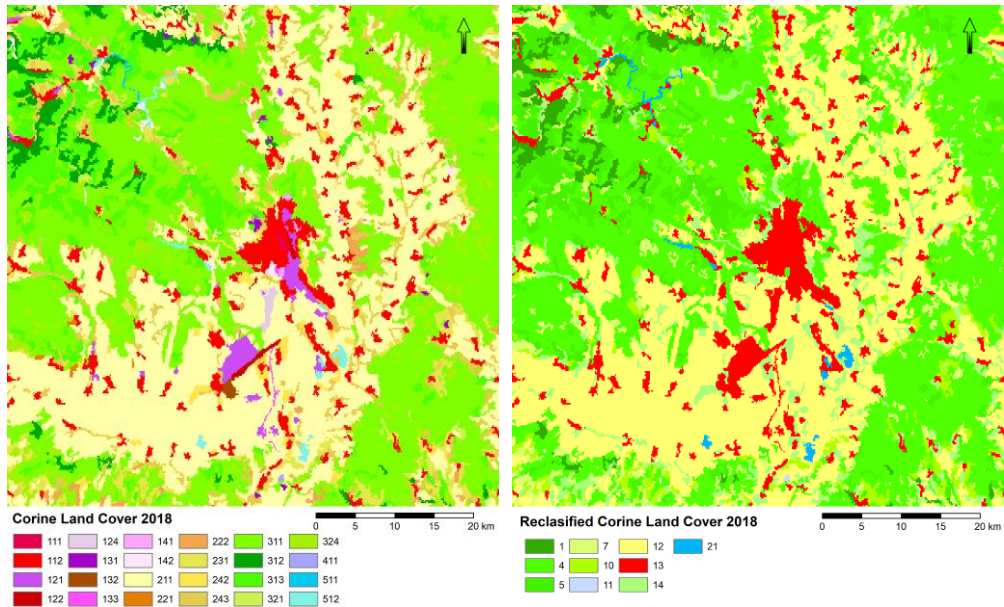
## The Corine Land Cover and JAXA data

The WRF model uses a land cover classification defined by the United States Geological Survey (USGS). These basic global data (MODIS LULC) have a horizontal spatial resolution of 1 km and the land cover data are categorised into 24 categories, of which 21 are included in the model. Although the data are regularly updated, their spatial resolution may not be entirely sufficient for more detailed simulations. For example, Siewert and Kroszczynski (2020) suggested that the Corine Land Cover dataset (CLC) is more suitable for regions with diverse land cover. These data are provided and regularly updated by the European Environment Agency (EEA 2021) for the territory of the European Union. The CLC data representing the year of 2018 were used at a spatial resolution of 100 m and consist of 44 classes of land cover. Since the WRF works with different classes of land cover, the integration of the data into the model required their reclassification by merging some classes based on the same physical properties. Data reclassification according to Siewert and Kroszczynski (2020) was used, in which the land cover classes are combined into the most similar classes used in the MODIS LULC classification (Tab. 2). The data were reclassified in ArcGIS, which appears to be the best choice for data pre-processing (Vladimirov et al. 2018). Most classes were merged from the category of anthropogenic surfaces since MODIS LULC classify them into one category as urban and built-up areas.

**Tab. 2.** CLC classes of land cover and their MODIS LULC reclassification

Number of CLC Class	Description of CLC Class	Number of MODIS LULC Class	Name of MODIS LULC Class
111–140, 142	Urban/Artificial Surfaces	13	Urban and Built-Up
141	Green Urban Areas	4	Deciduous Broadleaf Forest
211	Non-irrigated Arable Land	12	Croplands
221	Vineyards	14	Cropland/Natural Vegetation Mosaic
222	Fruit Trees or Berry Plantations	14	Cropland/Natural Vegetation Mosaic
231	Pastures	10	Grassland
242	Complex Cultivation Patterns	14	Cropland/Natural Vegetation Mosaic
243	Mixed Agriculture and Natural Vegetation	14	Cropland/Natural Vegetation Mosaic
311	Broad-Leaved Forest	4	Deciduous Broadleaf Forest
312	Coniferous Forest	1	Evergreen Needleleaf Forest
313	Mixed Forest	5	Mixed Forests
321	Natural Grasslands	10	Grassland
324	Transitional Woodland/Shrub	7	Open Shrublands
411	Inland Marshes	11	Permanent Wetlands
511	Rivers	21	Lakes
512	Inland Waters	21	Lakes

*Source: Siewert and Kroszczynski (2020)*



**Fig. 4.** CLC 2018 and MODIS LULC reclassification

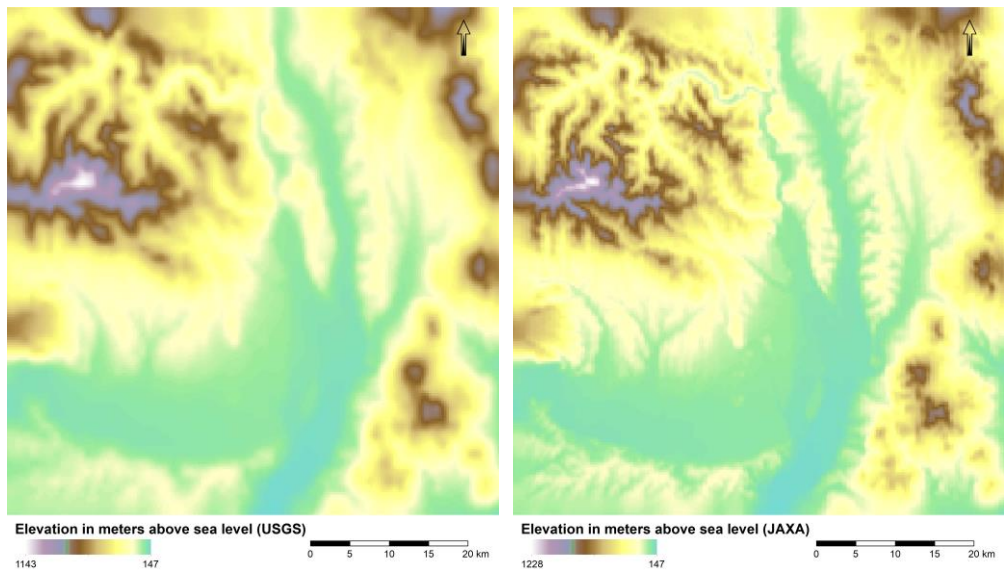
As a replacement for the original 1 km digital elevation model from the USGS, data from the Japanese Space Agency (JAXA 2019) with a spatial resolution of 30 m were selected. However, the freely available data from the Shuttle Radar Topography Mission (SRTM), at 30 m and 90 m spatial resolutions are also a very suitable and often preferred alternative (e.g. Jiménez-Estève et al. 2017, Bhavana et al. 2018, Vladimirov et al. 2018, Siewert and Kroszczynski 2020). The new input elevation data did not require any significant processing or modifications. We only resampled the data using bicubic convolution (Lillesand et al. 2004) to a lower resolution of 100 m from the original 30 m to reduce the volume of data and make the processing easier.

The pre-processed rasters were subsequently converted to a binary format used by WRF. The binary data consist of an index text file containing metadata about the dataset and the binary files (so-called tiles) dividing the data into regular grids with a predefined size with overlapping boundaries (Beezley et al. 2011). The advantage of this format is an efficient processing of large data files by splitting them into smaller chunks. When reading a smaller area within dataset, only the files required for the defined area are read instead of a whole dataset. Before the conversion, the input rasters were projected into the world geodetic system WGS84, since other coordination systems may not work properly. The data georeferenced to the WGS84 ellipsoid were transformed to a spherical projection, which was then transformed to a conical Lambert projection used by WRF. The geogrid process generates new, resampled input geographic data from binaries with a resolution corresponding to the defined grid resolution of domains. The nearest neighbour algorithm was used to resample the data in WPS, as it is fast to process and suitable especially for categorical data (Lillesand et al. 2004).

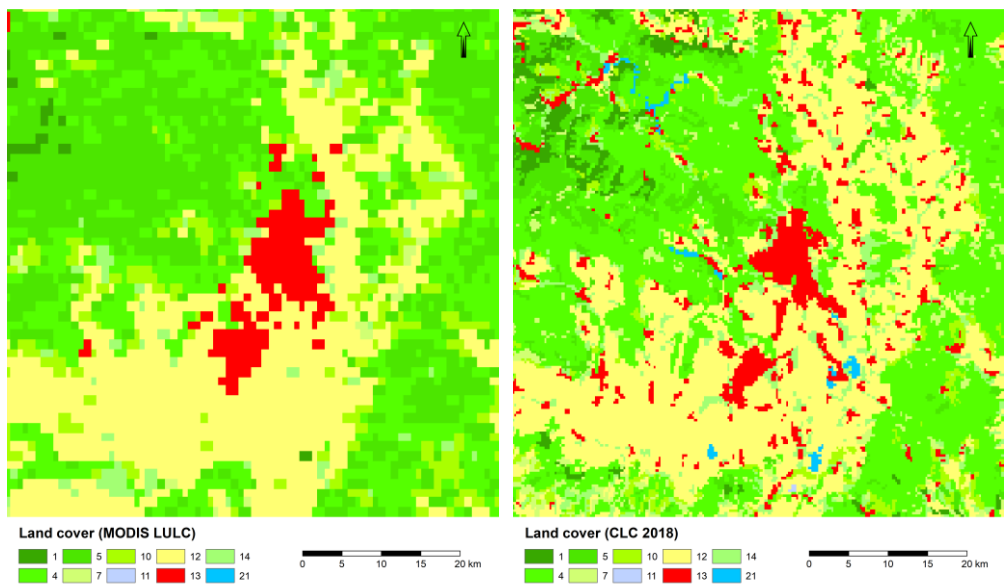
Fig. 5 and Fig. 6 show that JAXA and CLC data better represent the landscape in the study area, although it can be said that the WPS managed to resample the USGS elevation to a better resolution relatively accurately (Fig. 5) and the negative effect of a lower resolution digital elevation model on airflow would not be significant. However, the differences in the highest elevations are quite large, which may have a significant impact on the correct simulation



in the highest mountainous areas. In the case of land cover data, the situation is different. The MODIS LULC data resolution is very rough and many landscape features are missing such as water bodies and small built-up areas (Fig. 6).



**Fig. 5.** The USGS digital elevation model resampled to 250 m using WPS from the original spatial resolution of 1,000 m and the JAXA digital elevation model resampled to 250 m from the original spatial resolution of 100 m



**Fig. 6.** The MODIS LULC land cover resampled to 250 m using WPS from the original spatial resolution of 1,000 m and the CLC 2018 land cover resampled to 250 m using WPS from the original spatial resolution of 100 m



## Methodology for evaluating the accuracy of the model

To match all measurements of weather stations, the measurement values were compared with a model output in hourly steps. Three basic meteorological variables were compared – temperature and relative humidity at 2 m above ground and wind speed at 10 m above ground. The air temperature represented in Kelvins was converted to degrees Celsius and wind speed was calculated from U and V wind components.

Air humidity is defined in model output as a specific humidity, representing the weight ratio of water vapor in kilograms per kilogram of dry air.

$$es = 6.112 * e^{\frac{17.67 * T}{T + 243.5}} \quad (1)$$

The calculation consists of two parts. First of all, the determination of the partial pressure of water vapor for saturated air ( $es$ ). Where  $T$  is the air temperature and  $e$  is the Euler number (exponent). The second part of the calculation consists of the subsequent calculation of the specific humidity of the saturated air.

$$qs = \frac{0.622 * es}{\frac{p_{sfc}}{100} - (0.378 * es)}, \quad (2)$$

where  $qs$  is the specific humidity of the saturated air,  $es$  is the partial water vapor pressure for the saturated air,  $p_{sfc}$  is the air pressure at the surface in Pa.

The relative humidity ( $RH$ ) is then calculated simply by stating the ratio of the specific humidity calculated by the model and the specific humidity of the saturated air ( $qs$ ):

$$RH = \frac{q}{qs} * 100 \quad (3)$$

In the case of the Belanská Street station in Košice, it was necessary to make another adjustment of the measured quantities. The wind speed was measured at a different than standard height of 10 m. The height of the anemometer at the Belanská Street station in Košice with respect to the surroundings was estimated to be approximately 2 meters above ground and data were converted according to the formula:

$$\frac{v}{v_n} = \left(\frac{H}{H_n}\right)^\alpha, \quad (4)$$

where  $v$  represents the measured wind speed at height  $H$ ,  $v_n$  the wind speed at the height at which the observations take place  $H_n = 10$  m,  $\alpha$  represents the roughness coefficient with the value determined as  $\alpha = 0.25$  for the relatively open area without significant obstacles (Vučković et al. 2018).

Quantitative statistical methods were used to evaluate the overall accuracy of the simulation over the entire time series, determining the error of the model from the actual measured values. Due to the comparison of several relatively different quantities, it was more appropriate to include several statistical indicators (Botchkarev 2019), such as bias, mean error (ME), mean absolute error (MAE) and root mean squared error (RMSE), creating a detailed summary of the model accuracy:

$$bias = \frac{\sum_{i=1}^N f_i}{\sum_{i=1}^N o_i} \quad (5)$$

$$ME = \frac{\sum_{i=1}^N (f_i - o_i)}{N} \quad (6)$$

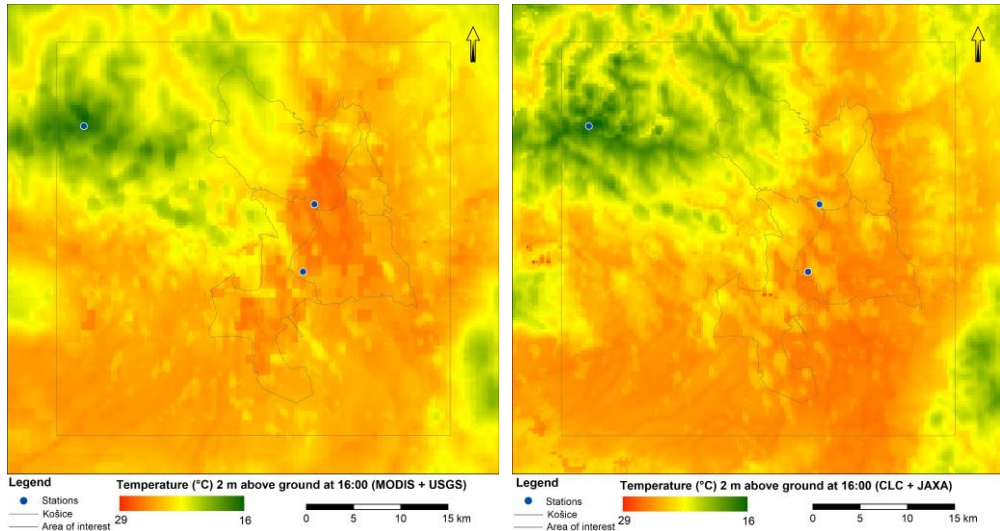
$$MAE = \frac{\sum_{i=1}^N |f_i - o_i|}{N} \quad (7)$$

$$RMSE = \sqrt{\frac{\sum_{i=1}^N (f_i - o_i)^2}{N}} \quad (8)$$

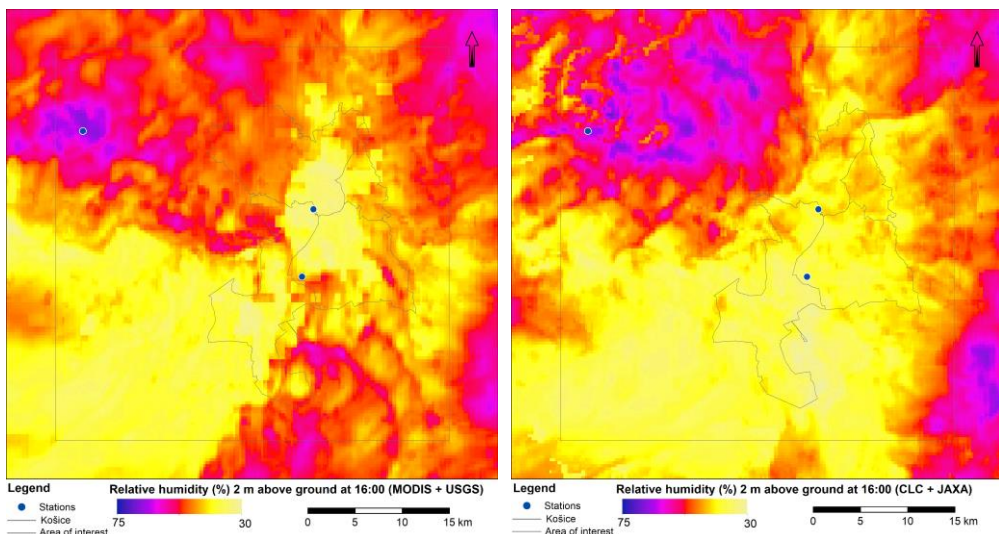
Where  $f$  represents the predicted value,  $o$  the value measured and  $N$  corresponds to the number of elements, i.e., the individual measurements in hourly steps over the entire time series.

## Results and discussion

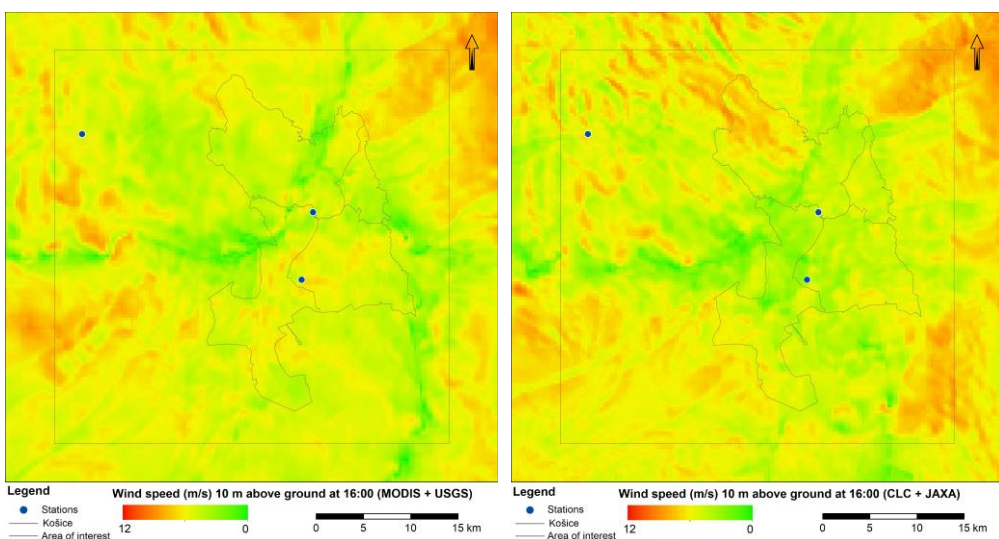
The results of the WRF-ARW model for standard, low-resolution data and new, high-resolution data are presented in Fig. 7–10 in the form of selected meteorological variables (temperature, humidity and wind speed). More detailed data better reflect the topography and transitions between the various types of land cover, which is mostly reflected in temperature and humidity (Fig. 7 and Fig. 8). With both variables, we can see that in the case of less detailed original data, the transitions between the built-up areas and other surrounding types of land cover are rougher and the model output shows a slightly scattered mosaic of the city, while in the case of using new, higher-resolution data the transitions between land cover types are smoother and temperature field is more stable and continuous. Such localised temperature anomalies in the case of original land cover data are caused by a combination of contrasting physical characteristics between urban and built-up area land cover category and surrounding greenery or cropland and their very rough spatial resolution. The urban and built-up area is prone to diurnal radiation heating and tends to overheat more, creating the rough temperature mosaic over this land cover category. For mountainous areas, the use of a more accurate digital elevation model has proven to be beneficial, mainly due to large differences in elevation. Such large differences in elevation were also projected in differences between modelled and observed temperature and humidity.



**Fig. 7. Predicted temperature at 2 m for MODIS+USGS and CLC+JAXA data at 16:00 CEST**



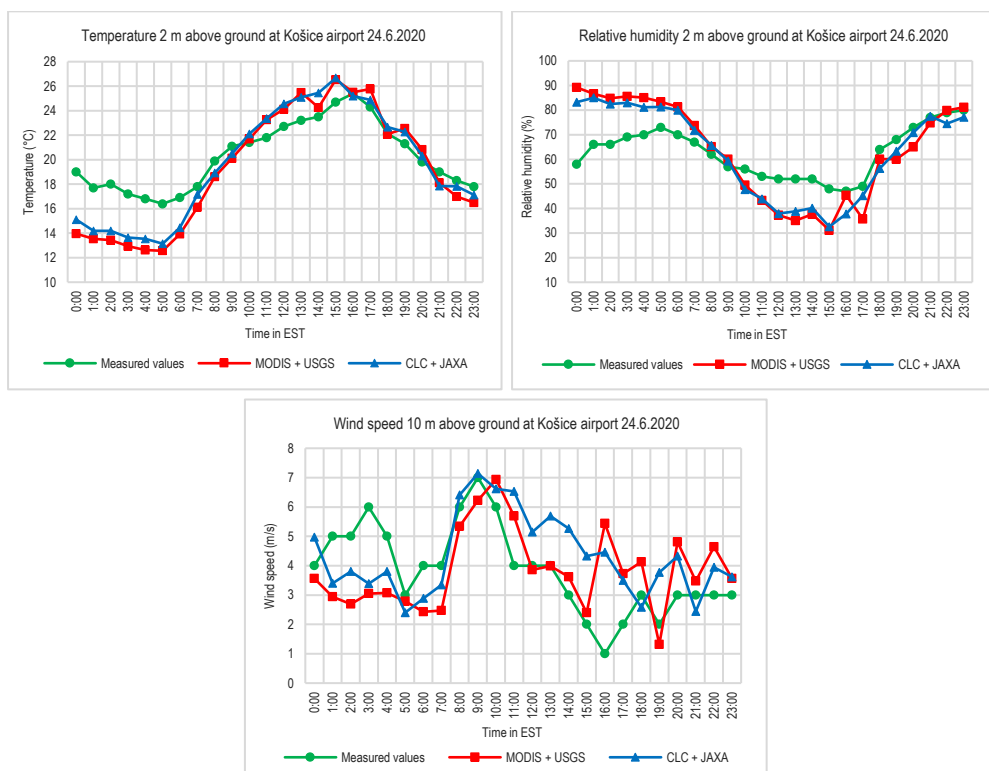
**Fig. 8.** Predicted humidity at 2 m for MODIS+USGS and CLC+JAXA data at 16:00 CEST



**Fig. 9.** Predicted windspeed at 10 m for MODIS+USGS and CLC+JAXA data at 16:00 CEST

The statistical indicators presented in Tab. 3 show that the model generally underestimated the temperature in both cases, which subsequently influenced the relative humidity. The wind speed was generally modelled quite accurately. In case of airport in Košice, which represents an open space area, the situation modelled with the old data was practically identical with measurements, but the newer data recorded a paradoxically slightly higher error rate. In the case of the Belanská Street station, the situation was different, and the new data showed improvement in wind speed modelling. However, it should be noted that the measurement of wind speed in this location was a bit inconsistent due to the different location of the anemometer and the values had to be recalculated to a height of 10 m. At the same time, wind data from the station on Kojšovská hoľa were missing for the given period, therefore the airport with

the most representative measurements was chosen for the overall comparison of the time series (Fig. 10). From the course of meteorological variables during the day, the model significantly overestimated the radiation cooling in the morning and expected a significantly lower temperature, up to 4 °C, which was also reflected in the high deviation of relative humidity. The same situation occurred in the morning at the Belanská Street station. However, Kojšovská hoľa positioned at a high elevation was not affected and the morning temperatures were practically identical to both model outputs – the deviation of the morning temperatures was mostly only up to 0.5 °C and the humidity, as the air was saturated in the morning, was identical (100%). During the day in all stations, the model expected a slightly higher air temperature, by about 1–2 °C, which was caused by the numerous cumilform clouds formed shortly before the noon, which slightly slowed down the radiation heating. The clouds were also considered in the model scenario, but in less extent than expected. Later in the afternoon, with decreasing convective activity, the convective clouds formation ceased, and temperatures rose higher in direct sunlight – after 15:00, the air temperature values were the same as was the model output. In both cases, the wind was simulated relatively accurately in all monitored sites. The highest wind speed was recorded identically in both scenarios in the morning between 9:00–10:00, while the wind gradually weakened during the day, which the model copied relatively accurately. However, the scenario with new data expected a slightly stronger wind in the second half of the day. In both cases, the air humidity was partly influenced by the humid air coming from the north in the second half of the day, but also by the ongoing convection, which increased the humidity locally in model situations, as seen, for example, south of Košice in the output using the original data.



**Fig. 10.** Time series of measured and predicted temperature, humidity and wind speed for the Airport Košice station

**Tab. 3.** Statistical comparison of the model accuracy in individual weather stations

Variable	Station	MODIS LULC + USGS				CLC 2018 + JAXA			
		bias	ME	MAE	RMSE	bias	ME	MAE	RMSE
Temperature in 2 m (°C)	Belanská S.	0.987	-0.268	1.934	2.399	0.970	-0.616	1.702	2.117
	Airport	0.950	-1.022	2.016	2.506	0.967	-0.660	1.708	2.091
	K. hoľa	1.029	0.363	1.076	1.526	1.011	0.139	0.977	1.409
Relative humidity in 2 m (%)	Belanská S.	0.893	-7.181	13.884	15.334	0.950	-3.351	9.327	11.578
	Airport	1.015	0.958	10.624	12.930	1.006	0.366	9.245	11.010
	K. hoľa	0.928	-6.827	6.827	12.569	0.934	-6.213	6.669	11.540
Wind speed in 10 m (m/s)	Belanská S.	1.599	1.825	1.953	2.243	1.069	0.315	1.706	1.974
	Airport	1.002	0.007	1.279	1.625	0.889	-0.482	1.319	1.551
	K. hoľa	-	-	-	-	-	-	-	-

Our presented results are similar to a study done by Siewert and Kroszczynski (2020) for Warsaw, where a higher simulation error was observed using standard, low-resolution geographic data (MODIS, GMTED2010) than new, higher-resolution data (CLC, SRTM). Vladimirov et al. (2018) in a study for the Sofia region in Bulgaria also used more detailed CLC data leading to a better representation of urban greenery and modelling of heat fluxes over the urban areas of Sofia. We observed the same effect for the Košice region where standard, low-resolution land cover data completely missed small built-up areas leading to a lower accuracy of simulations.

Vladimirov et al. (2018) and Jiménez-Esteve et al. (2017) report a better accuracy of simulations when using of a higher-resolution elevation data especially in mountainous weather stations. Our study confirms the necessity of more accurate, high-resolution elevation data especially for modelling atmospheric conditions in mountainous areas, such as Kojšovská hoľa, with lower errors (Tab. 3) in all evaluated meteorological variables.

## Conclusions

In this paper, we analysed the effects of more detailed geographic data on the WRF-ARW model at a microscale level. Higher-resolution geographic data representing elevation (JAXA) and land cover (CLC 2018) had a positive effect on the overall model accuracy. The use of a more detailed digital elevation model enabled more accurate modelling of the airflow in mountainous regions and in places with more fragmented topography. More accurate data on land cover have proven to be a key factor for increasing accuracy of the model, as they better represent the processes defining the interaction of the landscape with the atmosphere. The study showed that for the correct simulation of conditions at a microscale level it is necessary to use more detailed land cover data. The resampling of lower resolution land cover data during the WPS process to a higher resolution resulted in rougher and more extensive transitions between individual land cover classes strongly impacting the results of the model.

It should also be noted that a significant part of the accuracy is largely influenced by parameterisation of the model. Basic settings were used in this paper, but by optimising them, some processes could be modelled even with a higher accuracy. In such a case, for example, it would be appropriate to use other radiation schemes to simulate more accurately a morning temperature drop, in which the error rate of the simulation was the highest in both scenarios. Our future research focused on urban areas will require more precise parameterisation for large-eddy simulations together with new geographical data containing more detailed information about the diverse urban environment. Such microscale simulations would require a greater number of vertical levels and multiple domains to achieve better and more realistic results.

## References

- BEEZLEY, J. D., KOCHANSKI, A. D., MANDEL, J. 2011: Integrating high-resolution static data into WRF for real fire simulations. Paper 6.3. *Ninth Symposium on Fire and Forest Meteorology*. Palm Springs (American Meteorological Society). Retrieved from: <http://ams.confex.com/ams/9FIRE/webprogram/Paper192276.html>.
- BHAVANA, M., GUPTA, K., PAL, P. K., KUMAR, A. S., GUMMAPU, J. S. 2018: Evaluation of High Resolution Urban LULC for Seasonal Forecasts of Urban Climate using WRF Model. *The ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, IV-5, 303-310. DOI: <http://dx.doi.org/10.5194/isprs-annals-IV-5-303-2018>.
- BOLTON, D. 1980: The Computation of Equivalent Potential Temperature. *Monthly Weather Review*, 108, 1046-1053. DOI: [http://dx.doi.org/10.1175/1520-0493\(1980\)108%3C1046:TCOEPT%3E2.0.CO;2](http://dx.doi.org/10.1175/1520-0493(1980)108%3C1046:TCOEPT%3E2.0.CO;2).
- BOTCHKAREV, A. 2019: A new typology design of performance metrics to measure errors in machine learning regression algorithms. *Interdisciplinary Journal of Information, Knowledge, and Management*, 14, 45-76. DOI: <http://dx.doi.org/10.28945/4184>.
- CASE, J. L., CROSSON, W. L., KUMAR, S. V., LAPENTA, W. M., PETERS-LIDARD, CH. D. 2008: Impacts of High-Resolution Land Surface Initialization on Regional Sensible Weather Forecasts from the WRF Model. *Journal of Hydrometeorology*, 9(6), 1249-1265. DOI: <http://dx.doi.org/10.1175/2008JHM990.1>.
- EEA 2021: *Corine Land Cover 2018*. Copenhagen (EEA – European Environment Agency).
- JAXA 2019: *Digital elevation model from mission ALOS-2*. Tokyo (JAXA – Japanese Space Agency).
- JIMÉNEZ-ESTEVE, B., UDINA, M., SOLER, M. R., PEPIN, N., MIRÓ, J. R. 2017: Land use and topography influence in a complex terrain area: A high resolution mesoscale modelling study over the Eastern Pyrenees using the WRF model. *Atmospheric Research*, 202, 49-62. DOI: <http://dx.doi.org/10.1016/j.atmosres.2017.11.012>.
- LILLESAND, T. M., KIEFER, R. W., CHIPMAN, J. W. 2004: *Remote sensing and image interpretation (5<sup>th</sup> edition)*. New York (John Wiley & Sons Inc.).
- MEYER, D., RIECHTER, M. 2018: GIS4WRF module [cit. 2020-04-05]. Retrieved from: <https://gis4wrf.github.io/>.
- NCAR 2021: Weather Research and Forecasting – Advanced WRF. *National Center for Atmospheric Research* [cit. 2021-04-10]. Retrieved from: <https://www.mmm.ucar.edu/weather-research-and-forecasting-model>.
- NCEP 2020: NCEP GDAS/FNL 0.25 Degree Global Tropospheric Analyses and Forecast Grids. *Research Data Archive at the National Center for Atmospheric Research, Computational and Information Systems Laboratory* [cit. 2021-04-06]. DOI: <https://doi.org/10.5065/D65Q4T4Z>.
- NOAA 2020: Meteorological Assimilation Data Ingest System. *NOAA – National Oceanic and Atmospheric Administration*. Retrieved from: <https://madis.ncep.noaa.gov/>.
- SHMÚ 2020: *Surface synoptic observations (SYNOP)*. Bratislava (SHMÚ – Slovak Hydrometeorological Institute).
- SIEWERT, J., KROSZCZYNSKI, K. 2020: GIS Data as a Valuable Source of Information for Increasing Resolution of the WRF Model for Warsaw. *Remote Sensing*, 12(11), 1881. DOI: <https://doi.org/10.3390/rs12111881>.
- VLADIMIROV, E., DIMITROVA, R., DANCHOVSKI, V. 2018: Sensitivity of WRF Model Results to Topography and Land Cover: Study for the Sofia Region. *Annual of Sofia University „St. Kliment Ohridski“*, 111, 87-101.



- VUČKOVIĆ, M., HAMMERBERG, K., MAHDAVI, A. 2018: Urban weather modeling applications: A Vienna case study. *Building Simulations*, 13, 99-111. DOI: <https://doi.org/10.1007/s12273-019-0564-y>.
- WILLIAM, C. S. et al. 2019: A Description of the Advanced Research WRF Model Version 4. *NCAR Technical Notes* [cit. 2021-04-17]. Retrieved from: <https://opensky.ucar.edu/islandora/object/opensky:2898>.

***Acknowledgement: This research was funded by the Slovak Research and Development Agency (APVV) under the contracts No. APVV-18-0044 and by the Scientific Grant Agency of the Ministry of Education, Science, Research and Sport of the Slovak Republic and the Slovak Academy of Sciences (VEGA) under the contract No. 1/0300/19.***

---

***Authors' affiliations***

**Tomáš Fedor**

Institute of Geography  
Faculty of Science, Pavol Jozef Šafárik University in Košice  
Jesenná 5, 040 01 Košice  
Slovakia  
[tomas.fedor@student.upjs.sk](mailto:tomas.fedor@student.upjs.sk)

**Jaroslav Hofierka**

Institute of Geography  
Faculty of Science, Pavol Jozef Šafárik University in Košice  
Jesenná 5, 040 01 Košice  
Slovakia  
[jaroslav.hofierka@upjs.sk](mailto:jaroslav.hofierka@upjs.sk)