Simulation of organic aerosol, its precursors and related oxidants in the Landes pine forest in south-western France: Need to account for domain specific land-use and physical conditions

Arineh Cholakian1,2,4, Matthias Beekmann2, Guillaume Siour3, Isabelle Coll3, Manuela Cirtog3, Elena Ormeño5, Pierre-Marie Flaud1, Emilie Perraudin1, and Eric Villenave1

1 Univ. Bordeaux, CNRS, EPOC, EPHE, UMR 5805, F-33405 Talence Cedex, France
2 Université de Paris Cité and Univ Paris Est Créteil, CNRS, LISA, F-75013, Paris, France
3 Univ Paris Est Créteil and Université de Paris Cité, CNRS, LISA, F-94010, Créteil, France
4 LMD UMR CNRS 8539, ENS, Ecole Polytechnique, Institut Pierre Simon Laplace (IPSL), Palaiseau, France.
5 CNRS, Aix Marseille Univ, IRD, Avignon Univ, IMBE, 13397 Marseille, France.

Correspondence: Arineh Cholakian (arineh.cholakian@lmd.ipsl.fr)

Abstract. Organic aerosol (OA) still remains one of the most difficult components of the atmospheric aerosols to simulate, given the multitude of its precursors, the uncertainty of its formation pathways and the lack of measurements of its detailed composition. The LANDEX project (The LANDes Experiment), during its intensive field campaign in summer 2017, gives us not only the opportunity to compare biogenic secondary OA (BSOA), but also its precursors and oxidants obtained within and above the Landes forest canopy, to simulations performed with CHIMERE, a state of the art regional Chemistry-Transport Model. The Landes forest is situated in the south-western part of France, and is one of the largest anthropized forests in Europe (1 million ha), composed by a majority of maritime pine trees, strong terpenoid emitters, providing a large potential for biogenic SOA formation. In order to simulate OA build-up in this area, a specific model configuration set-up, adapted to the local peculiarities was necessary. As the forest is inhomogeneous, with interstitial agricultural fields, high-resolution 1 km simulations over the forest area were performed. BVOC emissions were predicted by MEGAN, but specific land cover information needed to be used, chosen from the comparison of several high-resolution land cover databases. Moreover, the tree species distribution needed to be updated for the specific conditions of the Landes forest. In order to understand the canopy effect in the forest, canopy effects on vertical diffusivity, winds and radiation were implemented in the model in a simplified way. The refined simulations show a redistribution of BVOCs with a decrease in isoprene and an increase in terpenoid emissions with respect to the standard case, in line with observations. Corresponding changes on simulated BSOA sources are tracked. Very low night-time ozone, sometimes near zero, remains overestimated in all simulations. This has implications to the night-time oxidant budget, including NO3. Despite careful treatment of physical conditions, simulated BSOA is overestimated in the most refined simulation. Simulations are also compared to air quality sites surrounding the Landes forest, reporting a more realistic simulation in these stations in the most refined test case. Finally the importance of the see breeze system which also impacts species concentrations inside the forest is made evident.
1 Introduction

Forested areas - either natural or artificially managed - induce different biosphere-atmosphere interactions, leading to complex physico-chemical processes that have a significant impact on the composition of the atmosphere on regional and global scales. They are responsible for 90% of volatile organic compound (VOC) emissions worldwide (Delmas et al., 2005), mainly isoprene, but also monoterpenes, and other VOCs (Guenther et al., 2012). The initial oxidation steps of these compounds consume OH and NO$_3$ radicals, and ozone (O$_3$) (Seinfeld and Pandis, 2016). When undergoing rapid oxidation processes, biogenic VOCs (BVOCs) generate various oxidation products, which may affect the atmospheric oxidative budget (Hallquist et al., 2009). Part of these products have sufficiently low vapor pressures or large enough Henry constants to further form biogenic secondary organic aerosols (BSOAs). Depending on the availability of NO$_x$, the associated oxidation pathways may lead to O$_3$ formation. Of course, BVOC oxidation processes are strongly affected by anthropogenic emissions, especially NO$_x$ availability (e.g. Sartelet et al. (2012); Shrivastava et al. (2019)). The complexity of this system and the eventual impact of their oxidation products on air quality and human health make forested areas a good environment to isolate and focus on only specific parts of the underlying chemical mechanisms.

While this system has been investigated in many studies for different types of forests (e.g. Hellén et al. (2018) for boreal forests; Shrivastava et al. (2019) for tropical forests), uncertainties still subsist in the quantification of the mixture of emitted BVOC species, the formation mechanisms and yields of secondary products (Hallquist et al., 2009), the interplay of biogenic with anthropogenic emissions (e.g. Xu et al. (2015)), as well as the impact of produced aerosols on regional climate (Gray Bé et al. (2017), via cloud activation of BSOA species) or on global climate (Kulmala et al. (2004); Sporre et al. (2019)).

Hantson et al. (2017) estimate that, for a 30-year period from 1971-2000, the global emissions of isoprene and monoterpenes are around 300 TgC yr$^{-1}$ and 26 TgC yr$^{-1}$ respectively. While this means that globally isoprene is the major precursor for the formation of BSOAs due to the sheer amount of its emission, monoterpenes have to be considered as well due to higher BSOA formation yields, varying from 10 to 60% (Griffin et al. (1999); Lee et al. (2006a); Lee et al. (2006b); Ng et al. (2007)). For example, for the SOAS study in summer 2013 in Southeastern US, Xu et al. (2015) found a more than 50% contribution of the monoterpane + NO$_3$ reaction pathway to night-time BSOA formation. As NO$_3$ is generated from the NO$_2$ + O$_3$ reaction, this latter pathway provides one of the possible links between anthropogenic and biogenic emissions. Of course, this distinction between precursors becomes more apparent locally, depending on the ecosystem characteristics in each region (types of trees, climate, etc.), and is also mediated by the oxidant availability.

This study aims to put into perspective that, while we dispose of all the information needed to simulate the global state of BVOCs and their oxidation products, locally this is not the case. At the local level, numerical models require the use of detailed information of land use, emission rates, chemical gas phase, aerosols schemes and aerosol – climate interaction mechanisms. We will be focusing on one specific ecosystem, in an effort to pinpoint the local characteristics of the Landes forest and their effect on the local air quality and atmospheric chemistry.
The Landes forest is one of the largest European forests, covering an area of about 1 million hectares in southwestern France. Its dominant tree species is maritime pine, Pinus pinaster Aiton., which is known to be a strong $\alpha$ and $\beta$-pinene emitter (Simon et al., 2001). Due to its rather homogeneous character, the Landes forest is an interesting region to study and to model BVOC-Oxidants-SOA-climate interactions. The LANDEX project (LANDEX, Landes Experiment: Formation and fate of secondary organic aerosols generated in the Landes forest) aimed at characterizing secondary organic aerosol formation observed in this monoterpane-rich environment. In the frame of LANDEX, a field campaign bringing together a dozen French and international partners was held in June and July 2017 at an instrumented site (Salles-Bilos) within the forest. Special care was taken to extensive characterization of BVOC species (Mermet et al., 2019), to measurements of radical species and of OH reactivity (Bsaibes et al., 2020), and to characterization of organic aerosol (to be published). (Mermet et al., 2021) analyzed BVOC and radical/oxidant species abundance at Salles-Bilos and concluded that a sesqui-terpene, $\beta$-caryophyllene, reaction with $O_3$ most strongly contributes to BVOC reactivity within the forest canopy. Frequent episodes of night-time new particle formation (NPF) have been observed during the campaign, but also in prior to it and has been linked to BVOC oxidation processes (Kammer et al. (2018); Kammer et al. (2020)). During the summer 2018 follow-up Cervioland experiment (a spin-off of LANDEX), Li et al. (2020) made evident the presence of very low volatile suitable precursors (diterpene species) by high resolution mass spectrometry analyses. Campaign results are still under active investigation. The LUCAS project (Land use, regional climate and atmospheric chemistry: the impact of forested surfaces on cloud enhancement and air quality in south-west France) within the LABEX COTE (https://cote.labex.u-bordeaux.fr/, last accessed: 29 july 2022) extends LANDEX objectives by aiming at investigating to which extent SOAs in the Landes forest may impact regional climate and precipitation through formation of cloud condensation nuclei. A second goal is to study the impact of future climate change and forest management practices on the regional BVOC-SOA system.

A prerequisite for addressing objectives of the LUCAS project is to set-up a suitable modelling framework for addressing BVOC emissions and subsequent SOA formation specifically tailored for the Landes forest. Our paper aims at constructing such a framework, by using the well referenced French and European CHIMERE model (Menut et al. (2013); https://www.lmd.polytechnique.fr/chimere/, last accessed: 31 july 2022), and to add subsequently detailed information for the Landes area. This includes specific information on land use, on tree species distribution, on BVOC emission factors, on detailed anthropogenic emissions, and on a specific dynamical treatment of the forest canopy. The paper carefully records the impact of these updates on atmospheric composition, BSOA formation pathways and how they improve (or not) agreement with observations from the LANDEX 2017 campaign. A second goal of this work is to use the CHIMERE model to provide 2D concentration fields at the surface, on the one hand to interpret the spatial representativeness of the Salles-Bilos measurement site, and on the other hand to provide insight into the link between transport and chemical processes.

The paper is organized as follows. After this introduction, section 2 provides information on the campaign measurements and air quality network observations used for the study. Section 3 describes successive inclusion of information specific to the Landes forest in the standard version of CHIMERE. Section 4 evaluates the induced changes in the time series of the different species and explores BSOA formation pathways. Section 5 presents a case study showing that both chemistry and transport processes affect species concentration at the measurement site. In the end, conclusions and perspectives are presented.
Figure 1. Measurements and monitoring stations in the Landes region and the high resolution simulation domain (horizontal resolution of 1 km). Black dots represent where the used air quality stations are located, the red circle shows the main measurement site in Salles-Bilos. © OpenStreetMap contributors 2022. Distributed under the Open Data Commons Open Database License (ODbL) v1.0.

2 Measurement sites and their characteristics

The data used here concern an area in and around the Landes forest in south-western France, in vicinity of the Atlantic coast. Detailed measurements were performed during the summer of 2017 (LANDEX episode 1) in a measurement site located at Salles-Bilos, the exact period of the measurements being from 20/06/2017 to 20/07/2017 depending on the availability of measurements for each studied component. More information about this site is given in the next section.

Additionally, 10 air quality monitoring sites – from the ATMO national network - located around the forest have been used for validation of more common species (NO$_x$, PM$_{10}$, PM$_{2.5}$, Figure 1). Meteorological fields (explained in detail in section 3.3) used in the simulations have been validated using the E-OBS database as well as the data provided by the measurement site at Salles-Bilos.

2.1 Characteristics of the main measurements site and measurements performed

The measurements site is located in Salles-Bilos (44°29’39.69”N, 0°57’21.75”W, 37 m a.s.l), about 50 km in the south west of Bordeaux. It is located in the middle of the Landes forest (Figure 1), a large forest with a majority of maritime pine (Pinus pinaster) trees. The forest parcels are intermittent with agricultural fields, and affectation can change during the years. The
Table 1. Summary of changes in each test case. Each change has been explained in more detail in sections 3.1 through 3.6. Treetype/EF column refers to changes in tree type inputs and their respective emission factors. The wind/swrd/Kz column refers to the test case in which the inside canopy wind speed, short wave radiation, and vertical exchange coefficient parameters are modified.

<table>
<thead>
<tr>
<th>Simulations</th>
<th>Meteorology</th>
<th>Land cover</th>
<th>MEGAN</th>
<th>Treetype/EF</th>
<th>Anthropogenic inputs</th>
<th>wind/swrd/Kz</th>
</tr>
</thead>
<tbody>
<tr>
<td>Base simulations</td>
<td>WRF</td>
<td>Globcover</td>
<td>MEGANv2.1</td>
<td>Not included</td>
<td>EMEP 10 km × 10 km</td>
<td>Not included</td>
</tr>
<tr>
<td>(section 3.2)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Case 1</td>
<td>ECMWF</td>
<td>Globcover</td>
<td>MEGANv2.1</td>
<td>Not included</td>
<td>EMEP 10 km × 10 km</td>
<td>Not included</td>
</tr>
<tr>
<td>(section 3.3)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Case 2</td>
<td>ECMWF</td>
<td>Theia 2018</td>
<td>MEGANv3</td>
<td>Not included</td>
<td>EMEP 10 km × 10 km</td>
<td>Not included</td>
</tr>
<tr>
<td>(section 3.4)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Case 3</td>
<td>ECMWF</td>
<td>Theia 2018</td>
<td>MEGANv3</td>
<td>Included</td>
<td>EMEP 10 km × 10 km</td>
<td>Not included</td>
</tr>
<tr>
<td>(section 3.4)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Case 4</td>
<td>ECMWF</td>
<td>Theia 2018</td>
<td>MEGANv3</td>
<td>Included</td>
<td>ATMO-NA 1 km × 1 km</td>
<td>Not included</td>
</tr>
<tr>
<td>(section 3.5)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Case 5</td>
<td>ECMWF</td>
<td>Theia 2018</td>
<td>MEGANv3</td>
<td>Included</td>
<td>ATMO-NA 1 km × 1 km</td>
<td>Included</td>
</tr>
<tr>
<td>(section 3.6)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

specific parcel where the measurements took place consists of pine trees planted in 2004. The height of the trees was around 10 m in 2017. Regarding biogenic emissions, the site is quite homogeneous since the majority of the trees in the surrounding area are maritime pines. However, as mentioned above, since the forest is parcellated the density of the forest and the geographical distribution of these emissions have a certain degree of heterogeneity.

The intensive field campaign was carried out just next to the Salles-Bilos ICOS station, part of the European ecosystem monitoring network, with detailed measurements of greenhouse gases and meteorological parameters are obtained at 15m above ground on a measurement tower, outside of the forest canopy (https://xylofront.pierroton.inra.fr/Salles2.html; last accessed July 2022; Moreaux et al. (2011)). During the LANDEX campaign, a large set of gaseous and aerosol measurements was added inside the forest canopy, at 6 m height (Bsibes et al. (2020); Mermet et al. (2021)). For the purpose of this study, among gaseous species the measurements of \( \text{O}_3 \), nitrogen oxides (\( \text{NO}_x \)) and biogenic volatile organic compounds (BVOCs) will be used, and among the particulate compounds, the concentrations of \( PM_{10} \), \( PM_{2.5} \) and organic aerosols will be considered. A summary of the measurements used at this site is shown in Figure 1.

\( \text{O}_3 \) measurements were performed inside the forest canopy, with the APOA-370 (HORIBA) UV absorption instrument using a temporal resolution of 5 minutes.

\( NO \) and \( \text{NO}_2 \) measurements were performed by chemiluminescence, using an APNA 370 (HORIBA) with a resolution of 5 minutes and the same height as the \( \text{O}_3 \) measurements. This measurement method being not specific for \( \text{NO}_2 \), resulting \( \text{NO}_2 \)
measurements may be overestimated especially, on a relative scale, for background conditions when much of NOy species are oxidized.

VOCs were measured with an ensemble of different instruments (Mermet et al. (2019), Bsaibes et al. (2020), Mermet et al. (2021)). A PTR-Tof-MS (Proton Transfer – Time of Flight – Mass Spectrometer) allowed to measure the total concentration of terpenoids (isoprene, mono and sesquiterpenes) below the forest canopy at the same height as the previous measurements (around 6 m). Another PTR-Tof-MS was utilized for measurements above the canopy. In addition, GC-MS measurements were also performed for characterizing BVOC speciation. In this work, the sum of the terpene species resulting from these measurements as well as the individual concentrations of α-pinene, β-pinene, limonene and ocimene are used. Since the purpose of the article is to focus on pollutant build-up from biogenic compounds, and because of their impact on atmospheric chemistry over the region, anthropogenic VOC measurements are not explored.

Attempts to measure radical species were also made, specifically for OH (using a FAGE instrument from the PC2A partner; Amedro et al. (2012)) and NO3 (IBBCEAS technique from the LISA laboratory – University Paris-Est Creteil, Fouqueau et al. (2020)). Because of technical problems, the measurements for OH were not available to be used in this paper. Since the observed NO3 concentrations were too low to produce a detectable signal on the instrument, they were considered for this entire study as below 3-5 pptv, which corresponds to the detection limit of the instrument.

Particulate matter was also measured during the campaign. PM10/PM2.5/PM1 concentrations were measured using a traditional TEOM-FDMS microbalance (implemented by EPOC) and a FIDAS 200 (deployed by INERIS) analyzer, which measures total aerosol concentration in the range of 0.18 - 100 µm (with 3 measuring ranges). INERIS also utilized an aethalometer at the same time as the FIDAS in order to measure the black carbon (BC) concentrations, identifying a strong biomass burning episode (on the 05/07 to 06/07) near the measurement site, producing a peak of BC with concentrations as high as 80 µg.m⁻³. The simulation of fire episodes using inputs from satellite data is implemented in CHIMERE, but this option was not activated for the runs performed in this study. Therefore, and since SOA build-up from biomass burning is not the aim of this paper, this peak has been removed from the PM data produced by the FIDAS.

A high-resolution time-of-flight AMS (HR-ToF-AMS, Aerodyne Research Inc.; DeCarlo et al. (2006)) was used in order to measure the bulk chemical composition of the non-refractory fraction of the aerosol operated under standard conditions (i.e., temperature of the vaporizer set at 600 C, electronic ionization (EI) at 70 eV) with a temporal resolution of 8 min. The concentration of organic aerosol in the PM1 fraction resulting from this instrument is used in this study.

2.2 Other measurement data

The latest version of the E-OBS data was used for large scale meteorological validations, for the European and the French domains (Cornes et al. (2018); https://www.ecad.eu/, last accessed 28/07/2022). These data provide daily regridded information for temperature (daily average, minimum and maximum), wind speed and direction, relative humidity, precipitation, and solar radiation for around 5000 stations for temperature related variables and around 1800 for other parameters covering the current day period going back to 1950s. The data can be downloaded in a regridded format (with a 0.25° × 0.25° horizontal resolution) or point data (as in data for each station); the second option is used in this study.
For more comparisons, data from 10 air quality monitoring stations from the ATMO network was used. The location of these stations, generally at the edge of the forest, is shown in Figure 1. The measurements available for most of these stations include nitrogen oxides (\(NO\) and \(NO_2\)) and PM (\(PM_{10}\) and \(PM_{2.5}\)) mass concentrations. These datasets were provided by the local air quality monitoring agency, the ATMO-NA (Atmo-Nouvelle-Aquitaine, https://www.atmo-nouvelleaquitaine.org/, last accessed: 19/07/2022), information on the instruments used for the measurements is provided on their site and historic data is provided upon request. The data covers the entire period of the campaign.

3 The modeling chain

In this section, the simulations and potential improvements or sensitivity tests will be presented in detail, in each case documenting the choices made for the changes. The parameters common to all sensitivity tests are explained in section 3.1. After the description of the base simulation (section 3.2), sections 3.3 through 3.6 describe sensitivity tests changing inputs/variable calculations. Keep in mind that each modification is added on top of the previous ones. The implemented improvements are indicated in Table 1.

3.1 CHIMERE chemistry transport model

The CHIMERE offline regional chemistry transport model (Menut et al., 2013) was developed initially in the early 2000s in order to simulate the concentrations of gaseous species (especially \(O_3\)). Later on, a module for the simulation particulate matter (Pun and Seigneur (2007); Bessagnet et al. (2008)) was added to the model. The model is currently used extensively in air quality monitoring and forecast in research as well as for operative purposes both in regional and hemispheric levels (Cholakian et al. (2019b); Lachatre et al. (2019); Trewhele et al. (2019); Lapere et al. (2020)). It has also been used in many model intercomparison studies, as well as model-observation comparisons within multi-model experiments, for example to investigate European particulate matter trends (Ciarelli et al., 2019). Mandatory input information includes meteorological fields, land cover parameters, biogenic/anthropogenic emissions factors and boundary/limit conditions. Having access to this data, CHIMERE then simulates 3D concentration and deposition fields for a list of gaseous and size resolved particulate species, depending on the selected chemical scheme. In this study, the 2017 version of the model has been used (Mailler et al., 2017). In the following section, information about the parameterization used for this model is provided.

3.2 Base case simulation

The simulations conducted in this study were performed on 3 domains: a continental domain with a horizontal resolution of 25 km, covering the whole Europe and the northern Africa (Figure SI-1), an intermediary nested domain focused on France with a horizontal resolution of 5 km (Figure SI-1) and a 1 km horizontal resolution domain focused on the Landes forest nested inside the intermediary domain (Figure 1). The vertical resolution of the simulations is the same for each domain, 15 levels starting from 13 m to about 12 km (a.s.l.). Only for the canopy parameterization case (section 3.6) vertical levels were modified. The simulations were performed for the period of 01-06-2017 to 20-07-2017, with a spin-up period of 5 days.
The SAPRC-07A (Carter, 2010) chemical scheme has been used for all the simulations since it provides more details for the terpenoid oxidation than the previously used Melchior scheme. The aerosol size bins are the same for all simulations, a 10-bin logarithmic sectional distribution in a range of 40 nm to 40 \( \mu \text{m} \); the chemical speciation of aerosols contains EC (elemental carbon), nitrates \((NO_3^-)\), sulfates \((SO_4^{2-})\), ammonium \((NH_4^+)\), primary organic aerosols \((POA)\) and secondary organic aerosols \((SOAs)\), dust, sea salt and PPM (Primary mineral Particulate Matter except the ones mentioned above).

The selected \(SOA\) scheme allows differentiating between about 10 families of semi-volatile \(BVOC\) oxidation products with different composition, volatility, and solubility and created in a single oxidation step (Pun and Seigneur (2007); Bessagnet et al. (2008)). It takes into account initial oxidant attack from \(OH\), \(O_3\) and \(NO_3\). The scheme has been used and evaluated in multiple occasions (Lemaire et al. (2016); Cholakian et al. (2018); Cholakian et al. (2019a)). Boundary/initial conditions are taken from climatological simulations of LMDz-INCA3 (Hauglustaine et al., 2014) for gaseous species and GOCART (Chin et al., 2002) for particulate matter.

For each CHIMERE simulation domain, the meteorological parameters were obtained by running the WRF model (version 3.9.1.1; Wang et al. (2015)) on the same domain and with the same horizontal resolution, and using NCEP large scale input data (for Environmental Prediction/National Weather Service/NOAA/US Department of Commerce, 2000). Several model configurations have been tested for WRF. A comparison of the model outputs with the E-OBS database (https://www.ecad.eu) and with the measurements obtained at the Salles-Bilos measurement site was conducted for all the performed meteorological runs. The results are presented in SI-2. For the base simulations, the WRF configuration showing the best comparison results (as detailed in the SI) with the Salles-Bilos site was chosen (see Figure 2 for time series and statistical information). All the parameterizations tested for WRF are presented in SI-2. The WRF parameterization that is retained is a two-way triple nested run with WSM 6-class graupel microphysics scheme, Kain-Fritsch cumulus scheme, RRTMG scheme for both shortwave and longwave radiation, Noah-MP land-surface, MYNN surface layer option, MYNN 3\(^{rd}\) level TKE boundary layer height option, topological wind, urban physics and lake physics options and nudging options activated. These options are all described in detail in the user guide for the Weather Research and Forecast (WRF) model (https://www2.mmm.ucar.edu/wrf/users/docs/user_guide_V3/user_guide_V3.9/ARWUsersGuideV3.9.pdf, last accessed: 19 july 2022).

The Globcover (Arino et al., 2008) dataset with a resolution of 300 m×300 m was used for the land cover parameters. Bearing in mind that the forest is parcellated and that some parcels can be transformed into agricultural land, changing this database to a more recent one was considered as beneficial for this study (see 3.4).

Biogenic emissions are provided by a reduced online version of the MEGAN (Guenther et al., 2012) model already having been integrated in the CHIMERE model. For the base simulations the version 2.04 of the MEGAN model has been used. This version of MEGAN comes with pre-calculated \(BVOC\) emission factors for isoprene, \(\alpha\)-pinene, \(\beta\)-pinene, limonene, ocimene, humulene and \(\beta\)-caryophyllene (also including the leaf area index (LAI) fields) with a horizontal resolution of 0.008 \(\degree \times 0.008\); therefore, any land cover related changes in the model will not impact the \(BVOC\) emission factors. These emission factors were calculated using the Globcover land cover database. In a later sensitivity test (section 3.4), a more recent version of the MEGAN model has been tested. Using this version, we have the possibility of recalculating the \(BVOC\) emission factors by
changing the land cover, the vegetation-specific emission factors and the LAI. A more detailed explanation of this model and its use is provided in section 3.4.

Anthropogenic emissions were taken from the EMEP (EEA) emission database with a 10 km resolution for the year 2014. The emissions in this data base for the years 2014, 2016 and 2017 were compared and no major differences were observed in the levels of anthropogenic emissions between these years for this specific region (the data can be directly compared on the EMEP emissions datacenter accessible here: https://www.ceip.at/webdab-emission-database/emissions-as-used-in-emep-models, last consulted in 13 July 2022). The reason for using the 2014 emissions is explained in detail in section 3.5, where a local high resolution emission data base is tested.

Figure 2. Meteorological comparisons for the 2 meteorological inputs: the blue series (SIM1) show the ECMWF simulations and the red ones (SIM2) the WRF outputs. Statistical information is provided in the legend for each figure (R correlation coefficient, B bias). The name of each parameter is written on the side of each panel, points on the right side of each panel showing the average data for each series. The comparisons are performed for the 1km domain and for the Salles-Bilos measurement site. The comparisons are performed on the surface level.
It is worth mentioning here that the CHIMERE model does not have an integrated forest canopy model; therefore, the model, while taking into account the biogenic emissions and the land cover information of the forest in different modules (for example for calculating the deposition of atmospheric species), will not take into account the dynamic physical changes introduced by the forest canopy into the concentrations of chemical species in the first two vertical layers of the simulations. These effects and the changes made in the model in order to represent them are presented in detail in section 3.6.

### 3.3 Meteorological inputs

In addition to the tested WRF parameterizations, another series of meteorological inputs was tested using the ECMWF high-resolution 3 hourly forecast dataset (Owens and Hewson, 2018). The global version of the ECMWF data (with lower resolution) is used for the continental domain, while for the other two domains the high-resolution (10 km) ECMWF run is used. The comparisons of the measurements from the Salles-Bilos site to a simulation run with the ECMWF meteorological inputs is shown in Figure 2, as well as the comparison for the WRF parameterization mentioned in section 3.2. Before commenting on this comparison, we first briefly present the meteorological situation during the campaign which is depicted by both models in a similar way. Rainfall events (not shown in Figure 2) occurred between June the 26th and July the 3rd and also in the July 8th to 13th 2017 period and were associated with low solar irradiance (< 750 W.m\(^{-2}\)), mild summer temperatures (< 25 °C), high relative humidity (> 75%) and moderate winds. These periods are defined as "low oxidation" periods. Two episodes of sunny conditions were observed during the campaign, i.e. from July the 4th to the 7th then from the 14th to the 18th. These periods were characterized by higher temperatures, relative humidity below 50% (during the day), low wind speeds (< 3 m.s\(^{-1}\)) and friction velocity (< 0.3 m.s\(^{-1}\)) and no precipitation. These periods were favorable to plant emissions and strong oxidant formation and are called periods of "strong oxidation". It appears that models adequately describe the broad meteorological situation during the campaign summarized by Mermet et al. (2021).

While both meteorological runs here represent the measurements site on the smallest domain quite accurately, the ECMWF run shows a better representation for the two bigger domains for all the tested parameters (mean, maximum and minimum temperature, wind speed and direction and relative humidity, see SI-2), thus providing more accurate meteorological patterns at the boundaries of the high-resolution domain. As for the high-resolution domain, the daily mean and maximum temperature values are well simulated in both meteorological inputs. However, both models have shortcomings in the representation of daily minimum temperature values: this problem is more significant in the WRF simulations (on daily minima, a bias of +0.22°C and +1.22°C is observed in the ECMWF and WRF data respectively). The wind speed simulations are accurate in both runs, while the wind direction is more representative of the domain in the ECMWF simulations. The friction velocity (u\(^*\)) shows the same pattern and is well represented in both datasets, although the night-time minimum is better represented in the WRF simulations. Precipitation and the Bowen ratio \(\beta\), the ratio of heat flux to moisture flux near the surface, were also compared, and are presented in SI-3.
3.4 Land cover, tree type and tree-specific emission factors

As mentioned above since the forest is parcellated and goes through annual changes, it is important to use an up-to-date source of information on forest cover and density when performing high resolution simulations. To this end, a detailed comparison of existing land cover data was conducted. It should be noted that, in most land cover databases, trees are lumped in two main functional groups: broadleaf and needleleaf trees. In addition to tree density (which varies upon the selected database), the broadleaf/needleleaf distribution is also a specific parameter of a land cover datasets. This is why, in this exercise, all databases have been examined with regard to their total tree density (this section), but also with respect to the BVOC emissions they induce, with respect to the broadleaf/needleleaf distribution (next section).

In particular, from this section on, a more recent version of MEGAN (MEGAN3.0, Guenther et al. (2018)) has been used, which provide the possibility to modify the land cover and the emission factors resulting from different tree types and different regions. Therefore, all the modifications mentioned in this section were made both in MEGAN and in CHIMERE simultaneously. We performed two separate test cases in this section: one for changing the land cover data and the second for changing the regional tree species and the emission factors tied to specific trees. Both these modifications result in changes in the general BVOC emission factors; therefore, they are presented in the same section.

In the first test case, we considered the datasets provided by CHIMERE and WRF and compared them with several other more recent or regionally specific data sets. CHIMERE allows the use of two datasets for land cover information: the Globcover data (Arino et al., 2008) as selected for the base simulation, and the USGS data (Gutman et al., 2008). As for the WRF model, it provides the option of using MODIS (Giri et al., 2005) or USGS data. Also, the forest cover density information made available in MEGAN was also included in our comparisons. In addition to these data, three databases of major interest were added to our comparison.

First, we considered the Copernicus database (https://land.copernicus.eu/pan-european/high-resolution-layers/forests/tree-cover-density, last accessed 29 July 2022), which provides (among other parameters) information about the tree cover density over Europe for the years 2012 and 2015 with a horizontal resolution of 20 m. The Copernicus database also provides forest changes across Europe between the years 2012 and 2015. These changes are, on the average, about +30% forest growth in cells where the changes were positive and about -40% in cells where the changes were negative. Given these important changes, only the most recent (2015) data was considered. This database was regridded to our 1km resolution domain to have a common grid for comparisons.

We also collected data from the IGN (Institut national de l’information géographique et forestière) which provides detailed information on forest density in France (via the DBForêt database) and on land cover information in general (via the DBTopo database). These data were downloaded from the IGN data portal (http://www.ign.fr/, last accessed 16 July 2022) for three administrative departments (Landes (40), Gironde (33) and Lot et Garonne (47)) in order to cover the entire Landes forest. Both the DBForêt and the DBTopo databases have high resolution (around 20 m); temporally, the DBTopo database is for the year 2017 while the BDForêt database was updated in the year 2016 with data from the year 2014. The data for both databases were processed into a format that is usable in CHIMERE (i.e. projection of shapefiles onto a regular netcdf grid) and then
There is also a land cover dataset provided by the French national institute of agriculture, alimentation, and environment (INRAE, Institut national de recherche pour l’agriculture, l’alimentation et l’environnement), called the Theia dataset (https://www.theia-land.fr/en/homepage-en/, last accessed 15 July 2022). This dataset includes several types (reflectance, snow mask, water level, etc.) of information gained from remote observation of the earth, including land cover data. The dataset was realized first in 2016, then updated in 2018, with a horizontal resolution of 20 m. Both datasets for years 2016 and 2018 are presented in this study (Figure 3). It is important to know that the base data used in order to generate the Theia datasets, apart
from satellite information is the BDTopo database presented in the previous paragraph. Comparing the two Theia datasets (2016 and 2018) a difference of 29% is seen when conifers tree density is involved (although differences in total tree cover are small, see below).

Since Theia-2018 is the most recent source of information we could get for the region at this time and since it is based on several independent data sources, it was taken as reference in Figure 3. The upper left frame shows the 2018 Theia tree cover density for the Landes forest in shades of green. It draws a strongly delimited triangle of dense forested area on the domain, which becomes irregular in the southern part of the domain and near the coastline. The other frames show the relative difference of each data base with respect to Theia-2018. It can be seen that there are practically no differences between Theia-2018, BDTopo and Theia-2016, and only slightly lower densities in the BDForêt and MEGAN databases, while strong differences appear with the MODIS, Globcover, Copernicus and USGS data. MODIS proposes a much higher forest density than Theia, especially in and around the edge of the Landes forest. A possible reason for this could be the data being outdated (the MODIS version used here dates from 2000) or the lower resolution of MODIS (1 km × 1 km) compared to the others. The Copernicus data base shows a consistent underestimation of the forested area over the Landes forest, as does the Globcover database to a lesser degree; probably because its resolution is higher than MODIS, but it is still outdated (2000s). The USGS data base shows either a strong overestimation or an underestimation for different 1km grid cells. Since this database is quite old, the changes between forest and agriculture could be responsible for these discrepancies. From this point on, the Theia-2018 land cover database has been used both in CHIMERE and in MEGAN.

The second step of this section is to analyze the estimation of the BVOC emission factors when tree types in the region and tree-specific emission factors are changed in MEGAN. The BDForêt database mentioned above provides more detailed information about distribution of tree species inside the Landes forest as well as their density. It was considered necessary to modify the tree type distribution in the MEGAN model since MEGAN 2.1 assumes 28% of maritime pine coverage for the Landes forest, while the BDForêt database shows around 95% of the same species. The reason for this discrepancy is that MEGAN considers an average tree species distribution for western European forests up to 300 km from the Atlantic coast, which does not take into account the specificity of the Landes forest. Therefore, modifications were made in the MEGAN v3 model inputs in order to i) create a specific ecosystem for the Landes forest containing 95% of maritime pines and ii) modify the emission factors for the maritime pine to experimental values obtained during the Landex campaign (E. Ormeno, personal communication). The resulting emission factors were integrated into CHIMERE and a sensitivity case was run with this new configuration. The resulting emissions of isoprene and the sum of terpenoids (mono and sesqui terpenes) using these new emission factors are shown in Figure 4 for both sensitivity tests detailed in this section. Figure 4 justifies the necessity of this modification, since as mentioned above the emissions of isoprene have dropped by about half, while the emissions of terpenoids have increased significantly, by about a factor of 2. Taking into account that maritime pine is terpene-emitting and that the forest is mostly covered by these species, these emissions are more realistic for the Landes forest than the original ones. Monoterpane emissions (averaged over the entire forest) increase around 33% when only land cover is changed, while emissions of isoprene decrease by around 30% by land cover data change. For the region being studied here it is important to notice that the effect of changing the type of trees and the experimental emission factors for these trees is more important for
the emission of monoterpenes, while the change caused by modifying the land cover seems to be more important for isoprene emissions.

It should be taken into account that, a test was run by replacing the LAI used in the model by the Copernicus 1 km horizontal resolution global LAI-v2 databases (accessible here: https://land.copernicus.eu/global/products/lai, last accessed July 2022) for the year 2017, leading to no significant changes in the results of the simulations compared to when the LAI implemented in CHIMERE (from MEGAN 3.0) is used.

3.5 Anthropogenic emissions

For the reference simulation, anthropogenic emissions from the EMEP inventory are used, with a horizontal resolution of 10 km. This is assumed to be precise enough because the region has, at least when it comes to the forest, a low anthropogenic footprint. However, the transport to the forest of compounds recently emitted by the anthropized areas around the forest (including large cities such as the Bordeaux metropole to the north and smaller ones such as Dax to the south, as well as industrial sites located to the west and northwest of the forest) as well as motorway traffic passing through the forest, need to be taken into account in more details. To this end, we have implemented in CHIMERE the kilometric emission data provided by the regional air quality agency (ATMO-NA). This 5th sensitivity case revolves around changing the anthropogenic emissions from the reference EMEP emissions to the emissions provided by ATMO-NA. This modification, in addition to providing anthropogenic
emissions with greatly increased horizontal resolution, has the advantage of being prepared from bottom-up data specific to this area. The comparison of these two emission datasets is shown in SI-4. In summary, the two databases show comparable emissions for all species, presenting the highest average bias for $SO_x$ of around 20%, dominant SNAP sectors being always the same for each species group between the two emission inventories. It should be noted that the emissions provided by ATMO-NA are relative to the year 2014; therefore, in order to have consistent emission inventories, we used the 2014 EMEP emission inventory for this test. For information, the total national $NO_x$ emissions differ around 9% between 2014 and 2017 on average (estimated from EMEP data).

3.6 Vertical diffusion, wind speed and shortwave radiation corrections

Since the CHIMERE model does not have a coupled canopy model, it cannot take into account the physical effects that the forest canopy might exert on concentrations of chemical species. The last sensitivity case is based on the corrections of the effects that the canopy can have on the simulation of the physical parameters such as vertical diffusion ($K_z$, $m^2.s^{-1}$), wind speed and shortwave radiation in the model (it will be referred as the final test case from here on).

While the effect of the canopy on BVOC emissions is already taken into account in the MEGAN scheme, the canopy effects on wind speed and vertical diffusion inside the forest are not simulated in the CHIMERE model. This normally causes no issues, since most measurements are performed above canopy level, but since the goal of this work was to understand the
atmospheric situation inside the canopy, almost all measurements are performed inside it. Therefore, we have added corrections
for the aforementioned physical parameters inside the canopy in order to simulate the physical conditions of the forest more
realistically and these modifications will be presented as the 6th and sensitivity test, supposed to be the most realistic one.

For this purpose, for the model cells where the forest density is higher than 70%, the vertical diffusion and the wind speed
correction suggested by Ogée et al. (2003) and Leuning (2000) implemented within the MUSICA canopy model have been
integrated into the fine domain of our simulation. The MUSICA canopy model is a 1D model developed in the early 2000s; it
has been detailed by Ogée et al. (2003), and the case study presented the aforementioned study is for the Landes forest and has
been validated by comparisons against measurements.

In the standard version of the CHIMERE model, the vertical diffusion ($K_z$) is calculated by the parameterization suggested
by Troen and Mahrt (1986). Depending on the atmospheric stability (determined by the Monin-Obukov length), the $K_z$ can
relate to $u^*$ in different ways depending atmospheric stability:

\[
K_z = k u^* \left(1 - \frac{z}{h}\right)^2
\]  
(1)

The factor $w_s$ is calculated differently depending on the atmospheric stability:

Stable conditions : $w_s = \frac{u^*}{(1 + 4.7 \frac{z}{L})}$
(2)

Unstable conditions : $w_s = (u^* + 2.8 \epsilon u^*_w)^{\frac{2}{3}}$
(3)

Where, $\frac{z}{L}$ is the altitude to height of the boundary layer, $u^*$ is the friction velocity and $L$ is the Monin-obukov length.
More explanation on the calculation of $K_z$ can be found in the Chimere model documentation (downloadable from https:

In this work we have added $K_z$ and horizontal wind correction factors ($\phi_w(\xi)$ and $\phi_h(\xi)$ respectively) using the references
mentioned above:

\[
\phi_w(\xi) = \begin{cases} 
1.25(1 + 3|\xi|)^{-\frac{1}{3}}, & -2 \leq \xi < 0 \\
1.25(1 + 0.2\xi), & 0 \leq \xi < 1 
\end{cases}
\]  
(4)

\[
\phi_h(\xi) = \begin{cases} 
(1 + 16|\xi|)^{-\frac{1}{3}}, & -2 \leq \xi < 0 \\
(1 + 5\xi), & 0 \leq \xi < 1 
\end{cases}
\]  
(5)

Where, $\xi$ the stability factor is calculated by:

\[
\xi = \begin{cases} 
\frac{b_p}{L}, & z < z_{ruf} \\
\frac{(z-d)}{L}, & z \geq z_{ruf} 
\end{cases}
\]  
(6)
Where, \( h_c \) is the canopy height (m), \( z \) is the altitude (m), \( d \) is the zero-plane displacement for momentum and \( z_{\text{ruf}} \approx 2.3h_c \).

The diffusion and wind corrections are affected only to the first level of the model, which has a height of 8 m specifically for this case, and for cells with a forest density greater than 70%; a stability correction is also included at the same time (explained in Ogée et al. (2003)).

The changes induced for the vertical diffusion are shown in Figure 5. The line shows the \( K_z \) in the measurement site, and the ribbon around the line shows the regional changes around the measurement site up to a distance of 5 km on each side. The upper panel shows the standard vertical diffusion and the lower one the corrected \( K_z \). Results indicate a strong decrease, by roughly one order of magnitude in the vertical exchange coefficient between the first and second model layer, when the canopy parameterization is taken into account.

The modification for shortwave radiation penetration inside the canopy comes from Hassika et al. (1997), which itself is a continuation of the work performed by Berbigier and Bonnefond (1995). Both studies revolve around the Landes forest, yet are general in their formulation. The following variation of Beer’s law is used to achieve this purpose:

\[
\text{swrd}(\lambda) = \text{swrd}(0)e^{-\frac{k \times LAI(\lambda)}{\cos \beta}}
\]

where \( \text{swrd}(\lambda) \) is the shortwave direct radiation at the altitude \( \lambda \), \( \text{swrd}(0) \) is the shortwave direct radiation above canopy level, \( LAI(\lambda) \) is the leaf area index at the altitude \( \lambda \), \( \beta \) is the zenithal angle and lastly \( k \) is an extinction factor, which was measured to be 0.33 for the Landes forest (calculated by Ogée et al. (2003) using experimental data). In the case of our simulation, we do not need a distribution of shortwave radiation for different altitudes within the canopy, but rather a bulk decrease that we calculate here for the middle of the lower part of the first model layer between 0 to 6 m height (because LAI is supposed to be distributed between 6 and 12 m above ground level). This simulated value can thus be observed at 6 m height at Salles-Bilos. According to visual inspection of the measurement site, most of the LAI is distributed in the last meters of the pine trees in that specific patch of the forest. This equation is also used on cells with higher than 70% tree density. It is also important to take into account that the swrd changes affect only the chemistry modules of the model, by reducing photolysis frequencies. On the contrary, emissions of BVOCs are not affected by these modifications, since the effect of the canopy on the diminution of swrd is already taken into account in the MEGAN model when calculating the BVOC emission factors.

These modifications were added in 2 steps: first the \( K_z \) and the wind speed modifications and second the swrd modification. Therefore, it was possible to determine the changes made by \( K_z \)/wind speed modifications compared to the changes induced by the swrd changes. The changes for \( K_z \) is shown in 5.

It is important to keep in mind that this modification has been performed as a sensitivity case and not a functional parameterization for the model. The use of correction factors that are specifically calculated for the Landes forest make their use unrealistic for other forested areas in Europe. Also, canopy induced modifications in the vertical diffusion are considered at the top of the first model layer fixed for this case at 8 m height, so still well within the canopy (of 12 m height at the measurement site). For the wind speed and radiation, modifications are considered for the middle of the first layer, so at 4 m height.
**Table 2.** Statistical information for test cases showing correlation to measurements when possible (R), bias (model – measurements), minimum, maximum and average values (noted as min, max and mean) and the percentage of change of the mean for each test case compared to the base simulation. When measurement data was available, variables the value is given in parenthesis.

<table>
<thead>
<tr>
<th>Species Tests</th>
<th>R</th>
<th>Bias</th>
<th>Min (ppb)</th>
<th>Max (ppb)</th>
<th>Mean (ppb)</th>
<th>% change</th>
</tr>
</thead>
<tbody>
<tr>
<td>NO₂ (ppb)</td>
<td>1:Base</td>
<td>0.43</td>
<td>-1.05</td>
<td>0.06(0.59)</td>
<td>5.75(12.17)</td>
<td>0.60(1.65)</td>
</tr>
<tr>
<td></td>
<td>2:1+meteo</td>
<td>0.56</td>
<td>-0.88</td>
<td>0.07</td>
<td>9.69</td>
<td>0.77</td>
</tr>
<tr>
<td></td>
<td>3:2+land cover</td>
<td>0.56</td>
<td>-0.89</td>
<td>0.07</td>
<td>10.08</td>
<td>0.76</td>
</tr>
<tr>
<td></td>
<td>4:3+megan</td>
<td>0.56</td>
<td>-0.91</td>
<td>0.07</td>
<td>9.91</td>
<td>0.74</td>
</tr>
<tr>
<td></td>
<td>5:4+emissions</td>
<td>0.59</td>
<td>-0.64</td>
<td>0.11</td>
<td>10.07</td>
<td>1.01</td>
</tr>
<tr>
<td></td>
<td>6:5+physical</td>
<td>0.61</td>
<td>-0.40</td>
<td>0.15</td>
<td>9.76</td>
<td>1.25</td>
</tr>
<tr>
<td>O₃ (ppb)</td>
<td>1:Base</td>
<td>0.67</td>
<td>6.80</td>
<td>19.19(1.97)</td>
<td>57.73(62.08)</td>
<td>34.15(27.35)</td>
</tr>
<tr>
<td></td>
<td>2:1+meteo</td>
<td>0.74</td>
<td>6.05</td>
<td>14.83</td>
<td>61.76</td>
<td>33.40</td>
</tr>
<tr>
<td></td>
<td>3:2+land cover</td>
<td>0.72</td>
<td>6.39</td>
<td>14.01</td>
<td>61.69</td>
<td>33.74</td>
</tr>
<tr>
<td></td>
<td>4:3+megan</td>
<td>0.72</td>
<td>2.97</td>
<td>9.44</td>
<td>56.73</td>
<td>30.32</td>
</tr>
<tr>
<td>Terpenes (ppb)</td>
<td>1:Base</td>
<td>0.30</td>
<td>-2.83</td>
<td>0.06(0.20)</td>
<td>1.93(24.86)</td>
<td>0.40(3.23)</td>
</tr>
<tr>
<td></td>
<td>2:1+meteo</td>
<td>0.28</td>
<td>-2.79</td>
<td>0.05</td>
<td>2.44</td>
<td>0.44</td>
</tr>
<tr>
<td></td>
<td>3:2+land cover</td>
<td>0.33</td>
<td>-2.64</td>
<td>0.07</td>
<td>3.82</td>
<td>0.59</td>
</tr>
<tr>
<td></td>
<td>4:3+megan</td>
<td>0.32</td>
<td>-1.89</td>
<td>0.15</td>
<td>9.13</td>
<td>1.34</td>
</tr>
<tr>
<td></td>
<td>5:4+emissions</td>
<td>0.31</td>
<td>-1.91</td>
<td>0.15</td>
<td>9.13</td>
<td>1.32</td>
</tr>
<tr>
<td></td>
<td>6:5+physical</td>
<td>0.31</td>
<td>-1.55</td>
<td>0.21</td>
<td>9.34</td>
<td>1.68</td>
</tr>
<tr>
<td>OH²⁻ (ppb)</td>
<td>1:Base</td>
<td>–</td>
<td>–</td>
<td>4.56×10⁻⁶</td>
<td>2.88×10⁻³</td>
<td>1.77×10⁻⁴</td>
</tr>
<tr>
<td></td>
<td>2:1+meteo</td>
<td>–</td>
<td>–</td>
<td>6.86×10⁻⁶</td>
<td>3.46×10⁻³</td>
<td>1.89×10⁻⁴</td>
</tr>
<tr>
<td></td>
<td>3:2+land cover</td>
<td>–</td>
<td>–</td>
<td>6.85×10⁻⁶</td>
<td>2.08×10⁻³</td>
<td>1.38×10⁻⁴</td>
</tr>
<tr>
<td></td>
<td>4:3+megan</td>
<td>–</td>
<td>–</td>
<td>6.23×10⁻⁶</td>
<td>6.18×10⁻⁴</td>
<td>6.61×10⁻⁵</td>
</tr>
<tr>
<td></td>
<td>5:4+emissions</td>
<td>–</td>
<td>–</td>
<td>8.86×10⁻⁶</td>
<td>5.17×10⁻⁴</td>
<td>8.18×10⁻⁵</td>
</tr>
<tr>
<td></td>
<td>6:5+physical</td>
<td>–</td>
<td>–</td>
<td>1.22×10⁻⁵</td>
<td>3.88×10⁻⁴</td>
<td>7.67×10⁻⁵</td>
</tr>
<tr>
<td>NO₃ (ppb)</td>
<td>1:Base</td>
<td>–</td>
<td>–</td>
<td>1.88×10⁻⁶</td>
<td>2.35×10⁻⁴</td>
<td>3.24×10⁻⁵</td>
</tr>
<tr>
<td></td>
<td>2:1+meteo</td>
<td>–</td>
<td>–</td>
<td>1.41×10⁻⁶</td>
<td>1.78×10⁻⁴</td>
<td>2.99×10⁻⁵</td>
</tr>
<tr>
<td></td>
<td>3:2+land cover</td>
<td>–</td>
<td>–</td>
<td>1.46×10⁻⁶</td>
<td>2.50×10⁻⁴</td>
<td>3.74×10⁻⁵</td>
</tr>
<tr>
<td></td>
<td>4:3+megan</td>
<td>–</td>
<td>–</td>
<td>2.62×10⁻⁶</td>
<td>2.10×10⁻⁴</td>
<td>3.31×10⁻⁵</td>
</tr>
<tr>
<td></td>
<td>5:4+emissions</td>
<td>–</td>
<td>–</td>
<td>2.35×10⁻⁶</td>
<td>2.98×10⁻⁴</td>
<td>4.54×10⁻⁵</td>
</tr>
<tr>
<td></td>
<td>6:5+physical</td>
<td>–</td>
<td>–</td>
<td>2.89×10⁻⁶</td>
<td>2.44×10⁻⁴</td>
<td>4.50×10⁻⁵</td>
</tr>
<tr>
<td>PM₁⁻⁻⁻⁻ (µg.m⁻³)</td>
<td>1:Base</td>
<td>0.47</td>
<td>6.25</td>
<td>3.71(0.17)</td>
<td>61.46(59.52)</td>
<td>11.13(4.88)</td>
</tr>
<tr>
<td></td>
<td>2:1+meteo</td>
<td>0.39</td>
<td>6.29</td>
<td>4.49</td>
<td>36.26</td>
<td>11.17</td>
</tr>
<tr>
<td></td>
<td>3:2+land cover</td>
<td>0.43</td>
<td>7.56</td>
<td>4.71</td>
<td>39.56</td>
<td>12.44</td>
</tr>
<tr>
<td></td>
<td>4:3+megan</td>
<td>0.46</td>
<td>8.37</td>
<td>5.14</td>
<td>45.54</td>
<td>13.25</td>
</tr>
<tr>
<td></td>
<td>5:4+emissions</td>
<td>0.47</td>
<td>8.54</td>
<td>5.33</td>
<td>48.10</td>
<td>13.42</td>
</tr>
<tr>
<td></td>
<td>6:5+physical</td>
<td>0.47</td>
<td>6.07</td>
<td>4.28</td>
<td>39.84</td>
<td>10.95</td>
</tr>
<tr>
<td>OA (µg.m⁻³)</td>
<td>1:Base</td>
<td>0.74</td>
<td>0.71</td>
<td>0.22(0.07)</td>
<td>13.80(20.35)</td>
<td>2.41(1.70)</td>
</tr>
<tr>
<td></td>
<td>2:1+meteo</td>
<td>0.69</td>
<td>0.56</td>
<td>0.21</td>
<td>16.15</td>
<td>2.26</td>
</tr>
<tr>
<td></td>
<td>3:2+land cover</td>
<td>0.70</td>
<td>1.00</td>
<td>0.23</td>
<td>18.14</td>
<td>2.70</td>
</tr>
<tr>
<td></td>
<td>4:3+megan</td>
<td>0.70</td>
<td>1.62</td>
<td>0.24</td>
<td>24.38</td>
<td>3.32</td>
</tr>
<tr>
<td></td>
<td>5:4+emissions</td>
<td>0.70</td>
<td>1.73</td>
<td>0.25</td>
<td>25.22</td>
<td>3.43</td>
</tr>
<tr>
<td></td>
<td>6:5+physical</td>
<td>0.70</td>
<td>0.99</td>
<td>0.22</td>
<td>17.68</td>
<td>2.69</td>
</tr>
</tbody>
</table>
4 Effects of the model configuration on simulations of SOA, precursors and oxidants

In this section, the results shown for the array of sensitivity cases, going from a regional standard configuration of CHIMERE to a one updated for the local Landes forest conditions, will be discussed. First primary gaseous species of biogenic and anthropogenic origin will be treated (section 4.1), then \(O_3\) as a secondary species (section 4.2), radicals (section 4.3), and finally particulate matter with focus on its organic fraction (section 4.4). Changes induced by sensitivity tests will be compared to measurements at the Salles-Bilos site. In section 4.5, the changes to the contribution of precursors and oxidants to \(BSOA\) formation will be assessed in the base case compared to the simulation with all the modifications. Keep in mind that whenever changes are discussed in this section changes of averages over the entire period of each test case compared to average of the entire period of the base case are being discussed.

Several of the species being discussed in the following sections have high local sensitivity depending on the forest coverage of a given location. As it is not clear, if the 1 km \(\times\) 1 km model grid cell corresponding to the Salles-Bilos measurement site exactly matches local conditions, the range of simulated concentrations in a 11 km \(\times\) 11 km square around the site will also be displayed.

In sections 4.1 to 4.4, observed and simulated time series at the Salles-Bilos site will first be analyzed. The percentage change of the concentration of different species caused by each sensitivity test compared to the previous one will be indicated. Finally, it will be addressed how much the complete set of sensitivity studies improved the comparison to observations in terms of bias, mean error and correlation.

4.1 Biogenic and anthropogenic primary gaseous species

The sum of monoterpenes is the first group of species that we consider here from PTR-Tof-MS measurements (m/z 137 peak as their main fragment and m/z 81 values as their second most important one) and sum of simulated \(\alpha\)-pinene, \(\beta\)-pinene, limonene, ocimene concentrations, see section 3.4) since these are main drivers of atmospheric chemistry for the given terpene emitting maritime pine forest. Figure 6 shows a comparison of \(O_3\) and terpene concentrations with measurements and different sensitivity tests, both shown in ppb (\(NO_2\) is shown in the SI). It should be noted that each of these \(BVOC\)s has been compared separately to measurements; the comparisons are presented in the SI (Figures SI7 through SI12).

Terpene concentrations are source dependent and being enhanced locally, being more important near emission sources throughout the forest area (Figures 4 and 10). Observed and simulated terpene concentrations show a marked diurnal cycle with increased night-time levels due to decreased sink processes (less \(OH\) and \(O_3\), but more \(NO_3\)) and decreased horizontal and vertical dispersion (night-time wind speeds and exchange coefficients are minimal), as pointed out by Bsaibes et al. (2020), and Mermet et al. (2021). Observed night-time terpene maxima range between 5 and 25 ppb. Largest concentrations are reached during nights of July 2 – 3, 4 – 5, and 16 – 17. These peaks generally correspond to lower friction velocity observations (below 0.1 m.s\(^{-1}\)), and larger temperatures as pointed out by Bsaibes et al. (2020).

Simulated terpene maximum levels with the base case configuration barely exceed a few ppb, and can reach up to about ten ppb in the 11 km \(\times\) 11 km area around the Salles-Bilos. The night-time peaks during July 4 – 5 and 16 – 17 are not
reproduced which may probably be explained by an overestimation of friction velocity during these nights. Increased day-time concentrations are simulated during the two hot weather periods (July 4 – 8, and July 16 – 18).

Some of the sensitivity cases explored in this study cause strong positive changes in the concentrations of terpenes: 11\% when changing land cover inputs, +103\% (always compared to the base case average) when updating biogenic emissions with local emission factors and +145\% when changing the $K_z$ parameterization. The 10\% increase while changing the land cover is related to the tree density increases on average in the region, although it is interesting to keep in mind that changing the land cover may also decrease the tree density in the forest locally, since some parcels were recently turned into agricultural fields. For the measurement cell, tree density increases by around 29\%. The increase of terpenes when updating the biogenic emissions occurs because both the tree types and the tree emissions rates for the forest have been rectified in the emissions produced by the MEGAN model from isoprene to terpene emitters. Coherently, this also leads to a decrease (-73\%) of isoprene concentrations, since changing the majority of the Landes forest tree type in MEGAN to maritime pine means that the density of isoprene emitting trees reduces significantly. Including the canopy parameterization in the model causes emitted terpenes to be less exchanged between inside the canopy and above. It also increases the variation of terpene concentrations locally,
seen from the larger envelope around for the canopy effects sensitivity case in Figure 6. The other two test cases show slightly negative effects on the terpene concentrations (-10% and -2% for the meteorological test and the anthropogenic emission test case respectively). Figure 7 shows the distribution of isoprene/terpene in base case and the final sensitivity cases for the measurement cell, as well as a more detailed distribution of different monoterpene/sesqui-terpene species. It is important to notice that this distribution changes heavily between the base case and the final case, being much closer to the observed distribution after the biogenic emission/land cover modifications. An overestimation of isoprene is still seen in the final case.

Finally, and importantly, the different test cases lead to a clear improvement in average and maximum simulated sum of terpene concentrations (see figure 6). In particular, the negative biases are strongly reduced for maximum terpene concentrations, from 14.0 ppb to 22.4 ppb as compared to 25.2 ppb in observations. It is also important to keep in mind that in the canopy test case, the changes seen in the concentrations arise almost entirely from the $K_z$ modifications and not from the swrd ones (as they do not affect the emission of BVOCs since their effect is already taken into account in the emission factor calculations).

Next, simulations of NO$_2$ as a tracer of anthropogenic emissions are discussed (Figure shown in the SI). Observed NO$_2$ level is maximum mostly during night and morning hours, reaching peaks between 3 and 7 ppb (for nights between July 2 – 3, 3 – 4, 6 – 7, 15 – 16, 16 – 17). Again, this is due to reduced local dispersion, in addition to suppressed NO$_2$ photolysis, and in spite of lower night-time NO$_x$ emissions.

The concentration of NO$_2$ (Figure shown in the SI) changes most in two cases: the anthropogenic emissions test case (+26%) and the meteorological test case (+20%). The refined distribution of emissions explains the major part of the first case, as average emissions over the inner model domain increase by only 4%. For the meteorological test case, the changes to the maximal values of NO$_2$ are attributed to the advection of air masses from the Bordeaux metropole (or the main highway crossing the south west of France) towards the measurement site. For other scenarios, changes are below 10%.
On the whole, average $NO_2$ concentrations of the specific Landes model configuration are in better agreement with observations compared to the standard version (-0.69 instead of -0.99 ppb for the mean bias and 0.62 instead of 0.43 for correlation compared to observations, Table 2). Please note, that due to the non-specificity of $NO_2$ measurements, they could overestimate actual concentrations. On a relative scale, this is expected to be less important for peak values indicating the presence of fresh $NO_x$, than for unpolluted periods when the $NO_x/NO_y$ ratio may be low.

4.2 Secondary gaseous species: $O_3$

Next, changes of sensitivity tests are discussed for $O_3$, as representative of gaseous secondary species, and directly and indirectly (via OH and $NO_3$ radical formation) governing oxidation of BVOC species leading to BSOA formation. $O_3$ concentration values show a strong diurnal variation due to photochemical formation during day-time and dry deposition and depletion by terpenes during night-time. Largest day-time $O_3$ maxima levels during the campaign were between 55 – 60 ppb during the so called "strong oxidation" periods characterized by elevated temperatures, low relative humidity, and low wind speeds (< 3 m.s$^{-1}$) from July the 4th to the 7th then from the 14th to the 18th Mermet et al., 2021). The higher values of $O_3$ are due to continental advection under easterly wind conditions (visible when looking at the continental domain simulations). Lowest night-time concentrations were observed for nights between July 2 – 4, 6 – 7, 14 – 17, corresponding to reduced local dispersion and enhanced $NO_2$ concentrations. In the early morning hours of such nights, observed in-canopy $O_3$ levels could even drop close to zero, due to depletion by mainly terpenes, dry deposition, and suppressed vertical mixing. While simulations with the standard model version correctly reproduce the day-time $O_3$ maxima, the significantly overestimate the $O_3$ minima, which never decrease below to 20 ppb.

Daily $O_3$ maxima are only slightly changed in sensitivity tests, the largest impact being noted for the test with refined emissions, leading to enhanced $NO_x$ emissions and an increase in the $O_3$ peak on July 7 from about 60 to about 70 ppb (Figure 6). The increase in terpene emissions does not lead to an increase in $O_3$ production, partly because it is compensated by a decrease in isoprene.

Night-time $O_3$ minima tend to decrease in different sensitivity tests. Here it is interesting to check if the simulations can reproduce the lowest night-time values obtained between July 2 – 4, 6 – 7, 14 – 17. Alternative meteorology from ECMWF instead of NCP/WRF leads to lower friction velocities during these nights and thus lower $O_3$ minima, from 19.2 ppb to 14.8 ppb for the average of these nights. Land cover and emissions factor changes do not affect $O_3$ minima to a significant degree (only 0.8 ppb of decrease for both cases), probably because terpene increases are compensated by isoprene decreases, resulting in rather unchanged $O_3$ depletion. Emission changes affect these minima slightly (1.5 ppb of decrease for the minima when emissions are changed), probably because of increased titration with $NO$ leading also to increased night-time $NO_2$. Including the canopy parametrization leads to a decrease in $O_3$ minima (inside the canopy), due to decreased vertical exchange and less $O_3$ transfer from above (from 12.9 ppb to 10.4 ppb).

However, even the refined model configuration still overestimates $O_3$ minima (10.4 ppb instead of 2.0 ppb observed for the three periods July 2 – 4, 6 – 7, 14 – 17), even if this overestimation has decreased with respect to standard model version (19.2 ppb instead of 2.0 observed). Also, the lowest simulated $O_3$ concentrations in the 11 km×11 km domain around Salles-Bilos,
are still above the observed ones. This is not without consequences for oxidant supply for night-time SOA build-up from terpenes, which is then overestimated for these nights. Reasons for this \( O_3 \) overestimation are unclear. Titration by terpenes and \( NO \) seems to be correctly taken into account as night-time peaks of these compounds seem well-simulated (see section 4.1 above). It can be because of missing sources for minor terpenoids, which even in low concentrations can be extremely reactive in the atmosphere making them capable of consuming practically the entirety of night-time \( O_3 \). Another plausible candidate for it might be an underestimation of deposition of \( O_3 \) over forested areas. However, the dry deposition speed of this species was compared to dry deposition speed calculated using measured \( O_3 \) fluxes; this speed is well simulated with an average of 0.43 \( m.s^{-1} \) and 0.62 \( m.s^{-1} \) for observations and simulation (Figure shown in SI).

Therefore, an additional unrealistic sensitivity test was performed, putting vertical exchange coefficients to zero during night-time (between 22h and 5h every night only for cells with tree density higher than 70%). This leads to closer \( O_3 \) night-time values compared to measurements (1.9 ppb, Table 2), but to unrealistically high terpene concentrations (concentrations of over 50ppb). Still an overestimation of canopy to outside exchange could be a reason for not displaying the near zero \( O_3 \) values.

### 4.3 Radical species

Radicals (shown in Figure 8) play a major role in atmospheric chemistry. \( NO_3 \) has been recognized as a major oxidant especially of terpenes (for example see Ng et al. (2017)). The \( NO_3 \) IBCEAS measurements available for this campaign were always below the detection limit of 3 to 5 ppt (M. Cirtog, personal communication), as shown in Figure 7. In the base simulation and the meteorological and the land cover test cases, night-time maxima can reach 1 – 3 ppt. When the concentration of terpenes is increased (i.e. in the biogenic emissions test case), night-time \( NO_3 \) concentration maxima are limited to below 0.5 ppt with one exception. The canopy parametrization leads to additional \( NO_3 \) decrease, because of increased terpene concentrations (and probably minor effects due to the \( NO_2 + O_3 \) source reaction flux). Interestingly, Mermet et al. (2021) calculated steady state in canopy \( NO_3 \) concentrations from \( NO \), \( NO_2 \), \( O_3 \), monoterpene, isoprene and radiation observations and found day-time maxima (up to 0.1 ppt) in contrast due to our simulated night-time maxima. This discrepancy is probably due to the night-time overestimation of \( O_3 \) in our simulations leading to an overestimated \( NO_3 \) formation. These changes and differences potentially affect SOA formation, as will be discussed in the next section.

Compared to \( NO_3 \), \( OH \) concentrations remain more similar in all sensitivity cases, changes being generally limited to several tenths of %. These changes are difficult to explain in absence of a dedicated budget study which is beyond the scope of this paper. For instance, the combined effect of land cover and MEGAN changes is relatively small, probably because increased loss reactions with terpenes are compensated by diminished loss with isoprene. Increases in daily maximum \( OH \) due to anthropogenic emission changes could be due to a more effective \( OH \) recycling via the \( NO + HO_2 \) reaction. When it comes to the canopy test case, the modification of swrd results in a diminution of \( OH \) by a factor of 1.4 on average, while the \( K_z \) and wind speed modification does not result in any significant change on the concentration of \( OH \). For more densely forested grid cells, the in-canopy radiation is reduced more and as a consequence, the \( OH \) concentration. Since no \( OH \) concentration measurements were available from this campaign (the FAGE instrument deployed by the University of Lille having unfortunately encountered technical problems), they have been estimated from in canopy global radiation, and
Figure 8. Time series for OH and NO₃ concentrations (both in ppb) for all the sensitivity cases explained above. The color schemes are shown on the right side. The name of each case is written above each time series. Note that the NO₃ columns have 2 different scales because of visibility issues.

a climatological (above canopy) maximum OH concentration $3 - 6 \times 10^6$ molecules cm$^{-3}$ (Mermet et al., 2021). Daily OH maxima estimated in this way were about a factor of 2 to 3 below our simulated in-canopy concentrations.

4.4 Particulate species

Figure 9 shows the comparison of all performed sensitivity cases to measurements for the Salles-Bilos site for particulate species, all values being in $\mu$g m$^{-3}$. $PM_{2.5}$ observations during the campaign period culminate at more than 20 $\mu$g m$^{-3}$ on July 7. Some higher hourly values (also visible in OA) in July 6 are due to a local fire which is not further analysed here. About half of this is composed of organic aerosol ($10 \mu$g m$^{-3}$). The simulated concentrations of secondary inorganic aerosol (sulfate + nitrate + ammonium), sea salt, black carbon and dust are about 3.0, 1.6, 0.9 and 7.6 $\mu$g m$^{-3}$ in the model, concluding on a significant dust contribution at this time at Salles-Bilos. The base case simulations show large $PM_{2.5}$ concentrations around this date (from July 6 to 9) at Salles-Bilos, reaching 60 $\mu$g m$^{-3}$. This simulation is very sensitive to the chosen meteorology; with ECMWF meteorology, the peaks decrease to 40 $\mu$g m$^{-3}$, however another dust related, yet unobserved peak at Salles-
Bilos occurs on July 13. Both of these episodes were seen in AOD observations at the Arcachon AERONET station (a city to the west of Bordeaux, near the coast line).

Besides dust, OA is the main driver of observed and simulated PM$_{2.5}$ variability. During the sunny and calm, high oxidation periods, from July the 4th to the 7th then from the 14th to the 18th observed OA concentrations reach about 10 µg.m$^{-3}$. This time pattern is rather well reproduced by the base simulation, OA peaks are however overestimated (see below). It is concluded from the model results and known OA sources, that most of it is biogenic SOA (84% at Salles-Bilos over the campaign period). For instance, the concentration of anthropogenic OA (primary and secondary is quite low at this site (average of 0.11 µg.m$^{-3}$ overall test cases).

Average OA concentrations (Figure 9) are mostly impacted by the biogenic emissions sensitivity case (+26%), the canopy test case (-24%), the land cover changes (+22%) and the meteorological test case (-7%). The test case least affecting the concentration of OA is the anthropogenic emission test by a +3% increase. Changing the biogenic emissions causes a significant increase in the concentrations of terpenoids (sum of mono and sesqui-terpenes), which then increase the formation of biogenic OA. The land cover changes also impact the concentration of OA for the same reason. Changing the $K_z$ parameterization and in canopy radiation decrease both OA and BSOA, as the concentrations of the three oxidants (OH, NO$_3$, and O$_3$) participate

Figure 9. Time series for PM$_{2.5}$ and organic aerosols (OA) both in µg.m$^{-3}$ for all the sensitivity cases explained above. The color schemes are shown on the right side. The name of each case is written above each time series.
in the biogenic OA formation process, while increasing the concentrations of terpenes (its major precursors). This test reveals an interesting feature: for BSOA formation around Salles-Bilos, changes in oxidation rates have a larger effect than changes in precursors (i.e. changes in terpenoids and isoprene).

Figure 10. 2D figures showing average concentrations over the campaign period for species of importance for this study (compound names and associated units written above each column) for 3h (first two rows) and 15h (second two rows) for base (first and third rows) and final simulation (second and forth rows). Note that the scales are different for each figure. The measurement site at Salles-Bilos is located in the middle of the white or black rectangle of 10km x 10 km extension.

The final test case shows enhanced OA concentrations with respect to the base case, an additional difference lies in the chemistry behind the production of OA as will be discussed in section 4.5; in the final test case the concentrations of precursors and radicals correspond better to observed data, as shown in the previous sections. The remaining overestimation can be due to the already stated oxidant overestimation for $O_3$ and $NO_3$ (noting that $OH$ cannot be controlled). It can also be due to the scheme used for the simulation of OA; thus, forthcoming work should test the sensitivity due to different aerosol scheme.

4.5 BSOA formation from different precursors and oxidant pathways

As pointed out in the introduction, different precursor/oxidant pathways for BSOA build-up have been found in the literature (e.g. Hallquist et al. (2009), Xu et al. (2015), Qin et al. (2018)) for different types of forest, and it is interesting to address this
question for a maritime pine forest. To do so, two sets of additional simulations, in which the formation of BSOA from each precursor and via each oxidant was separated in the process of the model run, were performed, leading to the estimation of the major pathways of BSOA formation in the base case scenario and also in the final one. These results are presented in Figures 11 and 12.

For the base case scenario, the formation of BSOA is achieved in majority from isoprene oxidation by OH. After the corrections made for the land cover, the tree types and the physical parameters (introduced in section 3.6) in the final test case, isoprene + OH stays the most important BSOA formation pathway, but monoterpenes become the dominant precursor group. The attack of β-pinene by NO₃ is outstanding in this group, in accordance for example with results found for the 2013 SOAS study in South-Eastern US (Xu et al., 2015). The reaction of sesqui-terpenes with O₃ is a third important reaction pathway.

Looking at which oxidant is responsible for BSOA formation, the majority of BSOA in the base test case is formed from OH radical reactions (by about nearly the half, i.e. 47%), followed by O₃ (nearly 30%) and NO₃ (more than 20%). In the final test case, the BSOA formation becomes more evenly distributed between the three precursors: OH, staying in majority (by 36%) while the other two oxidants follow closely (33% for NO₃ and 31% for O₃). These results can be compared to those reported by the BVOC reactivity study at Salles-Bilos (Mermet et al., 2021), based on measured BVOC species and O₃, and OH and NO₃ estimated from chemical equilibrium and solar radiation, even if this work addresses different features. In this study, the β-caryophyllene + O₃ reaction makes by far the largest contribution to BVOC reactivity. In our simulations, β-caryophyllene is part of the sesqui-terpene species family, which still makes an important contribution. Also, monoterpenes strongly contribute to BVOC reactivity, while isoprene is only a minor contributor. In general, O₃ is the major oxidant, followed by OH, while NO₃ only makes little contributions.
Differences with respect to our study can be explained by many factors which we do not attempt to quantify here: (i) the experimental study considers local oxidation rates, and those within the canopy are used here, while our study considers contributions within air masses arriving at the receptor site; (ii) while the experimental study looks at BVOC oxidation rates; (iii) our study overestimates night-time $O_3$ which leads to high $NO_3$ radical formation (see Figure 8), and could explain the larger share of $NO_3$ oxidation pathways; (iv) in the experimental study estimated $OH$ is smaller than that in our study (see section 4.3); (v) the initial share of BVOC species might be different even if Figure 7 shows a broad agreement for major BVOC species in the final case.

Interesting insight can also be obtained from analyzing the daily averaged concentrations of BSOA being formed from the different reaction pathways of BVOCs with oxidants presented in Figures 11 and 12. While the concentration of total BSOA stays quite similar between the two simulations (with the exception of some days and especially July 7) the chemistry behind their formation changes significantly. The contribution of isoprene in the formation of BSOA becomes lower in the final case, while the contribution of the $\beta$-pinene + $NO_3$ pathway, and of the oxidation of other terpenes increases significantly. Figure 12 shows that this is especially true for 07$^{th}$ of July when simulated BSOA was most overestimated. As terpene levels were well simulated on this day, this could most probably be due to an overestimation of radical concentrations ($NO_3$ and $OH$) and/or SOA yields. However, as many different pathways contribute to BSOA, it is difficult to pinpoint a specific one as responsible for the BSOA overestimation.

5 Spatial distribution of BSOA, precursors and oxidants over the Landes forest

5.1 Average spatial distribution during the Landex campaign

After this first analysis of time series at the measurement site where intensive campaign data was available, we now would like to extent the discussion to a broader spatial scale, covering the Landes forest and environments. To achieve this, we first present near surface 2D maps (of the high-resolution domain covering the Landes forest) of specific species pertaining to our study averaged over the campaign period (between 26$^{th}$ of June to the 20$^{th}$ of July 2017). The interest here is to analyze, how the changes in the model from the baseline test case to an updated local/regional configuration modifies the concentrations seen in these 2D fields.

Due to the short lifetime of terpenes during day-time (a few hours in most cases), the area of enhanced total terpene concentrations (at 15h UTC) closely corresponds to the high emission zone of the Landes forest. In the high-resolution simulation (1 km) with the canopy parameterization and all other local data bases, average terpene concentrations are heterogeneous varying from some ppb to up to 10 ppb. During night (3h UTC), terpene levels are maximum near the NE edge of the forest. This is due to larger emissions there (Figure 4c), but also to lower $O_3$ concentrations leading to diminished loss (Figure 10).

As expected, a strong increase is seen in the final case compared to the base case test over the whole domain. This is mainly due to enhanced terpene emissions in the local data-bases.
Isoprene concentrations in the local configuration are also enhanced over the Landes forest, but they are larger over forested areas in the north and east of Bordeaux, which include more isoprene emitting species like oaks, contrary to the Landes forest, which contains exclusively maritime pines.

The $NO_x$ 2D fields display largest $NO_x$ levels in the urban Bordeaux area, even larger at night due to enhanced vertical stability. In the simulation including high a resolution emission inventory, highways starting from Bordeaux are displayed, such as weak plumes in the surroundings of the city. It is also interesting to compare the difference in resolution in base and final simulations, noticeable in the 2D images shown in Figure 10. Daytime $NO_x$ concentrations in the Landes forest are relatively low (typically 2-3 ppb). According to this, photochemical $O_3$ production is not very active over the Landes forest,
since day-time $O_3$ levels are lower there (35 – 45 ppb) than in surrounding areas (45 – 55 ppb), mainly due to loss by reactions with terpenes or by dry deposition), with exception of an $O_3$ plume in the SW of Bordeaux. Night-time $O_3$ shows local minima in the surroundings of the metropole, below 20 ppb, due to $NO$ titration. If the results seen in Salles-Bilos can be extrapolated to the whole Landes forest, then simulated night-time $O_3$ levels are likely to be overestimated. Maximum values of $NO_x$ do not change between base case and final simulation; however, their spatial precision improves because of the increase in the resolution of emissions inventory in the final simulations.

The concentrations of the base sensitivity test and the final sensitivity case were compared to regional air quality station measurements discussed in section 2 (Figure 13). These comparisons show that the mean bias for $NO_2$ for all stations (regardless of the type of station) for the final test drops to -1.4 ppb (from an initial -2.9 ppb for the base simulation), while the averaged correlation gets to 0.42 in the final test from 0.37 for the base case. $NO$ concentrations show a bias and correlation of -2.4 ppb and 0.57 in the final test case compared to -3.1 ppb and 0.51 in the base simulation respectively. Unfortunately, the concentrations of $O_3$ (while being measured) in the air quality stations weren’t available for comparison for the period of simulations.

Night-time $OA$ in the updated configuration shows a hot spot at the NE edge of the Landes forest, corresponding to regions with enhanced terpene concentration levels. This effect is not seen in the baseline simulations, where concentrations are higher in the center of the forest (Figure 10). The night-time concentrations for the base case simulations are also lower. Day time $OA$ is generally enhanced over the Landes forest, showing that large terpene emissions there lead to local $OA$ build-up. In this averaged field, no $BSOA$ export from the Landes forest can be distinguished. Maximal concentrations occur within the forest at about 25 km from the coastline. It is hypothesized, that they are related to a sea breeze front forming along the Atlantic shoreline. This is also suggested by the 15h UTC temperature fields (not shown) which show large gradients (nearly 10K) between the coast and about 20 km inland. This phenomenon will be addressed in more detail in the next section. This sea
breeze is also seen in the baseline simulations, when looking at hour-by-hour changes in OA concentration. In the base case simulations concentrations are lower in day-time as well, similarly as what was seen at night-time.

It can be noted that simulated BSOA concentrations alone are above the WHO target of 5 µg.m\(^{-3}\) for PM\(_{2.5}\), even if they are expected to be lower during other seasons (however, we had shown above that they are overestimated with respect to measurements at the Salles-Bilos site). The spatial structure of PM\(_{2.5}\) and PM\(_{10}\) simulations follows broadly that of BSOA simulations with a major exception over the sea, where sea-salt makes an additional contribution to PM\(_{10}\).

Figure 13 illustrates the comparisons of PM\(_{2.5}\) and PM\(_{10}\) at air quality stations of the ATMO-NA network. It is interesting to notice that the profile seen in all stations is quite similar to each other. For PM\(_{10}\), the results show an average correlation of 0.3 and 0.42 and an average bias of -3.3 µg.m\(^{-3}\) and -2.9 µg.m\(^{-3}\) for PM\(_{10}\) for the base and the final sensitivity test respectively. For PM\(_{2.5}\) and for the base and the final sensitivity test respectively, a correlation of 0.4 and 0.58 and a bias of +4.8 µg.m\(^{-3}\) and +4.1 µg.m\(^{-3}\) is seen. Apparently, refining the simulations through the sensitivity tests adds physical information, as can be seen by the improved correlations (for PM, but also for NO\(_2\)).

5.2 A case study for the impact of see breeze fronts

In the previous section, mention was made of the crest of the high afternoon BSOA levels within the Landes forest about 25 km from the coastline, with the proposed hypothesis of a potential sea breeze front forming along the Atlantic coastline. Indeed, Planchon and Cautenet (1997) note that sea breeze fronts can go up to 40 km into continent at the french Atlantic coast, other references show similar values (i.e. Lerczak et al. (2001)) while Hughes and Veron (2018) show a maximum of 200 km inland movement in specific situations. In this section, we wish to address this feature by a case study during the low wind period from July 5 to 7, which should be prone to the development of see breeze systems. Figure 14 shows the time evolution of BSOA during both July 5 and 7 (a path to an animation is accessible in the SI-5). It shows a tendency of BSOA being transported to the coast by easterly winds during night-time (shown only in animations). Then they show the development of a see breeze front approximately starting at noon, and progressing into the land during the afternoon. This front is materialized by pronounced gradients in temperature (warmer on the land side), and the ridge in high BSOA in vicinity of the front. Wind fields well depict this see breeze system. These figures dramatically show that in addition to chemical BSOA formation, visible for example by comparing pictures for July 5 11h UTC and 14h UTC, such transport processes are major drivers for the BSOA variability over the Landes forest.

This phenomenon is also affecting the species time series at the Salles-Bilos site and needs to be taken into account to interpret the Landex campaign data, in addition to the general synoptic situation and local dispersion conditions on which previous studies have focused on already (Kammer et al. (2018), Bsaibes et al. (2020), Mermet et al. (2021)). Figure 15 shows the passage of the sea breeze front at Salles-Bilos, as a concomitant sharp decrease in BSOA and temperature during afternoon hours. Differences in the exact timing of the frontal passages are visible especially for July 7.
Figure 14. Sea breeze event for the 5th and 7th of July. The date and time for each panel is shown on the top of each panel, each row represents one variable/species, the name and the unity mentioned on the side.
Figure 15. Time series for BSOA (µg.m$^{-3}$), PM10 (µg.m$^{-3}$) and temperature (k) for the two sea breeze episodes compared to measurements. The episodes are highlighted on each panel. The confidence interval corresponds to the concentrations seen in a 11kmx11km square around the measurement site. The grey zones correspond to times of passage of the see breeze front.
6 Conclusions

This work aims to study the formation of BSOA in a forest canopy using the CHIMERE chemistry-transport model. For this purpose, we have examined the Landes forest, one of the largest pine forests in Europe located in south-western France. The forest is populated by a 95% majority of maritime pines, a mono and sesqui-terpene emitting tree species. It is also a region with rather homogeneous land use, and weak anthropogenic influences. The measurements used in this study were performed during the LANDEX episode 1 field campaign, carried out in the summer of 2017 with the goal of understanding the atmospheric chemistry related to biosphere-atmosphere interactions. At the Salles-Bilos site within the Landes forest, a set of detailed in-situ measurements has been obtained to illustrate these features.

In order to study the formation of BSOA we had to first focus on the entire physico-chemical system inside the forest canopy. Therefore, we performed multiple test cases improving the representation of the forest in the chemistry-transport model. It has to be kept in mind that while the goal is partially improving the performance of the model for this specific region, the sensitivity cases are also meant to show the effects of changing inputs or the simulation of different physical parameters on the concentration of different atmospheric species. Since the CHIMERE chemistry-transport model does not include an on-line canopy model, we have tried to mimic the simulation of a forest canopy by adding region-specific parameterizations.

The modelling chain consists of three nested domains, one covering the entire Europe, the intermediate one focused on France and the last nested domain focused entirely on the Landes Forest with a horizontal resolution of 1 km. In total, 6 sensitivity cases were tested, focusing on meteorological input fields, land cover inputs, improving the biogenic emission factors by correcting the type of trees present for this region in the MEGAN model, anthropogenic emissions and finally adding a parameterization prepared specifically for the Landes forest for the simulation of vertical diffusivity, wind speed and penetration of radiation inside the forest canopy.

Each of the studied scenarios shows effects on different species, for example not surprisingly NO\(_x\) are quite sensitive to changes in anthropogenic emission inputs, while terpenes show a double sensitivity to changing the emission factors as well as changing the vertical diffusivity and wind speed in the area. The scenario representing physical changes to the simulations of vertical diffusivity, wind speed and radiation penetration inside the canopy seems to have a more realistic view of what the measurements show about the atmospheric chemistry inside the forest, especially for terpenes, radicals and BSOA formation. Using the final test case, we also simulated the formation of BSOA from different precursors and oxidants. This showed us that the chemical pathways behind the BSOA formation for the base case scenario is inaccurate in the base case since the majority of BSOA is formed through the oxidation of isoprene, while this is changed to terpenes (specifically sesquiterpenes and \(\beta\)-pinene) in the final simulations. These final results are more in line with what is seen in the measurements for the Landes forest.

We also make evident spatial gradients between the forest and the surrounding areas especially for BSOA and short-lived species like mono- and sesqui-terpenes and isoprene. The formation of BSOA is quite localized inside the forest, consistent with high BVOC reactivity (which leads to BSOA formation in one step in our model, so in reality BSOA formation could take some more time as a multi-step process). We also could evaluate our simulations at regional scale and make evident an
improvement in almost all stations for all species measured at air quality stations located around the forested area. Still, for $PM_{2.5}$, a positive bias subsides, probably caused by the positive bias in OA, a major contributor to $PM_{2.5}$ concentration.

In addition, we analyzed the impact of transport processes within the forest area, and focused on sea breeze effects. We found for a period with low winds, sunny conditions and enhanced BSOA build-up, the sea breeze front passes over the Salles-Bilos measurement site and advects cleaner marine air masses. This manifests in a steep decline of temperature and BSOA at Salles-Bilos during afternoon. Thus, advection is important to be considered in addition to chemistry for well understanding observed time series, even for a measurement site carefully chosen in an environment with rather homogeneous land-use.

What needs to be kept in mind here is the necessity of a coupled canopy model or a sub-grid simulation scheme when it comes to simulating areas with particular physical representation like a forest. These types of sub-grid information exist for urban areas, but our results show that it is also important to put them into place for forested areas in CTMs. The next steps for this study are to investigate the reasons for the overestimation of the minima concentrations of $O_3$ inside the forest, as well as the overestimation by a factor of two of BSOA (for some peaks) after modifications. It is also important to understand the (still) overestimated concentrations of isoprene in the simulations, even after modifications in the tree types and land cover of the area. Further studies should also include the usage of different SOA simulation schemes in order to test their sensitivity to precursor/oxidant changes and the simulated concentrations of BSOA. Another interesting continuation of this work would be making hypothesis about the future of the forest: what effects would climate change modifying the land cover and potentially causing longer periods of drought have on a pine forest and its atmospheric chemistry? This question has become of an even greater importance and sad actuality given the massive forest destructions by the July 2022 fires.

**Code and data availability.** The model is available for download on the CHIMERE website (https://www.lmd.polytechnique.fr/chimere/2020_getcode.php). The WRF model is downloadable on their website (https://github.com/wrf-model/WRF). ECMWF data are available for download for registered members through request. The measurements can be provided by coauthors upon request. Simulation output will be available upon request. All the rest of the inputs/data used in the article are all downloadable free of charge on their respective websites mentioned in the text.

**Author contributions.** AC has performed the simulations and analyzed them. GS has provided meteorological and general inputs for the simulations. Main preparation of the article has been done by AC and MB. EV, EP, PF, MC, EO have provided measurements used in the article. EV is the responsible of the LABEX/COTE project, which the article is part of. All authors have contributed in rereading and editing the article.

**Competing interests.** The authors declare that they have no conflict of interest.
Acknowledgements. The authors would like to acknowledge the University of Bordeaux, the CNRS INSU LEFE program and the ANR in the frame of the Investments for the future Program, within the Cluster of Excellence COTE (ANR-10-LABEX-45) for their financial support. This work was granted access to the HPC resources of IDRIS under the allocation 2019-gen7232 and 2021-gen10274 made by GENCI. ATMO-NA (https://www.atmo-nouvelleaquitaine.org/, last accessed July 2022) is also acknowledged for both the air quality station measurements and the local anthropogenic emission dataset (Inventaire Atmo Nouvelle-Aquitaine 2014 – ICARE v3.2.1_rev1). They would also like to acknowledge the E-OBS dataset and the data providers in the ECA&D project (https://www.ecad.eu). Copernicus land data is acknowledged for the European forest dataset and also for leaf area index datasets which were generated using Copernicus Atmosphere Monitoring Service Information 2018. Neither the European Commission nor ECMWF is responsible for any use that may be made of the information it contains. IGN (https://www.ign.fr/, last accessed July 2022) is acknowledged for BDTopo and BDForet datasets.
References


