Abstract—Poor air quality is one of the main environmental causes of premature deaths worldwide, and mainly in cities, where the majority of the population lives. It is a consequence of successive land cover (LC) and use changes, as a result of the intensification of human activities. Knowing these landscape modifications in a comprehensive spatiotemporal dimension is, therefore, essential for understanding variations in air pollutant concentrations. In this sense, the use of air quality models is very useful to simulate the physical and chemical processes that affect the dispersion and reaction of chemical species into the atmosphere. However, the modelling performance should always be evaluated once the resolution of the input datasets largely dictates the reliability of the air quality outcomes. Among these, the updated LC is an important parameter to be considered in atmospheric models, since it takes into account the Earth’s surface changes due to natural and anthropic actions, and regulates the exchanges of fluxes (emissions, heat, moisture, etc.) between the soil and the air. This work aims to evaluate the performance of the Weather Research and Forecasting model coupled with Chemistry (WRF-Chem), when different LC classifications are used as an input. The influence of two LC classifications was tested: i) the 24-classes USGS (United States Geological Survey) LC database included by default in the model, and the ii) CLC (Corine Land Cover) and specific high-resolution LC data for Portugal, reclassified according to the new USGS nomenclature (33-classes). Two distinct WRF-Chem simulations were carried out to assess the influence of the LC on air quality over Europe and Portugal, as a case study, for the year 2015, using the nesting technique over three simulation domains (25 km², 5 km² and 1 km² horizontal resolution). Based on the 33-classes LC approach, particular emphasis was attributed to Portugal, given the detail and higher LC spatial resolution (100 m x 100 m) than the CLC data (5000 m x 5000 m). As regards to the air quality, only the LC impacts on tropospheric ozone concentrations were evaluated, because ozone pollution episodes typically occur in Portugal, in particular during the spring/summer, and there are few research works relating to this pollutant with LC changes. The WRF-Chem results were validated by season and station typology using background measurements from the Portuguese air quality monitoring network. As expected, a better model performance was achieved in rural stations: moderate correlation (0.4 – 0.7), BIAS (10 – 21 μg·m⁻³) and RMSE (20 – 30 μg·m⁻³), and where higher average ozone concentrations were estimated. Comparing both simulations, small differences grounded on the Leaf Area Index and air temperature values were found, although the high-resolution LC approach shows a slight enhancement in the model evaluation. This highlights the role of the LC on the exchange of atmospheric fluxes, and stresses the need to consider a high-resolution LC characterization combined with other detailed model inputs, such as the emission inventory, to improve air quality assessment.

Keywords—Land cover, tropospheric ozone, WRF-Chem, air quality assessment.

I. INTRODUCTION

LAND COVER (LC) and land use are two distinct terms which often appear associated but are commonly confused. LC identifies the different land types (e.g. urban, forest) of a given region, whereas the land use documents how people are using the land [1]. Combining information from both in a wide spatiotemporal spectrum is crucial for analyzing evolutionary dynamics of landscape patterns [2]–[4]. In this context, the use of remote sensing technology has been determined for obtaining these physical parameters and other surface-based satellite products over long time series at a global scale [5]. In turn, these changes on the Earth's surface and trends, largely related with the accelerated population growth and increasing needs of urbanization, impact the air quality due to changes in biogenic and anthropogenic emissions, heat and energy balances, urban climate, and dry deposition of air pollutants [3], [4], [6]–[10]. In order to evaluate the influence of LC and land use changes and their intensity (e.g. buildings density and their volumetry, green coverage), mesoscale atmospheric models [11], [12], regression models [2], [3], [7], [13], and statistical analyses [9], [14], have been employed for quantifying present and future impacts on air quality. The most of these studies are reported to indirect and direct effects of specific activities occurring urban areas (e.g. traffic emissions, energy use) on particulate matter (PM) concentrations, but the interference of these air pollution levels on the surface’s physical properties is also evaluated. Nevertheless, despite the efforts to quantitatively relate built-up areas and air quality, the relationship with other LC categories and other air pollutants has seldom been a concern. One of the exceptions is the work developed by [15], which evaluated the influence of the increased urban LC and associated changes in ozone concentrations in a future scenario, and they concluded that LC changes can lead to ozone variations. This variability in the ozone concentrations may be derived from numerous factors, such as, emission of the main ozone precursors (nitrogen oxides (NOx) and non-methane volatile organic compounds (NMVOC)) and increasing impervious surfaces in
cities expanding the urban heat island effect. Higher urban temperatures generally result in higher ozone levels due to an increased ground-level ozone production. In other words, the spatial distribution of ozone concentrations is the result of complex processes involving primary pollutant emissions, meteorology and photochemical reactions [16]–[18]. Conversely, increasing ozone concentrations can affect the vegetation, changing its composition and LC structure, and produce an indirect radioactive forcing effect [4].

This study aims to investigate the effect of two LC classifications: i) 24-classes USGS and ii) 33-classes USGS reclassified from CLC and specific high-resolution LC data for Portugal, on tropospheric ozone concentrations over Portugal. This highlight is based on two-fold conditions: due to ozone pollution episodes typically occurring in Portugal, in particular during the spring/summer, which leads, sometimes, to non-compliance with existing air quality standards; another important aspect is related with the detail and higher LC spatial resolution for Portugal (100m x 100m) than the reclassified CLC data (5000m x 5000m).

The paper is organized as follows. Section II describes the main configurations used to run the WRF-Chem modelling system and input datasets required, attributing greater relevance to the LC classifications to be tested. The ozone concentration results from two distinct WRF-Chem simulations for the year 2015, changing only the LC database, were compared and discussed in Section III A, using as support Leaf Area Index (LAI) and air temperature data, responsible for boosting NMVOC emissions and favoring the ozone production. To evaluate the behavior of the modelled ozone results, measurements from the Portuguese air quality monitoring network were used (Section III B). Lastly, a few concluding remarks and recommendations to improve the modelling performance are presented in Section IV.

II. MODEL SETUP AND INPUTS

A. WRF-Chem

With the purpose of evaluating the impact of LC changes on air quality, the WRF-Chem model version 3.6.1 was used. WRF-Chem is an online model developed by [19], which has been updated and is freely available to download from the WRF webpage [20], allowing a full integration and calculation in parallel of both components meteorology and chemistry. Hence, the same simulation grids (i.e. horizontal and vertical levels), physical parameterizations, transport schemes and vertical mixing are shared, favoring the meteorology-chemistry feedbacks. For example, chemical reactions are conditioned by the meteorological conditions, whereas the atmospheric aerosol chemistry can affect the radiative balance [21].

For this study, the WRF-Chem was configured to have three simulation domains using two-way nesting with horizontal resolution of 25 km², 5 km² and 1 km², covering from a large part of Europe and North Africa (d01) to a Portuguese region (d03) (Fig. 1). The vertical structure of the atmosphere was resolved with 29 vertical levels extending up to 50 hPa, being that the lowest level is located at an approximate altitude of 28 m above the surface.

Table I summarizes the main physical and chemical options used in the numerical WRF-Chem simulations.

The Earth’s surface works as an overriding driver of all interactions within the planetary boundary layer, where the
physical and chemical processes occur. Accordingly, of the options referenced in Table I, the Noah LSM scheme [22] assumes a prominent role in connection with the LC, since it incorporates vegetation parameters corresponding to annual minimum/maximum values (prescribed at the lookup table VEGPARM.TBL). LAI and emissivity vary in proportion to vegetation fraction, whereas the albedo varies conversely with it. However, there is no a clear evidence showing such linear relationships.

<table>
<thead>
<tr>
<th>Processes</th>
<th>Option</th>
</tr>
</thead>
<tbody>
<tr>
<td>Microphysics</td>
<td>Morrison double-moment</td>
</tr>
<tr>
<td>Short-wave radiation</td>
<td>RRTMG</td>
</tr>
<tr>
<td>Long-wave radiation</td>
<td>RRTMG</td>
</tr>
<tr>
<td>Surface layer</td>
<td>Monin-Obukhov Similarity</td>
</tr>
<tr>
<td>Land-surface model</td>
<td>Noah LSM</td>
</tr>
<tr>
<td>Boundary-layer scheme</td>
<td>MYNN 2.5 level TKE</td>
</tr>
<tr>
<td>Canopy</td>
<td>Grell 3D</td>
</tr>
<tr>
<td>Photolysis</td>
<td>Fast-J</td>
</tr>
<tr>
<td>Gas-phase mechanism</td>
<td>NOAA/ESRL RACM</td>
</tr>
<tr>
<td>Aerosol model</td>
<td>MADE/VBS</td>
</tr>
</tbody>
</table>

B. LC Classifications

The LC arises as the key parameter to be tested in this study, considering two distinct WRF-Chem simulations involving different LC classifications, succinctly described below:

1. 24-Classes USGS (Hereinafter Referred to as 'USGS')

By default, the WRF-Chem implements a 24-classes USGS LC classification, incorporating only one single urban class, called 'Urban and Built-Up Land'. Each LC class is characterized according to the vegetation properties included in VEGPARM.TBL (applicable only for Noah and RUC LSM). Otherwise, for other LSM options, the surface parameters by season (winter and summer) are specified at the lookup table LANDUSE.TBL, including the albedo, soil moisture availability, emissivity, roughness length, thermal inertia and surface heat capacity.

2. Reclassified 33-Classes USGS (Hereinafter Referred to as 'COS-CLC')

The new LC classification combines the CLC dataset containing 44 classes created based on the visual interpretation of 2012 satellite images with a 100m positional accuracy, and greater accuracy and specific LC data for Portugal [23]. These LC data were integrated in Geographic Information Systems, using the ArcGIS software, and reclassified according to the new 33-classes USGS nomenclature following the suggestions of [24]. Due to the high computational requirements for processing this information, the new LC data remapped from CLC were merged to a 5000 m-regular grid, whereas for Portugal, as a case study, a 100 m horizontal resolution was considered. In this LC reclassification process, the inclusion of three different urban classes: low-intensity residential (class 31), high-intensity residential (class 32) and industrial or commercial (class 33) should be highlighted. Moreover, when the goal is to use the WRF-Chem with urban modules, it is mandatory extending the LC database to these three urban classes.

Fig. 2 shows the result for Portugal based on the two LC classifications interpolated for the simulation domain 2 (5 km-grid resolution) by dominant class.

In a preliminary analysis, the USGS LC data commonly used in WRF-Chem (Fig. 2 (a)) showed that the LC for Portugal is not represented well. Concerning the COS-CLC dataset (Fig. 2 (b)), the spatial distribution seems to be more...
closer to reality, so that it is possible to observe the complex fragmentation associated to the landscape, as well as to clearly distinguish some points of interest. However, keep in mind that the interpolation from very fine resolution data to much coarser grid cells leads to considerable losses of detail, not taking the best advantage of the relevance of these inputs. This advice is particularly useful for studies over urban areas, where the need of adjusted urban parameterizations and higher input and output resolutions are essential to improve the modelling performance.

C. Emissions

Anthropogenic emissions from the EMEP (European Monitoring and Evaluation Programme) database [25] with a 0.1° x 0.1° horizontal resolution for the year 2015 were used. This annual emission inventory (IE) is available by GNFR (Gridding Nomenclature for Reporting) including estimated emissions of classic air pollutants (e.g. PM10, NOx, NMVOC), heavy metals and persistent organic pollutants for key activity sectors (e.g. road transport, industry). The horizontal and vertical interpolation for the simulation grids, temporal disaggregation by activity sector considering the seasonality, day of week and daily cycle, as well as speciation and aggregation of emissions into WRF-Chem species were performed using the emissions interface built by [26]. However, to run this emissions processing tool, the IE had to be prepared in the old EMEP format, being more evident the need to establish a correspondence between GNRF and SNAP nomenclatures.

Within the natural sources, biogenic emissions were calculated online, previously using the bio_emiss utility [27], that creates initialization fields (monthly LAI, fraction by Plant Functional Type and emission factors) for computing the MEGAN model (The Model of Gases and Aerosols from Nature – version 2.04) coupled to the WRF-Chem. Further information about this MEGAN model version is referenced in [28].

D. Initial and Boundary Conditions

Meteorological inputs for the coarser domain were forced by ERA-Interim’s global reanalysis data (0.5° x 0.5° horizontal resolution) provided from the ECMWF (European Centre for Medium-Range Weather Forecasts) [29], while the domains, 2 and 3 are driven by the meteorological initial and boundary conditions taken from the domains 1 and 2, respectively.

Time-variant chemical boundary conditions and idealized initial conditions (only used for the first simulation period) were extracted from the MOZART-4/GEOS-5 (The global Model for Ozone and Related Chemical Tracers) [30] with a 1.9° x 2.5° horizontal resolution and 56 vertical levels. Next WRF-Chem runs were initialized using chemical fields corresponding to the last hour of the previous simulation period.

In both global atmospheric models, meteorological and chemical fields were provided to the WRF-Chem at 6 h intervals.

III. RESULTS AND DISCUSSION

The WRF-Chem was applied every 2-days with an hourly resolution for the period of 24th December 2014 until 31st December 2015, taking into account that:

- model outputs for 2014 were discarded as model spin-up; and,
- chemical fields at the end of a 2-day simulation were passed as initial fields for the following simulation.

For the reasons already mentioned, Portugal was the case study, it means, therefore, which the modelling results based on LC changes were extracted from the domain 2 (5 km horizontal resolution). In order to capture vegetation dynamics and understanding how the biogenic NMVOC influence the production of tropospheric ozone, seasonal variations were examined: winter (Jan – Mar); spring (Apr – Jun); summer (Jul – Sep); autumn (Oct – Dec). In addition, also the influence of the Leaf Area Index (LAI) and ambient temperature is discussed, given the close relationship these variables with the LC and the ozone. For assessing the accuracy of the modelled ozone concentrations, observed data from background air quality stations were used.

A. Impact of LC Changes on Air Quality

In this section, sensitivity tests involving LC changes and evaluation of their impacts on air quality are presented. This analysis is focused on Portugal, considering the two LC classifications (see Section II.B) and the seasonally of the modelled variables resulting from the two WRF-Chem simulations for the year 2015.

Fig. 3 illustrates the spatial variations found in modelled ozone concentrations, comparing both LC approaches and seasons. On the right column, USGS-based average ozone concentrations are presented, working as a reference for quantifying the average and maximum differences in relation to COS-CLC methodology. Overall, average ozone levels are higher in spring, indicating for the period under analysis, that there are the most favorable conditions for the formation of this secondary pollutant.

To complement this analysis, as seen in Fig. 4, two main parameters interfering with the tropospheric ozone variability are presented: LAI and air temperature. Only results for the spring and summer are presented, because during these seasons typically occur higher ozone concentrations. As regards to the LAI, calculated as a function of the LC classes, higher values were recorded in spring, in full vegetation growth (i.e. maturity stage), favoring the emission of biogenic NMVOC. Correlating this spatial information with the average air temperature, only mapped for USGS because significant differences in relation to COS-CLC approach were not found, its influence on the average ozone values was not so evident, although it is widely recognized, particularly when analyzing daily profiles. Therefore, for a given region where higher values of temperature and LAI were estimated, the latter assuming a direct relationship with the NMVOC emission, the poor response in ozone production could be related with low NOx levels. In summer, the highest average air temperatures do not linearly correspond to increased ozone production,
because during this season there is a decrease of the vegetation fraction (lower LAI), which leads to lower albedo values, and consequently, higher near-surface air temperatures.

Crossing the spatial distribution of USGS-based average ozone concentrations with the dots representing observations from background monitoring stations, it can be concluded that the modelled seasonal averages reflect properly the reality.

Regarding the spatial differences comparing ozone results from both LC approaches, there is a supremacy of the average values estimated for the spring and summer using the COS-CLC database as an input (up to 3 $\mu$g.m$^{-3}$). During the autumn and winter, the USGS-based ozone data overcame in large part of the study domain those obtained from the COS-CLC approach. These differences are closely associated to the LC classes and annual variability of LAI. In terms of maximum differences, the USGS overestimates the seasonal ozone concentrations (up to 30 $\mu$g.m$^{-3}$) in whole domain. However, these values may reflect single occurrences, being mere indicators of upper extremes.

Analyzing the daily mean ozone profiles by season for the observations from background air quality stations, and USGS and COS-CLC simulation results based on those locations (Fig. 5), the diurnal cycle exhibits a similar trend to average ozone differences. It means that modelled ozone concentrations are slightly higher using COS-CLC LC inputs,
mainly during the spring and summer. In these seasons, the modelled ozone profiles have a constant BIAS throughout the cycle, probably due to the joint influence of the LC classes where the stations are located, and their associated physical parameters (e.g. albedo, LAI), as well as to relatively small variations in the diurnal air temperature. When observed and modelled ozone profiles are compared, as expected, smaller differences in daytime were found, given the preponderance attributed to the solar radiation and air temperature for the ozone photochemistry.

B. Model Evaluation

The model performance for Portugal was evaluated against observations from background air quality stations (rural, suburban and urban) (Fig. 6) with more than 75% of data availability, since these are the most adequate for comparing model results, in particular of ozone, over a 5 km-grid resolution (domain 2). Fig. 7 shows the behaviour of the statistical parameters analysed by station typology and season – year 2015: Pearson’s correlation coefficient, BIAS and Root Mean Square Error (RMSE).

![Fig. 4 (a) Average Leaf Area Index (m²·m⁻²) for the spring and summer using both LC approaches; (b) 2 m average air temperature (°C) for the spring and summer using the USGS LC approach](image)

![Fig. 5 Daily mean ozone profiles for the seasons including observations from background air quality stations and modelled concentrations (USGS and COS-CLC) for those locations](image)
As expected, higher ozone concentrations are typically associated to rural areas (intersect Fig. 3 with Fig. 6) and where the model performance was better: moderate correlation (0.4 – 0.7), and BIAS (10 – 21 µg.m$^{-3}$) and RMSE (20 – 30 µg.m$^{-3}$) are lower. Negative BIAS (Fig. 7 (b)) means that, in general, modelled ozone values in both approaches tend to be overestimated, such as verified in the daily mean ozone profiles (see Fig. 5). The RMSE (Fig. 7 (c)) is an indicator of the model absolute deviation, measuring the dispersion between the observations and modelled data. In turn, the validation for suburban stations had the worst performance, mainly in BIAS and RMSE. The location of these stations around the Greater Porto (Fig. 6), associated to an overestimation in emission of ozone precursors are determinant factors for high deviations in modelled ozone concentrations. For urban stations, mostly located in Lisbon region (Fig. 6), the reasons are similar to those indicated for suburban stations.

Concerning the model evaluation for the seasons, as referred to in Section III.A, emphasis should be attributed to spring and summer, periods where the model results had a better statistical performance.

IV. CONCLUDING REMARKS

In this work, the WRF-Chem model was employed to investigate the impacts of LC changes on tropospheric ozone concentrations using Portugal as a case study. Two distinct WRF-Chem simulations for the year 2015, changing only the LC database: i) 24-classes USGS; and ii) reclassified 33-classes USGS, were performed. The model results by season revealed small differences when comparing both LC approaches, using as support important physical parameters such as the LAI and air temperature, for substantiate those variations. However, it is not so easy to quantitatively model the relationship between LC and ozone concentrations, because the ozone varies at time and space and is influenced by many other factors, such as, emission of their precursors (NMVOC and NOx), transport and dispersion processes, ozone photochemistry and climate.

For a better understanding of the influence of LC changes on air quality, and consequent reduction of modelling uncertainties, some recommendations are presented:

- Making a comprehensive and integrated assessment of the overall impact of LC changes, starting by examining individual effects;
- Testing the model sensitivity to different spatial and thematic resolutions of land use and cover maps;
- Developing and enhancing land use and cover (LULC) models allowing to assess the predicted impact of land use changes on air quality at the city scale and, thus, to evaluate the trade-off among different landuse configurations.
In summary, further research involving land use and cover changes and their implications on the atmospheric environment is required. The use of LULC models could be a promising way for assessing potential human activity impacts on air quality under urban sprawl scenarios, through a cost-benefit balance that helps in urban planning decisions. However, improving the LULC database implies increasing the resolution of the air quality models, preferably for the same resolution of these inputs. This could be challenging, because mesoscale chemical transport models, usually used for regional and urban scales, are not able to go below 1 km resolution, due to the weaknesses in the turbulence schemes and the urban parameterizations are limited or inexistent. In this research line, this is perhaps one of the most challenging issues facing the scientific community. For an effective and immediate response, the use of Computational Fluid Dynamics models to evaluate small-scale atmospheric dynamics is recommended; however, common practice is to apply Gaussian models that even with new developments to simulate urban areas are still limited.

**ACKNOWLEDGMENT**

The authors would like to acknowledge the financial support of the FCT/MEC through national funds, and the co-funding by the FEDER, within the PT2020 Partnership Agreement and Compete 2020, for the research project FUTURAR (PTDC/AAG-MAA/2569/2014 - POCI-01-0145-FEDER-016752), CESAM (UID/AMB/50017 - POCI-01-0145-FEDER-007638) and the PhD grant of C. Silveira (SFRH/BD/112343/2015) and the post-doc grant of J. Ferreira (SFRH/BPD/100346/2014).
REFERENCES


