RAMSES PROJECT

Grant Agreement n° 308497

WP 4: Climate change scenarios for urban agglomerations

D4.2:

Agglomeration-scale urban climate and air quality projections

Reference code: RAMSES – D4.2

The work leading to these results has received funding from the European Community's Seventh Framework Programme under Grant Agreement No. 308497 (Project RAMSES).

Project Acronym: RAMSES

Project Title: Reconciling Adaptation, Mitigation and Sustainable Development for Cities

Contract Number: 308497

Title of report: D 4.2 Agglomeration-scale urban climate and air quality projections

Reference code: RAMSES – D 4.2

Short Description:

We describe the methods used to generate urban climate and air quality projections, and their results.

Authors and co-authors: Hans Hooyberghs, Dirk Lauwaet, Wouter Lefebvre, Bino Maiheu, Koen De Ridder (VITO); Iratxe González-Aparicio, Juan Angel Acero, Maddalen Mendizabal (TECNALIA)

Partners owning: VITO & TECNALIA

Contributions:

Made available to: Public

Versioning		
Version	Date	Name, organization
0.1	30/9/2014	Hans Hooyberghs (VITO)
0.2	21/11/2014	Koen De Ridder, VITO
1.0	27/02/2015	Koen De Ridder, VITO
1.1-1.2-1.3	09-10/03/2015	Koen De Ridder, VITO

Quality check

Internal Reviewers: **Alistair Ford (UNEW), Bin Zhou (PIK)**

Table of contents

List of Figures

List of Tables

List of Abbreviations

CMIP5	coupled model inter-comparison project phase 5
CORDEX	coordinated regional climate downscaling experiment
CPU	central processing unit (computer processor)
ECMWF	European centre for medium-range weather forecasting
GCM	global climate model
IPCC	intergovernmental panel on climate change (of the UN)
NDVI	normalized difference vegetation index
RCP	representative concentration pathway (IPCC climate scenario)
SW, LW	Short- and Longwave (radiation)
UHI	urban heat island
UrbClim	urban boundary layer climate model

1 Executive Summary

There is increasing concern regarding the impact of global climate change on cities. Together with drought and flooding, extreme heat stress is perceived as a problem that may increase considerably if no action is taken. Indeed, climate projections indicate that extreme heat waves, such as the one that occurred in 2003 in Europe, are expected to become more commonplace towards the end of the century.

Additionally, cities experience enhanced heat stress because of the urban heat island (UHI) phenomenon. Considering that cities are home to the majority of humans (in particular in Europe), and given also that cities are particularly vulnerable because of the concentration of infrastructure and economic activity, this is all the more relevant.

Yet, little or no information is available regarding future urban climate. Especially, climate projections at the scale of urban agglomerations are lacking, which is in part related to the computational constraints fine-scale climate models are facing. In order to remedy this, a new urban climate model (UrbClim) was developed.

During the first year of the RAMSES project, work was done to evaluate the UrbClim model for a number of target cities, such as Antwerp, London, and Bilbao. From this, we learned that the model achieves an accuracy comparable to that of existing traditional models, but at a speed that is more than a hundred times higher. As a result, UrbClim is one of the only urban climate models in the World capable of covering periods long enough (tens of years) to deduce relevant climate statistics.

During the second project year, this model speed was successfully exploited, coupling the UrbClim model to the Global Climate Models (GCMs) contained in the IPCC/CMIP5 archive, and conducting long simulations representing present (1986-2005) and future (2081-2100) climate conditions (and considering the RCP8.5 climate scenario). This was done for Antwerp, Bilbao, London, Rio de Janeiro, Hyderabad, New York, and Skopje; and also considering two additional cities (Berlin, and Almada in Portugal).

This was a major computational undertaking, considering that, for every one of the cities, we simulated two periods covering twenty years each. Moreover, all of these simulations were run 11 times, using forcings from 11 different GCMs in order to account for the uncertainty related to the climate projections. In summary, all simulations taken together required a staggering 45 CPU years (fortunately we had

access to several hundred CPU's simultaneously), and generated almost 7000 Gigabytes of simulation data.

Based on the results for the present period, it was found that, overall, and consistently within the city sample considered, **urban areas experience twice as many heat wave days than their rural surroundings**. This is problematic, as in most countries, so-called 'heat health action plans' are triggered using rural temperature forecasts only. Consequently, the (often vulnerable) urban populations undergo several heat wave days per year with no action taken.

Subsequently, when analysing the urban climate projections, it was found that, towards the end of the century, **the number of heat wave days is expected to increase by a factor of nearly ten**. This appears to be the case in both urban and rural areas but, given the higher number of urban heat wave days to start with, this means that many cities will be facing one month or more of heat wave conditions each year.

In another experiment, the UrbClim model was applied to 102 cities throughout Europe. Based on the results, it was shown that the average UHI intensity of these cities is well explained by population density and the mean regional wind speed. The main conclusion here is that **cities in Central and Eastern Europe are especially at risk for enhanced heat stress**. This is all the more relevant considering that these are also the areas for which climate models project a very large summer temperature increase.

Finally, we investigated the impact of climate change on atmospheric pollution. Focusing on the urban area of Bilbao, and on the harmful pollutant PM_{10} (particulate matter), we developed a statistical scheme linking PM_{10} concentrations to climatic parameters readily available in climate projections. Once established for present-day conditions, these relations were applied to output from an ensemble of regional climate simulation results.

From this, it was found that during the summer towards the end of the century **climate change is expected to increase in PM_{10} concentrations by 0.3 µg m^{-3}**, owing to increased temperature and humidity, and decreasing wind speed values. Conversely, the expected **change in winter is a decrease by a value of 1.1 µg m^{-3}**, mainly caused by enhanced precipitation. Note that for both winter and summer, the ensemble variance is larger than the change of the mean values.

2 Introduction

Global Climate Models (GCMs) predict an overall increase in air temperature, which for Europe in summer is expected to become very large (depending on the RCP scenario used), as shown in Figure 1. Notice in particular the extreme temperature increase over portions of Central Europe and the Balkans – as will become apparent this will have implications especially for the results obtained over Skopje, which is one of the RAMSES target cities.

Consistent with this increase, GCMs also predict a rise of the number, frequency, and intensity of heat waves (Meehl and Tibaldi, 2004; Diffenbaugh and Giorgi, 2012). Schär et al. (2004) estimate that, in the future, extremely hot conditions as occurring in Europe during the summer of 2003 could become more common in the future.

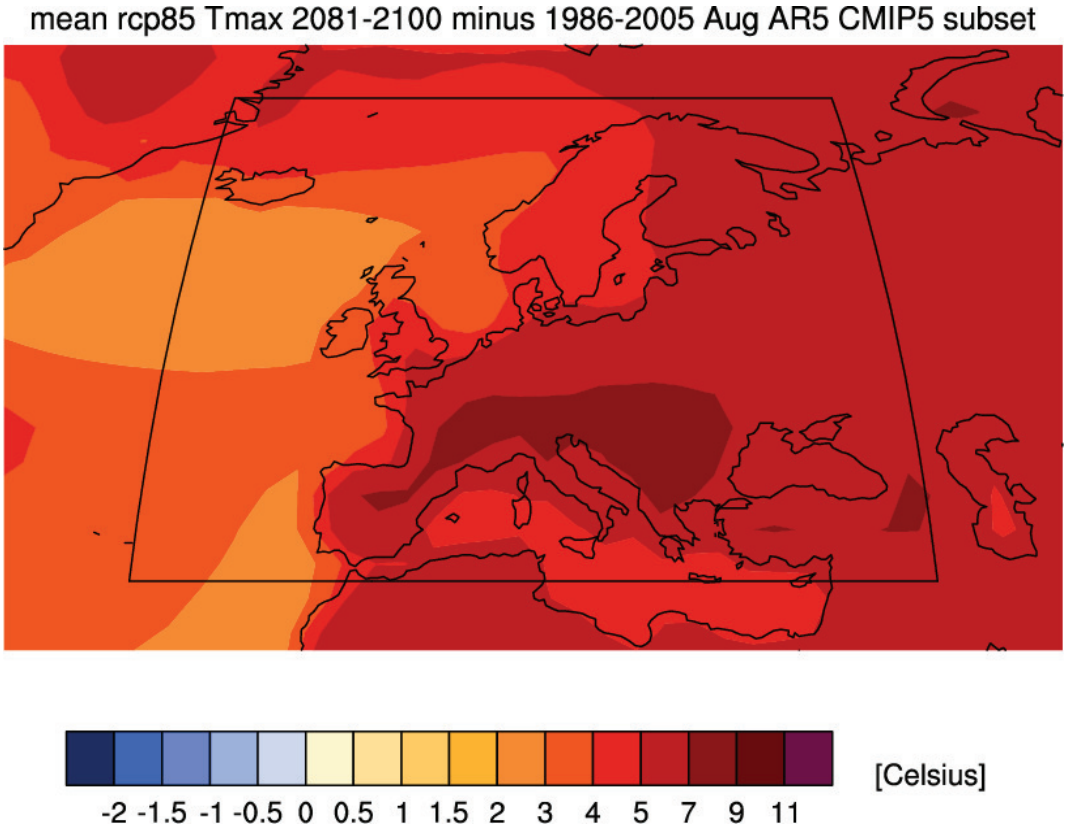


Figure 1. Projected mean increase of daily maximum temperature towards the end of the century compared to the 1986-2005 period, for the month of August, for IPCC scenario RCP8.5. (Source: KNMI Climate Atlas, http://climexp.knmi.nl/)

Cities experience an additional heat stress because of the urban heat island (UHI) effect. Moreover, the UHI intensity itself appears to increase during heat waves, as demonstrated by De Ridder et al. (2011) for Paris, and by Li and Bou-Zeid (2013)

for the Washington metropolitan area. This can be understood by the fact that the conditions giving rise to strong heat islands (clear skies, low wind speed) are generally the same as those that occur during heat waves. Hence, there is cause for concern regarding the urban effect in shaping future levels of heat stress, particularly as a large share of the European and global population lives in urban areas. Yet, while studies on global and regional (extreme) temperature changes abound, not much research has been devoted to the urban effect.

Within Task 4.2 of RAMSES (agglomeration-scale urban climate and air quality projections), we address this knowledge gap. Building on work done in the earlier stages of the RAMSES project, which focused on the capability of urban climate models to correctly reproduce observed meteorological quantities in and near cities, we have taken the research a step further by establishing urban climate projections. This was achieved by coupling the UrbClim urban climate model to output fields of GCMs, and by conducting ensemble simulations for present and future climate conditions, using the IPCC RCP8.5 scenario to do so. Since our focus is on human exposure to heat stress, we have analysed the simulation results in terms of the number of heat wave days. Moreover, we did several analyses to try and estimate the wind reduction factor over urban areas, trying to identify some 'universal' rules.

Apart from the impact of climate change on heat stress and human health (see, e.g., Patz et al., 2014, for a recent review), climate change is also expected to have an impact on air quality, through the effect of meteorological variables on atmospheric pollutant concentrations. To address this, we developed statistical relationships to explain concentrations of particulate matter by means of standard meteorological variables, including temperature, humidity, wind speed, and precipitation, among others. These relations were then applied to climate prognoses from an ensemble of regional climate models for Bilbao.

The remainder of this report is organized as follows: in Section 3, an overview is given of the urban climate projections, describing both the methodology as the main results. Related to this, in Section 4 we also briefly describe a simulation 'experiment' we conducted in the context of the RAMSES 'top-down' activities, considering more than hundred cities in Europe, and focusing on the explanation of the differences in UHI intensity observed in this sample. Section 5 addresses the wind climate in cities, and Section 6 then deals with the impact of climate change on air quality. We then close this report with the Conclusions (Section 7).

3 Urban climate projections

3.1 The UrbClim model

UrbClim is an urban climate model designed to simulate and study the urban heat island effect (UHI) and other urban climate variables (wind speed, humidity, …) at a spatial resolution of a few hundred metres (De Ridder et al., 2015). The model scales large-scale weather conditions down to agglomeration-scale and computes the impact of urban development on the most important weather parameters, such as temperature and humidity. UrbClim is composed of a land surface scheme describing the physics of energy and water exchange between the soil and the atmosphere in the city, coupled to a 3D atmospheric boundary layer module. The atmospheric conditions far away from the city centre are fixed by meteorological input data, while local terrain and surface data influence the heat fluxes and evaporation within the urban boundaries. The primary output consists of hourly air temperature and apparent air temperature maps with a spatial resolution of 250 m or 500 m. The simulated domain and the resolution are dependent on the city under investigation, but are always chosen such that the simulated domain comprises the city and its immediate surroundings.

The UrbClim model has been subjected to exhaustive validation. Model results have been compared with hourly temperature measurements for London, Bilbao, Athens, Almada, Toulouse, Berlin, Antwerp, Ghent, and Paris (De Ridder et al. 2014; De Ridder and Sarkar 2011; Keramitsoglou et al., 2012; Kourtidis et al., 2015). As an example, Figure 2 shows a comparison between model output and meteorological observations for London, for the summer (May-September) of 2012.

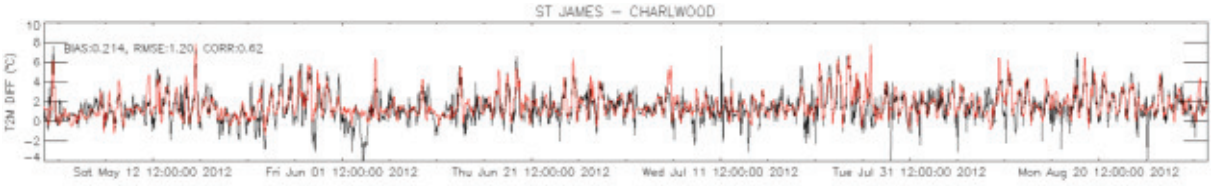


Figure 2. Urban-rural 2-m air temperature difference for London for the measurements (black) and the UrbClim model (red) for the period May-September 2012. Error statistics (model bias, root mean square error and correlation coefficient) are given in the top left of the plots.

The main advantage of the urban boundary layer climate model is that it is faster than high-resolution mesoscale climate models by at least two orders of magnitude. Because of that, the model is well suited for long time integrations, in

particular for applications in urban climate (adaptation) projections, and for a deployment involving a large number of cities. Note that a comprehensive description of the UrbClim model is provided in De Ridder et al. (2015).

3.2 UHI indicators: definition

Here, we define two indicators which satisfy the needs of the RAMSES project. The first indicator quantifies the difference in air temperature between the urban and the rural part (the definition of either will be given below) of the domains, while the second indicator is related to the number of heat wave days. In both cases, the UHI indicator depends only on the 2-m air temperature; effects of wind speed and (relative) humidity are not taken into account. Although physiological studies indicate that both these parameters are also important for the impact of heat stress on humans, the IPCC AR5 report (IPCC, 2013) questions their treatment in GCMs, especially when looking at local scales. Within the RAMSES consortium, an agreement was reached to use only indicators based on air temperatures.

In contrast with many other UHI studies, air temperatures are used instead of surface temperatures. While the latter are easily measured using remote sensing data, the impact of heat stress is more closely related to air temperature. Note that, while the surface UHI effect is largest during daytime, the air temperature UHI effect reaches a maximum in the late evening and early night. Therefore, we took care to include the (nocturnal) minimum temperatures in the indicators.

The first indicator quantifies the heat stress in the simulated domain using the temperature during warm summer nights. More specifically, we consider the temporal 95th percentile of the minimum temperatures for the summer period, i.e., the minimum temperature that is exceeded on average 7 to 8 times each summer (considering that summer extends over the period May-September). For cities in a hilly environment, the UHI effect is often blurred by topographical effects. Therefore, all air temperatures are henceforth corrected for terrain height by rescaling them to the average city height. In the remainder of the text, the indicator is referred to as the "temperature indicator" and we refer to the 'temperature UHI effect'. As an example, Figure 3 shows the result of such an indicator based on an UrbClim simulation for Antwerp, for the period 1986-2005.

The UHI effect is quantified as the difference between the urban and the rural value of the temperature indicator. Here, we define the rural value of a parameter as the

spatial 10th percentile of the concerned parameter in the domain under investigation. Analogously, the urban value is defined using the spatial 90th percentile. In order to only consider parameter values above land, we neglect the grid cells above water in calculating the percentiles.

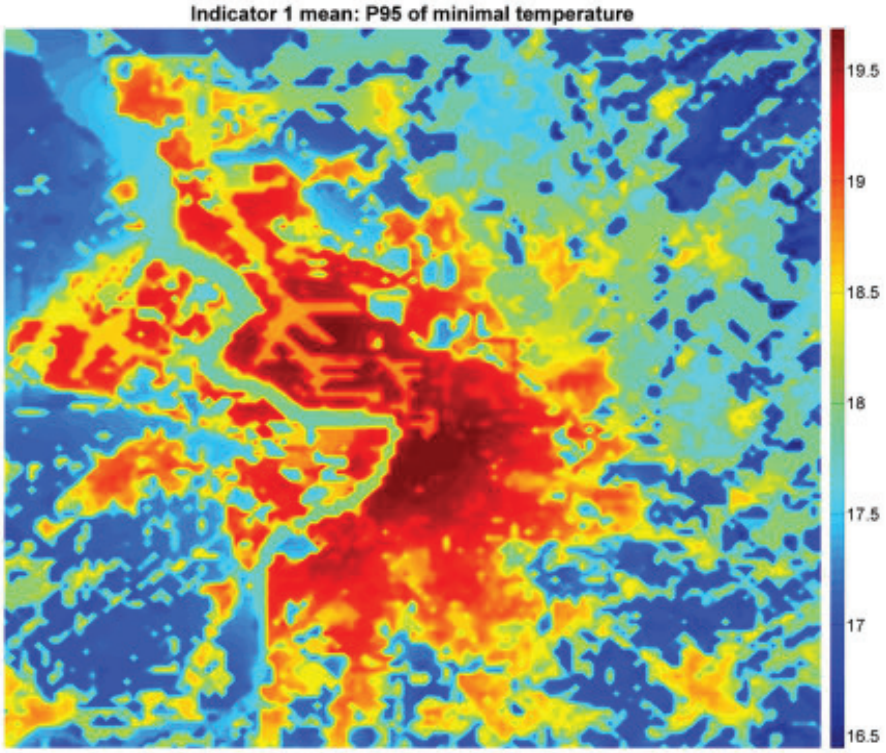


Figure 3. Temperature indicator for Antwerp for the period 1986 -2005.

The second indicator deals with the number of heat-wave days occurring in a given year. In the literature, a large number of definitions exists but, although the details of the different indicators are usually very distinct, all share some common properties. Since health effects become especially severe during prolonged periods of hot weather, most heat wave day definitions consider multiple consecutive days of extreme weather. Mostly, the definitions also deal with minimum and maximum temperatures, since lack of cooling during the night is an important factor in health effects related to heat stress. Finally, definitions generally make use of statistically determined thresholds. Within the RAMSES project, we have chosen to define a heat-wave day as a day during which both the three-day running average minimal and three-day running average maximal temperature exceed a threshold.

Concretely, a heat-wave day is observed in a grid cell if

- the three-day moving average minimal temperature of the grid cell is higher than the 98th percentile of the rural minimum temperatures during the summer period; and

- the three-day moving average maximal temperature of the grid cell is higher than the 98th percentile of the rural maximum temperatures during the summer period.

By setting the thresholds as above, i.e., taking the 98th percentile temperature values, the definition is consistent with the one used in Belgium (FPS Public Health, 2014). Indeed, the threshold values used in the Belgian definition, with values of 18.2°C and 29.6°C for minimum and maximum temperatures, respectively, correspond to the rural 98th percentile values observed near Antwerp. At the same time, this definition can be transferred to other areas, as it makes use of local climate statistics; i.e., rather than employing fixed values for the threshold temperatures, the 98th percentile temperature values are established based on local (rural) temperature time series.

To calculate the rural threshold values, ERA-Interim data for the period 1986-2005 is used. Note that this definition of a heat wave is rather stringent, hence, few heat waves are present in the reference runs for the years 1986-2005. The result for Antwerp is given in Figure 4, showing the number of heat-wave days per year. In the remainder, this indicator is referred to as the 'heat-wave day indicator' and we refer to the 'heat-wave UHI effect'.

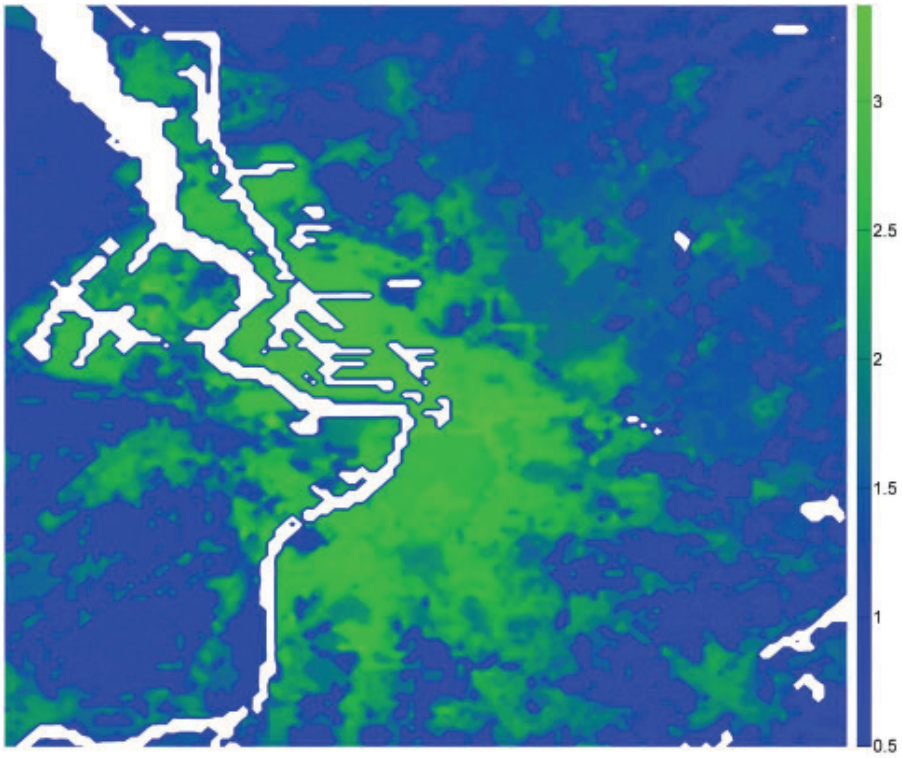


Figure 4. Heat-wave day indicator for Antwerp for the period 1986 -2005. Note that areas over water have been omitted.

The UHI effect, as before, is quantified as the difference between the urban and the rural number of heat-wave days, which are again defined using the spatial 90th percentile and the spatial 10th percentile over the domain under investigation, respectively.

3.3 Target cities

Within the current study, the focus is on nine cities, as shown in Table 1. These correspond to the RAMSES target cities, with the exception that

1. Bogotá is not included given the problems encountered to simulate that city (see below, in the section 'Technical issues and problems encountered');
2. conversely, we have added results for two other cities, Almada (PT) and Berlin (DE), which are target cities of the EU-FP7 NACLIM project in which VITO is a partner. Urban climate projections were produced following the same procedures (periods, resolution, etc) as in RAMSES, in order to allow comparison.

City	Lat (°)	Lon (°)	Domain size (km)	Resolution (m)	Summer period
Almada	*38.68*	*9.16*	*31*	*250m*	*May – Sep*
Antwerp	51.25	4.41	31	250m	May – Sep
Berlin	*52.52*	*13.38*	*51*	*250m*	*May – Sep*
Bilbao	43.26	-2.92	31	250m	May – Sep
Hyderabad	17.39	78.48	41	500m	Mar – Jul
London	51.51	-0.13	91	500m	May – Sep
New York	40.71	-74.01	121	1km	May – Sep
Rio de Janeiro	-22.84	-43.32	71	500m	Nov – Apr
Skopje	42.00	21.43	31	250m	May – Sep

Table 1. Specification of the domains and periods considered here.

The following sections describe the methodology for achieving the urban climate projections in detail. First, we provide an overview of the work flow, including a discussion of technical difficulties and problems encountered. Subsequently, the main results of the urban climate projections are given.

3.4 Towards the global applicability of the UrbClim model

In the RAMSES project, several of the target cities are outside Europe (Rio de Janeiro, New York, Hyderabad, Bogotá). The main difficulty in setting up UrbClim simulations for non-European cities is related to the availability of suitable terrain data. For European cities, we rely on the CORINE land cover data for Europe, the European Environment Agency Soil Sealing data, the Normalized Difference Vegetation Index acquired by the MODIS instrument on the TERRA satellite, and the Global Multi-resolution Terrain Elevation Data (GMTED) of the U.S. Geological Survey (USGS), respectively. Whereas the latter two are available globally, the former two are available for Europe only.

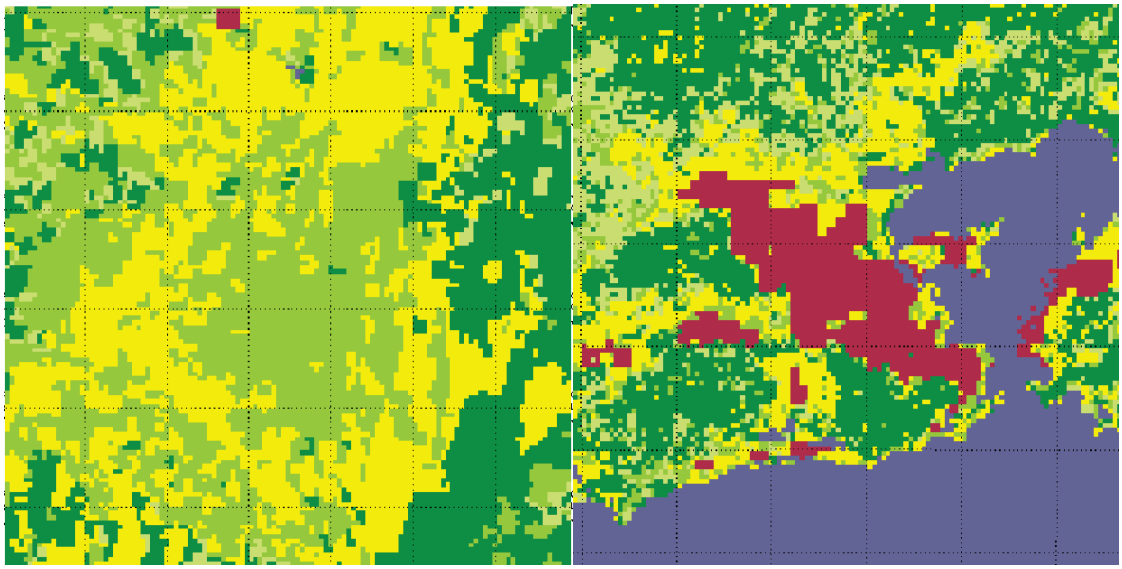


Figure 5. GlobCover land use data for Bogota (left) and Rio de Janeiro (right). Yellow colours indicate shrub land, green colors forest and woodland, and artificial structures are in red.

There are some alternatives for the CORINE land cover data, but none of them achieves the 100 m resolution of the CORINE data set. The highest resolution alternative is the GlobCover data set, which provides land cover data at 300 m resolution. Notwithstanding the global coverage of the data set, the quality of the output varies greatly between different continents, with the best data quality being observed for Europe (Bontemps, 2009). Among the main issues with the GlobCover data is its underestimation of the size of cities, which is obviously a major problem when using the data as input for UrbClim simulations. As an example, Figure 5 shows the GlobCover land use data for the Bogotá area. Man-made structures, indicated in red, are clearly absent.

As a second alternative, MODIS provides land use data at 500 m resolution. In contrast to GlobCover, MODIS appears to have a tendency to overestimate the size of urban areas, as indicated by Figure 6.

Moreover, there are differences in classifications between GlobCover or MODIS land use on the one hand and the CORINE land cover data set on the other hand. Most of these, as for instance the different treatment of harbours, which in MODIS and GlobCover are often classified as 'bare soil', while CORINE classifies them as 'industrial areas', are quite easily resolved by comparing the composed land cover with satellite data and modifying some class definitions by hand. A more subtle difference comprises the categorization of urban areas. Whereas CORINE makes a distinction between urban and suburban areas, GlobCover and MODIS contain only one class with artificial structures. To account for this, the definition of the classes in UrbClim is adapted. Finally, since the modelling of urban heat islands crucially relies on correct city delimitations, corrective measures are taken based on a visual comparison with satellite data. For instance, for Rio de Janeiro, we have made an overlay between the GlobCover and the MODIS land use data, where the city limits of the former are used as the boundary of the 'urban areas', and the city limits of the latter are used as the boundary of the 'Suburban area'. The result is shown in Figure 7.

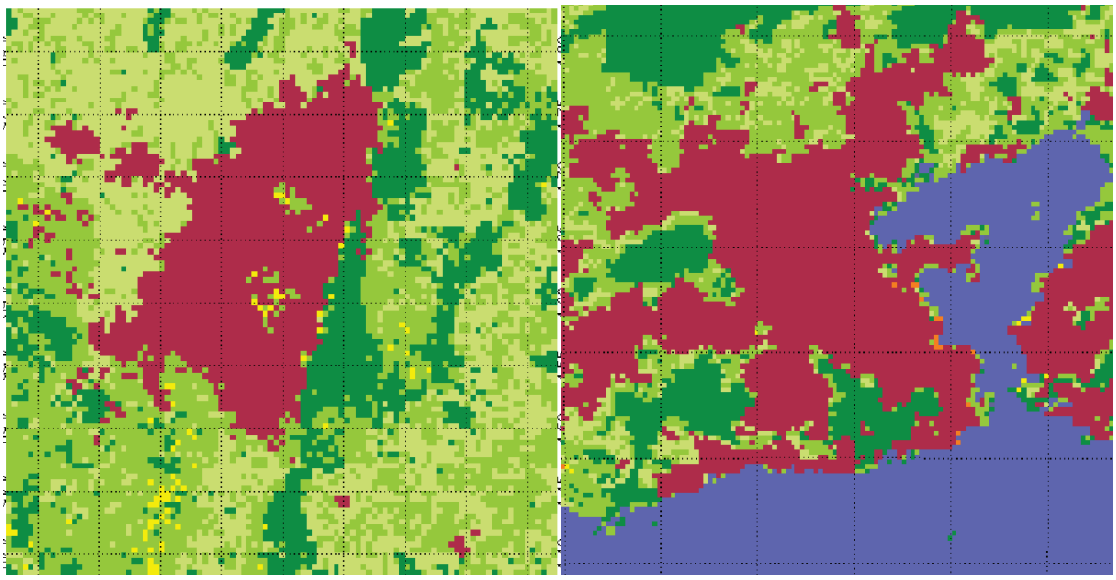


Figure 6. MODIS land use data around Bogota (left) and Rio de Janeiro (right). Colour coding is the same as in Figure 5.

With respect then to soil sealing information: whereas the EEA Soil Sealing data provides the covering of the ground by an impermeable material on a 100 m scale for Europe, similar data sets are not available on a global level. Although the VIIRS

Impervious Surfaces data set do provide similar data, the resolution of this data set is far from sufficient.

Hence, we chose to apply a different strategy, trying to establish a relation between the land use data, the NDVI and the soil sealing. In detail, the soil sealing in an urban grid cell is determined as a function of the GlobCover and MODIS land use and the NDVI in the cell and the neighbouring cells, where the neighbouring cells under consideration form a 2×2 km^2 square around the cell. Since the soil sealing in non-urban cells is typically much lower than the soil sealing in urban cells, the soil sealing of non-urban cells is set to zero.

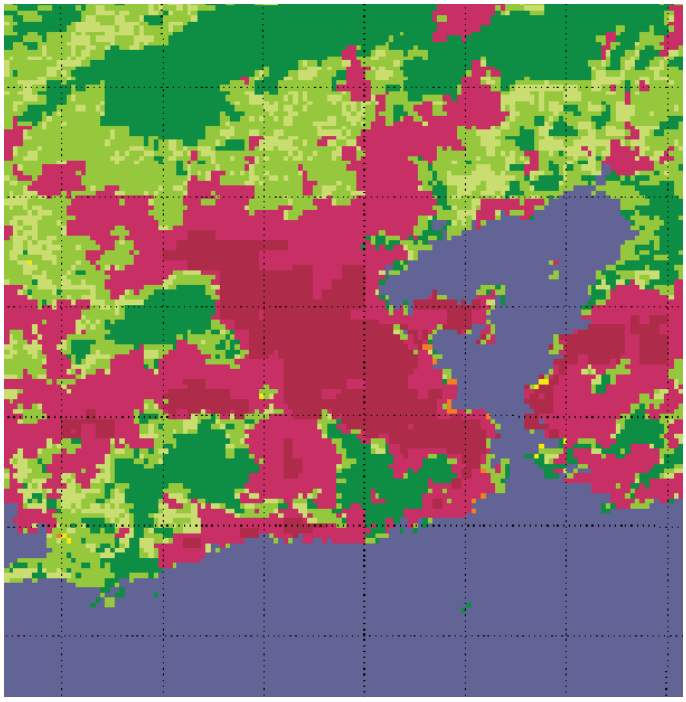


Figure 7. Combination of GlobCover and MODIS land use data around Rio de Janeiro. Yellow colours indicate shrub land, green colours forest and woodland, and artificial structures are indicated in red, with sub-urban areas being light-red and urban areas dark red.

The regression has been constructed and validated for European cities via a leaving-one-out principle, i.e., the regression parameters were derived using the data for Almada, Berlin, Bilbao, London and Skopje, and validated using the data for Antwerp. The regression acquires an explained variance of $R^2 = 0.67$; a visual representation of the reconstructed soil sealing for Antwerp can be found in Figure 8. Although some visible differences between both soil sealing sets exists, the distributions of the soil sealing is more or less equal and the obtained map is good enough to use as input for UrbClim simulations, as we illustrate in the next section. Note also that the reconstructed soil sealing contains less details than the original

EEA soil sealing, which is not surprising since the regression input data (the NDVI and the GlobCover / MODIS land use category) is provided on a much coarser grid.

Following this evaluation based on independent data, the regression can be applied to non-European cities, and now for any city the terrain input for UrbClim can be constructed. As an example, we show the soil sealing data for Rio de Janeiro in Figure 9.

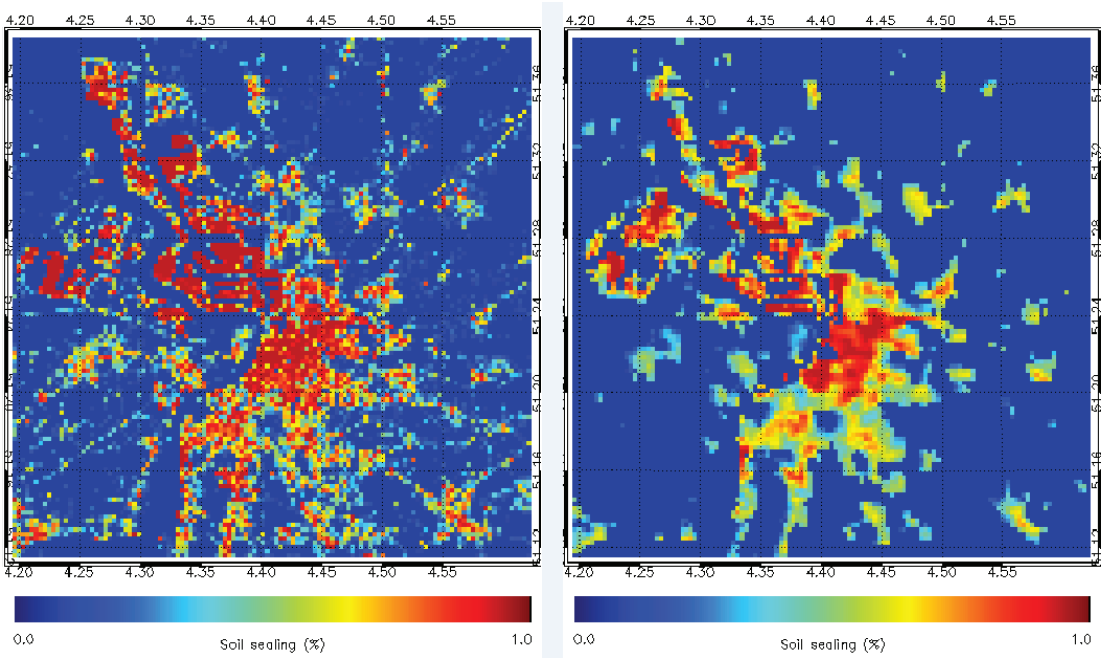


Figure 8. EEA soil sealing (left) and reconstructed soil sealing (right) for Antwerp.

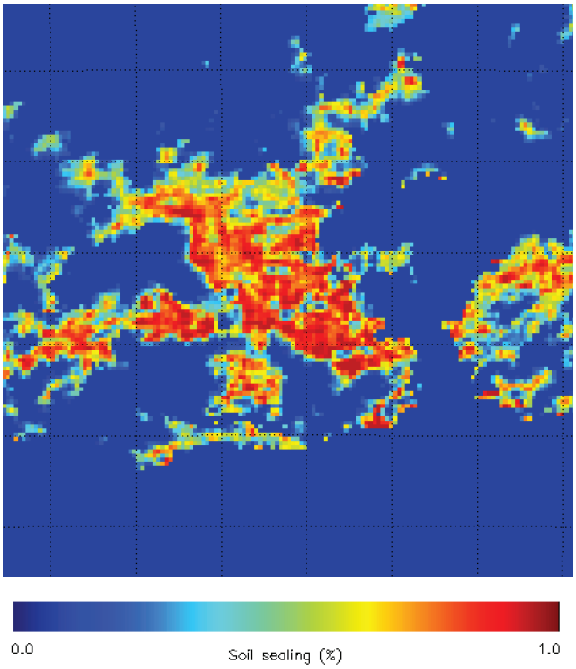


Figure 9. Reconstructed soil sealing degree for Rio de Janeiro.

3.4.1 Evaluation of UrbClim with the new input data

The procedures to reconstruct land cover and soil sealing described above were validated using UrbClim simulation results for Antwerp. Since the soil sealing regression described above was established without using any data for Antwerp, this method provides an independent validation of the complete procedure outlined in the previous paragraphs.

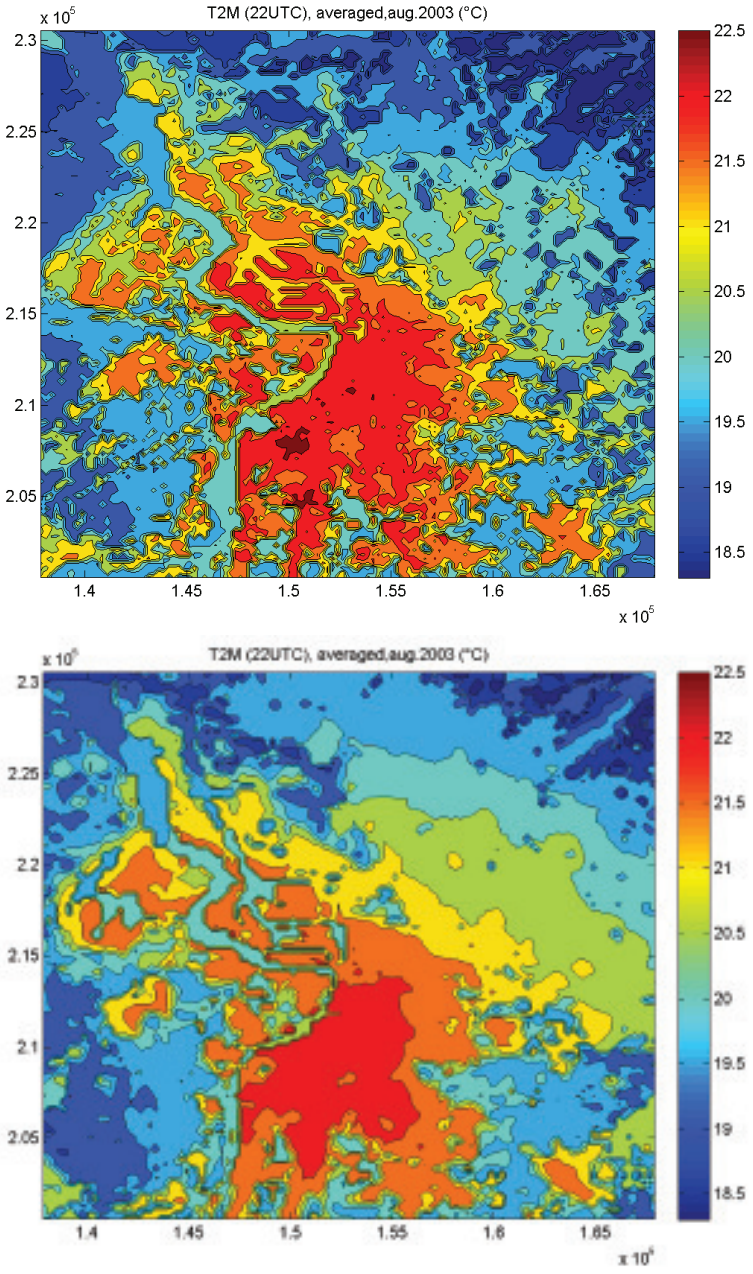


Figure 10. Mean temperature at 22 UTC for the Antwerp domain for August 2003. The upper figure shows the results for an UrbClim simulation using EEA soil sealing and CORINE land cover data, while the lower figure uses GlobCover and MODIS land use data and reconstructed soil sealing data.

UrbClim runs with GlobCover / MODIS land cover and reconstructed soil sealing were performed for the period 1986-2005, and the results of these runs were

compared with output data for the same period, but obtained with CORINE land use and EEA soil sealing data. As an example, Figure 10 shows the mean urban temperature (as defined previously) at 22 UTC in Antwerp for August 2003. Although the level of detail is a bit less in the runs with the reconstructed terrain fields, the differences are smaller than the error between UrbClim and meteorological observations.

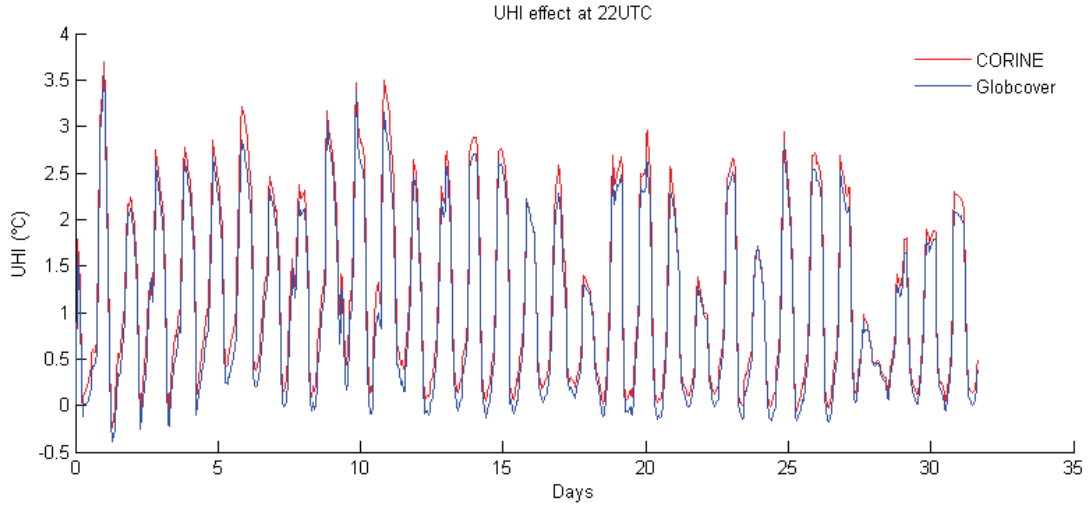


Figure 11. UHI intensity at 22UTC in Antwerp for August 2003 for the CORINE (red) and GlobCover / MODIS (blue) runs.

Both the spatial distribution of the temperatures and the time series of the simulations exhibit a good degree of similarity. Figure 11 compares the difference between the rural and the urban temperature at 22 UTC. As described above, the rural and urban temperature are defined as the 10th respectively the 90th spatial percentile occurring in the domain. Once more, the differences between both simulations are smaller than the uncertainty of UrbClim output. The largest error found for the period 1986-2005 amounts to 0.47 °C, the root mean square error being 0.10 °C. The correlation between both time series amounts to 0.997. Finally, the heat wave day statistics are also similar for both simulations. We thus conclude that the scheme to specify terrain parameters based on GlobCover and MODIS data has been successfully validated.

3.4.2 Demonstration

In the previous section, we introduced techniques to upgrade UrbClim for a world-wide deployment. Since the MODIS input data are only available at a resolution of 500 m, the resolution of the UrbClim runs will also be limited to this. In principle, a smaller grid increment is possible, but in practice the gain is negligible. As an

example of UrbClim's global usability, this section provides results for Doha (Qatar) and Rio de Janeiro (Brazil).

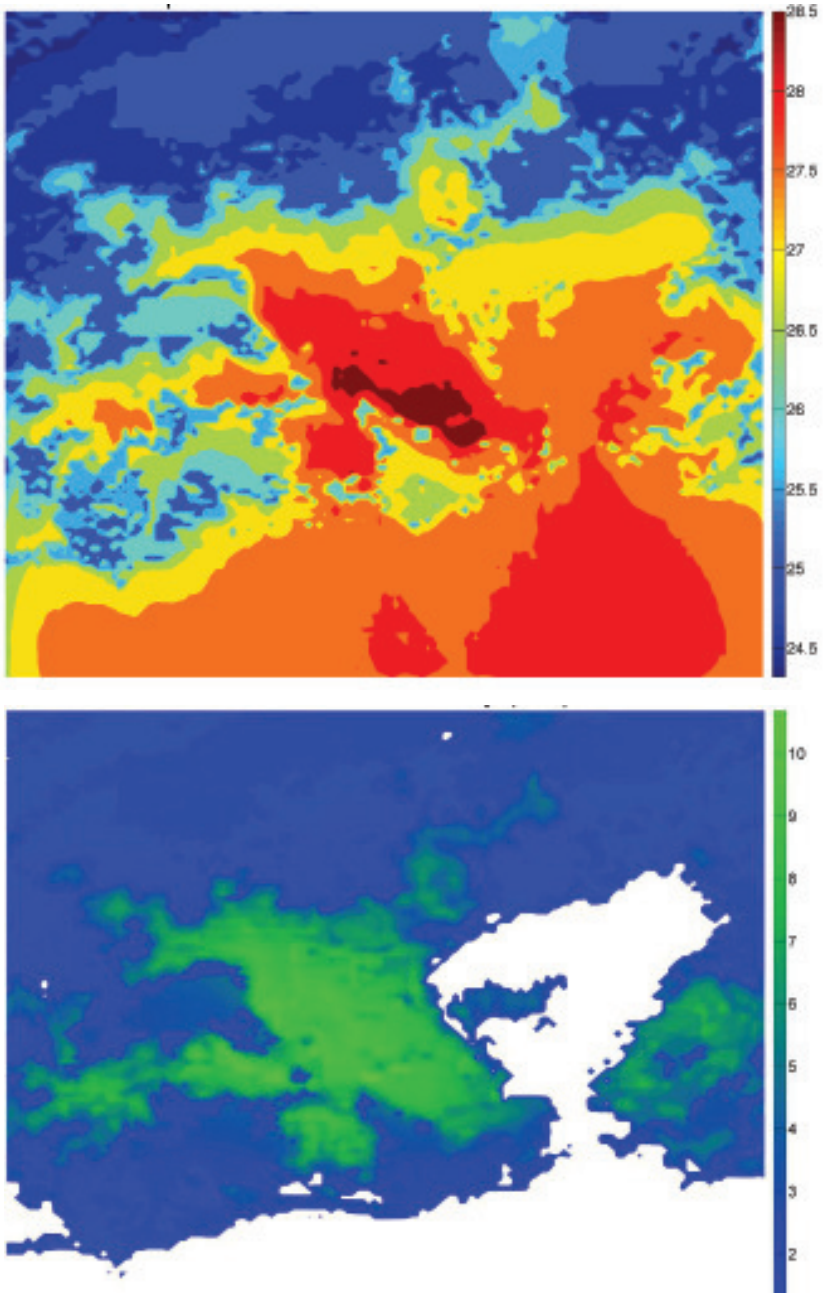


Figure 12. Temperature indicator (upper panel) and heat wave day indicator (lower panel) for Rio de Janeiro for the summer (November–April) in the period 1986-2005.

The meteorological forcing data was taken from the large-scale (70 km resolution) ERA-interim re-analysis data set of the European Centre for Medium-Range Weather Forecasts (ECMWF). As for European cities, the vegetation cover and the elevation data are taken from the Normalized Difference Vegetation Index acquired by the MODIS instrument on the TERRA satellite, and the Global Multi-resolution Terrain Elevation Data (GMTED) of the U.S. Geological Survey (USGS), respectively.

The land use data is a combination of the GlobCover Land Use and the MODIS Land Use data, and the soil sealing has been inferred from the land use and the vegetation cover, using the method explained above.

For Rio de Janeiro, the summer is defined as the period November-April, which coincides with the months with the highest maximum temperatures. In Figure 12, the UHI-effect for the summer period from 1986-2005 is illustrated using the temperature and the heat-wave day indicator, as introduced previously.

UrbClim is deployable on a global scale and for a wide range of climates. As an illustration, we performed simulations for the summer climate in Doha (Qatar) for the summer of 2005, Figure 13 showing the mean temperature at 0 h local time.

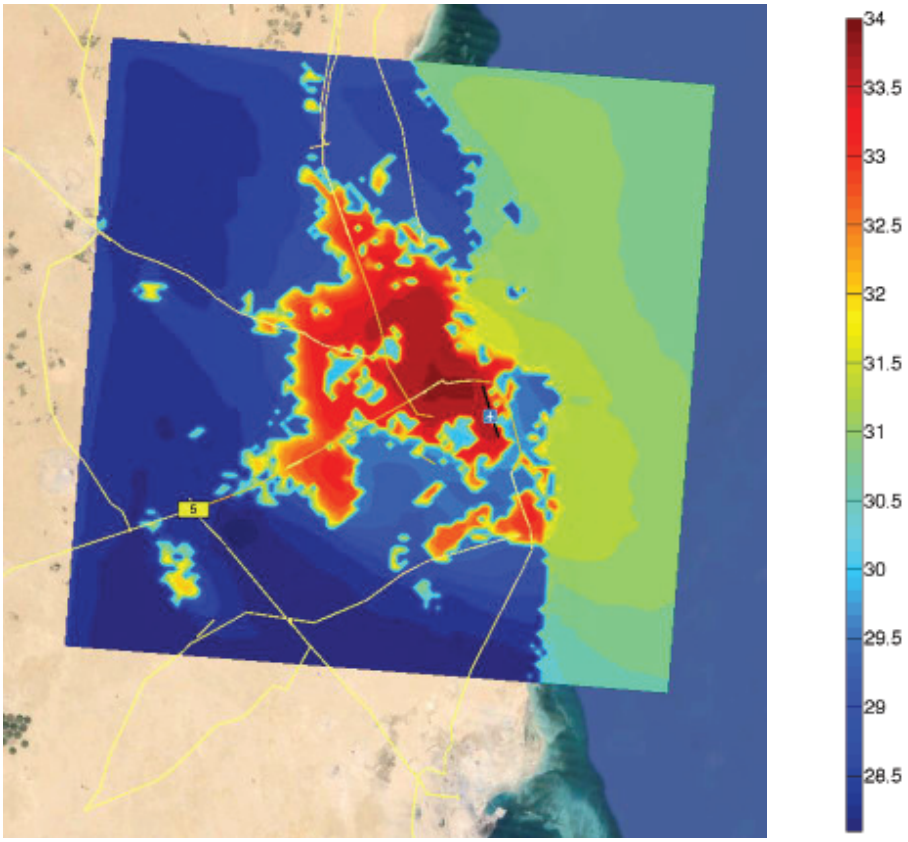


Figure 13. Average temperature at 0 h local time for the summer (May–September) of 2005 in Doha in Qatar.

3.5 Coupling to GCM output fields

One of the major research topics of WP4.2 comprises a study of future urban climate. For these urban climate projections, obviously, we need to rely on Global Climate Model (GCM) output fields instead of the ERA-Interim forcing which has been used as a forcing for the UrbClim model. Note that, ideally, we would have

used Regional Climate Modelling results from the CORDEX initiative, but given that the CORDEX archive is currently far from complete, we have had to rely on CMIP5 results instead. In addition, we cannot force UrbClim using just one GCM. Rather, the model requires forcing by an ensemble of GCM outputs, in order to properly represent the uncertainty associated with the global climate projections. This will make it possible to obtain mean values and meaning tendencies, together with an uncertainty range around them. The latest revision of the IPCC AR5 WG I report (IPCC, 2013) also substantiates this as good practice, highlighting that the agreement between ensemble means and climate data exceeds the agreement with any single climate model by a large amount.

The IPCC 5th Assessment report (IPCC, 2013) identifies four climate scenarios, ranging from very strong mitigation scenarios (RCP2.6) to a business-as-usual scenario (RCP8.5). Due to CPU-limitations, we have currently only considered climate projections using the RCP8.5 scenario. (Based on what is shown in AR5, it is reasonable to assume that, to first order, results for other RCP scenarios can be estimated simply by scaling between the reference (current) state and the RCP8.5 scenario. In any case, the reference case and the RCP8.5 scenario define a range for presenting climate projection results.) AR5 also identifies three time frames: a reference period (1986-2005), a near future period (2026-2045) and an end-of-century time frame (2081-2100). To be consistent with the IPCC, the same time frames have been used in our simulations. Note that, for the urban climate projections, the ERA-interim runs are considered as the benchmark for all future climate projections. (Below, we will explain how the ERA-Interim reference period is made to match with the GCM-based reference period.)

3.5.1 Input from the GCMs for climate projections

One of the main difficulties when switching from using ECMWF (Re-)Analysis data to CMIP5 GCM output fields is that of data availability. Indeed, while the ECMWF data are quite comprehensive, e.g., providing vertical profiles with a great level of detail (fine layers, many variables) every three hours, data available in the CMIP5 archive is much more limited. The minimal input requirements of the UrbClim model sets a constraint on the GCMs of which outputs can be used to make urban climate projections. Table 2 provides a list with the necessary meteorological input parameters and the minimal time increment required for these variables when using them as forcing input for UrbClim.

Category	Variable	Required time increment
Surface	Latent heat flux	3h
	Surface upward sensible heat flux	3h
	Downwelling short-wave radiation	3h
	Downwelling long-wave radiation	3h
	Surface air pressure	3h
	Air temperature at surface	3h
	U component of wind velocity at surface	3h
	V component of wind velocity at surface	3h
	Specific humidity at surface	3h
Sea	Sea surface temperature	monthly
Precipitation	Precipitation	3h
	Convective precipitation	3h
Vertical profiles	U component of wind velocity	6h
	V component of wind velocity	6h
	Temperature profile	6h
	Specific humidity profile	6h
Soil	Soil temperature	monthly
	Soil moisture content	monthly

Table 2. List of required meteorological input data.

Model	Institute	ΔLon	ΔLat
ACCESS1.0	CSIRO / BOM	1.875	1.25
BNU_ESM	BNU	2.8125	2.79
CCSM4	NCAR	1.25	0.94
CNRM_CM5	CNRM	1.40625	1.4
FGOALS-G2	CAS	2.8125	2.79
GFDL-ESM2M	GFDL	2.5	2.02
GISS-E2-R	NASA	2.5	2
HadGem2-ES	MOHC	1.875	1.25
IPSL-CM5A-MR	IPSL	2.5	1.2676
MIROC-ESM-CHEM	Univ. Tokyo	2.8125	2.79
MRI-CGCM3	MRI	1.125	1.1121

Table 3. List of GCMs used to drive the UrbClim model with.

Following a detailed scrutiny of the CMIP5 archives it was found that 11 climate institutes provide model output in which all required parameters are available at a suitable time increment. In order to reduce the computational complexity, for each participating institute a single model is selected according to the data availability and the model performance derived from the IPCC Fifth Assessment report (AR5). The list of GCMs used and their latitude and longitude grid size is shown in Table 3.

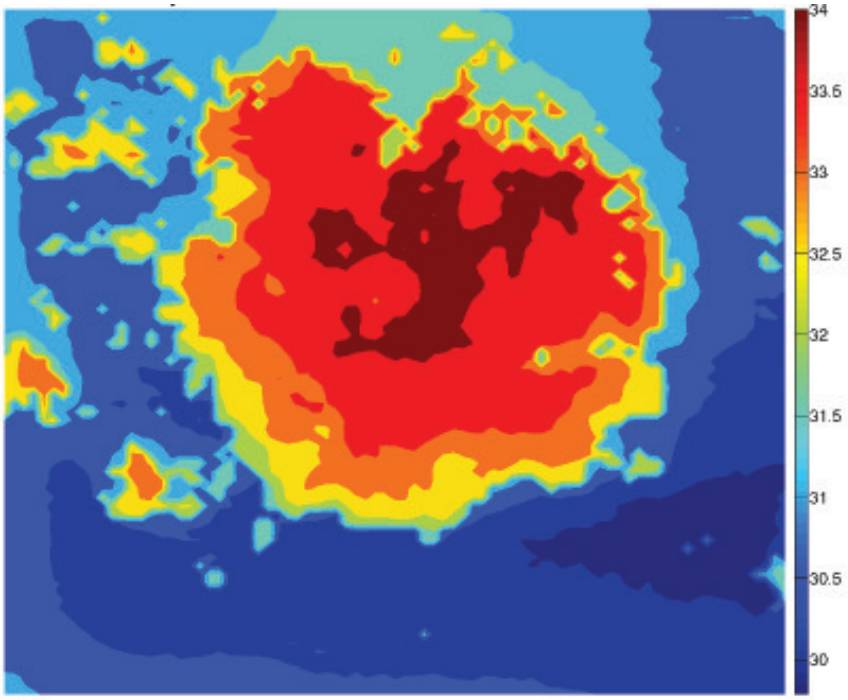


Figure 14. Temperature UHI-indicator for Hyderabad for UrbClim runs driven by the CNRM-CM5 model.

Even though only 11 models are retained, downloading and pre-processing the GCM output data is still a complicated and time consuming task. Apart from the technical difficulties related to the different conventions (e.g., related to variable names), there are two main issues connected to the use of the GCM output as driving meteorological input for UrbClim. Firstly, vertical profile data are available only at a 6-hourly time step, which is insufficient to properly resolve the diurnal cycle, and in particular the daily minimum and maximum temperatures that are required to construct the heat stress indicators (see above). To remedy this, a neural network routine was set up, using as input ERA-interim data, to bridge the gap between the required time frequency of the vertical profiles in UrbClim (3-hourly) and the best resolution present in the CMIP5 archives (6-hourly). Secondly, UrbClim requires precipitation data at a relatively fine scale level (i.e., the city and its direct surroundings), while the GCM output is provided at a much coarser resolution, thus neglecting local thunderstorms among other things. Hence a downscaling routine based on ERA-interim data was established.

Data from the selected GCMs was interpolated and downscaled for each of the cities listed above. These results were then used as meteorological input for the UrbClim model. Generally speaking, for all cities except Hyderabad (which exhibited some results which deviated from their expected values), the overall simulation results are in accordance with expectations. Figure 14 shows the appearance of strong boundary effects in the southern portion of the domain, and also very strong temperature gradients at the city boundaries. At this moment, it is yet unclear why these effects are observed for Hyderabad, while they are absent for the other cities. Possibly, the urban boundary issue is related to the specification of terrain parameters, characterized by a lack of diversity in the land use categories and the satellite NDVI (a measure of vegetation abundance) in the rural area. Because of these unfamiliar effects, the simulation results for Hyderabad are for the moment neglected in most of the data analysis.

3.5.2 Bias correction

A first test for the UrbClim simulations forced with input from the CMIP5 GCM fields consisted of a comparison with simulation results with ERA-interim as a large-scale driver. Of course, a day by day comparison is nonsensical but, in principle, the 20-year statistics should agree. However, as can be seen for the IPSL-CM5A-MR model in Figure 15, there are large discrepancies, even when considering averages over the full period of 20 years.

There are many reasons why the UrbClim simulations with model output of the GCMs do not match the simulations with ERA-interim as meteorological driver. Although the GCMs are improving constantly, they are still far from perfect, especially if one concentrates on a single location in the world. Moreover, the grids of the GCMs are very coarse compared to the ERA-Interim grid, which of course may cause major differences. For instance, the height of the grid cell in which the city is located compared to the ERA-Interim grid cell, is often very different to the height of the cell in which the city lies according to the GCMs, especially in mountainous areas. Furthermore, the way in which UrbClim treats the input variables was tailored for ERA-Interim data, and hence differences should be expected for simulations when using GCM data as input.

Yet, it is reasonable to enforce that the ERA-interim runs should be considered as the benchmark for all future climate projections. Therefore, we introduce a bias correction which reduces the differences between (1) the urban climate simulated

with ERA-interim meteorological input, and (2) the urban climate simulated with the GCM as a driver, for the reference period.

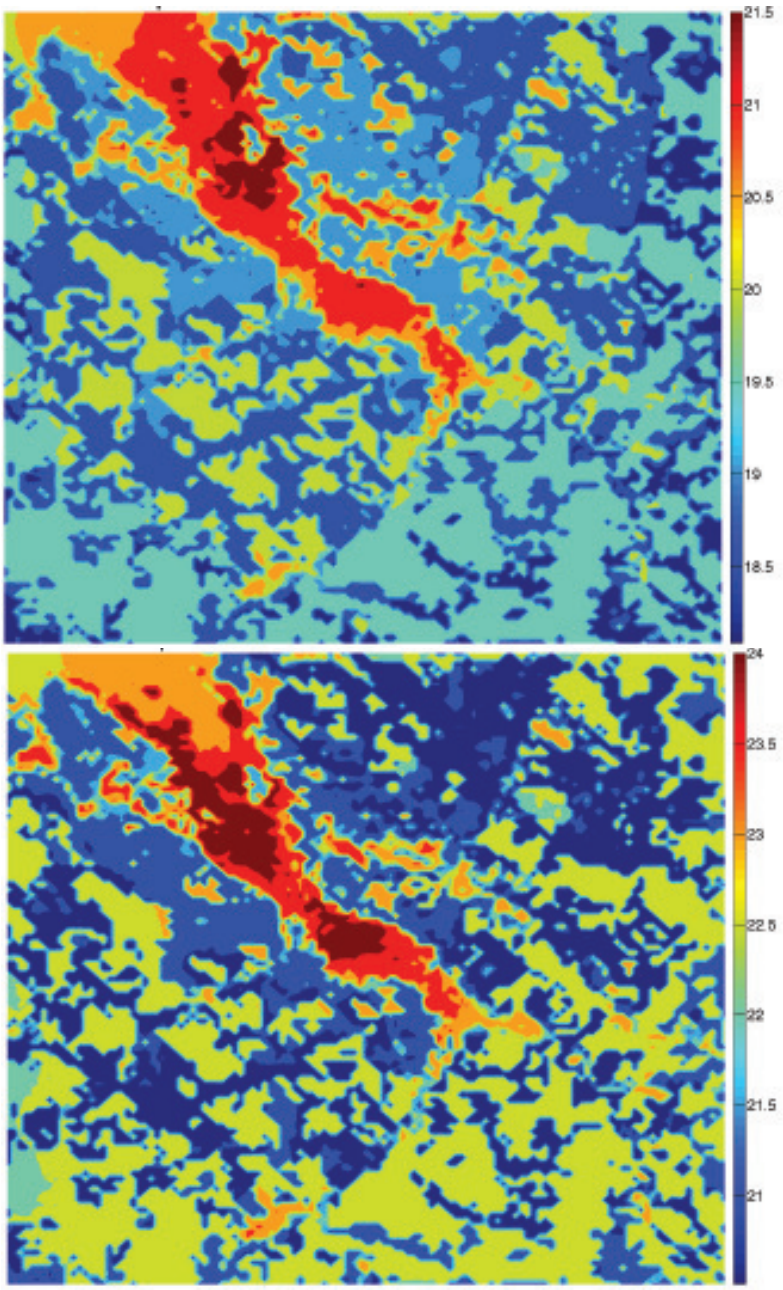


Figure 15. Temperature UHI-indicator for Bilbao for the summers of 1986-2005 for UrbClim simulations coupled to ERA-Interim data (upper panel) and to the IPSL-CM5A-MR global climate model (lower panel).

The bias correction rescales, for each grid point separately, the mean and the variance of the (apparent) hourly temperatures of the CMIP5 reference run to those of the ERA-interim runs for the reference period. A different scaling is applied for each summer month. Hence, we rescale $T(i,j,t)$, the temperature in grid cell (i,j) at time t corresponding to month M and hour h, by

$$T(i,j,t) \rightarrow \left(T(i,j,t) - \langle\langle T_{i,j}^{M,h}\rangle\rangle_{GCM_{ref}}\right).\frac{\sigma_{i,j}^{M,h}|_{ERA-int}}{\sigma_{i,j}^{M,h}|_{GCM_{ref}}} + \langle\langle T_{i,j}^{M,h}\rangle\rangle_{GCM_{ref}} \quad (1)$$

where $\langle\langle T_{i,j}^{M,h}\rangle\rangle_{GCM}$ is the temperature according to the UrbClim simulations using the GCM, $\langle\langle T_{i,j}^{M,h}\rangle\rangle_{ERA-int}$ is the same quantity for the ERA-interim simulation, and $\sigma_{i,j}^{M,h}|_{GCM}$ and $\sigma_{i,j}^{M,h}|_{ERA-int}$ are the corresponding standard deviations. Using (1), bias corrected temperature data are calculated for the reference period for all considered CMIP5 models. The result of the bias correction is illustrated in Figure 16 for the IPSL-CM5A-MR model. The correspondence between the bias corrected GCM run and the UrbClim simulations driven by the ERA-Interim data is much better than that between the latter and the GCM run without the bias correction.

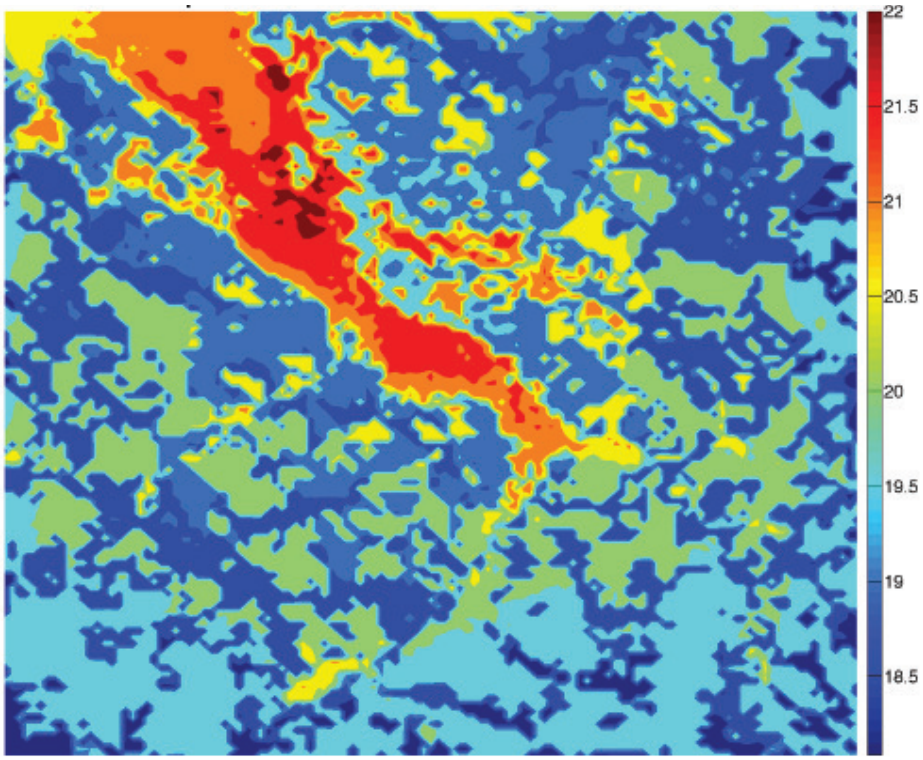


Figure 16. Temperature UHI-indicator for Bilbao for the summers of 1986-2005 for the bias-corrected UrbClim simulations coupled to IPSL-CM5A-MR global climate model.

Also the output temperatures for the scenario runs (i.e., climate projections, for which no reference benchmark is available) are rescaled according to the bias correction for the reference period. In detail, the output temperatures of the scenario runs are rescaled such that the increase in mean temperature and standard deviation between the original (non-bias corrected) is mimicked in the bias corrected versions. Mathematically, this implies

$$T(i,j,t) \rightarrow$$
$$\left(T(i,j,t) - \langle\langle T_{i,j}^{M,h}\rangle\rangle_{GCM_{scen}}\right).\frac{\sigma_{i,j}^{M,h}|_{ERA-int}}{\sigma_{i,j}^{M,h}|_{GCM_{ref}}} + \langle\langle T_{i,j}^{M,h}\rangle\rangle_{ERA-int} + \langle\langle T_{i,j}^{M,h}\rangle\rangle_{GCM_{scen}} - \langle\langle T_{i,j}^{M,h}\rangle\rangle_{GCM_{ref}} (2)$$

In the remainder of this report, the urban climate projection results presented are always based on a comparison between bias corrected data for the reference period and bias corrected data for the scenario runs.

3.5.3 Technical issues & problems encountered

While implementing the scheme described above, many technical problems and computational issues were encountered. In the remainder of this section, we highlight some of the more important problems and discuss their solutions. Special attention is devoted to the pre-processing of the CMIP5 input data and the computational challenges involved in the urban climate projections. The combination of all the issues described below, makes the pre-processing of the CMIP5 data a lengthy and, unfortunately, inefficient task, since every CMIP5 model requires a different, custom-made, processing.

(i) Acquiring and pre-processing the CMIP5 data

Although only 11 models are retained, downloading and pre-processing the GCM output data is still a complicated and time consuming task. At first, in total, approx. 18 TB of data distributed over more than 16.000 files has been downloaded. Next, since all institutes use different conventions for the meteorological parameters and since some files contain incompatibilities and errors, we have composed a different pre-processing routine for all GCMs. Apart from the technical difficulties related to the different conventions, there are two main issues related to the use of the GCM output as driving meteorological input for UrbClim. Firstly, UrbClim in principle requires vertical profiles with a 3-hr resolution, while the CMIP5 database only provides 6hr profiles. Secondly, UrbClim requires precipitation data at a local (i.e. the city and its surroundings) level, while the GCM output is provided at a much coarser resolution. In the next sections, we elaborate in detail on these issues.

(ii) Construction of vertical profiles

Although the CMIP5 archives provide 6-hourly vertical profiles, it should be noted that we cannot simply and solely use the 6-hourly profiles as external driving fields for the UrbClim model. Indeed, the 6-hourly frequency is far from sufficient, as UrbClim requires a forcing with a well-resolved diurnal cycle. This is particularly important in the context of urban climate modelling, which attaches particular

importance to the daily minimum and maximum temperatures, which clearly cannot be properly contained in data available at 6-hourly intervals only. Even the three-hourly interval we aim for is perhaps a bit limited, but no finer temporal resolution is available within CMIP5 (at least not for a sizeable ensemble of GCM results), and from our experience with the ECMWF's 3-hourly forcing frequency, we know at least that UrbClim results obtained with this forcing frequency compare favourably with observations. One of the challenges is therefore to reconstruct the vertical profiles of the required quantities on a 3-hourly time frequency, based on the 6-hourly information available in the CMIP5 archives.

Since many of the GCMs contained in CMIP5 do archive 3-hourly surface variables (air temperature, wind speed, humidity, turbulent energy and momentum fluxes), we first attempted to reconstruct boundary-layer profiles from this surface information, using the relevant boundary-layer conservation equations to do so, including consideration of the large-scale pressure gradient as an external forcing. However, several issues arose related to the stability of the obtained vertical profiles. In fact, as also pointed out in Press et al. (1992 – see their Section 9.7), it was found that the strongly non-linear character of the system of coupled equations describing the turbulent fluxes, together with a sometimes poor initial guess for the considered quantities, may hamper the convergence of the method. Based on this finding, it was decided to abandon this physically-based avenue, in favour of a statistical approach, which will be briefly described below.

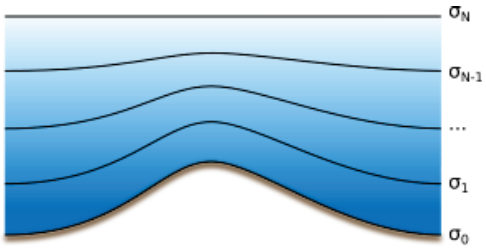


Figure 17. Hybrid sigma pressure levels.

A first step in the construction of 3-hourly vertical profiles based on the 6-hourly vertical profiles, was to make the 6-hourly profiles uniform, since there are two different types of vertical level definitions in the output data of the GCMs. Firstly, for most models the data is provided on *hybrid sigma pressure level systems*. In these systems, the output data is provided on scaled pressure levels. Typically, the pressure on a level is defined either as $p = ap_0 + bp_s$ or $p = a + bp_s$, where a, b and p_0 are (level-dependent) parameters, and p_s is the surface pressure. The parameter b typically decreases with the height. Hence, close to the surface, the

pressure levels closely follow the terrain and its elevation, while higher levels are constant with respect to the mean sea surface height. The situation is sketched in Figure 17.

In these hybrid sigma pressure level systems, the (geopotential) height of the pressure levels is not part of the output of the GCMs. It must therefore be calculated using the hypsometric equation,

$$h = \frac{R}{g}\overline{T}\,\ln\left(\frac{p_s}{p}\right) \tag{3}$$

where h indicates the height above the surface, p the pressure at height h, p_s is the surface pressure, T the layer mean virtual temperature and R and g denote the gas constant for air and the gravitational acceleration, respectively. Other models make use of a *hybrid height level systems*. Similarly to the definition of the pressure in hybrid pressure systems, the level height is now defined as a function of the terrain height and some constants, such that the lowest levels closely follow the terrain height, while the higher levels are more or less constant with respect to the mean sea surface height. In these hybrid sigma height level systems, the pressure of the model levels is not provided, and hence the hypsometric equation is once more used to provide a link between the level pressure and the level height.

In summary, using the hypsometric equation, the formats of all the vertical profiles can be equalized. These rescaled 6-hourly vertical profiles for temperature, humidity, and wind velocity can then serve as input for the statistical tool to compose 3-hourly profiles. The statistical approach however not only uses the 6-hourly profile, but also 3-hourly surface quantities (in particular the turbulent energy and momentum fluxes) are used. Indeed, while for the selected GCMs the vertical profile data are available at 6-hourly time step, surface quantities (including the turbulent surface fluxes) are available at 3-hourly time step. It is precisely this extra information that provides an aid in reconstructing the vertical profiles at the 3-hourly time steps in between a pair of 6-hourly data.

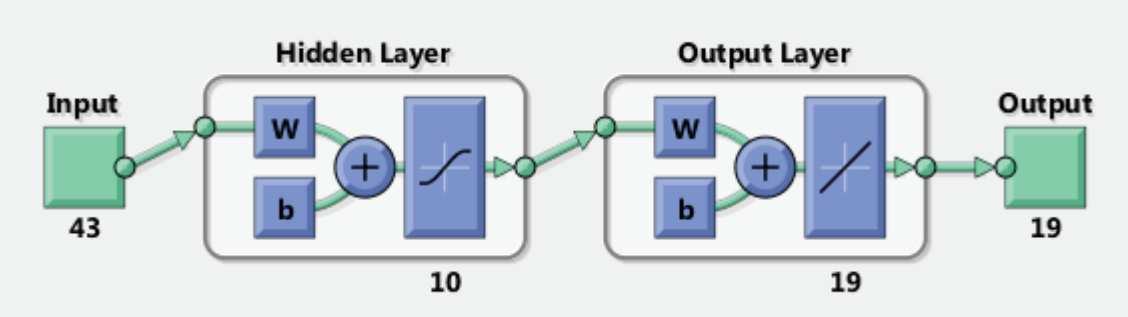


Figure 18. Set–up of the neural networks

Using these input data, 3-hourly forcing profiles were obtained by means of a regression based on a neural network approach. The neural networks were trained based on ERA-Interim data from the period 1986-2005, and the method was then validated with independent data, using ERA-Interim data for the period 2006-2013.

In the neural networks we use 6-hourly profiles of temperature, wind speed, and humidity, and 3-hourly surface data of the same parameters as explanatory variables to interpolate the 6-hourly profile data with. For each variable, a separate neural network was established. We opted for a two-layer feed-forward network with sigmoid hidden neurons and linear output neurons (Figure 18). To keep computer memory burden within bounds, the networks were trained using a conjugate-gradient method.

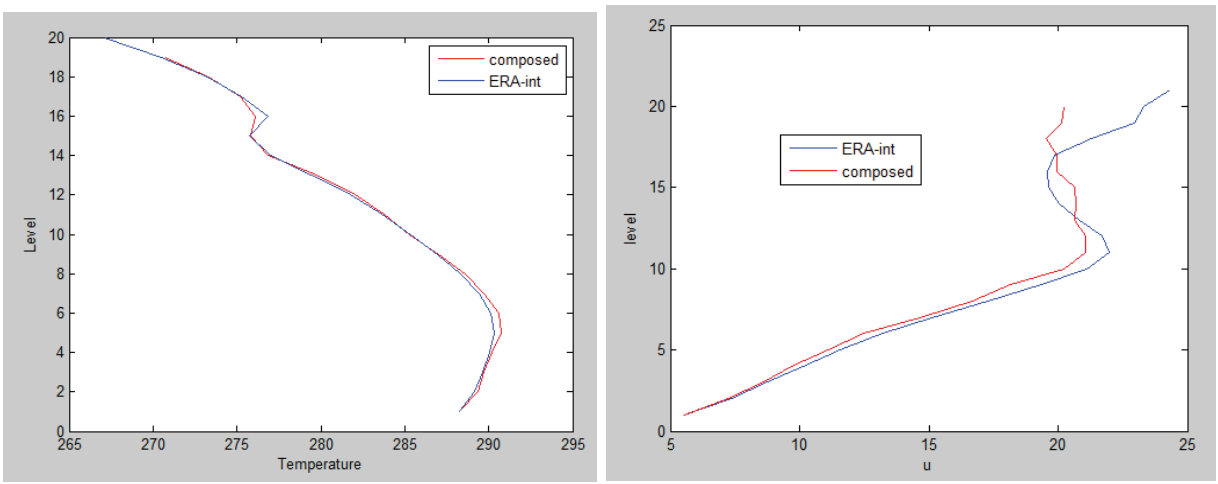


Figure 19. Vertical temperature profiles from ERA-Interim (blue), together with the reconstructed (interpolated) profiles (red).

Figure 19 gives an example of the performance of the neural network-based interpolation method. This is of course just a demonstration for a temperature and a wind speed profile at one particular moment in time, but the error statistics based on a large number of profiles confirm the good performance of this method.

A validation was also undertaken involving the UrbClim model: we ran the model once with regular 3-hourly ERA-Interim data as external forcing, and then we ran the model again in identical conditions, but this time using ERA-Interim data sampled at a 6-hourly time step, and filling up towards a 3-hourly step, using the method described above. As can be seen from Figure 20, the differences between the results obtained with the two types of forcing are negligible. Hence, we now have established a (novel) method of creating the 3-hourly meteorological profiles required by UrbClim as external forcing, which opens the way towards the establishment of urban climate projections.

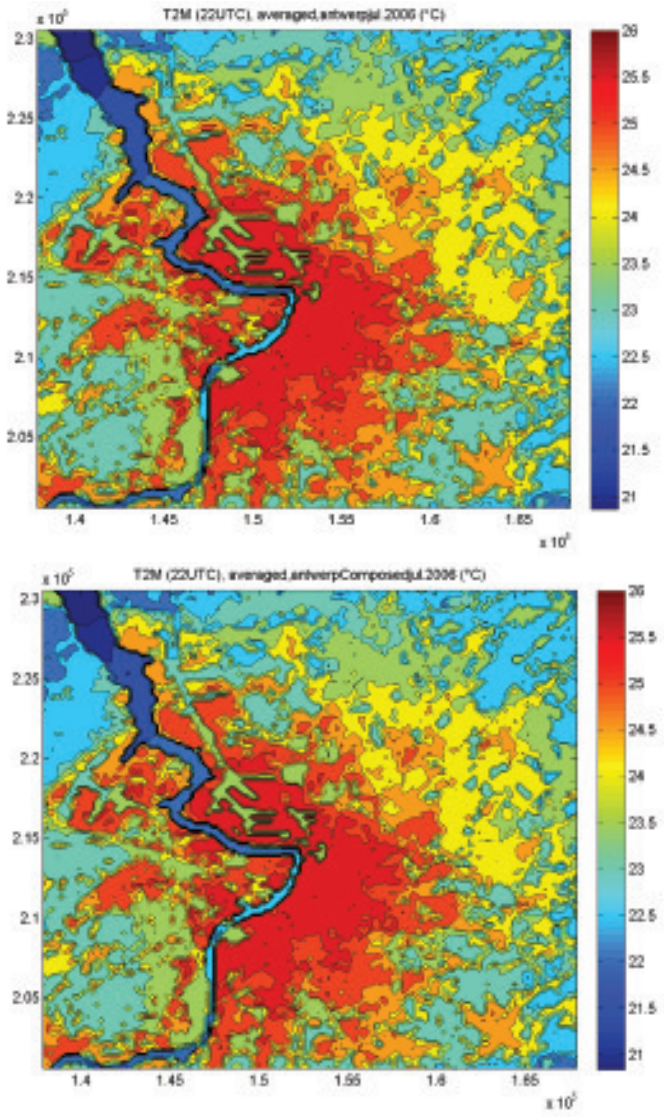


Figure 20. Simulated 2-m air temperature at 22:00 UTC for Antwerp, averaged for the month of June 2006, obtained with a regular 3-hourly ERA-Interim forcing (left) and with a 6-hourly forcing complemented with 3-hourly interpolated profiles (right).

(iii) Downscaling of precipitation

Thunderstorms producing convective precipitation in summer have typical length scales of a few kilometres. This resolution cannot be resolved by the GCMs (>1-2° grid cells) or ERA-Interim (0.75° grid cells). Therefore, all the models apply precipitation parameterization schemes. These schemes produce correct total rainfall amounts for the grid cell area, but they do not capture the local peaks in rainfall, which are important when modelling the local climate. Furthermore, the larger the grid cells, the more drizzle is produced by the model, since there will

always be a small part of the grid cell where it is raining. Therefore, it is necessary to downscale the rainfall from the GCMs to a suitable resolution.

Since UrbClim is focusing on temperatures and not on runoff and flooding, the peak rainfall amounts are perhaps less important, as long as the total rainfall amount is correctly reproduced for the model domain. UrbClim has delivered good validation results while being fed with raw ERA-Interim rainfall data, so we decided to take the spatial resolution of ERA-Interim as a reference, and downscale the GCM rainfall data to this resolution. Doing so also enhances the comparability of the results. In the subsequent paragraphs, the downscaling method is briefly described.

First, a number of ERA-Interim grid cells is selected surrounding the UrbClim domain, corresponding to the resolution of the GCM to be downscaled (e.g. a 3° GCM grid cell corresponds to a 4 × 4 grid cell domain in ERA-Interim). Then all the ERA-Interim reference rainfall data (May-September of the period 1986-2005) is used to derive a linear relation between the rainfall amount in the 4 × 4 subdomain and the centre grid cell covering the UrbClim domain. This is done separately for convective and large-scale precipitation (see Figure 21). The slope of this relation is subsequently used.

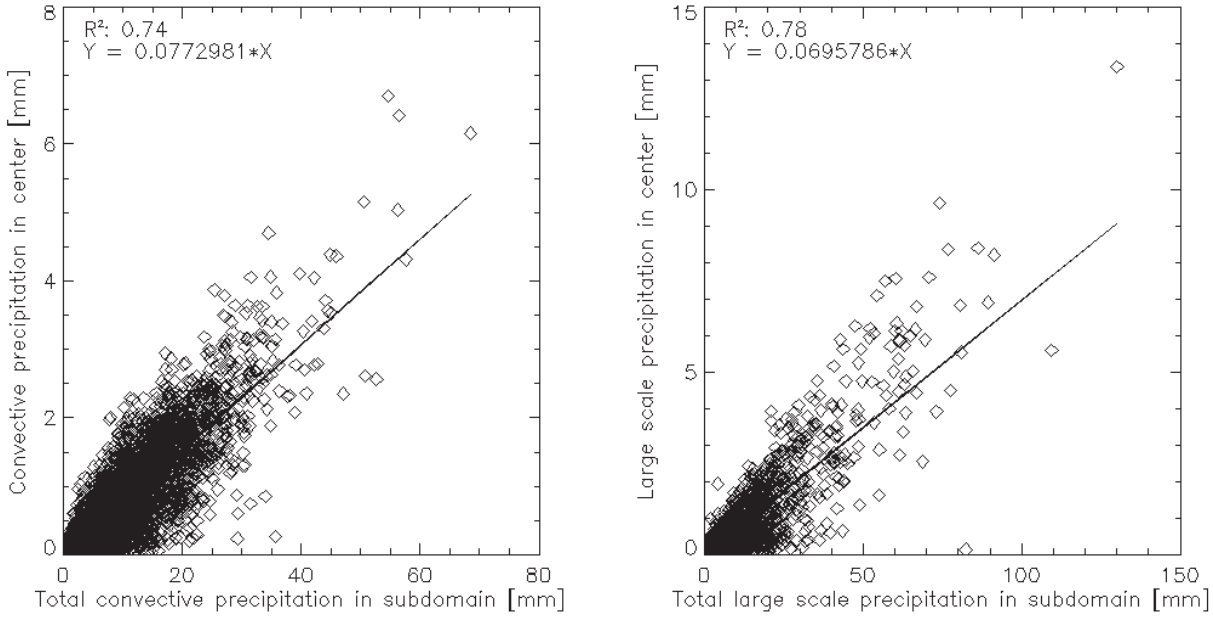


Figure 21. Statistical relations established between GCM and ERA-Interim precipitation, for convective-only precipitation (left) and total precipitation (right).

Subsequently, a histogram is created for precipitation values, using a value resolution of 1 mm, reflecting the number of times the total precipitation in the subdomain had the considered value while no precipitation occurred in the centre

grid cell. As can be imagined, the lower the total precipitation, the more '0'-values are present in the centre cell. For every class in the histogram, the probability of a '0'-value in the grid centre is calculated.

Finally, the GCM rainfall data for the grid cell covering the UrbClim domain are selected. For every time step when rainfall is present, a random number between 0 and 1 is created. If this random number is below the probability of the corresponding histogram class, the rainfall is set to 0. If the random number is above the probability threshold, then the rainfall is multiplied with the slope factor derived previously.

In this way, we are able to significantly reduce the amount of unrealistic rainfall drizzle, while increasing the peak rainfall to ERA-Interim standards, and at the same time preserving the timing and amounts of rainfall from the GCM. In the example below, we show results from the MIROC-ESM-CHEM model for the Antwerp domain during the summer of 1986. The peak rainfall is increased and the number of very small rainfall events is reduced significantly, which make the precipitation values more realistic and suitable to force UrbClim simulations.

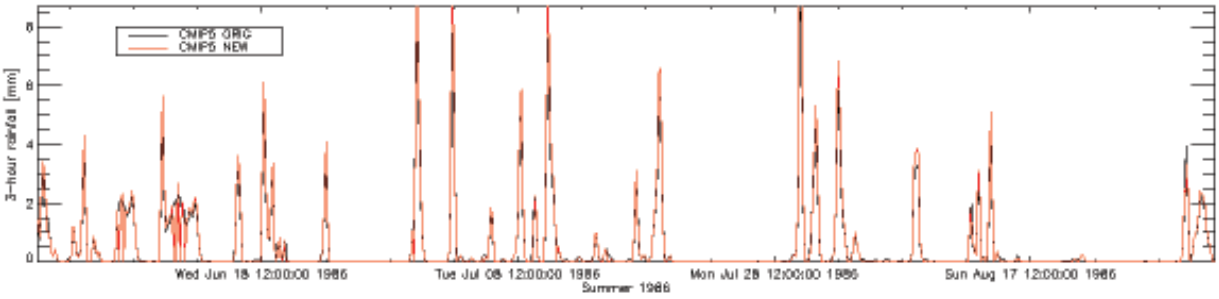


Figure 22. Precipitation time series from the MIROC-ESM-CHEM GCM after the processing step described in the main text.

(iv) Data conventions in the CMIP5 Archive

Most difficulties in the pre-processing of the output data of the GCM have their origin in the different data formats used, with many of those issues being related to the time formats, since there is a plethora of calendars formats in use. Not only different calendars (Gregorian, 365 days, 360 days, etc), but also different reference dates and time parameters (days, hours, months) are encountered. Moreover, the time frequency for sea and soil parameters often differs from those of the surface parameters and vertical profiles. Another time frequency related problem is related to the actual time at which the 6-hourly and 3-hourly data are provided. Finally, data files in the CMIP5 database cover periods of different lengths. Most models

provide annual or penta-annual (covering five years) files, but others stick to (very big) 50-year files, or (a large number of) monthly files. Also the definition of the soil variables varies greatly, since the soil layers in CMIP5 are model-dependent. Hence, for every GCM a different conversion is required.

Another recurring problem comprises the incompatibility between the terrain height specified in the GCMs and the surface pressure specified in the output data of the runs. As an example, if we use this erroneous terrain height to calculate the pressure at 1800m above the surface, nonsensical results are obtained, as is shown in Figure 233, where the mean annual pressure at 1800 m above Europe is shown for the BNU-ESM model. Not only is the terrain clearly visible, the pressure differences between the grid points are also excessive.

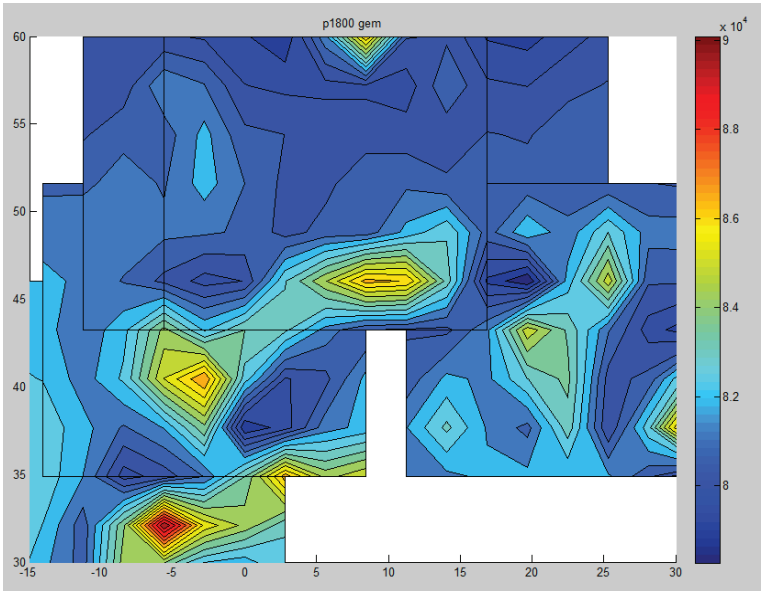


Figure 23. Pressure at 1800m calculated using the original height file of the BNU-ESM model.

We remedy this problem by recalculating the terrain height using the hypsometric equation, the pressure at the surface and the sea surface pressure contained in the GCM data. If the pressure at 1800 m is calculated using this new terrain height, we get much more reasonable results, as is shown in Figure 24. Note also that the difference between both height files varies between -1200 m and +1500 m, as can be seen in Figure 25. We surmise the difference to be related to the interpolation of the original (irregular) grid of the CMIP5 model to the output (regular) grid in the CMIP5 archive. This example also shows that interpolating GCM data in areas with considerable topography may be problematic, as will also be apparent for Bogotá (see below).

Many other smaller problems with the CMIP5 data were also addressed during the pre-processing. For instance, the soil temperatures of one model were found to be over-estimated by 273.15 degrees (!). The combination of all the issues mentioned above, makes the pre-processing of the CMIP5 data a lengthy and inefficient task, since every model requires a different routine.

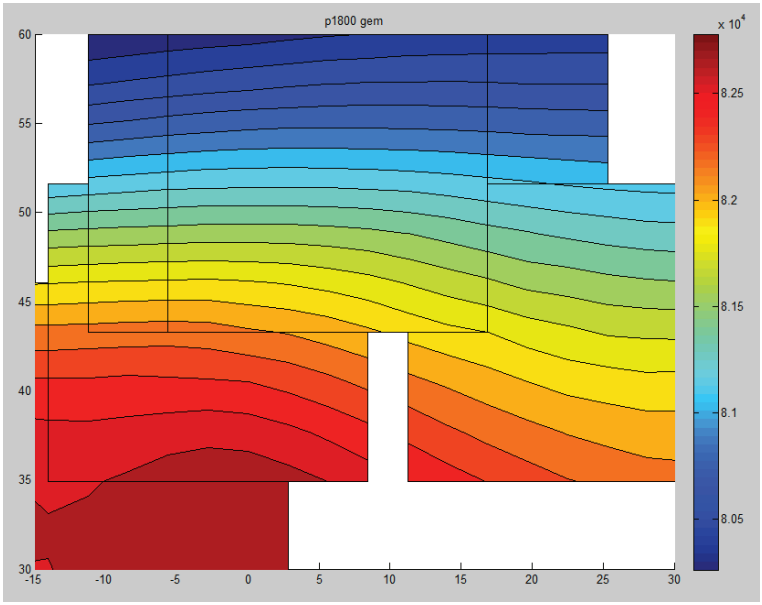


Figure 24. Pressure at 1800m calculated using the corrected height file for the BNU-ESM model.

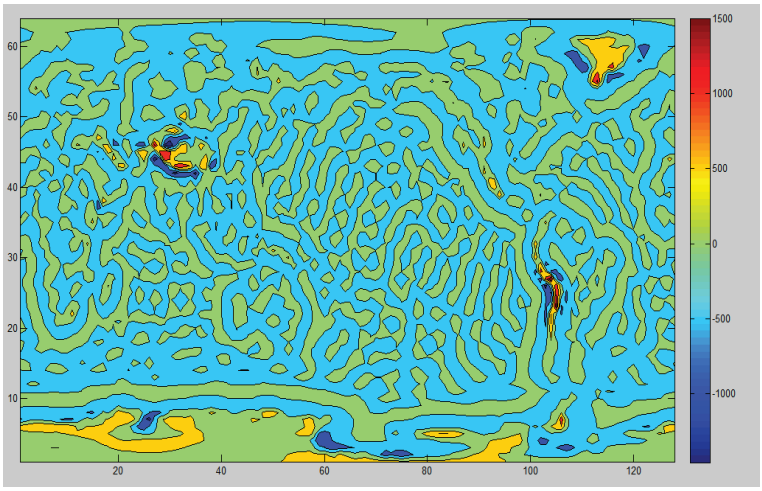


Figure 25. Overall difference between the original height and the corrected height for the BNU-ESM model

(v) CPU usage and storage

The urban climate projection simulations represent a major computational challenge. For every city under consideration, 660 model runs have to be completed (11 (models) × 3 (periods) × 20 (years) = 660 (runs)), not including the reference

ERA-Interim runs (20 in total). The duration of a single run is dependent on many factors, including the number of grid points, the climate of the city and the local terrain. Since the duration of a single runs scales more or less cubically with the number of grid-points (considering a fixed domain size), the simulations become very slow for the larger domains under consideration (Berlin and London), but also the climate of a given area, and in particular high wind speeds (requiring shorter time steps) can increase the simulation time substantially (e.g., Bilbao). Table 4 summarizes the CPU time required for all simulations, the total amounting to around **45 CPU years**. Fortunately, for these simulations, we had access to a multi-processor computer cluster (with up to 200 CPUs simultaneously, thus resulting in approximately 100 continuous simulation days. *Even though this may appear rather long, it is the merit of UrbClim, having been designed to run fast, that allows to conduct such simulations at all.*

	time for one run (h)	Total run time (y)	UrbClim data volume (Gb)	GCM data volume (Gb)
Almada	*50*	*3.88*	*400*	*110*
Antwerp	50	3.88	400	110
Berlin	*185*	*14.36*	*800*	*110*
Bilbao	65	5.05	325	110
Hyderabad	35	2.72	200	150
London	75	5.82	2400	110
New York	25	1.94	350	110
Rio de Janeiro	45	3.49	450	200
Skopje	50	3.88	325	110
Total		**45.02**	**5650**	**1120**

Table 4. Overview of the required CPU and storage resources used in the simulation of the urban climate projections.

Regarding the storage requirements: UrbClim's primary output consists by default of the air temperature and the humidity at two metres above the surface, for all grid points in the simulated domain. In addition, also wind speed and land surface temperature can be stored, and a few other variables. In the raw output files, all data is provided on an hourly basis. The size of these output files depends on the

grid used in the simulations, but for a typical grid with 121 × 121 points, the size of the output for one summer is around 250 Mb per surface variable saved. If both the temperature and the humidity are stored for a single summer, the total output for a run on a grid with 121 × 121 points comprises a little more than 500 Mb. As mentioned above, the climate projections require 11 (models) x 3 (periods) x 20 (years) = 660 runs per city. Hence, in total, around 400 Gb of output data is produced which, including the CMIP5 input data augments, leads to approximately 510 Gb of data for one city. Within the study at hand, nine cities are considered; the storage demand of all the cities combined amounts to approximately **6.7 Tb of data stored**, as shown in Table 4.

(vi) Bogotá

We encountered several issues related to the modelling of the urban climate in Bogotá. A first, data-related problem concerns the availability of land cover data for this city. As mentioned before, the GlobCover data-set, which is being used for non-European cities, has major quality issues for the North-Western part of South-America. In particular Bogotá is completely omitted in the GlobCover data, see Figure 5.

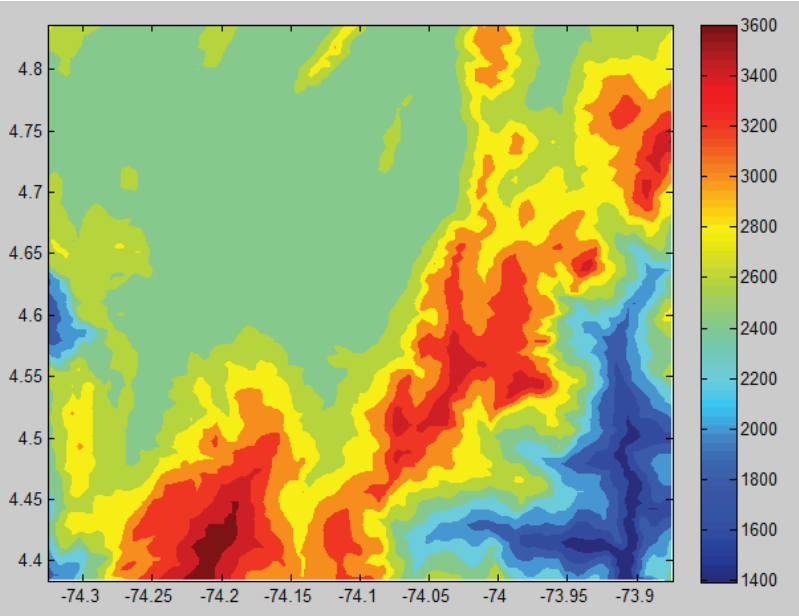


Figure 26. Terrain height (in metres) in the 50km × 50km domain around Bogota.

In addition, Bogotá is located in an area characterized by extreme topographical variation, as can be seen from Figure 26. While the model has been shown to be able to handle complex terrain, including steep topography (e.g., for Bilbao, see De Ridder et al, 2015), the problem of reconstructing upper-air pressure fields (see

above for details) apparently hits the ceiling in terms of the uncertainty it generates on these fields. Since UrbClim, as any atmospheric model, is extremely sensitive to (even slight) horizontal pressure disturbances (see, e.g., Pielke, 2002), we found that no stable model run could be achieved for Bogotá. While the model did not crash, the results it produced for this city were found to be prone to spurious features, including unexpected warm or cool air 'bubbles' entering the domain, most likely related to a lack of equilibrium between the pressure and the momentum fields.

On the other hand, an inspection of the climate features of Bogotá has shown that most likely, urban heat stress is not the first concern for this region. Therefore, we have decided to omit Bogotá from the simulations with UrbClim. While it is of course regrettable that we couldn't apply the model on this city, it would take a major effort to remedy the problem related to the uncertainty on the GCM pressure fields, which we believe is beyond the scope of the RAMSES project, and would at least put the simulations for the other cities at risk.

3.6 Urban climate projection results

In the previous section, the methodology for generating the urban climate projections was introduced. Here, we present the main results and discuss the future urban climate more in detail. The focus lies on the two UHI-indicators introduced before: the temperature indicator, which yields the minimal temperature during the five percent warmest summer nights, and the heat-wave day indicator, giving the number of heat wave days occurring in a single summer. Below, by way of demonstration we first present results for Skopje. Subsequently we focus on the comparison between the nine cities under investigation.

As mentioned before, in this study UrbClim was driven by meteorological input of 11 GCMs, hence the final result will consist of an ensemble mean for each of the considered UHI indicators, together with an uncertainty range. (Of course, proper attention has to be given to the communication of this uncertainty towards users of this information.) Nevertheless, this uncertainty is an important piece of information. Also, environmental and planning professionals in cities generally know very well how to deal with uncertainty, as it is an inherent aspect of their work (at least so they tell us). In the remainder of this text, we will therefore consider ensemble means and variances over the 11 GCM forcings under consideration, although some plots show the results of individual GCMs for guidance (and for visual aspects). Uncertainties are indicated using the 'likely range' defined in the IPCC AR5 report.

This likely range constitutes a 5% to 95% range (± 1.64 standard deviation) across the ensemble distribution.

Table 5 summarizes the mean temperature change between the reference period (1986-2005) and the end of the century (2081-2100), based on the 11 GCMs considered here. Note that the results of this table are in accord with the results of the most recent IPCC report (AR5). Whereas the warming in Rio and Hyderabad appears to be less than the expected average (regional-scale) warming, Skopje experiences, as most of the Mediterranean region, a very considerable temperature increase.

City	Mean temperature change (° C)
Almada	*4.0*
Antwerp	3.9
Berlin	*4.3*
Bilbao	4.2
Hyderabad	3.4
London	3.8
New York	5.1
Rio de Janeiro	3.3
Skopje	5.9

Table 5. Mean temperature change between the reference period (1986–2005) and the end of the century (2081–2100), based on the 11 GCMs considered here, for all cities.

In the next paragraphs, the results for Skopje are discussed in detail: we show the evolution of the heat wave day and the temperature UHI indicator, and compare maps of the reference period with maps of the end-of-the-century scenario (RCP8.5) for both indicators. While Skopje may perhaps not be representative for the group of cities considered, it is of special interest since it is the subject of an intense collaboration with WP 6 (Health, collaboration with Gerardo Sanchez of WHO, and with involvement of the Skopje City Authorities).

3.6.1 Skopje

The mean temperature and number of heat wave days in Skopje are greatly influenced by climate change (Figure 27). Both the ensemble mean rural and urban

95th percentile of the minimal temperatures increase by approximately 7 °C. These temperature changes are in agreement with results presented in the appendices (regional climate atlases) of the IPCC AR5 report.

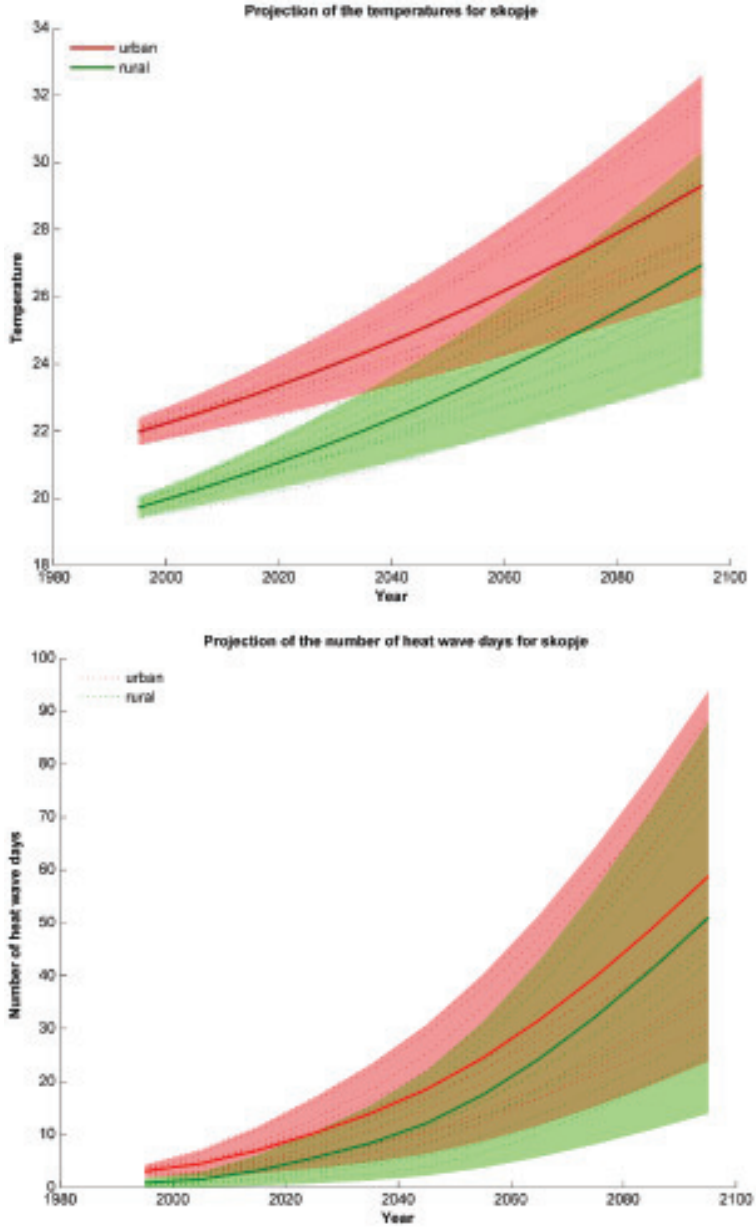


Figure 27. Evolution of the rural (green) and urban (red) temperature indicator in Skopje (upper panel) and the number of heat wave days (lower panel). The thick lines shows the ensemble mean, while the shaded areas indicate the 5% to 95% range based on assuming a normal distribution for the ensemble. Dotted lines indicate the results obtained with the individual GCM forcings.

Due to the large spread in the GCM results, also the uncertainty on the urban and rural temperatures is rather large, which implies that the lower error margin for the urban temperature is lower than the mean for the rural temperature, and that the upper error margin for the rural temperature is higher than the mean for the urban temperature. Also the number of heat-wave days increases considerably (Figure 27)

increases by a staggering number of days. The uncertainty spread on the number of heat wave days is also significantly larger than the spread on the temperatures, which is because the former quantity is very sensible to a small increase in temperature.

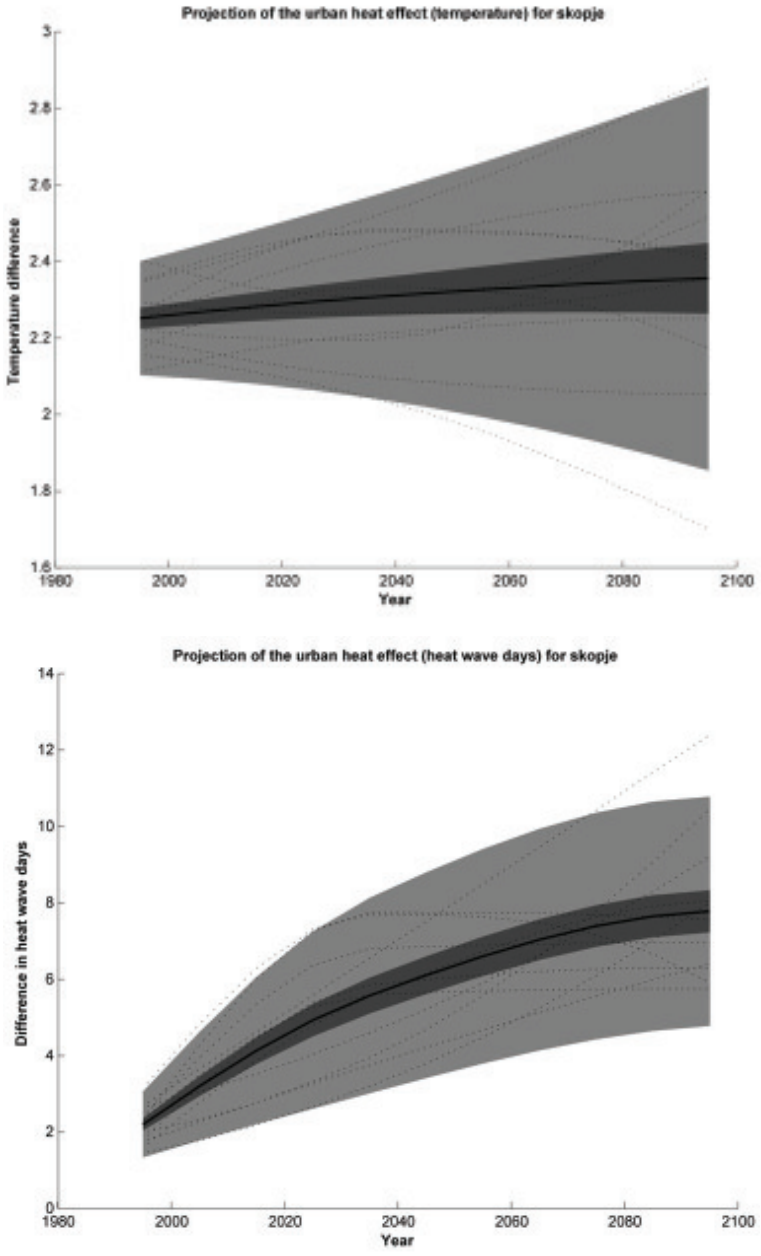


Figure 28. Upper panel: evolution of the temperature UHI indicator in Skopje (UHI intensity, i.e., average urban-rural temperature difference). Lower panel: the evolution of urban-rural difference in the number of heat wave days. The different types of lines and shadings are as in Figure 27.

Figure 28 shows the evolution of the urban-rural differences of the temperature and heat wave day indicators. A striking feature in this figure is that the UHI intensity barely changes during the considered period. Below, we will discuss possible reasons for this phenomenon. Conversely, the difference in the number of

heat wave days between Skopje and its rural surroundings does increase, despite the fact that the UHI intensity barely changes. This is related to the use of thresholds in the calculation of the number of heat wave days.

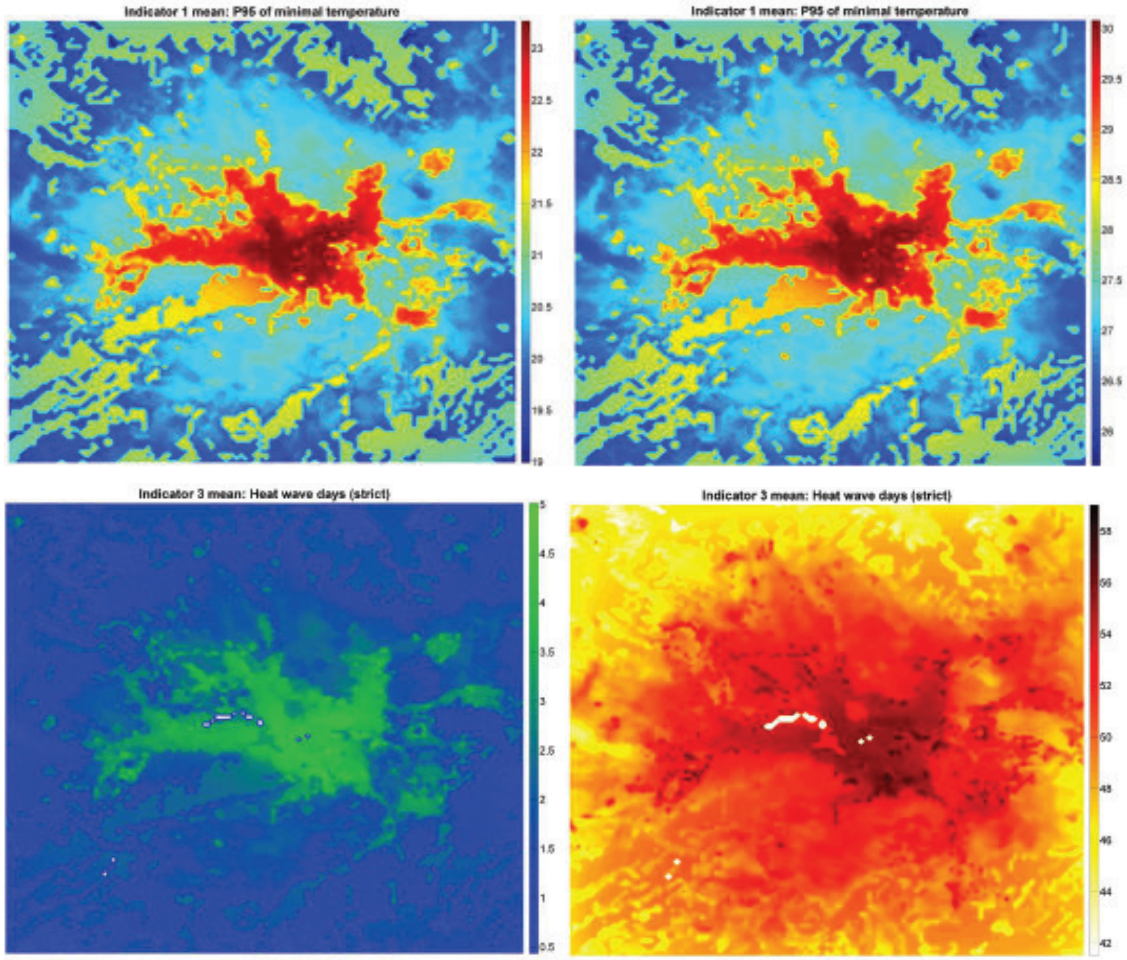


Figure 29. Temperature (upper panel) and heat wave day (lower panel) indicators for the reference period (1986-2005) (left) and for the end-of-century (2081-2100) RCP8.5 scenario (right). For the upper panel, notice the different colour legends.

The same quantities as those just discussed are now shown under the form of maps (Figure 29). Here again it is shown that, while the urban and the rural temperature increase considerably between the current-day situation and the end-of-the-century scenario, there is no appreciable difference in the spatial patterns between the map of the temperature indicator for the reference period and the temperature indicator for the future period. The mean increase in the UHI-effect is 0.10 ± 0.09 °C, and the overall temperature increase is thus much larger than the increase in the temperature UHI effect. Hence, we conclude that the temperature increase of the background dominates the urban and rural temperature changes.

On the other hand, there is a large increase in the difference between the number of heat wave days in the city and the surrounding countryside. During the reference

period, there are 2.2 annual heat-wave days more in the city center than in the rural surroundings, but by 2081-2100, this difference increases to 7.5 days. Although the temperature difference between the urban and the rural part of the domain is negligible in comparison with the change in urban and rural temperatures, the difference in the number of heat-wave days does increase significantly. This is because of the sensibility of the calculation of the number of annual heat-wave days to small increases in the temperature, and also because these calculations take both the minimum and the maximum temperatures into account.

Hence, it is fair to conclude that the number of annual heat wave days increases much more rapidly within the city of Skopje than in the nearby rural areas.

3.6.2 Comparison of results obtained for the different cities

In the previous section, we highlighted the results for a single city, Skopje. In this section, we shall see that the main results concerning the change in the UHI strength are easily transferred to the other cities in the study. Note that, as discussed above, the results for Hyderabad and Bogotá are omitted from this analysis.

In the remainder of this section, we focus on the change in both UHI-indicators between the reference period (1986-2005) and the end of the current century (2081-2100) for the other eight cities under consideration. The results for the change in the 95th percentile of the minimum rural and urban temperature and the change in temperature UHI-indicator are summarized in Table 6. This table also gives the standard error on the change in the UHI-indicator, as calculated using the standard deviation of the UHI-indicator for the reference period and the end-of-the-century and the standard rules for the propagation of uncertainty.

For all cities except Rio de Janeiro, the UHI-indicator increases slightly, although the increase often is below the error margin. For Rio de Janeiro, the temperature UHI-indicator decreases, but the magnitude of the decrease is again negligible in comparison with the increase of the urban and rural temperatures. As was the case for Skopje, the change in the UHI temperature indicator is much smaller than the change in the urban and rural temperatures themselves.

City	Absolute change in 95th percentiles of the minimum temperatures			
	Rural (°C)	Urban (°C)	ΔUHI (°C)	Error (°C)
Almada	*4.51*	*4.58*	*0.07*	*0.07*
Antwerp	4.12	4.20	0.08	0.07
Berlin	*4.53*	*4.64*	*0.11*	*0.09*
Bilbao	4.56	4.63	0.08	0.06
London	4.31	4.32	0.01	0.08
New York	4.87	5.08	0.21	0.12
Rio	3.30	3.18	-0.12	0.10
Skopje	6.74	6.85	0.10	0.09

Table 6. Change in the temperature UHI indicator between the reference period (1986-2005) and the end of the century future (2081-2100) under RCP8.5 conditions. The first two columns indicate the increase in the 95th percentile of the urban and rural minimal temperatures, while the third column indicates the change in the UHI temperature indicator (with a positive sign indicating an increase in the temperature UHI strength). The rightmost column gives the error on the change in the UHI temperature indicator, calculated as the ensemble standard deviation divided by the square root of the number of ensemble members (which amounts to 11 in this exercise).

The distinct behavior for the different cities raises the question whether or not it is possible to predict the evolution of the temperature UHI-effect based simply on the meteorology contained in the GCM output. A multiple linear regression was established in order to test the hypothesis that the change in the temperature UHI-indicator can be explained using the change in meteorological parameters in the GCMs. As explanatory variables, the *changes* in the mean near-surface temperature and wind speed, and in the total precipitation, are used. Note that the explanatory variables are always taken from the ensemble means over the 11 GCMs, and that we consider the difference of the mean meteorological quantities for the end-of-century scenario (2081-2100) and the reference period (1986 – 2005). The regression is thus

$$\Delta UHI_T = a_0 + b_0\Delta\overline{T} + c_0\Delta\overline{w} + d_0\Delta rain \quad (4)$$

where ΔUHI_T indicates the change in UHI strength according to the temperature indicator, $\Delta\overline{T}$ and $\Delta\overline{w}$ denote the change in mean temperature and surface wind speed, and $\Delta rain$ the change in total precipitation. The regression yields $R^2 = 0.65$, thus a fairly good relation is found between the meteorology contained in the output of the GCMs and the change in the temperature UHI-indicator, which

indicates that the change in the input meteorology dominates the terrain input in the urban climate projections.

The evolution of the heat-wave indicator between the reference period (1986-2005) and the end-of-the-century scenario (2081-2100) for all cities under investigation is summarized in Table 7. The table shows the relative increase in the heat-wave UHI-effect, where the relative increase is defined as the absolute increase divided by the magnitude of the UHI-effect in the reference period. Note that for the heat-wave day UHI-indicator, it only makes sense to compare relative increases, since the absolute increases are influenced excessively by the number of heat-wave days in the reference period. Once more, the results for Skopje can be readily transferred to the other cities. Indeed, although the change in temperature UHI-effect is negligible in comparison with the change in the urban and rural temperatures, the increase in the heat-wave day UHI-indicator is significant for all cities. Therefore, the number of heat-wave days within the city limits increases more rapidly than the number of heat-wave days in the rural surroundings, and this for all cities under investigation.

City	Relative change in the heat-wave day UHI-indicator	
	UHI (°C)	Error (°C)
Almada	3.15	0.25
Antwerp	3.00	0.54
Berlin	3.27	0.70
Bilbao	3.36	0.63
London	3.49	0.49
New York	5.13	0.37
Rio	2.76	0.33
Skopje	2.51	0.32

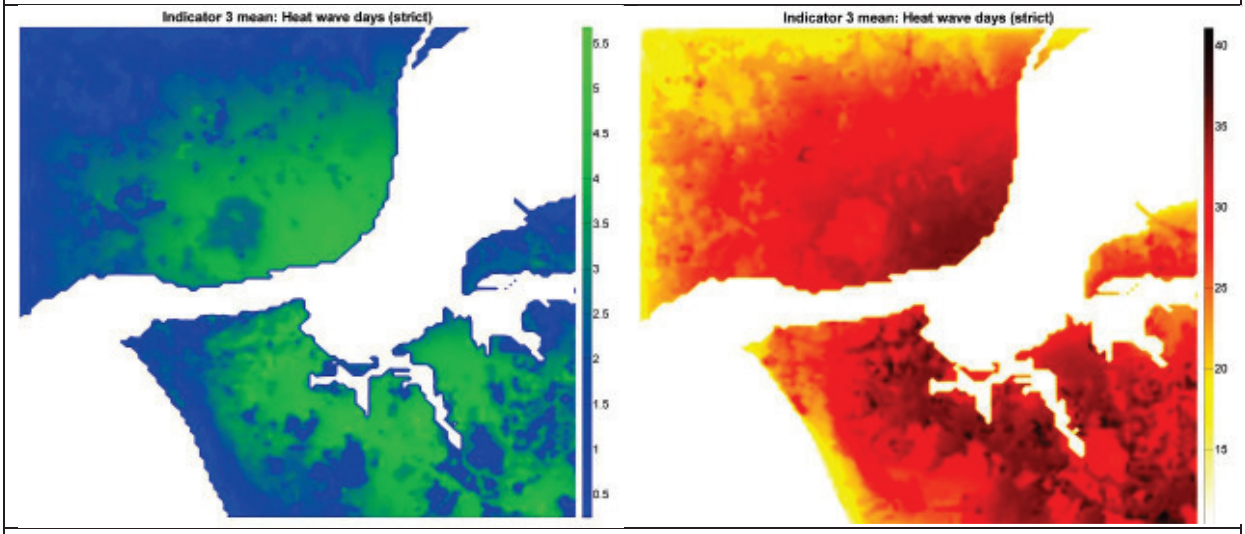
Table 7. Change in the heat-wave day UHI-indicator between the reference period (1986-2005) and the end of the century (2081-2100). The first column indicates the relative change in the heat-wave UHI indicator, while the second column indicates the error on this quantity.

Neglecting Skopje for the moment, the order of the cities for the change in the heat wave days UHI indicator and the change in the temperature UHI indicator is more or less the same. The greatest increase in the UHI effect is in both cases observed for New York, while (neglecting Skopje) Rio de Janeiro has the smallest increase in heat wave UHI effect, whereas it was also the only city with a decreasing temperature UHI effect. The other three cities are situated somewhere in between.

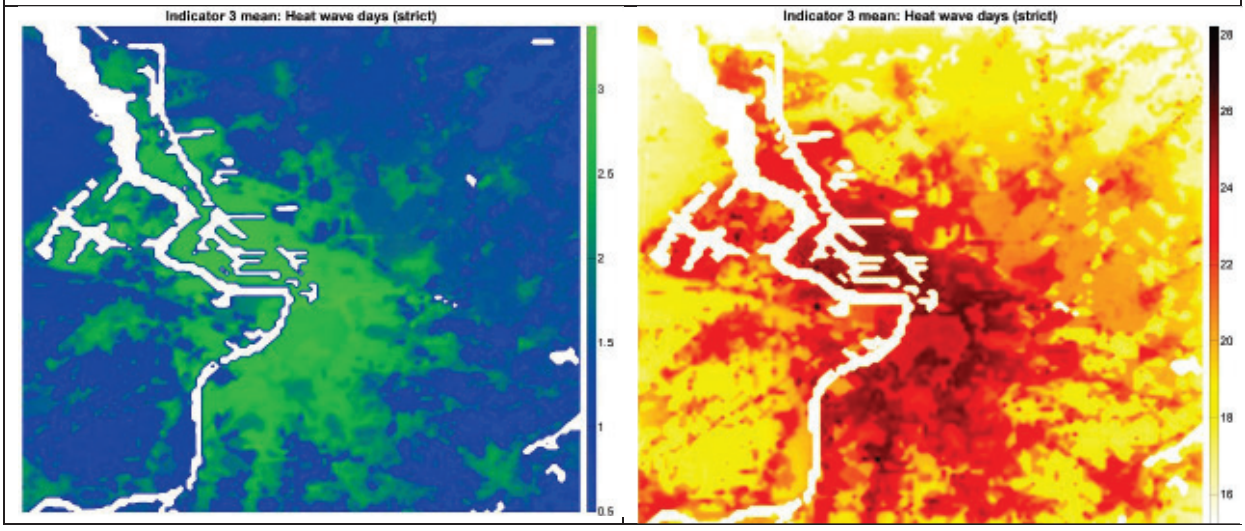
Only Skopje does not appear to follow this pattern. However, Skopje is slightly peculiar in our selection of cities, since its warming is extreme and consequently almost every day becomes a heat wave day in the city. Therefore, the number of heat-wave days in the city center reaches an artificial limit, which causes a slowed down increase in the heat wave UHI indicator.

Below, we show maps containing the number of heat wave days for current (1986-2005) and end-of-century (2081-2100, RCP8.5) conditions, for all the cities except Skopje (which is available already in Figure 29). Overall, it is found that the number of heat wave days increases by up to a factor ten in the considered urban areas, and this happens in a rather consistent way across all cities.

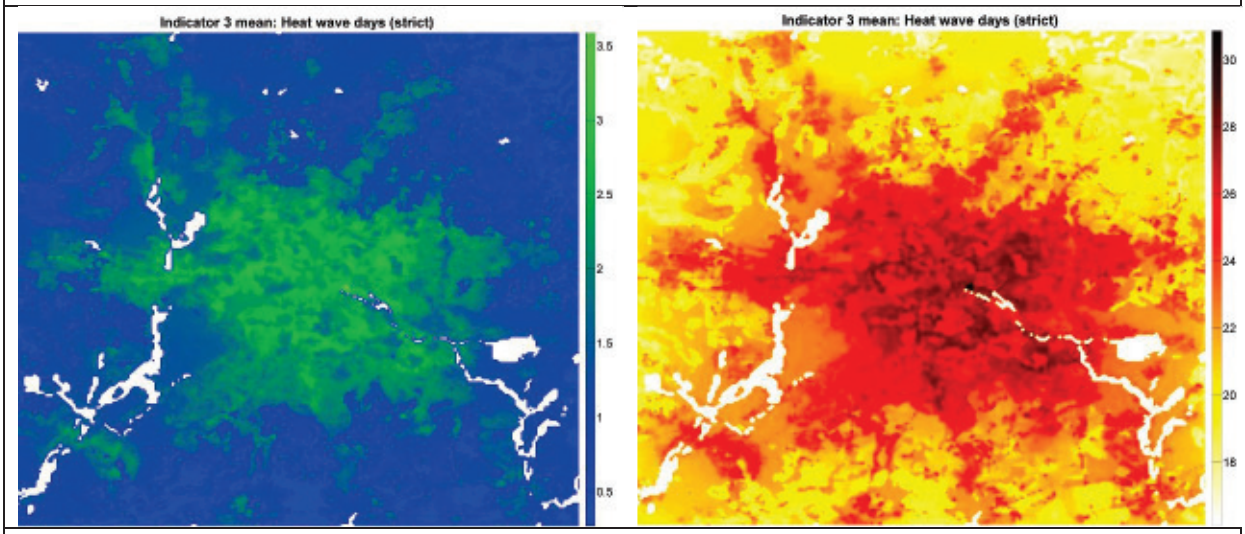
Almada



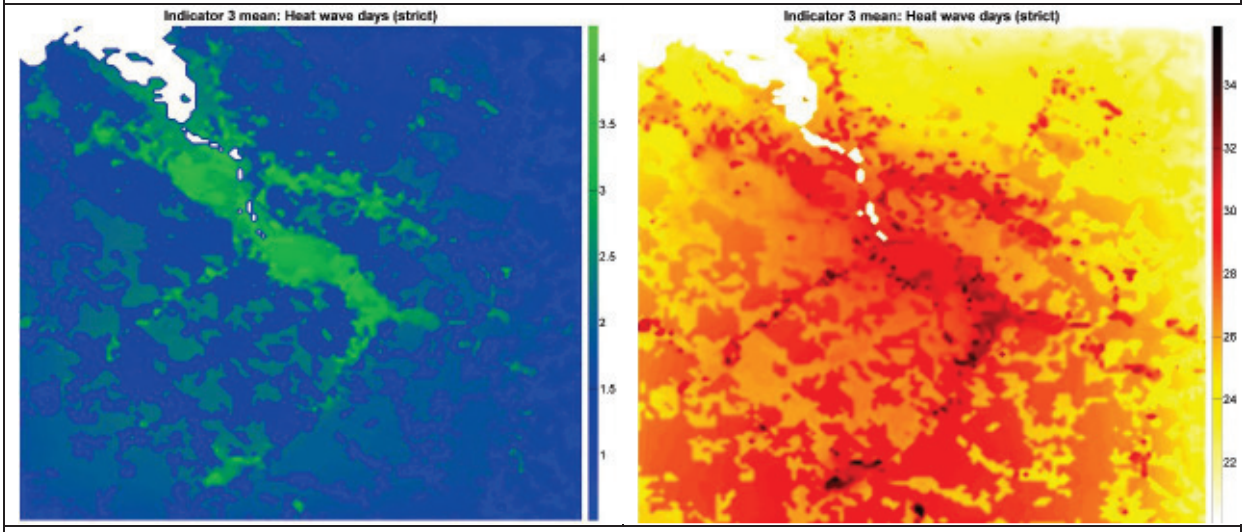
Antwerp



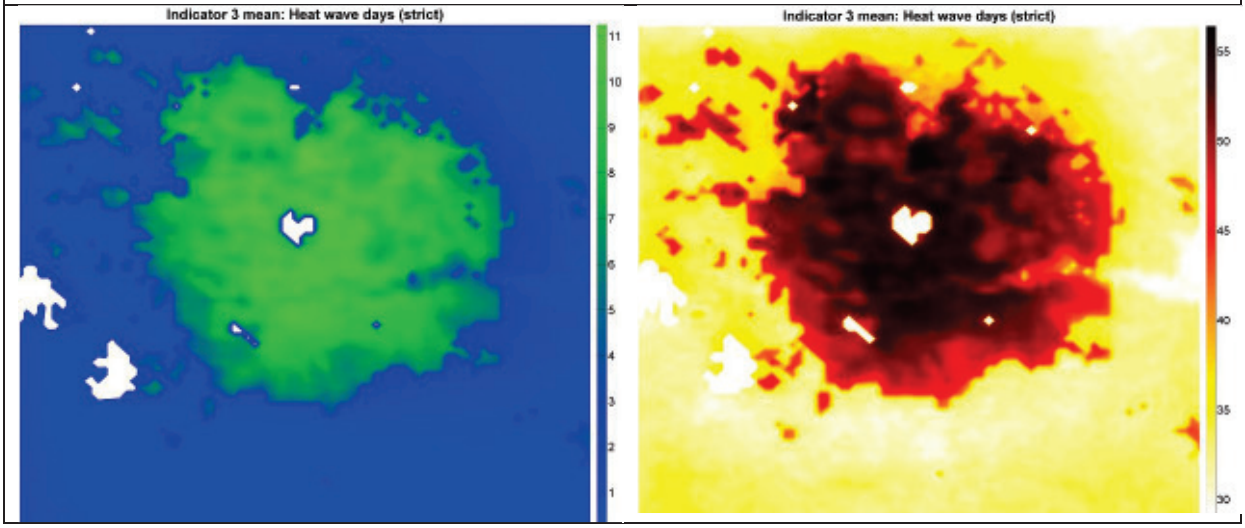
Berlin



Bilbao



Hyderabad



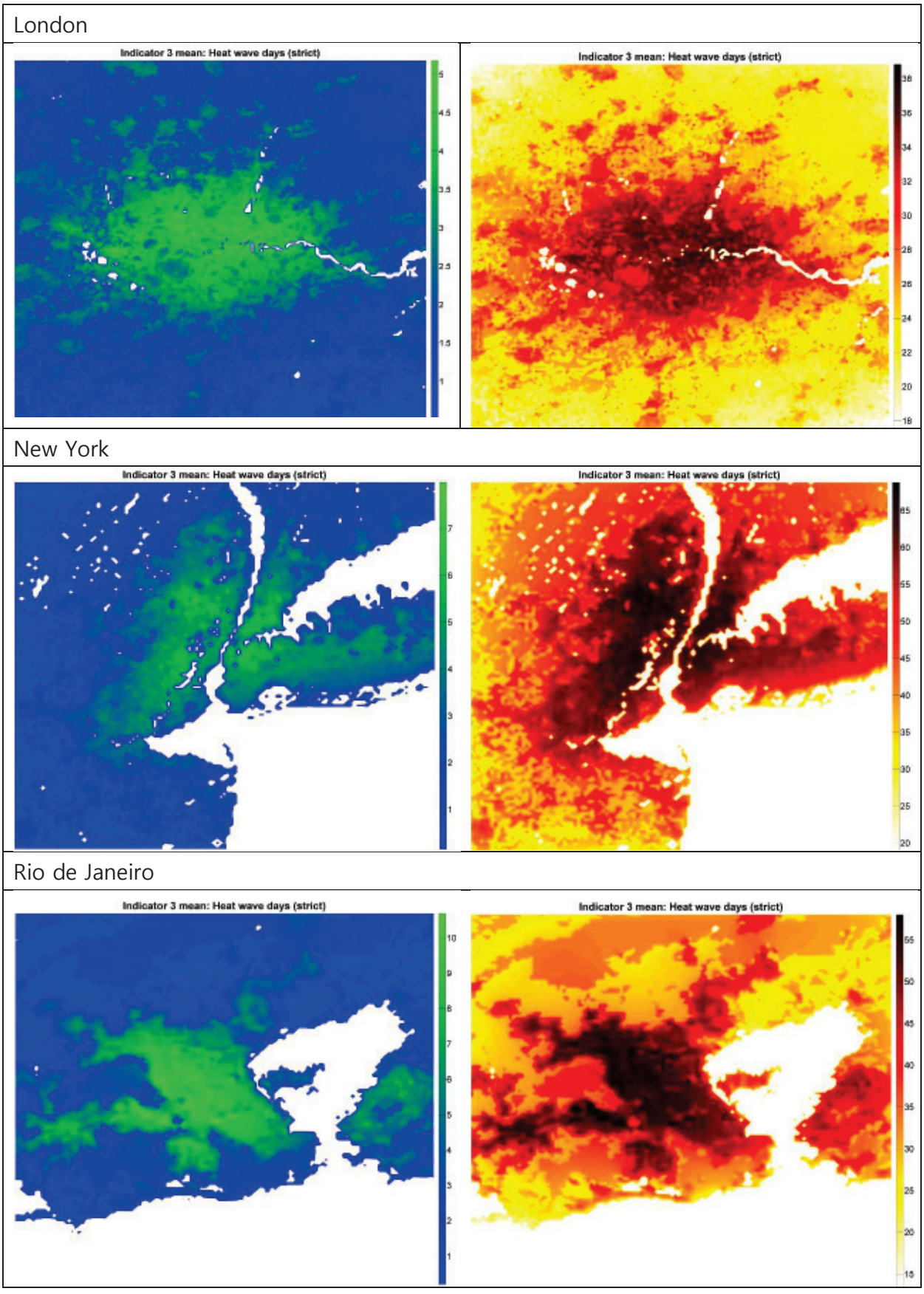


Figure 30. Number of heat wave days for (left) current (1986-2005) and (right) end-of-century (2081-2100) RCP8.5 conditions, for selected cities.

3.6.3 Discussion regarding the UHI intensity changes

A striking aspect of the results presented above is that, while overall temperatures increase quite a lot in the RCP8.5 scenario, the UHI intensity does not appear to change significantly with climate change. Initially, we found this a puzzling result. Indeed, as mentioned at the start of Section 3, given (1) that GCMs predict an increase in the intensity, duration, and frequency of heat waves, and (2) that UHI intensity increases during heat waves, it would appear straightforward to assume an increase also in the average UHI intensity.

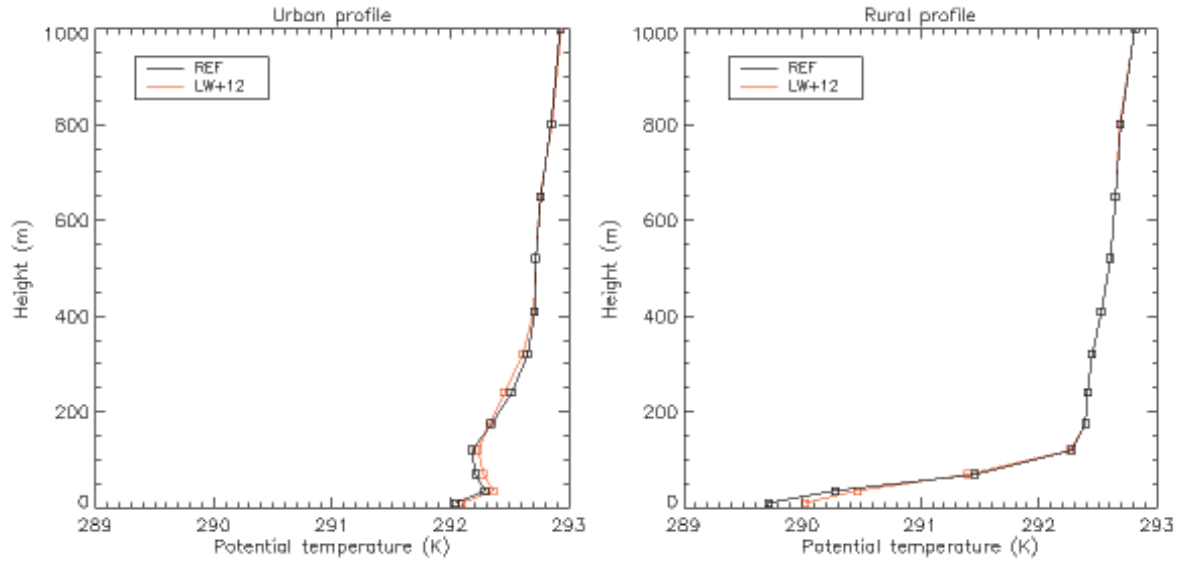


Figure 31. Effect of enhanced downwelling longwave radiation on the temperature profile over the city of Brussels (left) and over a nearby rural area (right). The black line corresponds to the reference simulation result, the red line shows the result of the simulation with enhanced longwave radiation.

Yet, as shown in the previous section, this increase of the UHI strength does not take place, as observed rather consistently among all cities considered. We hypothesised as a possible explanation a mechanism involving downwelling longwave radiation. Indeed, with an increasing amount of greenhouse gases in the atmosphere, this downwelling longwave radiation is also projected to increase. This radiation increase also normally increases near-surface air temperature. However, as demonstrated by recent work we did on this topic, it turns out that the night-time temperature increase following a downwelling radiation increase behaves differently over urban and rural areas. (We consider the night mainly because then the UHI effect is at its strongest.) In fact, in a sensitivity study done for the Brussels area (Figure 31), Lauwaet et al. (2015), demonstrated that the increase of air temperature near the ground following enhanced radiation is much more effective over rural areas than over urban areas. In other words, the extra downwelling longwave

radiation (of the order of 12 W m^{-2}) that results from the enhanced amount of greenhouse gases typically yields a higher temperature increase over rural areas than over a city. The reason for this behaviour is that, at night-time, vertical temperature profiles in rural areas are generally characterized by a more or less strong inversion, the stability of which inhibits much vertical mixing. The extra heat at the surface, caused by the enhanced longwave radiation absorption, is therefore distributed over a relatively small vertical extent, thus resulting in a rather strong increase in temperature. Over urban areas, which have a rather neutral or even slightly positively buoyant temperature profile, the extra heat is distributed over a much larger vertical domain, i.e., the heating is much 'diluted', so to say, yielding a much lower temperature increase.

3.7 Connections to other work packages

While working on the development of the heat indicators, and on the definition of the periods, scenarios, and urban domains to be covered, a very intense collaboration has been established with WP6 ('City specific health impacts of climate change, damage and adaptation costs'), which is led by partner WHO. In particular, many discussions were held regarding the simulations for Skopje. This is one of the RAMSES target cities, and one for which the local administration carries a particular interest for the problem of heat stress. Perhaps a bit coincidentally, Skopje turned out to be one of the most affected cities in terms of excessive (expected future) heat stress.

In addition, more recently we started an initiative involving VITO and LSE/Seneca, on estimating the costs of urban heat stress related to a reduced labour efficiency. This will be based on empirical relations between labour efficiency and the wet bulb globe temperature (WBGT), which is a measure for heat stress. From the urban climate modelling point of view, one of the main challenges is to estimate indoor WBGT values based on the outdoor variables simulated by UrbClim. Some preliminary screening of the literature has revealed that there might be approaches allowing to make such a connection (Lemke and Kjellstrom, 2012; Nguyen et al., 2014; Smargiassia et al., 2008).

Nevertheless, a more detailed and in-depth assessment of such techniques is required to evaluate their usefulness for our purposes. This work has been planned for early 2015.

4 Top-down exercise: '102-cities' UHI simulation

As mentioned before, one of the main strengths of UrbClim is its capability to be configured very flexibly (for almost any city in the World) and also to run very fast. In the previous section, this property was taken advantage of in the simulation of the many years needed to establish climate statistics.

This section describes another way of taking advantage of UrbClim's flexibility and running speed, by conducting a simulation for a large number of cities. This work is to be seen as a part of the RAMSES endeavour to tackle urban climate adaptation also through a top-down approach, in addition to the bottom-up (detailed case studies) work we already do. Forcibly, running the model on 100(+) cities does not yield as high a level of detail as what we obtained for the detailed case studies of the previous section. Conversely, the larger sample available in the top-down approach has the potential of producing generic insights.

Amsterdam	Budapest	Graz	Lodz	Nottingham	Skopje
Antwerp	Cardiff	Hamburg	London	Nuremberg	Sofia
Arlon	Charleroi	Hannover	Luxembourg	Oslo	Stockholm
Barcelona	Den_Haag	Hasselt	Lyon	Palermo	Strasbourg
Basel	Dortmund	Helsinki	Maastricht	Paris	Stuttgart
Belgrado	Dresden	Katowice	Madrid	Podgorica	Tirana
Berlin	Dublin	Koln	Malaga	Porto	Torino
Bern	Dusseldorf	Kopenhagen	Manchester	Portsmouth	Utrecht
Bilbao	Eindhoven	Krakow	Marseille	Poznan	Valencia
Birmingham	Essen	Leeds	Milan	Prague	Vienna
Bordeaux	Firenze	Leipzig	Mol	Riga	Vilnius
Bratislava	Frankfurt	Leuven	Mons	Rome	Warschau
Bremen	Geneva	Liege	Munchen	Rotterdam	Wavre
Bristol	Genua	Lille	Namur	Saarbrucken	Wroclaw
Bruges	Ghent	Lissabon	Napoli	Sarajevo	Zagreb
Brussels	Glasgow	Liverpool	Newcastle	Sevilla	Zaragoza
Bucharest	Goteborg	Ljubljana	Nice	Sheffield	Zurich

Table 8. List of the 102 cities that were considered in the exercise.

For the 102 cities listed in Table 8, we simulated the UHI effect for the summer (May-September) of 2009 using UrbClim, at 250 m resolution. It took less than one

year of CPU time, i.e., this is the total CPU time, and since we distribute this simulation over a large number – typically ~ 100 – of processors, the total wall-clock time was much less, totalling a few days only) to complete this task. The meteorological input for this study was taken from the large-scale (70 km resolution) ERA-interim re-analysis data set of the European Centre for Medium-Range Weather Forecasts (ECMWF). The terrain input consisted of the spatial distribution of land use types, the degree of soil sealing, the vegetation cover fraction with a spatial resolution of 250m and detailed elevation data. These quantities were all taken from publicly available data sets, specifically, the 2006 CORINE land cover data for Europe, the European Environment Agency soil sealing data, the Normalized Difference Vegetation Index acquired by the MODIS instrument on the TERRA satellite, and the Global Multi-resolution Terrain Elevation Data (GMTED) of the U.S. Geological Survey (USGS), respectively.

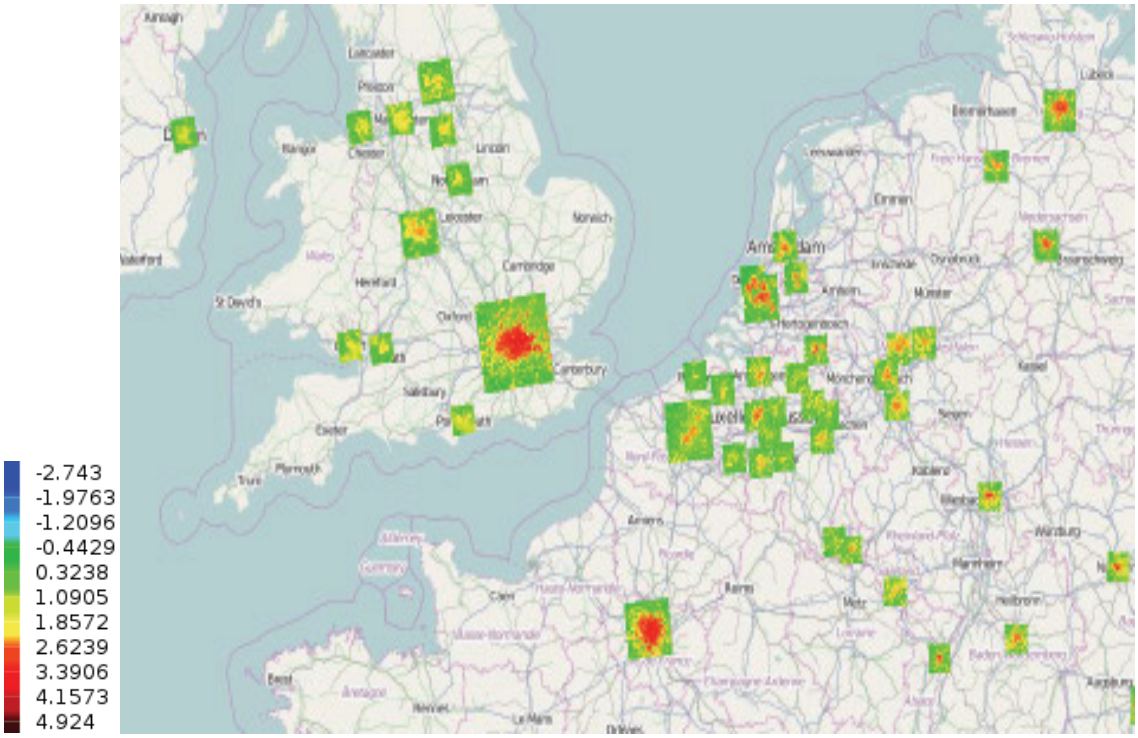


Figure 32. Map with a subset of the results of UrbClim for 102 cities (see www.urban-climate.eu/services/eu_cities for the map covering the whole of Europe).

The UHI effect for the considered cities is visualised using the temperature indicator, i.e. the 95th percentile of the minimal temperature during the period May – September 2009. Figure 32 and Figure 33 show some results. It is noticeable that there is a large range of different UHI intensities found across the city sample considered. In order to try and explain these differences, we did a simple regression analysis. As independent variables, we considered (1) population density, and (2)

wind speed, as in the past these quantities have been empirically related to UHI intensity (Oke, 1982).

More concretely, the average UHI strength of each city was related to the logarithm of the number of inhabitants in a radius of 20 km around the city centre and the mean observed wind speed. The former was taken from the Population Density of Europe (Gallego, 2010), and the latter as taken from the ERA-Interim archive of ECMWF. The result of a simple multi-linear regression is visualised in Figure 34. The coefficient of determination R^2 = 0.65 indicates the presence of a good relationship between the UHI intensity and the population density and the mean wind speed, although a more detailed analysis should include other variables.

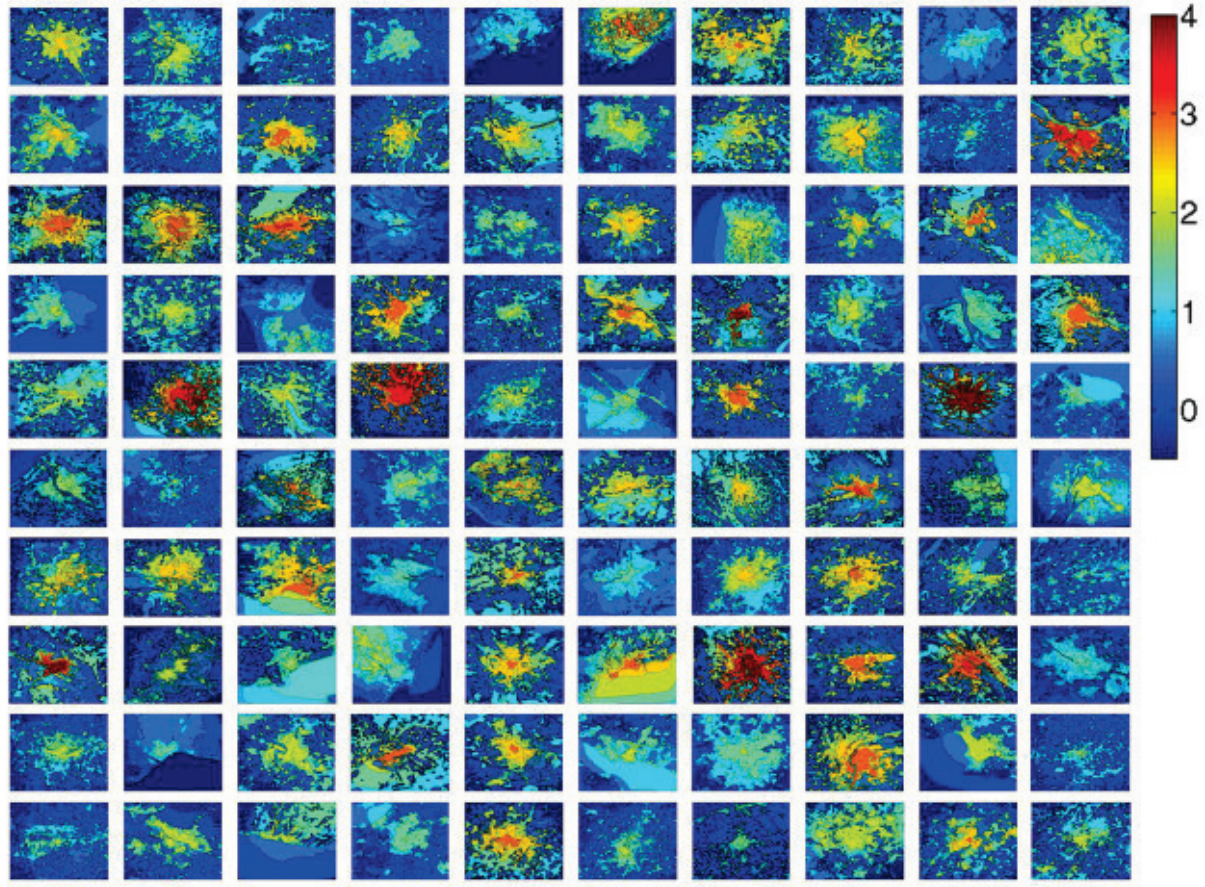


Figure 33. Maps containing the mean summer UHI effect for the considered cities, for the year 2009.

The main conclusion from this exercise, which we believe of relevance when considering adaptation measures related to urban heat stress, is that (1) larger cities tend to have a stronger UHI intensity, and (2) lower ambient wind conditions also lead to a stronger UHI. Now, while conclusion (1) may not be utterly surprising, it is also important to realize that the relation between the UHI strength and population density is logarithmic, meaning that even relatively small cities exhibit a (relatively seen) large UHI effect. Therefore, adaptation efforts should also be directed towards

relatively smaller cities and towns. Conclusion (2) is equally not really surprising, i.e., cities that are characterized by lower ambient wind speed values have a higher UHI intensity.

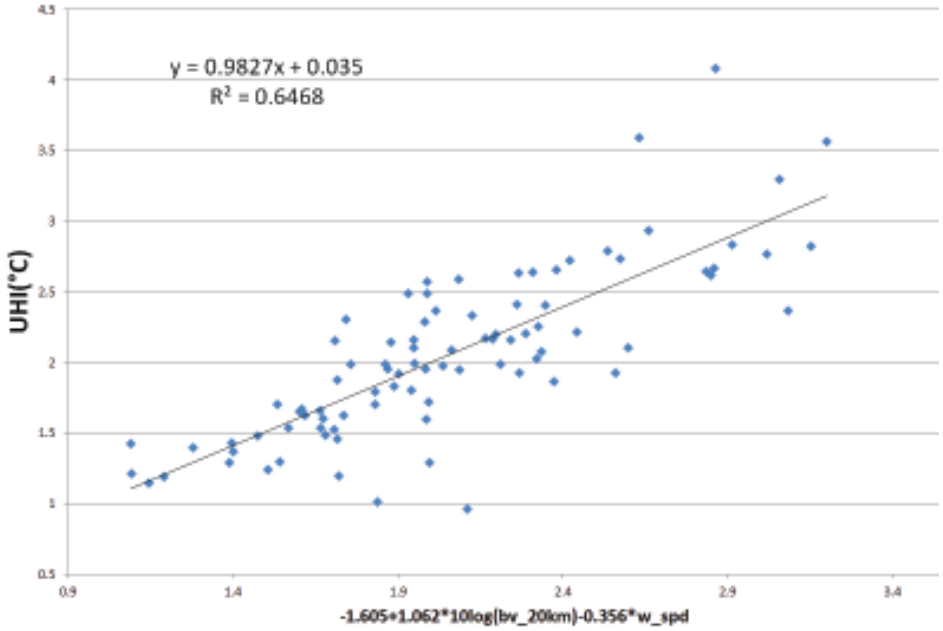


Figure 34. The average UHI intensity as a function of population density and mean wind speed.

As a final conclusion, from the sample of cities considered, it turns out that large cities located in continental climates (away from coastal areas), particularly cities in Eastern and Central Europe, are more at risk to excessive heat stress. As shown in the Introduction, these areas are also particularly prone to a rise in the background temperatures with climate change.

5 Wind climate in cities

Even though the focus of the urban climate modelling work is on heat stress and the urban heat island effect, we also investigated some aspects of wind climate in urban areas.

In the sections below, we will first describe the application of the UrbClim model to the simulation of the wind field near the ground, including the validation of simulated wind compared to observed values, and a brief analysis of the urban wind field reduction (in comparison to rural areas, that is). Subsequently, we provide a simple analysis of the urban-vs-rural wind speed, based on a simple model, and broadly confirming the ratios of the urban to rural wind speed values previously obtained with UrbClim. Finally, considering that many of the GCMs available in the CMIP5 archive contain daily mean and maximum (10-minute gust) wind speed values (rather than 3- or 6-hourly values), we have established and validated a method to construct daily wind speed probability functions based on these two quantities.

5.1 UrbClim wind speed simulations

We conducted a wind simulation and validation with UrbClim for the area of Antwerp, which previously (see deliverable D4.1) already was the domain for the 2-m temperature validation. The reason for selecting Antwerp is the availability of wind measurements (conducted by us) obtained with a sonic anemometer (see below), which is capable of performing measurements also at the very low wind speed values typically found in the urban canopy.

The city of Antwerp is located in Northern Belgium. The number of inhabitants in the city is approximately 500,000, but the wider metropolitan area counts nearly double that. Antwerp is characterized by a dense urban core surrounded by built-up areas of medium density. The area surrounding Antwerp is relatively flat, and is dominated by grassland and crop areas, and some scattered forest. The city is intersected by the Scheldt River, and has a major harbour area in the North.

For the validation of simulated wind speed, use is made of observations made at two locations. The urban site is in on top of a small building in the court yard of a school ('Koninklijk Lyceum Antwerpen') located near the city centre, the instrument

being positioned at 1.8 m above the roof. This small building being much lower (4 m above the ground) than the surrounding buildings (15 to 25 m above the ground), we consider it to be within the urban canopy layer. It must be noted that the wind measurement was set up at approximately 10 m distance from a building wall, meaning that local flow distortion may limit the representativeness of this point measurement. The rural site is located on a stretch of grassland at a farm in the nearby village of Vremde, at a distance of around 12 km from Antwerp's centre and fully within a rural zone. The nearest obstacles are trees that are located at a distance of 100 m.

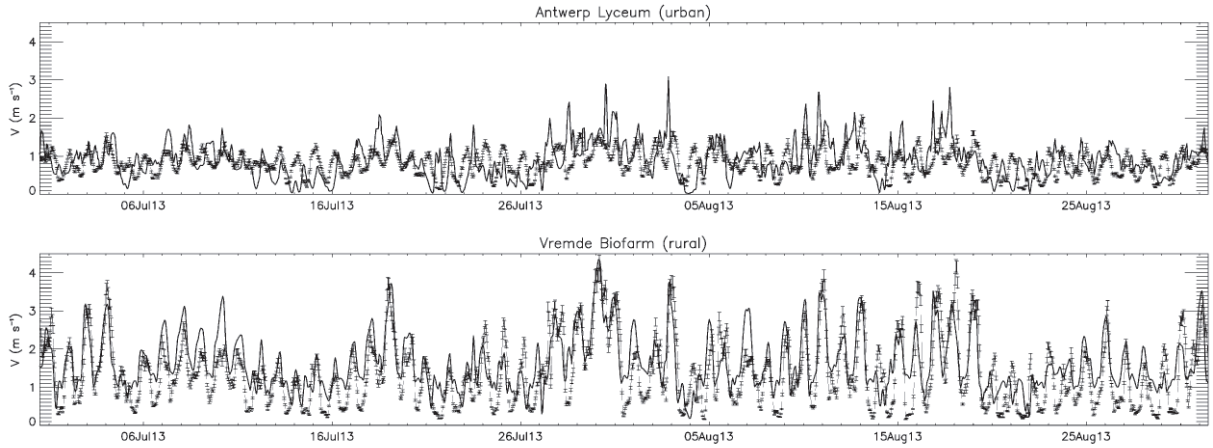


Figure 35. Simulated (solid line) and observed (error flags) wind speed for the urban and rural sites in the Antwerp domain, for the period July-August 2013.

At both locations, the instrument used to measure wind speed is a WindSonic, which is a 2-D ultrasonic anemometer manufactured by Gill Instruments. Without any moving parts, this sensor is particularly well suited for the observation of the low wind speed values that typically occur in the urban canopy. In order to determine instrumental accuracy, prior to deploying the sensors in the field, two WindSonic devices were set up a few metres distant from each other, and their measurement response was compared during an exercise lasting several days. This yielded an average between-instrument difference of around 4 % of the observed wind speed values, the latter covering the range 0.5-1.5 m s^{-1} (which is typical of the urban values observed). At higher wind speed values (of 12 m s^{-1}), the manufacturer specifies an accuracy of 2 %. Considering that we are measuring in the rather lower speed range (see Figure 35), we will adopt an accuracy of 4 %.

An UrbClim simulation was conducted for the months of July-August 2013, again using a spatial resolution of 250 m, and hourly time series of wind speed were extracted from the grid cells corresponding to the two wind measurement sites. Figure 35 shows the simulated versus the observed wind speed values for the urban

and rural sites. At the rural site, we found a bias of 0.25 m s^{-1}, a root mean square error of 0.64 m s^{-1}, and a correlation coefficient of 0.75. From Figure 35 it can also be seen that the rural wind speed simulation reproduces the diurnal cycle relatively correctly, apart from the daily wind speed minima, which are overestimated by the model.

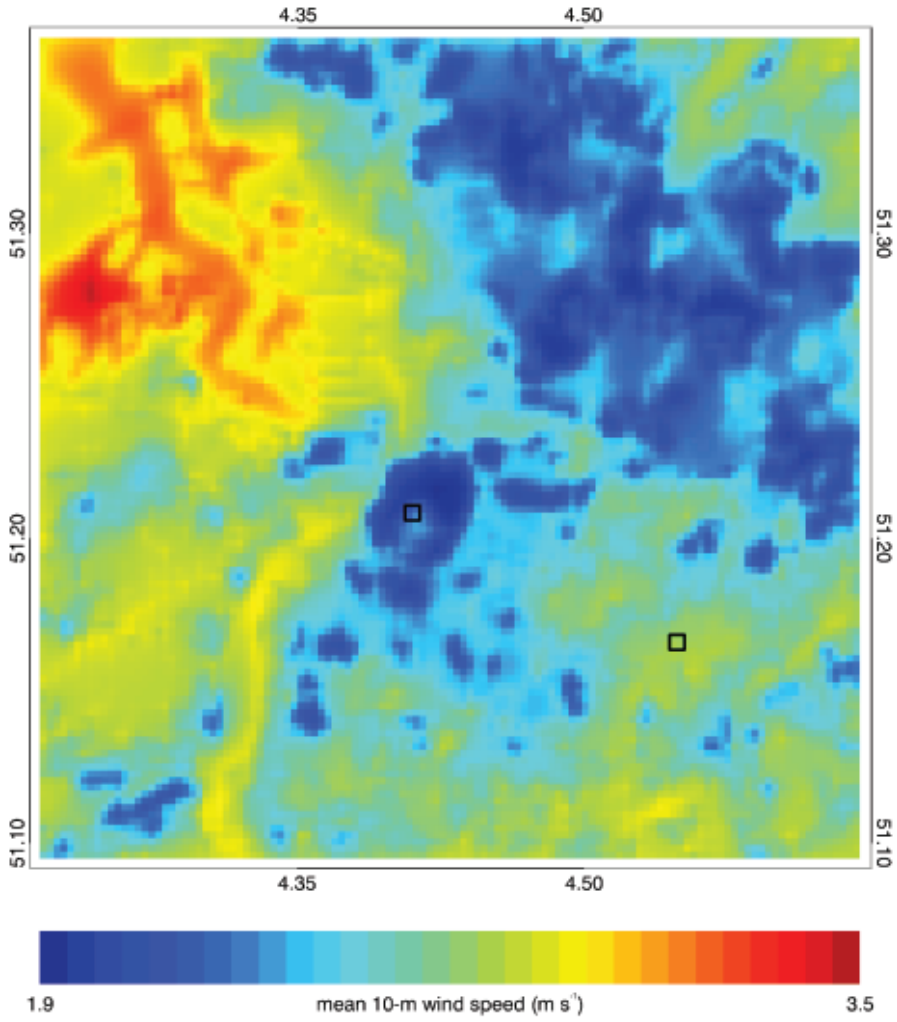


Figure 36. Simulated mean wind speed for the Antwerp domain, July-August 2013. The squares denote the positions of the wind measurement sites. The dark blue area (approximately horse-shoe shaped) in the middle of the domain corresponds to the most densely built-up areas of the city.

At the urban site, the bias was found to be 0.07 m s^{-1}, the root mean square error 0.49 m s^{-1}, and the correlation coefficient was found to have a low value of 0.30. It is also clear from Figure 35 that, while the wind speed magnitude is fairly well captured, most of the time the model is not capable of simulating the phase of the observed wind speed correctly. This is most likely related to microscale wind climate effects in the urban measurements, induced by the small-scale heterogeneity of the urban fabric, and which cannot be captured with 250-m size grid cells. Still, when

considering error statistics obtained with full mesoscale models, as e.g. presented in Chemel and Sokhi (2012) and Loridan et al. (2013) for London, the UrbClim results are at least of the same level of accuracy. This is of course in part related to the quality of the forcing data. Still, UrbClim also reproduces fairly well the observed ratio of the urban-over-rural wind speed, a quantity which is entirely determined by the model itself, giving a simulated ratio of 0.55, against an observed value of 0.60.

Figure 36 shows the mean 10-m wind speed for the simulation domain for July-August 2013. The city of Antwerp, as well as the forest areas in the North-East, clearly stand out as areas of low wind values. On the contrary, the upper section of the Scheldt River, as well as most of the harbour and docks area in the North-Western part of the domain, are characterized by higher wind speed values, as expected.

5.2 Urban-to-rural wind speed conversion: empirical approach

To convert wind speed measured at a rural site to an urban location, Oke (2006) suggests the use of Wieringa's (1986) formula

$$u_u(z_u) = u_r(z_r)\frac{\ln\left(Z_{ref}/z_{0r}\right)\ln(z_u/z_{0u})}{\ln\left(Z_{ref}/z_{0u}\right)\ln(z_r/z_{0r})}, \qquad (5)$$

with $u_u(r)$ the wind speed at level $z_u(r)$, $z_{0u}(r)$ the local roughness length at the urban (rural) location, and Z_{ref} a 'reference' height which, according to Wieringa (1986) is set to a value of 60 m.

In fact, this expression is obtained by (1) extrapolating the wind speed at the rural location from the measurement height z_r to the reference height Z_{ref} (using the log profile for wind speed in the atmospheric surface layer – and ignoring stability effects), (2) assuming that at $z = Z_{\text{ref}}$ taken high enough, the wind speed does not change when moving from a rural to a nearby urban location, and (3) extrapolating back down from Z_{ref} to the required height z_u.

Note that in the above a capital 'Z' refers to height above the surface and a small 'z' to the height above the displacement height d, which is taken as $d = 0.7\ h$, with h the height of the roughness elements. The roughness lengths are taken as $z_0 = 0.3\ (h–d)$. In case of standard WMO wind measurements, which in principle are

taken over short grass, a value of z_{0r} = 0.01 m is recommended, together with z_r = 10 m.

The problem with the above formula is that z_u (and also z_r) must be located in the inertial sublayer, i.e., avoiding the roughness sublayer (RSL), for the log wind profile to be valid. The RSL extends to about twice the height of the average roughness element's (buildings) height, so obtaining roof-level wind speed from (1) is problematic.

De Ridder (2010) presents a correction term to the logarithmic wind profile that accounts for RSL turbulence, and that describes the wind profile down to the top of the roughness elements (i.e., roof level). With this correction term, Wieringa's expression becomes

$$u_u(z_u) = u_r(z_r)\frac{\ln\left(Z_{ref}/z_{0r}\right)[\ln(z_u/z_{0u}) + \psi_M^*(z_u/z_*)]}{\ln\left(Z_{ref}/z_{0u}\right)\ln(z_r/z_{0r})}, \quad (6)$$

with

$$\psi_M^*(z/z_*) = \frac{1}{\lambda}\left(1 + \frac{\lambda}{\mu\, z/z_*}\right)e^{-\mu z/z_*}, \quad (7)$$

which is a function of the RSL height z_*. The latter is taken as z_* = 1.55 h. The coefficients appearing in the last expression are given by μ = 2.59 and λ = 1.5.

The above expressions were tested using data presented in Manning et al. (2000), which were acquired close over roof tops in the town of Leek in the UK, in an area with an estimated mean street canyon height of 10-14 m. They took measurements at heights of 14.0 and 16.2 m, and compared locally observed wind speed with 10-m values observed at the nearby Manchester airport. The wind speed measured at 16.2 m was found to equal approximately 63 % of the 10-m wind at Manchester; for the wind measured at 14.0 m the figure was 54 %. These values are shown in Figure 37 as asterisk symbols, for the two different heights.

We estimated these urban-to-rural wind speed ratios at the experimental measurement heights using above expressions, letting the canyon height h (hence the roughness and displacement height as well) vary, between 10 and 14 m, the range given by Manning et al. (2000). The effect of this variation on the calculated wind speed ratios is given by the vertical bars in Figure 37, and the intermediate result (for h = 12 m) is shown by the square symbol. This exercise was done once with and once without taking RSL effects into account.

From Figure 37 it can be seen that the simple set of equations above captures the observed urban-to-rural wind speed ratio fairly well. Yet, there is a large uncertainty (related to an inaccurate knowledge of the street canyon height – it is fair to say that this quantity is in fact not even well defined, especially when considering slant roofs, differing building heights, complex building configurations, ...). Also, it is difficult to say which of the approaches – i.e., with or without RSL effect – is the better.

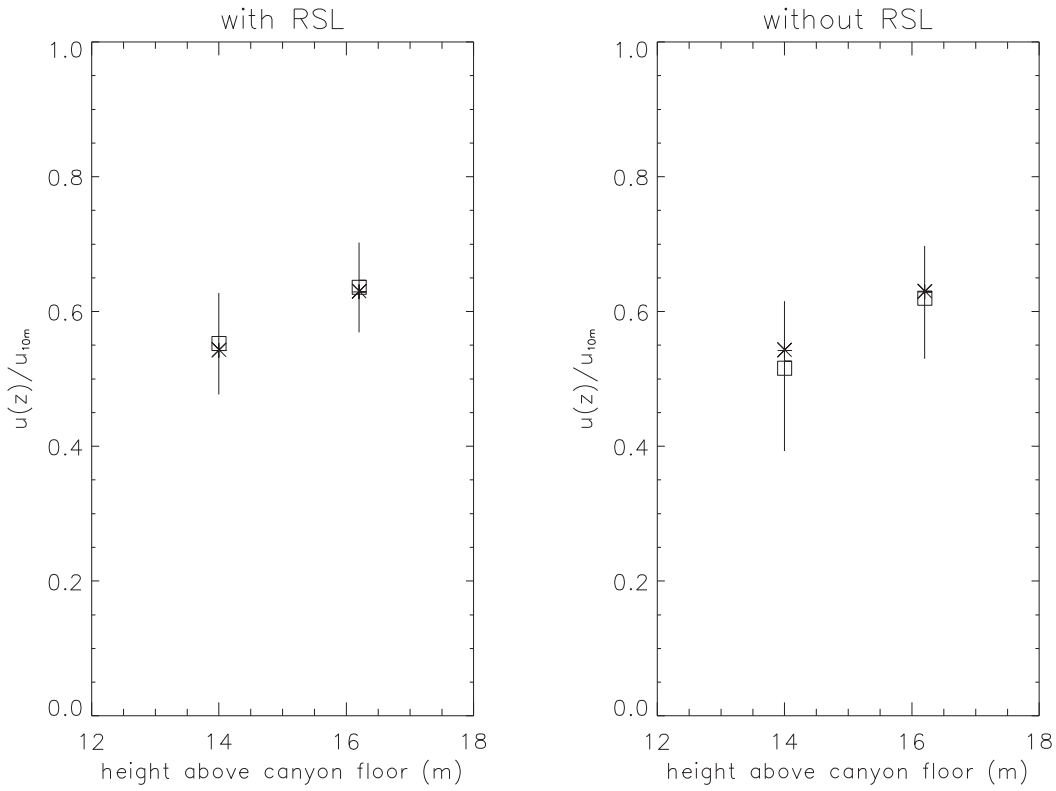


Figure 37. Ratio of the wind speed at the experimental site compared to the 10-m wind speed observed at Manchester airport, as a function of height above the canyon floor (14.0 resp. 16.2 m). The square symbol gives the calculated result assuming a street canyon height of 12 m, and the vertical bars represent the variability related to the uncertainty in this height between 10 m (lower end of the bar) and 14 m (upper end). The asterisk symbols correspond to observed values. The left panel shows results obtained with the RSL correction, and the right panel without this correction.

It should be noted, though, that the observations and estimates here were related to levels above the roof level. Wind speed values at the rooftop level itself, being affected more by the RSL, will exhibit larger differences between the calculations with/without RSL effects. This can be seen in Figure 37, where the lower end of the vertical bars positioned at an abscissa value of 14.0 m (leftmost bar in each panel) correspond to 'rooftop conditions'. The ratio of the estimated wind speed under

these conditions to the 10-m value measured at Manchester is 48% and 39%, for the case with and without RSL effects, respectively.

In conclusion, urban wind speeds are markedly lower than those observed at rural locations (or at locations with a low roughness length in general such as a airports), and either approach (i.e., with or without RSL correction) will in principle yield an improved estimate of roof level wind speed as compared to simply using rural wind measurements. Even though the comparison with the Manning et al. (2000) data does not allow to favour either approach, the one with RSL correction should in principle be the more correct, as it is based on the currently 'best available knowledge'. Roof level wind speed will not become as low with the latter approach.

5.3 Diurnal wind statistics from daily values

Inspection of the CMIP5 archive revealed that, for many of the models included, daily values of mean and maximum ('gust') wind speed are available. Hence, below we describe a method to generate diurnal probability density functions from observed values of mean and maximum wind speed. Using the Weibull distribution as a basis, the main work described here consists of the theoretical derivation of the parameters that characterize this distribution, using mean and maxima wind speed as input. Subsequently, using experimental data from three different sites, it is shown that the theoretically derived distribution function does indeed represent the diurnal variability of observed wind speed well.

5.3.1 Method

It is assumed that the cumulative probability density function (PDF) of the wind speed in the course of a day can be represented through the Weibull distribution, defined as

$$F(V) = 1 - e^{-\left(\frac{V}{\lambda}\right)^k} . \quad (8)$$

The objective of this section is to specify the shape parameter k and the scale parameter λ from observed daily values of mean and maximum wind speed. Note that the 'maximum wind speed' employed here refers to the highest ten-minute mean value occurring in the day under consideration. Note that the Weibull distribution is often used for the representation of wind speed data. As required, it has a positive domain only. Furthermore, its two parameters provide sufficient

flexibility to accommodate many different shapes, while still only needing only two conditions or equations to determine the two parameters.

The mean wind speed can be expressed as

$$V_0 = \int_0^\infty V \, dF(V) = \lambda \, \Gamma\left(1 + k^{-1}\right), \tag{9}$$

with Γ the gamma function (see, e.g., Abramowitz and Stegun, 1965).

Another equation is easily obtained, by expressing a ten-minute period as the fraction of a day during which occur the wind speeds that establish the maximum daily wind, as follows:

$$\varepsilon = \int_{V_t}^\infty dF(V) = e^{-\left(\frac{V_t}{\lambda}\right)^k}, \tag{10}$$

with ε = 600 s / 86400 s ≈ 0.0694, and with V_t a threshold wind speed, the mean of the values above which constituting V_{max}.

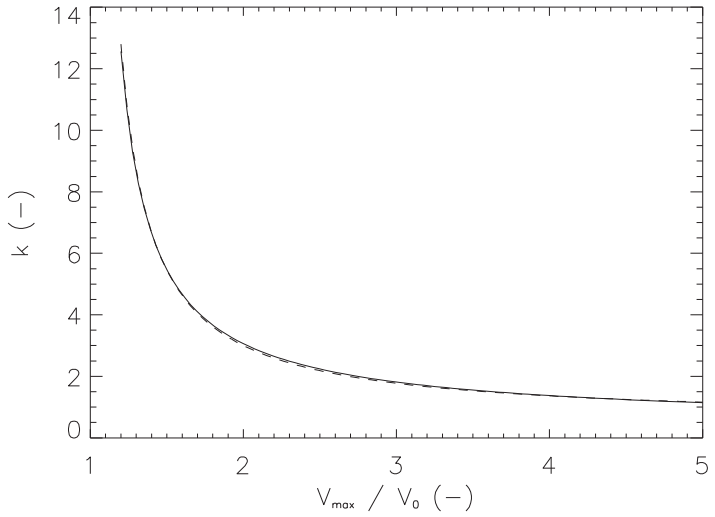


Figure 38. Weibull shape parameter k calculated as a function of the ratio V_{max}/V_0, (solid line). The dashed line represents the curve fitted to the exact solution.

Note that the last expression introduces an extra unknown, bringing their total number, together with k and λ, to three. Therefore, an extra equation is required to close the system and allow a solution. This is obtained by expressing the ten-minute maximum wind speed as the weighted average of those wind speeds occurring above the earlier introduced threshold speed V_t, that is,

$$V_{\max} = \frac{1}{\varepsilon}\int_{V_t}^{\infty} V\, dF(V) \\ = \lambda\,\Gamma\left(1+k^{-1}\right)\left\{1-P\left[1+k^{-1},\left(\frac{V_t}{\lambda}\right)^{k}\right]\right\}, \tag{11}$$

with P the incomplete gamma function (Abramowitz and Stegun, 1965). Inserting (2)-(3) in (4), one obtains

$$1-P\left(1+k^{-1},\ln\varepsilon^{-1}\right)-\varepsilon\frac{V_{\max}}{V_0}=0, \tag{12}$$

which is a transcendental equation in the variable k, with (known) parameters V_0, V_{max}, and ε. This equation was solved iteratively for a range of ($V_{\max}/V_0$) ratios. The resulting values for k as a function of this ratio are shown in Figure 38.

Since it is time consuming and impractical to recur to an iterative solution of (4) in applied studies, a curve was fitted to the exact solution. It was found that, using

$$k \approx a + b\left(\frac{V_{\max}}{V_0}-1\right)^{-1}, \tag{13}$$

with a = 0.55 and b = 2.45, the fitted values of k differed from the exact solution by less than 2.5 % over the entire range of the domain. Now, with k known, the cumulative PDF for diurnal wind speed is fully established.

5.3.2 Results

In order to verify our approach, it was applied to observed wind data. These were taken from the archive of wind measurements obtained from a meteorological tower near the town of Mol in the northern part of Belgium, at 51.2186 °N and 5.0887 °E, and located in flat country a few meters above sea level. The terrain around the tower is characterised by the presence of pine trees with a height of approximately 20 m and medium-sized buildings, most of which are lower than the trees. Wind observations are obtained from continuously operated cup anemometers at different heights, the measurements of which are averaged and archived over ten-minute periods.

As a first verification, daily measured 48-m wind speed PDF's were compared with theoretically obtained PDF's, using the relations above with the theoretically derived shape and scale parameters as a function of measured values of mean and ten-minute maximum wind speed. The latter were obtained from the ten-minute values of measured wind speed, for each day. Figure 39 shows such a comparison, for data

collected over the period 1-30 August 2005. It can be seen that the theoretical curves very closely match the curves established using the observations at ten-minute temporal resolution on most days. In particular, this figure shows that the method described in the previous section is capable of capturing very well the different wind regimes occurring on different days.

Since the period of one month considered so far is rather limited in the context of applied wind studies, a longer period was considered, of a full year. In this exercise, we looked at daily mean values of the cube of the wind speed, as e.g. wind power varies according this quantity. (Alternatively, it might be useful to consider the square of the wind speed, as this determines the drag force on the built environment.)

Again using the ten-minute wind speed values recorded at the Mol site, but for the whole year 2005 this time, the observed mean of the wind speed cubed for a given day was calculated as

$$\left\langle V_{obs}^3 \right\rangle = \frac{1}{N}\sum_{i=1}^{N} V_i^3, \tag{14}$$

with the V_i the N (= 144) measured wind speed values occurring in the day considered. The theoretical value, based on the results of the previous section, was calculated as follows:

$$\begin{aligned} \left\langle V_{\mathrm{mod}}^3 \right\rangle &= \int_0^\infty V^3 \, dF(V) \\ &= \left[\frac{V_0}{\Gamma\left(1+k^{-1}\right)} \right]^3 \Gamma\left(1+3k^{-1}\right). \end{aligned} \tag{15}$$

The comparison of the observed versus the theoretical daily wind speed cube is shown in Figure 39. Again, the correspondence is remarkably good, the root mean square error having a value of 8.2 $m^3 s^{-3}$, despite the fact that for the calculation of the theoretical value, only two wind speed values (i.e., the daily mean and maximum) were used. When using observed wind speed at a height of 114 m instead of 48 m, the error increased somewhat, as expected given the higher wind speed values measured at that level, but still being very low. At both measurement heights, the fractional bias was found to be extremely low, with values below one percent.

Finally, even though our method was shown to work well on the site of Mol, we wanted to verify whether data from another observational site, representing

drastically differing conditions, could also be represented well. Therefore, data were selected from two other sites:

- The Koltur site, which is located at the western side of the Faroe island group between Iceland and the United Kingdom at 61.98 °N and 6.96 °W at an altitude of 0 m, dominantly influenced by conditions prevailing in the Northern Atlantic. Data were taken from the 10-m wind speed archive for the year 2005.
- The Alpine Met site of the Cold Land Processes Field Experiment (Elder and Goodbody, 2004) in, Colorado, USA, at 39.848 °N and 105.861 °W, altitude 3585 m, representing vastly different compared to the other sites. Use was made of 10-m wind speed, for the period 20 September 2002 to 1 October 2003.

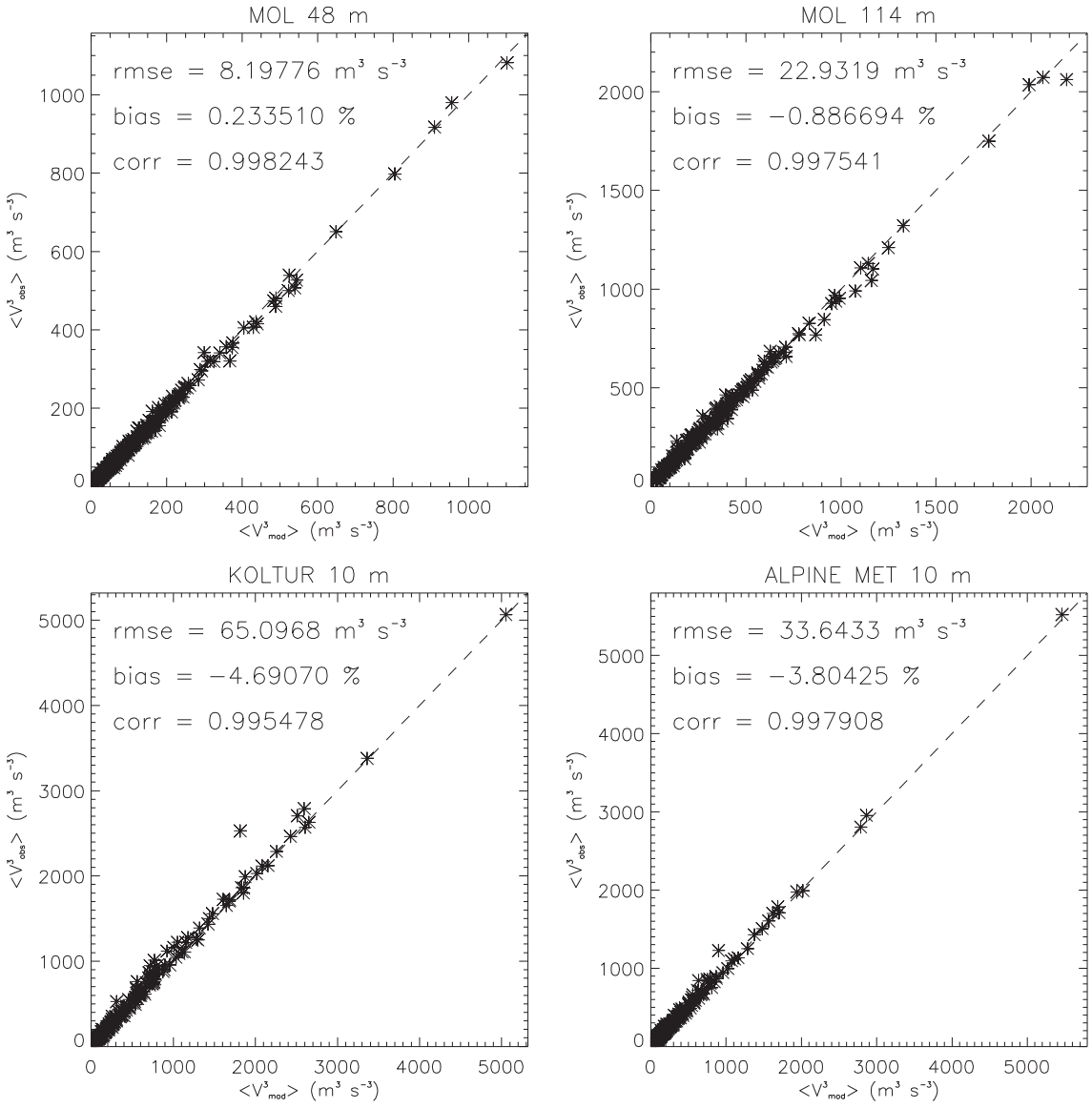


Figure 39. Modelled (horizontal axis) versus observed (vertical axis) values of the daily mean of the cube of the wind speed, for the site of Mol at 48 m (upper left) and 114 m (upper right), for Koltur (lower left), and for Alpine Met (lower right).

Despite the vastly different conditions these two sites are exposed to as compared to Mol, our methodology works as well over these areas. The RMSE values are higher but, again, that can be explained by the higher wind speeds. In fact, the normalized RMSE, obtained by dividing the RMSE by the maximum $<V^3>$ value, exhibits values in the much narrower range of 0.6-1.1 %. The fractional bias was higher than for Mol, though, but still exhibiting very acceptable (low) values of a few percent.

In conclusion of this section, it is fair to say that diurnal wind statistics, constructed using the Weibull distribution and only using values of the daily mean and ten-minute maximum wind speed, represent well the full statistics established using the ten-minute data themselves. Data were used from three different measurement sites, between them representing drastically different wind climates. Our method was found to reproduce observed wind statistics with correlation coefficients approaching unity, with normalized root mean square error values of the order of one percent or less, and with values for the fractional bias ranging from less than 1 % to approximately 5 %. This implies that all statistically relevant information of a diurnal wind time series can be captured fairly well using only these two parameters, that is, the daily mean and maximum values.

6 Air quality

6.1 Introduction

Present and future air quality assessments must account for changes in human activities. Strategies are developing with the aim of reducing emissions and using new clean technologies, however atmospheric composition is also affected by the regional climatic variability. The Basque Country region is not an exception and it is also suffering the consequences of climate change. According to projections of the Regional Climate Models (RCM) provided by the EU FP6 ENSEMBLES project (http://www.ensembles-eu.org/), which run under the A1B IPCC greenhouse gas emission scenarios (Nakicenovic et al., 2000), a warming of the mean surface air and sea over the region is expected by the end of the 21st century (González-Aparicio and Hidalgo, 2011; Chust et al., 2011; Moncho, 2009).

The most vulnerable area of the region in terms of air pollution is the Bilbao city. Its coastal and complex terrain and its natural resources favoured a significant industrialization since the 1800's. Particularly, the PM_{10} concentrations are one of the main air quality concerns in the area. Viana et al (2003) analysed the PM_{10} influence by the climate and González-Aparicio et al (2013) studied the relation between the pollutant and the meteorological variability, developing a diagnosis model of PM_{10}.

In this context, the objective of this study is to evaluate the PM_{10} changes in Bilbao over the 21st Century based on the regional climate changes. Below, we describe how the climatic projections of the RCMs are used with the PM_{10} diagnosis model in order to obtain the PM_{10} expected changes. We then evaluate and discuss the most influential climatic changes on the PM_{10} and consequently, the future PM10 changes for summer and winter seasons, independently.

6.2 Data and methodology

The statistical PM_{10} diagnosis model developed by González-Aparicio et al. (2013) was employed to generate the future scenarios in Bilbao. The hourly input variables are quantitative: air temperature (temp, ^{o}C) and specific humidity (q, kg water kg^{-1} air) at 2 m, incoming short-wave radiation (solar, W m^{-2}) and precipitation (prep,

mm) at the surface, wind speed (wspd, m s^{-1}) and direction (wdir, º) at 10 m, the traffic intensity (number of vehicles $hour^{-1}$) and the NO2 emissions (µg m^{-3}). There are also qualitative variables: winter (December, January and February) or summer (June, July and August), day (4-20 UTC) or night (21-3 UTC), working days (Monday-Friday) or non-working days (Saturdays and Sundays) and wind direction quarters (North, South, East or West). The model is adjusted to three urban environments:

(1) the city centre (see Eq (16) below);

(2) a coastal sub-urban area characterised by residential neighbourhoods where the PM10 concentrations are considered 10% (in summer) and 18% (in winter) less than in the city centre;

(3) The inland sub-urban area, characterised by high traffic density and industrial activities, where the PM10 concentrations are considered 35% (in summer) and 20% (in winter) higher than in the city centre.

$$PM_{10} = -14.682 + 0.724 * (NO_2) + 0.019 * (Temp) + 0.005 * (Solar) - 295.9 * (q) + 0.493 * (wspd) + 0.079 * (wdir) - 0.005 * (Traffic) + \begin{cases} -5.449(forWinter) \\ 0.0(forSummer) \end{cases} + \begin{cases} 0.487(forWorkingday) \\ 0.759(forSaturday) \\ 0.0(forSunday) \end{cases} + \begin{cases} 0.0(forDaytime) \\ 5.6(fprNighttime) \end{cases} + \begin{cases} 22.707(forNorthwind) \\ 12.684(forSouthwind) \\ 6.824(forEastwind) \\ 0.0(forWestwind) \end{cases} \quad (16)$$

The input variables are of hourly frequency. However, testing the model with daily variables, it was observed that the error committed was not significant (up to 5 µg m^{-3}) comparing the calculated daily PM_{10} based on hourly input variables. Since the purpose of this analysis was to generate PM_{10} future scenarios in a seasonal basis, the accuracy of the model was considered suitable for this use. Also, the future changes in PM_{10} levels calculated in this analysis are only subjected to climatic change, so that the NO_2 and the traffic intensity were considered the same as the present period (2000-2010).

The meteorological data used to force the diagnosis model were daily outputs of the Regional Climate Models (RCM). The RCMs selected were six from the thirteen included in the EU-FP6 ENSEMBLES project: CLM (Steppeler et al, 2003), HIRHAM (Christensen et al., 1996), RACMO (Lenderik et al., 2003), REMO (Jacob, 2001), PROMES (Castro et al., 1993) and ALADIN (Spiridonov et al., 2005). These RCMs were driven by three different Global Climate Models (GCM): ARPEGE (Gibelin and Déqué 2003), HadCM3 (see Rowell 2005) and ECHAM5 (Roeckner et al., 1996). Each pair of RCMs was driven by the same GCM allowing an assessment of the global

model influence on the future projections. The RCMs selected were the same as those employed by González-Aparicio and Hidalgo (2011) in order to assess the expected changes including the same range of uncertainties. One RCM (HIRHAM) at hourly based frequency forced by the ECHAM5 GCM was also included with the aim of testing the adequacy of the PM_{10} results based on daily frequencies.

The results of the models were evaluated for summer (from June to August) and winter (from December to February) during the period 2000 - 2100. The projected changes were given at seasonal basis by the difference between the future (2010-2100) and the present period (2000-2010) to avoid the systematic bias; i.e. the discrepancy between the observations and the model outputs.

6.3 Results and discussions

6.3.1 Projected meteorological changes.

The following analysis describes the climatic changes expected over the Bilbao estuary. Figure 40 and Figure 41 show in detail the trends of the model outputs along the 21st century and Table 9 summarizes the statistics for averaged 30-year-window periods extracted from the future projections, in 2020-2050 and 2070-2100. The results are compared with similar studies carried out over the Basque Country at a regional scale. The end of the section summarizes a review of the climatic changes over the area.

Mean air temperature at 2 m: the models showed a positive slope along the 21st century for both seasons. In summer, the ensemble future projections showed an increase of 0.72 ºC in 2020-2050 and the spread of the change, i.e. minimum and maximum changes, was 0.12-0.54 ºC, given by REMO-ECHAM5 and CLM-HadCM3. In 2070-2100, the increase was 1.60 ºC with respect to the present and the spread was 0.56-3.36 ºC given by HIRHAM-ARPEGE and CLM-HadCM3. In winter, the ensemble models indicated a change of 1.12 ºC in 2020-2050 and 2.5 ºC in 2070-2100 with respect to the present. The spread was given by HIRHAM-ECHAM5 and CLM-HadCM3 (0.80-1.40 ºC) for the middle of the century and by HIRHAM-ARPEGE and RACMO-ECHAM5 (1.59-3.40 ºC) for the end of the century.

These results were comparable with the regional temperature change studied in González-Aparicio and Hidalgo (2011): for summer, the 90th percentile of the maximum temperature showed an increase of 1.9 ºC in 2020-2050 and 3.5 ºC in

2070-2100. In winter, changes of the 10th percentile of the minimum temperature were 2.5 ºC in 2020-2050 and 3.1ºC in 2070-2100.

Specific Humidity: There was a general upward trend in both seasons. In summer, the ensemble average revealed an increase of 70% for the 2020-2050 and 135% for 2070-2100 with respect to the present period. In winter, the increase was more dramatic: 120% and 169% more for the middle and the end of the century, respectively. The specific humidity is the variable in role with the major change. The consequences of higher specific humidity could be due to an increase of the water vapour caused by the temperature rise and evaporation capacity.

Incoming short wave radiation: In summer a positive trend (from 5.14 to 5.92 W m^{-2} from the mid to the end of the scenario) was exhibited by the ensemble models except for the HIRHAM-ARPEGE model which showed downward behaviour (from -9.37 to -12.65 W m^{-2} for 2020-2050 and 2070-2100). In winter the models were in agreement depicting an ensemble negative trend (from -0.49 for 2020-2050 and -5.64 W m^{-2} for 2070-2100). That meant that in summer, the cloudiness decrease with higher income short-wave radiation. On the contrary, the decrease of the radiation in winter implied more cloudiness and hence, more probability of precipitation.

Precipitation (pr): Precipitation (pr): The projections for the 21st century over the area did not reach a consensus over all the models selected. For example, in summer, the ensemble average showed a decrease of -0.096 mm in 2020-2050 and a smaller change (-0.0032 mm) in 2070-2100 with respect to the present time. These results were in agreement with the study carried out by Moncho and Chust (2012) who worked with four of the RCMs selected in this study (HIRHAM, PROMES, RACMO and ALADIN). They highlighted that the precipitation distribution did not show a clear behaviour pattern over the 21st century. Chust et al. (2011) enhanced also that for the Iberian Peninsula there was no agreement between the models, arguing that the location within a transitional climate regime between oceanic temperate and the dry subtropical climate (Barry and Chorley, 2003) makes it difficult to predict future rainfall over the area.

Wind speed and direction: The patterns of the wind showed a future shift in the speed but it was insignificant in the direction. In summer, the ensemble models indicated a slight decrease of the wind speed: -0.03 and -0.02 m s^{-1} for the middle and the end of the century with respect to the present. This is because the RCMs driven by ECHAM5 and HadCM3 indicated a reduction up to -0.17 m s^{-1} but the RCMs driven by ARPEGE (ALADIN & HIRHAM) showed an increase up to 0.18 m s^{-1}

by 2100. In winter there was greater decrease up to -0.1 m s^{-1} in 2070-2100. All the models showed the same downward trend with the exception of HIRHAM-ECHAM5 that expected an insignificant increase of 0.09 m s^{-1} in 2070-2100. It could not be stated that the models show an important change in the wind direction: the ensemble average show up to 6.3 º and -9.5 º of change for the end of the century.

Therefore, in general, the models showed similar trends over the 21st century although some of them did not expect significant shifts. The results indicated high dependency between the RCM and the GCM driver; the trends of the RCM forced by the same GCM were very similar as González-Aparicio and Hidalgo (2011) pointed out.

In summer, the models expected a rise in the air mean temperature leading to an increase in the kinetic energy of the gas-phase equilibrium: i.e. the capacity to hold water vapour in the air increases. This was directly linked with the positive slope in the specific humidity. Also, the modelled income short-wave radiation expected a reduction which meant a decrease in the cloudiness, being able to explain the decrease of the modelled precipitation. A slight decrease in the wind speed with no significant changes in the wind direction was identified. Summarizing these shifts, it could say that these patterns of behaviour are characterized by more calm conditions. In winter, the results were different. The solar radiation evolution indicated a decrease, implying more cloudiness and hence, more probability of rain. The models also showed a rise in the air mean temperature, associated with an upwards trend in the specific humidity due to the more water vapour mass. The change of the specific humidity in winter was greater than in summer due to the greater moisten atmosphere by higher temperatures and precipitations and thus, greater evaporation. These winter meteorological situations depicted in the future projections could be associated with more frequent front's arrival.

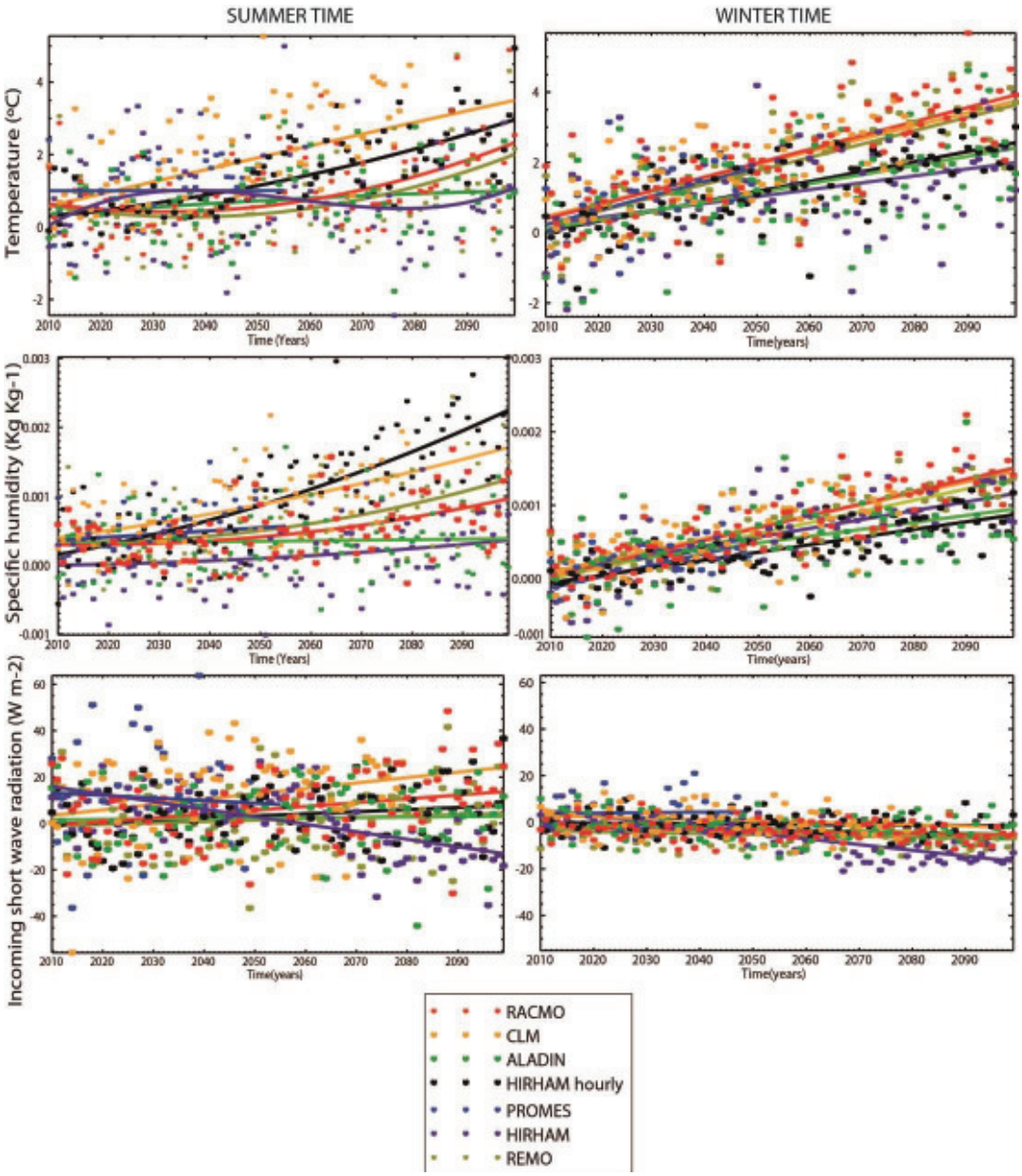


Figure 40. Projected mean air temperature, specific humidity and radiation changes for future periods (2010-2100) with respect to the present period (2000-2010) for summer.

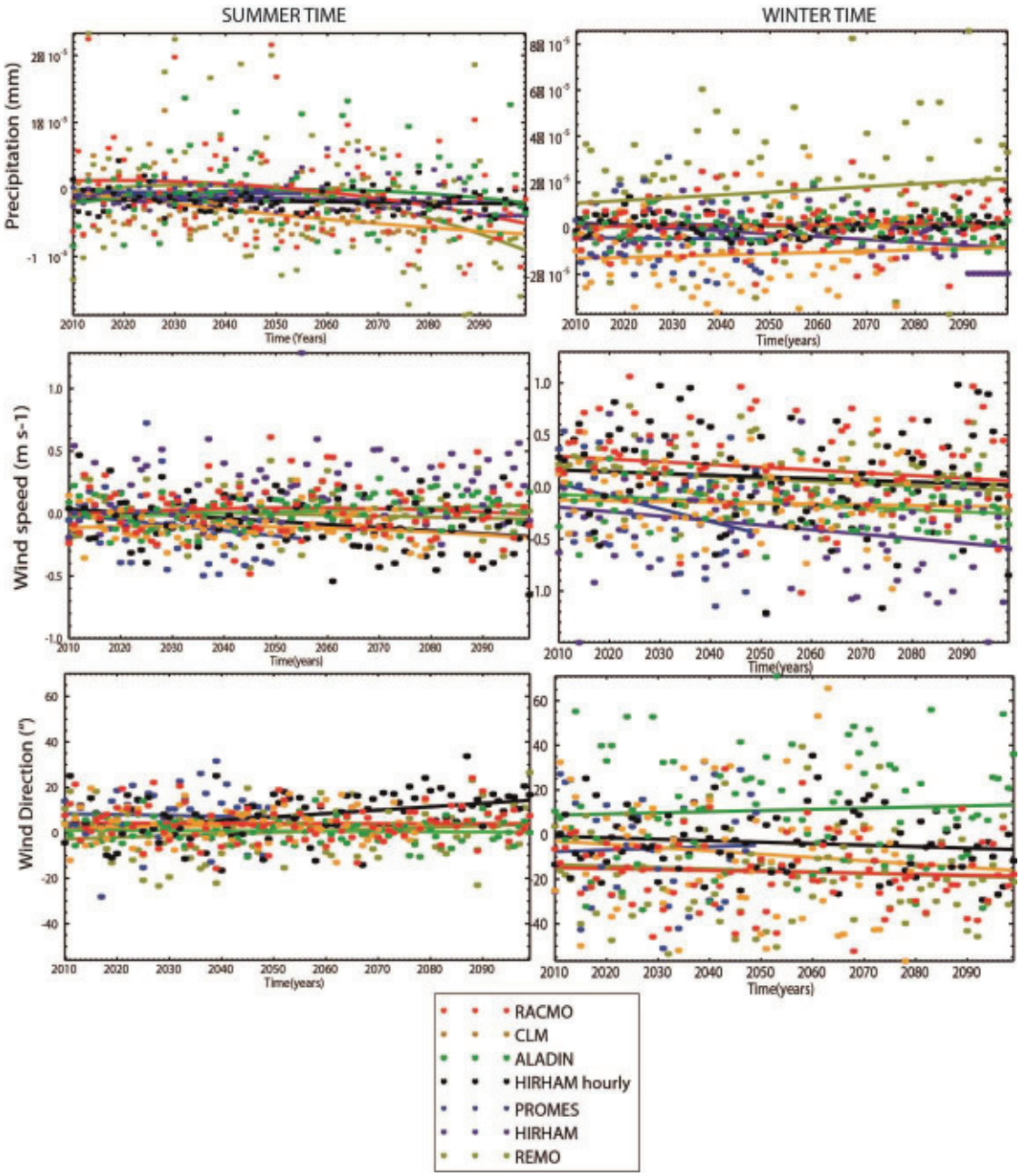


Figure 41. Meteorological expected changes for future periods (2010-2100) with respect to the present period (2000-2010) for winter.

SUMMER		ECHAM5			ARPEGE		HadCM3		ENSEMBLE
Variable	Time frame	RACMO	REMO	HIRHAM -h	ALADIN	HIRHAM	CLM	PROMES	
Temperature	2020-2050	0,309	0,114	0,577	0,384	0,746	1,542	1,352	0,718
	2070-2100	1,535	1,040	2,282	0,839	0,550	3,366	NaN	1,602
Specific Humidity	2020-2050	0,00028286	0,00033838	0,00041729	0,0003139	5,10862E-05	0,00068228	0,00063295	0,00038839
	2070-2100	0,00073199	0,00096996	0,00175612	0,00037065	0,000338375	0,0013283	NaN	0,0009159
Radiation	2020-2050	0,76	-2,43	0,98	-0,12	9,37	13,16	14,23	5,14
	2070-2100	13,49	6,06	6,81	2,45	-12,65	19,35	NaN	5,92
Precipitation	2020-2050	1,43E-06	1,83E-06	-0,0187496	1,39E-07	-1,15E-06	-2,59E-06	-0,652566	-9,59E-02
	2070-2100	-3,52E-06	-6,08E-06	-0,0198818	-8,70E-07	-2,21E-06	-5,64E-06	NaN	-3,32E-03
Wind speed	2020-2050	0,00543001	-0,0035384	-0,0222226	-0,0346151	0,12954	-0,126331	-0,164527	-0,03089487
	2070-2100	0,00575504	-0,0484687	-0,152339	0,0556919	0,182003	-0,16887	NaN	-0,02103796
Wind Direction	2020-2050	3,207	-3,823	2,950	0,878	2,950	2,957	10,503	2,803
	2070-2100	5,055	3,024	13,450	0,031	13,450	3,092	NaN	6,350
WINTER									
Temperature	2020-2050	1,312	1,125	0,796	0,943	1,007	1,391	1,217	1,113
	2070-2100	3,392	3,093	2,110	1,965	1,590	2,821	NaN	2,495
Specific Humidity	2020-2050	0,0004323	0,00035997	0,0001988	0,00038332	0,000434978	0,00048387	0,00025431	0,00036393
	2070-2100	0,00127845	0,00113832	0,00072245	0,00076138	0,000938674	0,00103887	NaN	0,00097969
Radiation	2020-2050	-3,582	-4,988	-0,667	-2,434	2,124	1,809	4,316	-0,489
	2070-2100	-4,928	-6,807	-1,582	-3,139	-14,798	-2,562	NaN	-5,636
Precipitation	2020-2050	1,97E-06	1,58E-05	1,94E-05	8,71E-07	-2,99E-07	-1,39E-05	-0,52 E-07	-7,02E-02
	2070-2100	1,98E-06	1,96E-05	2.62E-05	-2,43E-07	-8,64E-06	-7,12E-06	-4,29E-06	-6,09E-01
Wind speed	2020-2050	0,237495	0,130532	0,252726	-0,0784076	-0,211694	-0,199846	-0,289001	-0,02259937
	2070-2100	0,127454	0,0487135	0,0965457	-0,233364	-0,517107	-0,139241	NaN	-0,10283313
Wind Direction	2020-2050	-18,336	-16,701	-3,452	9,991	-3,452	-5,890	-6,866	-6,387
	2070-2100	-18,784	-17,965	-6,967	11,792	-6,967	-18,558	NaN	-9,575

Table 9. summer and winter disaggregating of the meteorological changes expected changes into three time frames for each model selected.

6.3.2 PM_{10} future variability

As the analysis above showed, the main variables suffering a significant future change were the mean air temperature, specific humidity, wind speed and precipitation.

Meteorological influence on the PM_{10}: The upward trend of the temperature had a positive effect on the PM_{10} variability and the increase in the specific humidity could lead to higher water vapour content in the future climate. This could lead to higher H_2O_2 concentrations, the principal SO_2 oxidant. Thus, the enhanced chemical reaction in the atmosphere increases the sulphate aerosols generation, one of the main components of the PM_{10}. The wind speed is one of the main dispersive patterns of the air pollution. When the winds are weak, there is stagnation and the pollution increases. When the winds are higher, the pollutant dispersion is larger. In

summer seasons, the downward ensemble trend of the wind speed let the PM_{10} future concentrations increase.

The behaviour for winter season was more complex. At the first third of the century it showed an increase up to 0.50 m s^{-1}, causing lower PM10 concentrations. During the middle of the century, there was a substantial decrease up to -0.20 m s^{-1} which leads to an increase in the PM_{10} concentration. These represent a potentially significant effect for polluted regions. However, the expected increase in the regional extreme precipitation response to climate change (Moncho 2009) could be the major cause of the PM_{10} response (Racherla and Adams, 2006; Jacob et al. 2009).

PM_{10} expected changes over the 21st century: For summer, there was an ensemble upwards trend of the models (Figure 42). In 2020-2050, the expected increase was 0.15 µg m^{-3} and the spread of the models indicated a range between -0.414 µg m^{-3} to 0.658 µg m^{-3}. Two of the models had a downward trend: REMO-ECHAM5 (-0.414 µg m^{-3}) and ALADIN-ARPEGE (-0.034 µg m^{-3}). In 2070-2100, the ensemble average showed a positive change (0.30µg m^{-3}) with respect to the present and the spread was between -0.071 µg m^{-3} to 1.020 µg m^{-3}. Although CLM-HadCM3, ALADIN-ARPEGE and REMO-ECHAM5 had negative shifts, there were insignificant: -0.071 µg m^{-3}, -0.052 µg m^{-3} and -0.017 µg m^{-3}, respectively. The positive increase had greater impact: 0.282 µg m^{-3} (RACMO-ECHAM5), 0.545 µg m^{-3} (HIRHAM-ECHAM5) and 1.020 µg m^{-3} (HIRHAM-ARPEGE). For winter, the general trend was negative due to worse meteorological conditions and the expected increase of the extreme precipitation. The ensemble average showed a decrease of -0.650 µg m^{-3} for the mid of the century and the spread of the model results were -1.498 µg m^{-3} to 0.597µg m^{-3}. For the end of the century, the ensemble averaged was -1.124 µg m^{-3} with the spread between -1.826 µg m^{-3} to 0.562 µg m^{-3}. In both time frames, only ALADIN-ARPEGE indicated the positive change.

Extrapolating to the two additional urban environments, for summer seasons, the increase of the PM10 was 0.135 µg m^{-3} in 2020-2050 and 0.27 µg m^{-3} in 2070-2100 at the coastal site. At the inland site, the increase was greater: 0.20 µg m^{-3} in 2020-2050 and 0.405 µg m^{-3} in 2070-2100. For winter, the change was -0.27 µg m^{-3} in 2020-2050 and -1.32 µg m^{-3} in 2070-2100 at the coastal stations. At the inland station the decrease was smaller: -0.52 µg m^{-3} and -0.89 µg m^{-3} in 2020-2050 and 2070-2100, respectively.

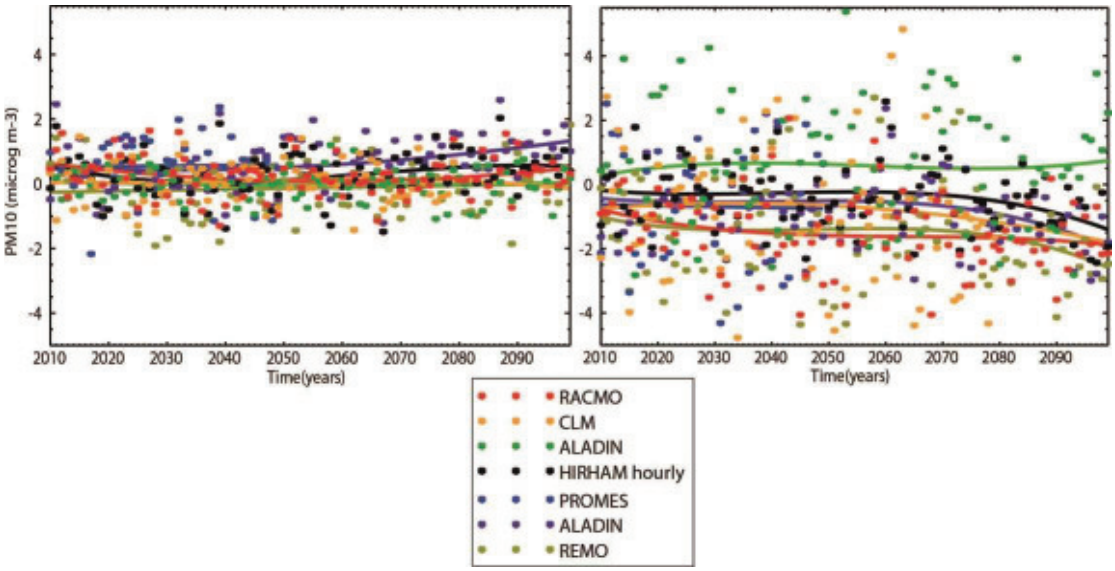


Figure 42. PM10 predicted future projections (2010-2100) based on the regional climate models selected for summer (left) and winter seasons (right).

6.3.3 Conclusions regarding future PM_{10} concentrations levels

Regarding the physical influence of the climatic variables in the PM_{10} variability, the upward trend of the temperature has a positive effect on the PM_{10} concentrations and the increase in the specific humidity may lead to higher water vapour content in the future climate. The weaker wind speed enhances the pollutant stagnation. For both seasons, there is an upward trend of all the models to exceed the daily current PM_{10} mean. In particular, during winter there is no clear pattern of increasing the PM_{10} concentrations, associated with the complex climatic projections results over the area, but during summer it will be more frequent to exceed the current mean than for winter seasons.

Finally, the results presented indicate that climate change could have substantial effects on air quality in urban areas and as a consequence on human health, agriculture, and natural ecosystems. Changes in local climatic variables such as temperature, radiation, specific humidity, wind patterns and precipitation affect the different components of the pollutant life cycle and therefore affect the pollutant concentration. In general, climate change can be highly variable from place to place, depending on the urban development as well, so that the effects of climate change on air quality need to be evaluated from case to case. Finally, consideration is that climate change is opposite in sign to that of currently envisioned pollution control measures. Although changes in pollutant emission might actually modify the influence of climate change, a global perspective of emission-control policies that consider both climate change and air quality should be integrated.

7 Conclusions

This report provides an overview of the main research activities, and the results obtained within WP4 ('Agglomeration-scale urban climate and air quality projections') during the second year of the RAMSES project.

The main focus was on the simulation of future levels of urban heat stress and associated human exposure. After the implementation of the (sometimes very cumbersome) coupling of the UrbClim urban climate model to GCM output fields contained in the CMIP5 archive, long-term multi-model ensemble (i.e., using several GCMs to force UrbClim) simulations were conducted, yielding urban climate projections for most of the RAMSES target cities, including mean projection values as well as an uncertainty range.

The main outcomes from this exercise are that (1) urban areas experience around twice as many heat wave days compared to their rural surroundings, and (2) the expected ten-fold (!) increase in the number of urban heat wave days towards the end of the century, which is projected to reach one month per year (not as a single event, but rather, all days taken together will add up to a month).

Moreover, from an experiment involving UHI simulations for 102 cities, we learned that especially the larger cities in Eastern and Central Europe are at risk, all the more as the background (regional) temperatures in these regions are expected to undergo drastic changes (see also Figure 1, in the Introduction).

With respect to the wind climate in cities, we conducted a validation of the UrbClim model (based on our measurements in Antwerp), and we performed an analysis of the ratio of urban to rural wind speed values, which was substantiated by a simple semi-empirical analysis. Moreover, we developed a new method to obtain diurnal wind speed statistics based on daily mean and maximum values, and for which projections are available in the CMIP5 archive, and with an accuracy of a few percent.

Finally, we conducted an investigation regarding the impact of climate change on air quality. Based on statistical relationships between observed concentrations of particulate matter (PM_{10}) and meteorological quantities, we made projections of the concentration of this pollutant. It was found that towards the end of the century, when considering the effect of climate change alone (i.e. assuming no changes in the emissions), PM_{10} concentrations are expected to rise by approximately 0.3 μg m^{-3} during the summer, and to decrease by 1.1 μg m^{-3} during winter. Even though the

analysis was done for the sole city of Bilbao, the method used here is fairly generic and could be transferred to other urban areas which have access to the required input data (i.e. measured pollutant concentrations, and concurrent meteorological data).

With respect to urban climate, the aim is now to employ the UrbClim model to evaluate the effect of certain adaptation measures, in particular those related to changes in the physical infrastructure of a city. This work will be conducted within Task 4.3.

8 References

Abramowitz, M., and I.A. Stegun, 1965. Handbook of Mathematical Functions: with Formulas, Graphs, and Mathematical Tables. Dover Publications, 1046 pp.

Barry, R., and R. Chorley, 2003. *Atmosphere, weather and climate*. Eight edition, Routledge, London, 421 pp.

Bontemps S. et al, GLOBCOVER 2009: Products Description and Validation Report, 2009.

Castro M, Fernandez C, Gaertner MA. 1993. Description of a mesoscale atmospheric numerical model. Diaz JI, Lions JL (eds). In Mathematics, Climate and Environment. Rech.Math. Appl. Ser. Mason 230–253 pp.

Chemel, C., and R.S. Sokhi, 2012. Response of London's urban heat island to a marine air intrusion in an easterly wind regime. *Boundary-Layer Meteorol.*, **144**, 65-81.

Chust, G., Borja, Ángel, Caballero, A., Irigoien, X., Sáenz, J., Moncho, R., Marcos, M., et al, 2011. Climate change impacts on coastal and pelagic environments in the southeastern Bay of Biscay. Climate Res. 48, 307–332.

Christensen JH, Christensen OB, Lopez P, van Meijgaard E, Botzet M.(1996). The HIRHAM4 regional atmospheric climate model. Danish Meteorological Institute (DMI). Technical Report 96-4.

De Ridder, K., 2010. Bulk transfer relations for the roughness sublayer. *Boundary-Layer Meteorol.*, **134**, 257-267.

De Ridder, K., Acero, J.A., Lauwaet D., Lefebvre, W., Maiheu B., Mendizabal, M. 2014. Validation of agglomeration-scale climate projections, RAMSES project deliverable report D4.1.

De Ridder, K., Lauwaet D., Maiheu B. 2015. UrbClim – a fast urban boundary layer climate model. *Urban Climate*, **12**, 21-48.

Diffenbaugh, N.S., and F. Giorgi, 2012. Climate change hotspots in the CMIP5 global climate model ensemble. *Climatic Change*, **114**, 813–822.

Elder, K. and A. Goodbody. 2004. CLPX-Ground: ISA Main Meteorological Data. Boulder, CO: National Snow and Ice Data Center. Digital Media.

FPS (Belgian Federal Public Service) Public Health, 2014. Hittegolf- en ozonpiekenplan. (*Heat wave and ozone peaks plan*). Report available through

http://www.health.belgium.be/filestore/8108413/plan%20HOP_NL2014.pdf (in Dutch, link accessed on 27 February 2015).

Gibelin AL, Déqué M. 2003. Anthropogenic climate change over the Mediterranean region simulated by a global variable resolution model. Climate Dynamics 20: 327–339.

González-Aparicio, I., and Hidalgo, J., 2011. Dynamically based future daily and seasonal temperature scenarios analysis for the northern Iberian Peninsula. Int. J. Climatol. 32 (12), 1825–1833.

Gonzalez-Aparicio, I, Hidalgo, J, Baklanov, A, Padro, A, Santa-Coloma, O (2013). An hourly PM10 diagnosis model for the Bilbao metropolitan area using a linear regression methodology. ENVIRON SCI AND POLL RESEARC, 20, 4469-4483.

IPCC, 2013. Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change [Stocker, T.F., D. Qin, G.-K. Plattner, M. Tignor, S.K. Allen, J. Boschung, A. Nauels, Y. Xia, V. Bex and P.M. Midgley (eds.)]. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, 1535 pp, doi:10.1017/CBO9781107415324.

Jacob D. 2001. A note to the simulation of the annual and inter-annual variability of the water budget over the Baltic Sea drainage basin. Meteorology and Atmospheric Physics 77: 61–73.

Jacob, D.J., Winner, D.A., 2009. Effect of climate change on air quality. Atmos Environ 43 (1), 51-63.

Keramitsoglou, I., et al, 2012. Evaluation of satellite-derived products for the characterization of the urban thermal environment. *J. Appl. Remote Sens.*, **6**, 061704.

Kourtidis, K., A.K. Georgoulias, S. Rapsomanikis, V. Amiridis, I. Keramitsoglou, H. Hooyberghs, B. Maiheu, D. Melas, 2015. A study of the hourly variability of the urban heat island effect in the Greater Athens Area during summer. *Science of The Total Environment*, **517**, 162-177.

Lauwaet, D., De Ridder K., Saeed S., Brisson E., van Lipzig N.P.M., Maiheu B., Hooyberghs H., 2015. Assessing the current and future urban heat island of Brussels. *Submitted to Urban Climate*.

Lemke, B., and T. Kjellstrom, 2012. Calculating workplace WBGT from meteorological data: A tool for climate change assessment. Industrial Health, **50**, 267–278.

Lenderink G, van den Hurk B, van Meijgaard E, van Ulden A, Cuijpers H. (2003). Simulation of present-day climate in RACMO2: first results and model developments, KNMI, Technical Report 252. 24 pp.

Li, D., and E. Bou-Zeid, 2013: Synergistic Interactions between Urban Heat Islands and Heat Waves: The Impact in Cities Is Larger than the Sum of Its Parts. *J. Appl. Meteor. Climatol.*, **52**, 2051–2064.

Loridan, T., F. Lindberg, O. Jorba, S. Kotthaus, S. Grossman-Clarke, C. S. B. Grimmond, 2013. High Resolution Simulation of the Variability of Surface Energy Balance Fluxes Across Central London with Urban Zones for Energy Partitioning. *Boundary-Layer Meteorol.*, DOI 10.1007/s10546-013-9797-y.

Manning A J, Nicholson K J, Middleton D R, Rafferty S C, 2000. Field study of wind and traffic to test a street canyon pollution model. *Environmental Monitoring and Assessment*, **61**, 283-313.

Meehl, G.A., and C. Tebaldi, 2004. More intense, more frequent, and longer lasting heat waves in the 21st century. *Science*, **13**, 994-997.

Moncho R. 2009. Validación y calibración de los modelos regionals del clima para el estudio de las precipitaciones intensas del País Vasco. MSc. thesis, Physics department, Valencia University (Spain).

Moncho, Roberto, Vicente Caselles, and Guillem Chust. 2012. "Alternative Model for precipitation Probability distribution: Application to Spain." Climate Research 51 (1) (January 24): 23–33. doi:10.3354/cr01055.

Nakicenovic, N., J. Alcamo, G. Davis, B. Vries, J. Fenhann, S. Gaffin, K. Gregory, et al. 2000. "IPCC Report on Emissions Scenarios." Cambridge University Press: Cambridge, United Kingdom,.

Nguyen, J. L., J. Schwartz, D. W. Dockery, 2014. The relationship between indoor and outdoor temperature, apparent temperature, relative humidity, and absolute humidity. *Indoor Air*, **24**, 103-112.

Oke, T.R., 2006. *Initial Guidance to Obtain Representative Meteorological Observations at Urban Sites*. World Meteorological Organization, Instruments and Observing Methods, IOM Report No. 81, WMO/TD-No. 1250.

Patz, J.A., H. Frumkin, T. Holloway, D.J. Vimont, A. Haines, 2014. Climate change challenges and opportunities for global health. *Journal of the American Medical Association*, **312**, 1565-1580.

Pielke, R.A, Sr., 2002. *Mesoscale meteorological modeling*. 2nd Edition, Academic Press, San Diego, 676 pp.

Racherla, P. N.; Adams, P. J. (2008) The response of surface ozone to climate change over the Eastern United States ATMOS CHEM AND PHYSICS Volume: 8 Issue: 4 Pages: 871-885.

Rowell DP. 2005. A scenario of European climate change for the late 21st century: seasonal means and interannual variability. Climate Dynamics 25: 837–849.

Roeckner E, Bengtsson L, Feichter J, Lelieveld J, Rodhe H. 1999. Transient climate change simulations with a coupled atmosphere ocean GCM including the tropospheric sulfur cycle. Journal of Climate 12: 3004–3032.

Sarkar, A., and K. De Ridder, 2011. The urban heat island intensity of Paris: a case study based on a simple urban surface parameterisation. *Boundary-Layer Meteorology*, **138**, 511-520.

Schär, C., P.L. Vidale, D. Luthi, C. Frei, C. Haberli, M.A. Liniger, and C. Appenzeller, 2004. The role of increasing temperature variability in European Summer heat waves. *Nature*, **427**, 332-336.

Smargiassia, A., M. Fournier, C. Griot, Y. Baudouin, and T. Kosatsky, 2008. Prediction of the indoor temperatures of an urban area with an in-time regression mapping approach. *Journal of Exposure Science and Environmental Epidemiology*, **18**, 282–288.

Spiridonov V., Somot S., Déqué M. (2005) : ALADIN-Climate : from the origins to present date. ALADIN Newsletter n. 29 (nov. 2005)

Steppeler J, Doms G, Sch¨attler U, Bitzer HW, Gassmann A, Damrath U, Gregoric G. (2003). Meso-gamma scale forecasts using the nonhydrostatic model LM. Meteorology and Atmospheric Physics 82: 75–96.

Viana M, Querol X, Alastuey A, Gangoiti G, Menéndez M (2003) PM levels in the Basque Country (Northern Spain): analysis of a 5-year data record and interpretation of seasonal variations. Atmos. Env. 37:2879-2891

Wieringa, J., 1986. Roughness-dependent geographical interpolation of surface wind speed averages. *Quarterly Journal of the Royal Meteorological Society*, **112**, 867–889.