Country and species-dependent parameters for the Heating Degree Day method to distribute NO_x and PM emissions from residential heating in the EU-27: application to air quality modelling and multi-year emission projections

Antoine Guion[1], Florian Couvidat[1], Marc Guevara[2], and Augustin Colette[1]

[1]French National Institute for Industrial Environment and Risks (INERIS), Verneuil-en-Halatte (60550), France
[2]Barcelona Supercomputing Center (BSC), Barcelona (08001–08042), Spain

Correspondence: Antoine Guion (antoine.guion@ineris.fr)

Abstract. The combustion of fossil and biofuels in the residential sector can cause high background levels of air pollutants in winter, but also pollution peaks during cold periods. Its emissions are dominated by space heating and show strong daily variations linked to changes in outside temperatures. The Heating Degree Days (HDDs) approach allows to represent daily variations in space heating emissions. The method depends on a temperature threshold ("Tb") below which building heating is activated, and a fraction ("f") considering the relative contribution of space heating to total residential combustion emissions. These parameters are fixed in the literature. However, they are likely to vary according to the country and pollutant. Using statistics on household energy consumption, we provide country- and species-dependent Tb and f parameters to derive daily temporal factors distributing PM and NO_x emissions from the residential sector in the EU-27. Tested in the CHIMERE model, the simulations show better performance scores (temporal correlation and threshold exceedance detection) in winter, especially for PM, when compared to the simulation with a monthly temporal factor, or based on HDDs but using fixed parameters from the literature. Finally, the HDDs with fitted parameters are used as a method to project official annual residential combustion emissions in subsequent years, as these are typically reported with a 2-year time lag. Results show that this method performs better regarding the persistence method and remains within emission uncertainties for both PM and NO_x emissions, indicating the importance of considering HDDs for air quality forecasting.

1 Introduction

Among the various anthropogenic emission sectors contributing to the deterioration of air quality, the residential sector is particularly important. This sector can cause high background levels of pollutants during winter but also pollution peaks during specific cold periods (e.g. Juda-Rezler et al., 2011; Denier Van Der Gon et al., 2015; Cincinelli et al., 2019; Mbengue et al., 2020; Rudziński et al., 2022; Navarro-Barboza et al., 2024). Activities emitting air pollutants in the residential sector are directly linked to the energy consumption, mainly for heating and cooking. Composition of emissions due to combustion processes varies according to the type of fuel consumed (e.g. liquid fuel, solid biomass, gas). Nitrogen oxides (NO_x), carbon monoxide (CO), sulphur dioxide (SO_2) and particulate matter (PM) are identified as the main primary pollutants emitted from

the residential sector (Tammekivi et al., 2023).

Despite significant improvement over the last decades, the residential sector is still identified as a major contributor to pollution affecting most of urban areas in Europe. In 2021, residential combustion is estimated to contribute to 59% of total $PM_{2.5}$ emissions and 83% of total Benzo(a)pyrene (BaP) emissions in EU-27 countries, excluding shipping and aircraft emissions, according to official inventories (CEIP, 2023). Applying the source apportionment to measurements of ambient particulate concentrations in several European cities, Chen et al. (2022) estimated that biomass burning contributes to 12.4%±6.9% of organic aerosol concentrations on an annual average, and 16.9%±8.4% in winter. Based on modelling, the residential sector would be responsible for 22.7% and 10.3% on average of the mortality in European cities attributed to $PM_{2.5}$ and NO_2 pollution respectively (Khomenko et al., 2023). Urban centres are particularly affected by air pollution because of the high levels of population and emissions (e.g. Crippa et al., 2021). Nevertheless, through the advection of air masses and the long-distance transport of pollutants, the residential sector has an impact on suburban areas and represents a challenge for air quality on a regional scale (Mbengue et al., 2020; Stirnberg et al., 2021).

The impact of residential emissions on air quality can be estimated by using Chemical Transport Models (CTMs), which can be used to better understand the processes controlling air pollution and to evaluate targeted and effective emission reduction strategies. However, significant uncertainties remain regarding the quantification and spatio-temporal distribution of emissions from residential wood combustion (e.g. Denier Van Der Gon et al., 2015). An accurate representation of gas and particle emissions from the residential sector is necessary for use as input data in CTMs. CTMs rely on the use of gridded and temporally resolved emissions. Spatialized emissions are based on inventories calculated on an annual basis for each country and distributed spatially by using spatial proxies. The annual (or sometimes monthly) emissions from these inventories need to be distributed temporally at an hourly frequency by using temporal profiles (e.g. Guevara et al., 2021b; Kuenen et al., 2022). The use of these inventories raises two issues: (i) the temporalisation of annual residential emissions (via time factors) and the estimation of day-to-day variations accounting for the influence of meteorological conditions throughout the year, and (ii) the projection of annual emissions from past years to the current year, given that emission inventory submissions are reported with a 2-year time lag.

The temporal profiles for the residential combustion sector in the scientific literature remain generally simple and do not take into account the influence of meteorological conditions. For example, temporal profiles of the Netherlands Organisation for Applied Scientific Research TNO (Denier van der Gon et al., 2011) include monthly, weekly and hourly profiles with no spatial variation or weather dependency for this sector. The same applies to GENEMIS profiles (Ebel et al., 1997), but with a country dependency. The temporal profiles in the Emissions Database for Global Atmospheric Research EDGARv5 (Crippa et al., 2020) are provided as monthly climatologies. Therefore they only account for the impact of meteorology on residential combustion with a monthly temporal resolution. However, emissions from the residential sector, dominated by heating activity, have strong day-to-day variations, and should depend on the outdoor temperature (Considine, 2000). Based on measurements of Black Carbon (BC) and specific gaseous tracers of biomass burning (levoglucosan and mannosan), Mbengue et al. (2020) found a pronounced seasonal change in the contribution of residential heating, with significant daily variations. The residential emissions can even in some countries be subject to strong movement of population during the weekend or holidays. López-

Aparicio et al. (2022) showed that in the case of Norway the use of secondary homes can lead to important differences between weekdays and weekends in the spatialization of residential emissions. Modelling the day-to-day evolution of emissions from the residential sector is expected to improve the performance of CTMs simulations (e.g. Baykara et al., 2019). To this end, the Heating Degree Days (HDDs) approach was proposed to represent daily heating emission variations by using outdoor temperatures (Guevara et al., 2021b).

HDD is a concept initially used in the energy sector since it was demonstrated that variations in residential energy consumption can be inferred from weather conditions (Quayle and Diaz, 1980). HDDs are calculated from the difference between the current outdoor temperature and a given threshold (Thom, 1954). The latter threshold (hereafter referred to as "Tb") corresponds to an ambient temperature at which building heating is activated. The Tb value is critical because it directly modifies the threshold at which HDD are accumulated. This parameter should mainly ~~depends~~ depend on housing characteristics, the climate and local heating habits, and therefore may vary spatially. Over Europe, this value is usually set to 15.5°C as suggested by the MET-Office (weather forecast institute for the United Kingdom) or applied by Spinoni et al. (2015) for computing European HDD climatologies. However, values of Tb tested in different studies over Europe (Stohl et al., 2013; Mues et al., 2014) vary between 15°C and 18°C. To the best of our knowledge, there is no European data-set of Tb recommended for the use of HDDs yet. Another critical parameter combined with the HDD for calculating the temporal distribution of emissions is the relative contribution of space heating to residential emissions (hereafter referred to as "f") or in other words, the fraction of total emissions whose temporal variability is assumed to be driven by changes in the temperature. The value of f should vary according to the type of fuel consumed (e.g. gas or wood) and therefore on the type of species emitted.

Projecting annual emissions from past years to the current year for use in air quality forecasts can be an important issue, as meteorological conditions lead to inter-annual variability in emissions, particularly from heating. Within the Copernicus Atmosphere Monitoring Service (CAMS, Peuch et al. (2022)) regional air quality forecasting system, forecasts are based by running several CTMs with emission inventories from a past year. As the emission inventories used to produce air pollution forecasts for year n are generally not available until September of year n_{+2}, it is common practice to use emission inventories from a "reference" year (at least two years before) for the current year, extended to the following years by persistence. The disadvantage of this method is that the forecasts cannot take into account sudden changes in emissions, which could be due to societal (e.g. lockdown during the COVID-19 pandemic) or meteorological causes (e.g. milder winter). As the residential sector is sensitive to outdoor temperature, HDDs could be used to modulate these emissions from year n to n_{+2}, and therefore improve the forecasts (Guevara et al., 2022).

The work presented in this article aims to fulfill the two following objectives: (i) to assess the sensitivity of $PM_{2.5}$, PM_{10} and NO_2 surface concentration to different HDD-based experiments and identify the best parameterization ~~by comparision~~ in comparison to in situ observations, and (ii) use HDDs as a method to project national emission totals from the "other stationary combustions" sector C according to the Gridding Nomenclature for Reporting ($GNFR\_C$) and compare them with emissions assuming persistence and reporting uncertainty.

The article is divided into several sections. First, a set of country and species-dependent parameters for the HDD method (Tb and f) are determined based on national statistics on household energy consumption. Information on the different HDD formu-

lations as well as the spatialization of Tb and f parameters are provided in the Section 2. The modelling experiments carried out as part of this article are then detailed in the Section 3. Air quality simulations carried out using the CHIMERE regional CTM (Menut et al., 2021) over Europe for the full year 2018 are presented in Sect. 4.1. Finally, projections of total annual emissions from $GNFR\_C$ using HDDs are made between 2009 and 2018, and assessed against persistence and estimated emission uncertainty for the EU-27 countries (Sect. 4.2).

2 Heating degree day (HDD) method to distribute emissions

2.1 The HDD methodology

The HDD method and the description of its parameters (Tb and f) to infer a temporal factor (TF_{HDD}) used to distribute the total annual emissions on a daily basis for each country are presented in this section. The HDDs of the year n and the day d are computed for each grid cell of latitude i and longitude j by calculating the temperature difference between the daily average of the outdoor temperature at 2 meters $T2m$ and the ambient temperature Tb above which a building is no longer heated (fixed threshold) (Eq. 1). A minimum value of 0 is set for HDDs, assuming that there ~~is~~ are no space heating emissions when $T2m$ exceeds Tb. TF_{HDD} is then computed with the ratio between the daily HDDs and the annual cumulation of HDDs over the number of days N in the corresponding year (equal to 365 for non-leap year and 366 for leap year) (Eq. 2). The parameter f accounts for the fraction of household activities that are not sensitive to temperature variations, such as cooking and water heating, which are considered constant throughout the year. $Tb(c)$ and $f(c)$ are calculated by country c (see Sect. 2.2). Lastly, the total annual emissions $E(i,j,n)$ are distributed daily by applying $TF_{HDD}(i,j,d,n)$ (Eq. 3).

$$HDD(i,j,d,n) = max(Tb(c) - T2m(i,j,d,n), 0) \tag{1}$$

$$TF_{HDD}(i,j,d,n) = \frac{1}{N} \times f(c) + \frac{HDD(i,j,d,n)}{\sum_{d=1}^{N} HDD(i,j,d,n)} \times (1 - f(c)) \tag{2}$$

$$E(i,j,d,n) = E(i,j,n) \times TF_{HDD}(i,j,d,n) \tag{3}$$

2.2 Calibration of parameters by country

2.2.1 The non-temperature dependent fraction ($f(c)$)

As presented in the Introduction (see Sect. 1), the parameters $Tb(c)$ and $f(c)$ are critical in the formulation of $HDDs$, influencing the daily distribution of total pollutants emitted. f is generally fixed at a constant value in the scientific literature (e.g.

Mues et al., 2014; Spinoni et al., 2015), with no variation by country or species. The reference value of f is defined here at 0.2, following Guevara et al. (2021b) (hereafter referred to as $f_{ref.}$). In this work we propose a spatialisation of these parameters for the 27 member countries of the European Union (EU) based on national statistics on household energy use from countries with necessary information.

The fraction of emissions from the residential and commercial sector that is not related to heating, designated as the $f(c)$ parameter, is calculated for each country on the basis of the "Disaggregated final energy consumption in households" data-set (Eurostat, 2023). This data-set provides the quantity of energy consumed in households in European countries, disaggregated by type of fuels (according to the Standard International Energy Classification, SIEC) and activity (mainly space heating and cooling, water heating, cooking and lighting). The statistics for 2018 are used (as that year is selected in Section 3 for the simulations with the CHIMERE model).

Several energy sources are generally used in European households and their ~~proportion~~ proportions allocated to space heating are different. As PM and NO_x emissions come from very different energy ~~type~~types, the parameter $f(c)$ can be estimated for the two pollutants ($f_{PM}(c)$ and $f_{NOx}(c)$). PM emissions from residential heating come mainly from the consumption of solid fuels and oil that are used in fireplaces, stoves and oil-fired boilers. Based on the Eurostat data-set, the average fraction of the energy classes "solid fossil fuels, peat products, oil shale and oil sands" ($SFF - P1000 - S2000$ SIEC code) and "primary solid biofuels" ($R5110 - 5150 - W6000RI$ SIEC code) consumed for space heating in relation to all household activities are calculated for each EU-27 country, to derive $f_{PM}(c)$. The same fraction is calculated for the energy type "natural gas" ($G3000$ SIEC code) to represent the fraction of NO_x emissions from space heating ($f_{NOx}(c)$) that comes mainly from the use of gas boilers. As biogas is mainly used for transport, it is not included. Other pollutants, such as carbon monoxide (CO) and volatile organic compounds (VOCs) for example, are characterised by the reference value (0.2, Guevara et al. (2021b)).

Figure 1 presents $f_{PM}(c)$ and $f_{NOx}(c)$ for each EU-27 country. With the exception of a few countries, $f_{NOx}(c)$ remains between 0.1 and 0.4. Latvia (0.48), Poland (0.47), Romania (0.41) and Spain (0.54) use more than 40% of natural gas for household activities other than heating (e.g. water heating, cooking). Portugal stands out from the other countries, with a high $f_{NOx}(c)$ value (0.94). Portugal uses very little gas in its energy mix for domestic activities (9% compared to 32% for the EU-27 average) and this small fraction is mainly used for heating water (62% compared to 19% for the EU-27) and for cooking (35% compared to 6% for the EU-27). For most countries, the $f_{PM}(c)$ is less than 0.10. It is even equal to 0.01 for Belgium, Greece, Hungary, Malta and the Netherlands, which means that all solid fossil fuels are used exclusively for heating. With Portugal (0.29), Slovakia (0.25) and Finland (0.24) the highest, $f_{PM}(c)$ does not exceed 0.3 in the EU-27 countries.

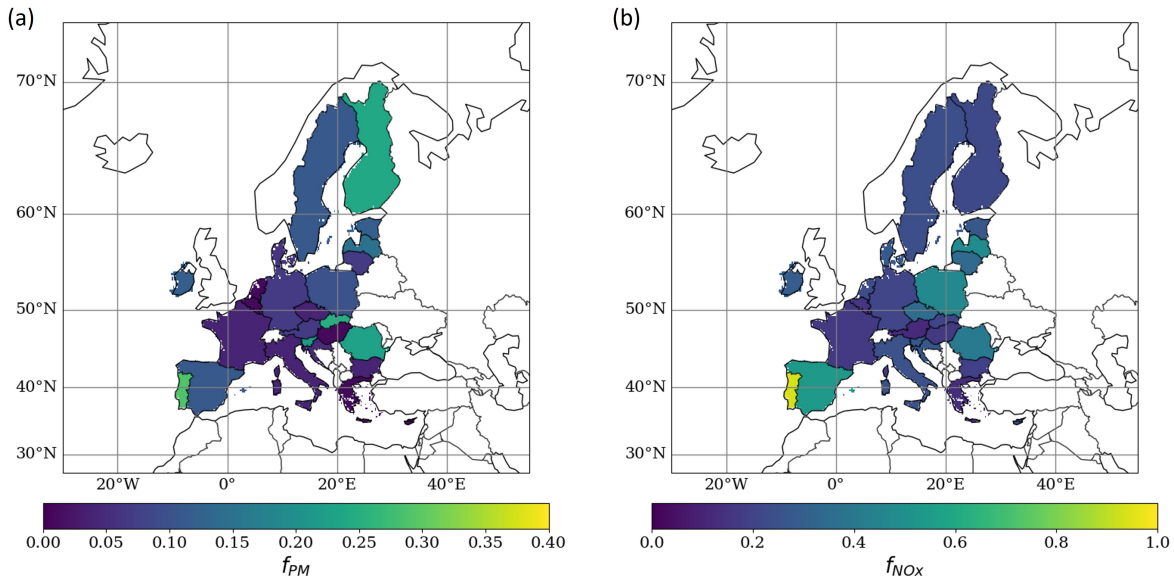

Figure 1. Parameter $f_{PM}(c)$ (panel (a)) and $f_{NOx}(c)$ (panel (b)) based on Eurostat (2023). The list of numerical values per country can be found in the Supplementary Material (SM, Tab. S1). For countries without data to compute $f(c)$, the available European average is assigned (i.e. 0.25 for $f_{NOx}(c)$ and 0.09 for $f_{PM}(c)$).

2.2.2 The temperature threshold ($Tb(c)$)

$Tb(c)$ is calculated by fitting it to the national domestic gas consumption statistics used as a proxy of heating use (hereafter referred to as $Tb_{fit}(c)$) taken from the Transparency Platform provided by The European Network of Transmission System Operators for Gas (ENTSOG, 2023). In this article, $Tb_{fit}(c)$ is compared with the reference threshold of 15.5°C, which does not vary by country (hereafter referred to as $Tb_{ref.}$).

The Transparency Platform is a web tool that provides technical and commercial data on gas transmission systems for several countries. Measurements of physical gas flow at a daily frequency are available at interconnection points and connections to different types of infrastructure: liquefied natural gas terminals, production facilities, storage facilities, ~~transmissions~~ transmission systems, distribution systems and consumer metering systems. As industrial consumers require large quantities of high-pressure gas for their plants, they are generally directly connected to the pipelines. This ensures that the distribution network (carrying low-pressure gas) is used for domestic purposes (both commercial and residential). For the purposes of this study, only data related to distribution systems and consumer metering systems are retained. In addition, only countries for which the proportion of gas used by households for space heating is at least 60% are kept, which ensures that daily gas consumption is an appropriate proxy for assessing the temperature threshold. Based on these criteria and available data, data are gathered for 8 countries (Hungary, Roumania, Italy, France, Belgium, Netherlands, Latvia and Estonia). These countries cover the different regions of Europe, being as representative as possible of the diversity of weather conditions and building construction. The

operator and the temporal coverage of gas flow data for each country can be found in SM (Tab. S2).

$Tb_{fit}(c)$ is calculated by fitting the country-averaged $TF_{HDD}(c)$ (see Eq. 2) to the temporal factor of domestic gas consumption ($TF_{gas}(c)$, see Eq. 4) by country using an optimisation solver by machine learning. Based on non-linear optimisation, the Nelder-Mead algorithm (Gao and Han, 2012) provides the minimised RMSE as the successful solution. The daily $TF_{gas}(c,d)$ for the year n is calculated as follows :

$$TF_{gas}(c,d,n) = \frac{Q(c,d,n)}{\overline{Q(c,n)}} \tag{4}$$

where $\overline{Q(c,n)}$ is the yearly average of distribution gas flow [kWh/day]. In order to reduce dependence on daily data for a specific year, which could be not representative of the country's domestic consumption, the fitting was based on data over several years (from 2 to 6 years over the 2016-2021 period depending on the country). $TF_{HDD}(c)$ is calculated ~~using surface temperature~~from the daily average temperature, using hourly $T2m$ data from the ERA5 reanalyses (Muñoz-Sabater et al., 2021).

An example with Belgium and Romania is shown in Figure 2 where $TF_{gas}(c)$ shows a marked seasonal cycle with a maximum in winter and a minimum in summer. It presents significant day-to-day variations during the winter period. Using the calculated $Tb_{fit}(c)$ (15.92°C and 16.63°C for Belgium and Romania respectively), $TF_{HDD}(c)$ manages to closely follow the evolution of $TF_{gas}(c)$. The peaks of gas consumption, corresponding to colder periods, are reproduced by the HDD parameterization. The time series of $TF_{gas}(c)$ and $TF_{HDD}(c)$ (for both $Tb_{fit}(c)$ and $Tb_{ref.}(c)$) for the other 6 countries are available in the SM (Fig. S1).

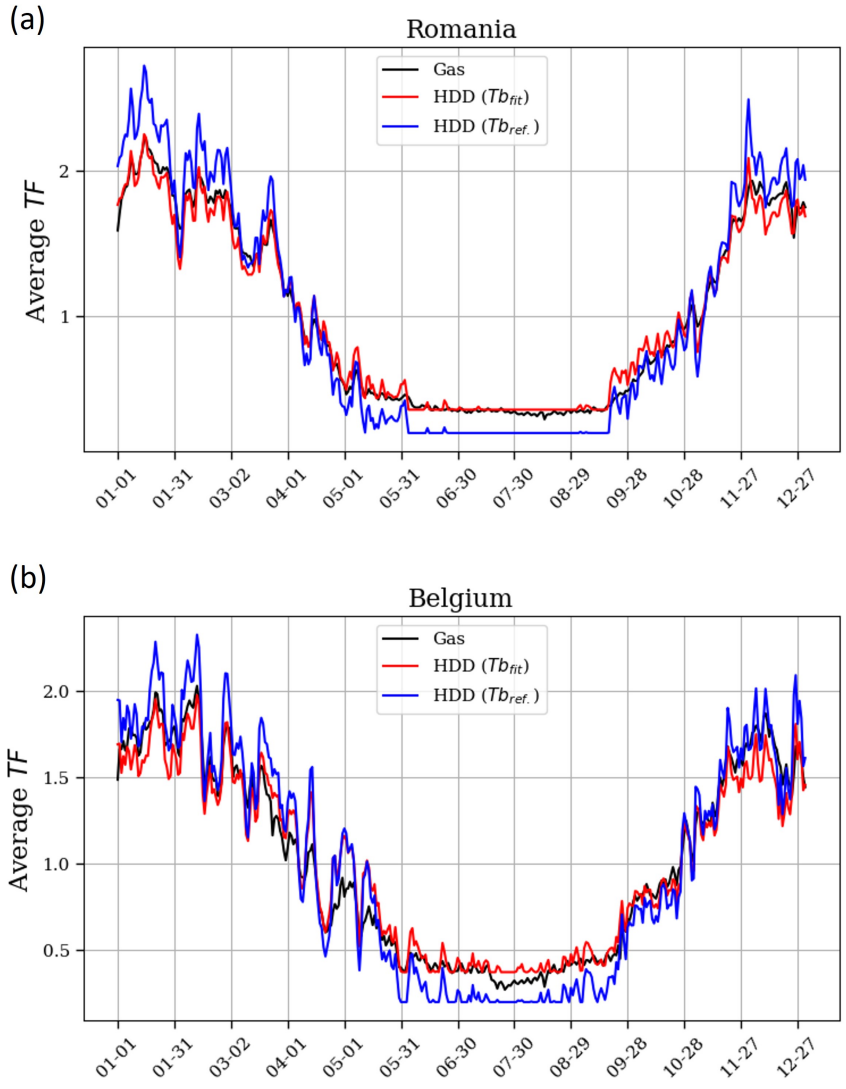

Figure 2. Daily evolution of $TF(c)$ (unitless) for gas consumption (in black), for HDDs with Tb_{fit} (in red) and for HDDs with $Tb_{ref.}$ (in blue). $TF(c)$ is illustrated for Romania (a) and for Belgium (b), averaged over the period 2018-2021 and 2016-2021 respectively.

Calculated $Tb_{fit}(c)$ values by country are shown in the Table 1. $TF_{HDD}(c)$ using $Tb_{fit}(c)$ and $Tb_{ref.}$ can be compared. The mean RMSE between $TF_{gas}(c)$ and $TF_{HDD}(c)$ is lower by -36% using $Tb_{fit}(c)$ compared to $Tb_{ref.}$ on average over the 8 countries. The $Tb_{fit}(c)$ values found for the 8 European countries highlight a latitudinal gradient decreasing from 15.68° for Italy to 11.49°C for Latvia. A lower $Tb_{fit}(c)$ means that heating in a given country is activated for a lower threshold of ~~ambiant~~ ambient temperature. Ciais et al. (2022) obtained a similar south-north gradient (ranging from 16.4°C to 13.5°C) by applying linear regressions between natural gas consumption and temperature data over 2016-2019. Grythe et al. (2019) used observed

BaP in situ concentrations, as a proxy of wood burning emissions, at five urban sites in Norway to determine a Tb around 10°C for Norway rather than the usual value of 15°C. The Finnish Meteorological Institute applies a Tb value for calculating HDDs that varies according to the season: 10°C from spring onwards and 12°C from autumn onwards (StatFin, 2023).

Table 1. Value of $Tb_{ref.}$ and $Tb_{fit}(c)$ for each country calculated from the national supplier's gas data using an optimisation solver (columns 1 to 3). The 3rd and 4th column compare the average RMSE between $TF_{gas}(c)$ and $TF_{HDD}(c)$ using $Tb_{ref.}$ and $Tb_{fit}(c)$.

Country	Value of $Tb_{ref.}$	Value of $Tb_{fit}(c)$	RMSE of TF_{HDD} with $Tb_{ref.}$	RMSE of TF_{HDD} with $Tb_{fit}(c)$
Hungary	15.50°C	17.69°C	0.23	0.08
Romania	15.50°C	16.63°C	0.18	0.08
Italy	15.50°C	15.68°C	0.20	0.12
France	15.50°C	15.96°C	0.11	0.09
Belgium	15.50°C	15.92°C	0.16	0.11
Netherlands	15.50°C	15.43°C	0.17	0.12
Latvia	15.50°C	11.49°C	0.26	0.22
Estonia	15.50°C	12.49°C	0.14	0.06

Finally, the European map of Tb_{fit} is derived by interpolating the values of the 8 fitted countries to the other EU-27 countries using the Inverse Distance Weighting (IDW) (Fig. 3). As Sweden and Finland are not in the interpolated domain, they have been assigned the averaged value of Estonia and Latvia (11.99°C). The interpolated value is calculated as the distance-weighted average of the central coordinates of the 8 neighbouring fitted countries, as follows :

$$x_p = \frac{\sum_{c=1}^{Z} \frac{x_c}{d_c}}{\sum_{c=1}^{Z} \frac{1}{d_c}} \tag{5}$$

where x_p is the interpolated value, x_c the $Tb_{fit}(c)$ value of the Z (=8) neighbouring countries c, and d the distance in km between the national central coordinates. For countries outside the EU-27, the value assigned is the average Tb_{fit} of the 8 countries, namely 15.16°C.

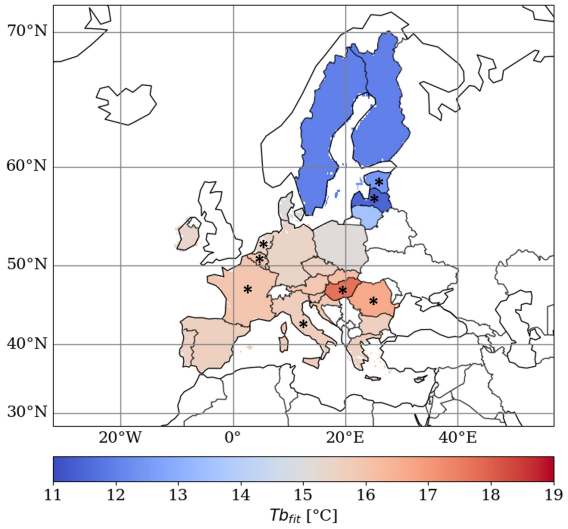

Figure 3. Estimated $Tb_{fit}(c)$ for EU-27 countries calculated on the basis of domestic gas consumption from the ENTSOG platform. The list of numerical values per country for $Tb_{fit}(c)$ can be found in SM (Tab. S1). Countries marked with a star are those for which the $Tb_{fit}(c)$ has been fitted. The other countries are interpolated using the IDW approach.

3 Setup of the numerical experiments to assess the impact of HDD methodology

3.1 Daily variability of emissions in the residential sector

Table 2 shows the different simulation experiments carried out in this work. Different configurations of HDD-based temporal factors are used to distribute anthropogenic emissions from $GNFR\_C$ over the year. The different experiments are as follows:

- "$MonthTF$" distributes emissions temporally with monthly profiles (m) for each grid cell (i, j) for the corresponding year (n), based on the calculation of monthly average HDDs. Therefore, $TF(i,j,m,n)$ is left constant for each month without accounting for day-to-day variation, as possibly found in the literature (e.g. Ebel et al., 1997; Denier van der Gon et al., 2011).
- "$DayTF\_ref.$" is based on HDDs to derive the daily $TF_{HDD}(i,j,d,n)$ (see Eq. 2). The reference parameters $f_{ref.}$(=0.2) and $Tb_{ref.}$(=15.5°C), as described in Guevara et al. (2021b), are used.
- "$DayTF\_Tb_{fit}$" is the same as "$DayTF\_ref.$" but uses $Tb_{fit}(c)$ (country dependent, see Fig. 3) in the calculation of $TF(i,j,d,n)$.
- "$DayTF\_Tb_{fit}\_f_{spec.}$" is the same as "$DayTF\_Tb_{fit}$" but uses $f_{spec.}(c)$ (country and species dependent, see Fig. 1) in the calculation of $TF(i,j,d,n)$.

Hourly profiles are taken from Guevara et al. (2021b) and are not modified between the different experiments. Hourly profiles vary according to the day of the week, particularly at weekends, and generally peak at 8am and 8pm, depending on the species emitted.

Figure 4 shows the seasonal variation of the daily temporal factors for the different experiments (European average) used to distribute $GNFR\_C$ emissions. There are significant variations on both a daily and seasonal scale between the experiments. The "$DayTF\_Tb_{fit}\_f_{spec.}$" experiment for PM shows the greatest seasonal variation, reaching a maximum of ∼2.2 in winter and a minimum of ∼0.2 in summer. The resulting effect of the experiments on the total anthropogenic emissions of fine and coarse particles, and NO_x can be found in SM (Fig. S2). Each HDD-based experiment is tested in order to both simulate air quality with the CHIMERE CTM model (see Sect. 4.1) over the year 2018 and to calculate multi-year projection of emissions (see Sect. 4.2) over the 2009-2018 period.

Table 2. Characteristics of the experiments conducted to distribute annual anthropogenic emissions from the "other stationary combustions" sector ($GNFR\_C$) over the grid domain i, j.

Experiment name	Temporal resolution	Tb parameter	f parameter
$MonthTF$	monthly	$Tb_{ref.}$ = 15.5°C	$f_{ref.}$ = 0.2
$DayTF\_ref.$	daily	$Tb_{ref.}$ = 15.5°C	$f_{ref.}$ = 0.2
$DayTF\_Tb_{fit}$	daily	$Tb_{fit}(c)$ (count. dependent, see Fig. 3)	$f_{ref.}$ = 0.2
$DayTF\_Tb_{fit}\_f_{spec.}$	daily	$Tb_{fit}(c)$ (count. dependent, see Fig. 3)	$f_{spec.}(c)$ (count. and spec. dependent, see Fig. 1)

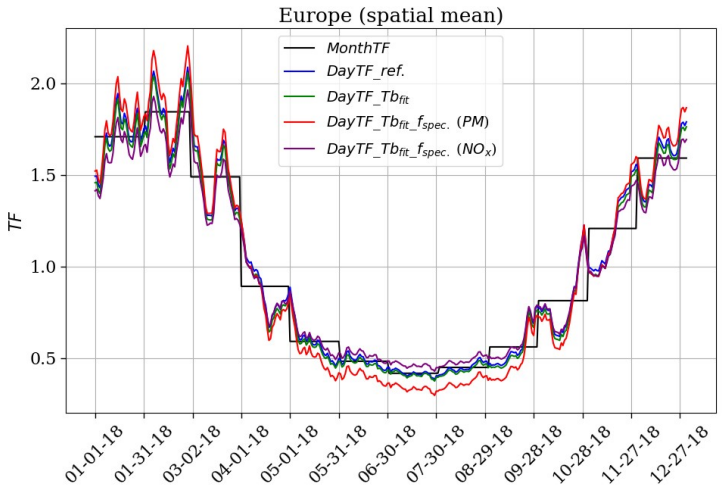

Figure 4. Daily temporal factor for $GNFR\_C$ for the different experiments detailed in Table 2 averaged over Europe for 2018.

3.2 Air quality simulations for the year 2018

3.2.1 Anthropogenic emission data-sets

Several sets of emissions data are used to meet the objectives of this work. Firstly, annual gridded totals of anthropogenic emissions are provided by the Regional Inventory for Air Pollutant "CAMS-REG-AP-v5.1" (Kuenen et al., 2022). The national totals by country come from the EMEP Centre on Emission Inventories and Projections (CEIP, 2023) which is responsible at European level for compiling the emissions of the State Parties to the Convention on Long Range Transboundary Air Pollution (LRTAP) for official publication. The annual emission totals reported for a given year are recalculated each subsequent year in accordance with any updated guidelines and/or data sets.

Based on specific spatial proxies for the distribution of emissions, "CAMS-REG-AP-v5.1" covers the European domain at a $0.05° \times 0.10°$ grid resolution. Specifically designed for air quality modellers, it provides emission for the main pollutants (NO_x, SO_2, non-methane VOCs (NMVOCs), NH_3, CO, PM_{10}, $PM_{2.5}$ and CH_4) at a wide range of sector levels . Following Denier Van Der Gon et al. (2015), condensable PM are represented in this inventory in a consistent way for all European countries. In addition to emissions from stationary residential combustion, the 'other stationary combustion' sector ($GNFR\_C$) also includes stationary commercial combustion (which can be assumed to behave in the same way as the residential sector) and stationary combustion from agriculture, forestry and fishing, which may not behave in the same way, but whose contribution to the total $GNFR\_C$ is negligible. Emissions from power stations and other industries are not covered by sector C, but by $GNFR\_A$ and B, respectively. The annual average spatialized contribution of sector C to total anthropogenic emissions of NO_x and PM over the period 2009-2018 (based on CAMS-REG-AP-v5.1) can be found in SM (Fig. S3). On average over EU-27 countries (national totals from CEIP (2023)), the contribution of $GNFR\_C$ is 9.1% for NO_x, 33.4% for PM_{10} and 51.6% for $PM_{2.5}$. Emissions of PM are ~~splitted~~ split into primary organic aerosol (POA), elemental carbon (EC) and other primary PM based on the CAMS speciation table.

Temporal profiles for other sectors than $GNFR\_C$ are taken from the "CAMS-TEMPOv3.2" product (Guevara et al., 2021b) that provides temporal profiles of European emissions of the main atmospheric pollutants. Those gridded temporal factors are available for monthly, daily, weekly and hourly cycles. Identifying the main emission drivers for each sector, the profiles are calculated on the basis of statistical information linked to the variability of the emissions (e.g. traffic counts) and parameters dependent on meteorology.

3.2.2 The CHIMERE configuration

The CHIMERE v2020r1 model (Menut et al., 2021), a regional three-dimensional Eulerian CTM, is used to simulate air quality over the year 2018 by testing different HDD parameterizations. CHIMERE represents the processes of gas-phase chemistry, aerosol formation, atmospheric transport and deposition. The chemical scheme used is MELCHIOR2 (Derognat, 2003) which includes 44 species and around 120 reactions. Aerosol microphysics and thermodynamics as well as secondary aerosol formation mechanisms are represented in the inorganic and organic aerosol module from Couvidat et al. (2012, 2018). POA are

assumed to be semivolatile and can partition between the gas- and particle- phase as a function of temperature and the concentration of organic aerosols. The gas-phase fraction can react with the hydroxyl radical and form lower volatility compounds via aging. Secondary organic aerosols (SOA) are formed with the Hydrophilic/Hydrophobic Organics (H^2O) mechanism.

The Fast-Jx module version 7.0b (Bian and Prather, 2002) ~~caclulates~~ calculates online the photolysis rates accounting for the radiative impacts of aerosols. Vertical and horizontal advection follow the scheme of Van Leer (1977). The physical and chemical time steps are 10 minutes. The simulated domain covers the Europe, with the following coordinates as corners: 30°N-72°N,-25°E-45°E. The spatial resolution is 0.2°x0.2° (around 20 km). 9 vertical layers from 998 up to 500 hPa are used. The meteorological fields are from the ~~operationnal~~ operational analysis of the Integrated Forecasting System (IFS) model of the European Centre for Medium-Range Weather Forecasts (ECMWF). Time varying boundary conditions for gas and dust aerosols are also provided by the IFS (Flemming et al., 2015; Rémy et al., 2022).

Natural emissions are calculated online: biogenic emissions using the MEGAN model (Guenther et al., 2006, 2012), sea salts and dimethyl sulfide marine emissions following the scheme of Mårtensson et al. (2003) and Liss and Mervilat (1986) respectively, and finally mineral dust emissions based on parameterization of Marticorena and Bergametti (1995); Alfaro and Gomes (2001). Emissions from forest fires are not included.

3.2.3 In situ observations for validation

In situ measurements of surface concentration are used to evaluate air quality simulations based on the CHIMERE model. These observations are taken from the European air quality observation database AQ e-Reporting (EEA, 2023), which gathers air quality data provided by EU members states and other EEA collaborating countries.

The analyses presented in this work focus mainly on NO_2, PM_{10} and $PM_{2.5}$ species. The stations selected are background stations (urban and rural type). For the calculation of daily averages and maximums, stations that do not cover an availability ratio of at least 75% of hourly measurements are excluded. Over the whole domain, our study considers 1831 stations for NO_2, 1070 for PM_{10} and 540 for $PM_{2.5}$ for the year 2018.

3.3 Modelling annual emissions between 2009 and 2018

3.3.1 Formula

Total annual emissions from $GNFR\_C$ can be projected over several years by calculating HDDs. While emission inventories take at least two years to be officially published, temperature reanalyses are available in a much shorter timeframe (from a few weeks to a few months). By assuming that changes in emissions from residential and commercial heating are dominated by the effect of the meteorology (rather than changes in emission factors and changes in heating habits), projected emissions ($E_{proj}(i,j)$) can be calculated by using the total gridded emissions $E(i,j)$ of the reference year n_{ref} and comparing HDD

factors between the projected year n_{ref+x} and the reference year n_{ref} (Eq. 6).

$$E_{proj}(i,j,n_{ref+x}) = E(i,j,n_{ref}) \times \left(\frac{1}{N} \times f + \frac{\sum_{d=1}^{N} HDD(i,j,d,n_{ref+x})}{\sum_{d=1}^{N} HDD(i,j,d,n_{ref})} \times (1-f) \right) \quad (6)$$

Gridded emission totals from the CAMS-REG-AP-v5.1 inventory (Kuenen et al., 2022) are used to model emissions of $PM_{2.5}$, PM_{10} and NO_x from $GNFR\_C$ between 2009 and 2018. The multi-year projections using the HDD method ~~is~~ are calculated for x=2 (equivalent to the delay in the official publication of emissions) and for x=3 (to assess the feasibility of the method in a longer publication scenario). They are compared with persistence (also for x=2 and x=3), which assumes that emissions in subsequent years are identical to those in the reference year.

3.3.2 Estimated uncertainty of nationally reported emissions

In the process of emission reporting of the LRTAP Convention, countries have to update emissions for several past years when they introduce a change in the emission calculation methodology. Therefore, to assess if the uncertainties introduced in the modelling of emission with HDD are comparable to emissions uncertainties, we compare those different reporting years for the same target year. As suggested in the Informative Inventory Report from CIEP/EMEP (Schindlbacher et al., 2021), the magnitude of the recalculation can provide a general estimate of the uncertainty of released emissions. This uncertainty (U_{+z}), expressed as a percentage, is calculated based on the average relative difference between total national emissions ($E(c)$) of the year (n) reported for a given year (yr) and the total for the same year, but reported in subsequent years (z) up to 2 years (U_{+2} with z = {1,2}) and 3 years (U_{+3} with z = {1,2,3}), as follows:

$$U_{+z}(c,n) = \frac{E_{yr+z}(c,n) - E_{yr}(c,n)}{E_{yr}(c,n)} \times 100 \quad (7)$$

$U_{+2}(c,n)$ and $U_{+3}(c,n)$ are calculated in this work for the official publication of national emissions totals for each year n between 2009 and 2018, then averaged over this period. Finally, the national totals of EU-27 countries in $GNFR\_C$ supplied by CEIP are used. This method enables uncertainty to be calculated quickly for a specific sector and by country.

4 Results

4.1 Impact on air quality modelling skills

4.1.1 Overview of the year 2018

In order to provide a general validation of the baseline simulation, surface concentrations simulated with the "$MonthTF$" experiment are compared with ~~AQ eReporting~~ AQ e-Reporting observations over the whole of 2018, with a focus on winter

(defined here from January to March) and autumn (from October to December). Table 3 shows the scores on daily concentrations averaged over all the European background stations. On annual average, the bias (model - observations) is negative for $PM_{2.5}$ (-3.09 $\mu g/m^3$), PM_{10} (-9.07$\mu g/m^3$) and NO_2 (-7.45 $\mu g/m^3$) and positive for O_3 (+8.27 $\mu g/m^3$). In winter, the bias for $PM_{2.5}$ becomes positive (+1.99 $\mu g/m^3$). The RMSE calculated over winter is lower (e.g. 4.30$\mu g/m^3$ for $PM_{2.5}$) than over the whole year (5.95$\mu g/m^3$ for $PM_{2.5}$) for each species analyzed. Conversely, it is higher when calculated over autumn (5.95$\mu g/m^{"}$ for $PM_{2.5}$). The correlation coefficient R on annual average is almost equal to or greater than 0.5 for all species (between 0.47 and 0.73). It increases considerably when calculated for winter, between 0.63 and 0.81.

It should also be pointed out that these mean values show considerable spatial and temporal variability. The temporal distribution of the daily bias (see Fig. S4 in SM) and the spatial distribution of the stations with their corresponding annual bias (see Fig. S5 in SM) are presented in the supplementary material. The majority of background stations included in the validation calculation are of the urban or suburban type (between 69% and 77% depending on the species). A representativeness bias may lead to a reduction in PM and NO_2 peaks at station points when the regional CTM simulates average concentrations over 20km grids. Nevertheless, these scores show an overall agreement with those presented in the latest articles using the 2020 version of CHIMERE (e.g. Menut et al., 2021; Guion et al., 2023), as well as with other regional CTMs (e.g. Bessagnet et al., 2016).

Table 3. Validation scores for the CHIMERE reference simulation ("$MonthTF$") calculated from ~~AQ-eReporting~~ AQ e-Reporting observations for $PM_{2.5}$, PM_{10}, NO_2 and O_3 species, averaged over Europe in 2018.

	Observations [$\mu g/m^3$]	Model [$\mu g/m^3$]	Bias (mod.-obs.) [$\mu g/m^3$]	RMSE [$\mu g/m^3$]	Pearson corr. (R)
			$PM_{2.5}$		
Annual	13.38	10.29	-3.09	5.95	0.57
JFM	15.47	17.46	1.99	4.30	0.78
OND	11.80	13.87	2.07	6.93	0.55
			PM_{10}		
Annual	20.25	11.18	-9.07	13.50	0.47
JFM	22.97	16.17	-6.80	8.76	0.72
OND	20.66	12.97	-7.69	13.87	0.55
			NO_2		
Annual	16.23	8.78	-7.45	9.27	0.58
JFM	19.87	14.42	-5.45	8.82	0.63
OND	18.32	10.42	-7.90	10.23	0.56
			O_3		
Annual	56.58	64.85	8.27	17.55	0.73
JFM	48.93	59.12	10.19	15.76	0.81
OND	43.24	57.72	14.48	20.07	0.70

The effect of the different HDD parameterizations integrated in the emissions used in CHIMERE is analysed by comparing the Spearman correlation and RMSE scores in concentration with the available observations (stations averaged by country) for each experiment and more specifically for each country for which the Tb parameter has been calculated based on the national gas data. Using ~~a~~ an HDD-based temporal profile, therefore with daily variation by comparison to the baseline simulation "$MonthTF$" experiment, mainly affects the simulated concentrations of $PM_{2.5}$, PM_{10} and NO_2 from January to March and from October to December (see Fig. S4 in SM). As a result, the spring and summer seasons, when wood heating is almost non-existent in both southern and northern Europe, are not included in the analysis here. The case of the beginning of 2018, characterized by negative temperature anomalies (corresponding to cold spells occurring on a European scale) and high residential heating emissions, will be detailed in the subsection "Threshold level exceedance during the cold spells at the beginning of 2018" (4.1.2) with a specific analysis on peak concentrations.

Fig. 5 shows the performance metrics for $PM_{2.5}$. There are no observations available from AQ e-Reporting for this pollutant in 2018 for Hungary, Romania, Estonia and Latvia. Analyzed individually by country, performance between experiments can

vary significantly. For Italy, France, Belgium and Netherlands in winter, the "$DayTF\_Tb_{fit}\_f_{spec.}$" experiment presents the best temporal correlation, with an increase of the Spearman coefficient by +8.7%, +7.9%, +10.5% and +8.6% compared to the "$MonthTF$" experiment. In terms of RMSE scores, "$DayTF\_Tb_{fit}\_f_{spec.}$" remains close to the "$MonthTF$" experiment, although there is a slight decrease for France and the Netherlands in winter (-0.5 and -1.3$\mu g/m^3$ respectively). The "$DayTF\_Tb_{fit}$" experiment ~~(with $f_{ref.}$ = 0.2)~~ leads to a significant decrease in the RMSE for Belgium and the Netherlands (~~-1,6~~ -1.6 and -3.0$\mu g/m^3$ compared to "$MonthTF$" respectively), which can be explained by a lower overestimation of the modelled concentration peaks in these countries. Indeed, the f_{PM} parameter is lower in "$DayTF\_Tb_{fit}\_f_{spec.}$" (0.01 for Belgium and Netherlands) than $f_{ref.}$ (0.20). As the simulations show a positive bias for these two countries (already the case with "$MonthTF$", see Fig. S5), the RMSE scores are higher for "$DayTF\_Tb_{fit}\_f_{spec.}$" than for "$DayTF\_Tb_{fit}$". The validation scores obtained also depend on the accuracy of the values provided by the Eurostat dataset. The bias scores can be found in the SM (see Fig. S11).

For the autumn season, the HDD parameterization ~~do~~ does not appear to have any beneficial effect on performance scores. While the correlation varies very little (about -0.03 for R compared to "$MonthTF$"), the RMSE increases, whatever the parameterization (+~~2,9~~2.9$\mu g/m^3$ on average over the four countries compared to "$MonthTF$"). However, autumn 2018 does not appear to be the most relevant period for assessing the impact of HDD, as it was relatively warmer than average, with no major european cold spell (see Fig. S12 in SM).

Performance scores can vary considerably from one country to another (and even more so than between experiments which, as a reminder, differ only in the parameters of the temporal distribution of heating emissions). This is discussed in more detail in the discussion (see Sect. 5), but national differences in the calculation of emissions and the representation of certain processes such as long-distance transport have an impact on the simulation scores for each country.

The equivalent figure for PM_{10} can be found in the SM (see Fig. S10), with results similar to those for $PM_{2.5}$.

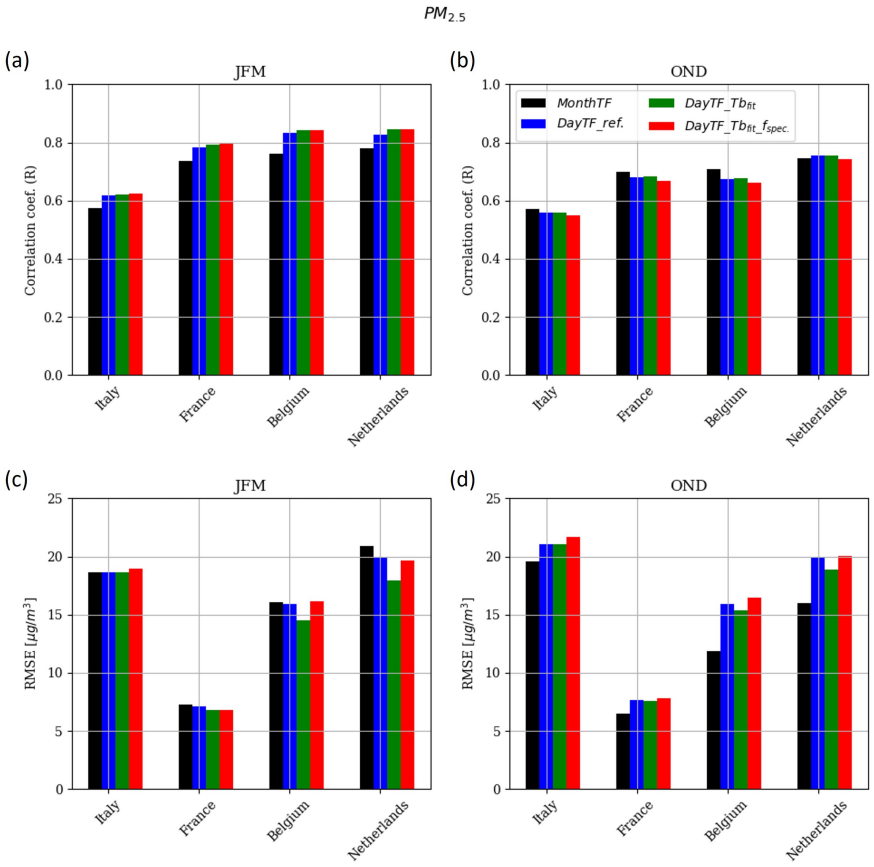

Figure 5. Spearman correlation (R coefficient) and RMSE ($\mu g/m^3$) of hourly $PM_{2.5}$ concentrations for the months JFM (panel (a) and (c) respectively) and OND (panel (b) and (d) respectively), averaged over stations in countries that have been fitted with gas consumption data and for which concentration measurements are available.

The spatial distribution of score variations induced by the "$DayTF\_Tb_{fit}\_f_{spec.}$" experiment (shown in Figure 6 for $PM_{2.5}$ in JFM) allows us to identify the regions most sensitive to the adjusted HDD parameters. The Spearman temporal correlation is considerably better (compared to "$MonthTF$"), relatively uniformly between and within countries (up to +0.2), with the exception of Belgium, Southern Netherlands and Western Poland, where the increase in the R coefficient is smaller (about +0.05). This analysis also shows that the improvement in scores also concerns countries for which the Tb_{fit} has been interpolated. Being larger in Eastern Europe (about -2$\mu g/m^3$), the decrease in RMSE with experience "$DayTF\_Tb_{fit}\_f_{spec.}$" does not concern all countries. The change in RMSE is almost zero in Belgium, the Netherlands, Southwestern Germany, Eastern Poland, and Portugal. The spatial variations in scores for PM_{10} are similar to those for $PM_{2.5}$, while they are very small for NO_2 (as discussed below).

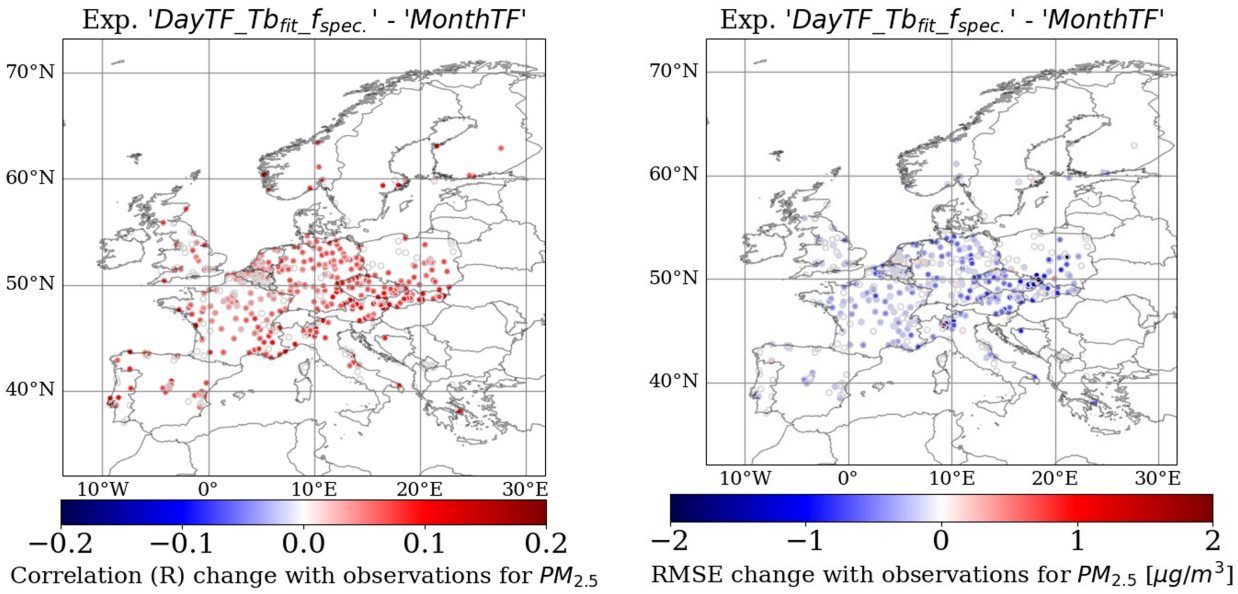

Figure 6. Average change in Spearman correlation (R coefficient) and RMSE ($\mu g/m^3$) (left and right respectively) at stations between the experiment "$DayTF\_Tb_{fit}\_f_{spec.}$" and "$MonthTF$" simulated in JFM for $PM_{2.5}$.

Fig. 7 shows the performance metrics for NO_2. With the exception of the Netherlands, Romania and Latvia, the correlation coefficient in winter varies slightly between the CHIMERE experiments. However, "$DayTF\_Tb_{fit}\_f_{spec.}$" has the lowest RMSE in winter for Italy, France, Hungary, Romania, Estonia and Latvia, with an average decrease of -4.4% compared to "$MonthTF$". For the autumn season, HDD-based experiments do not improve the correlation coefficient (for all parameterizations tested), but the RMSE decreases by -3.6% on average for Hungary, Romania, Estonia and Latvia for "$DayTF\_Tb_{fit}\_f_{spec.}$" compared to "$MonthTF$".

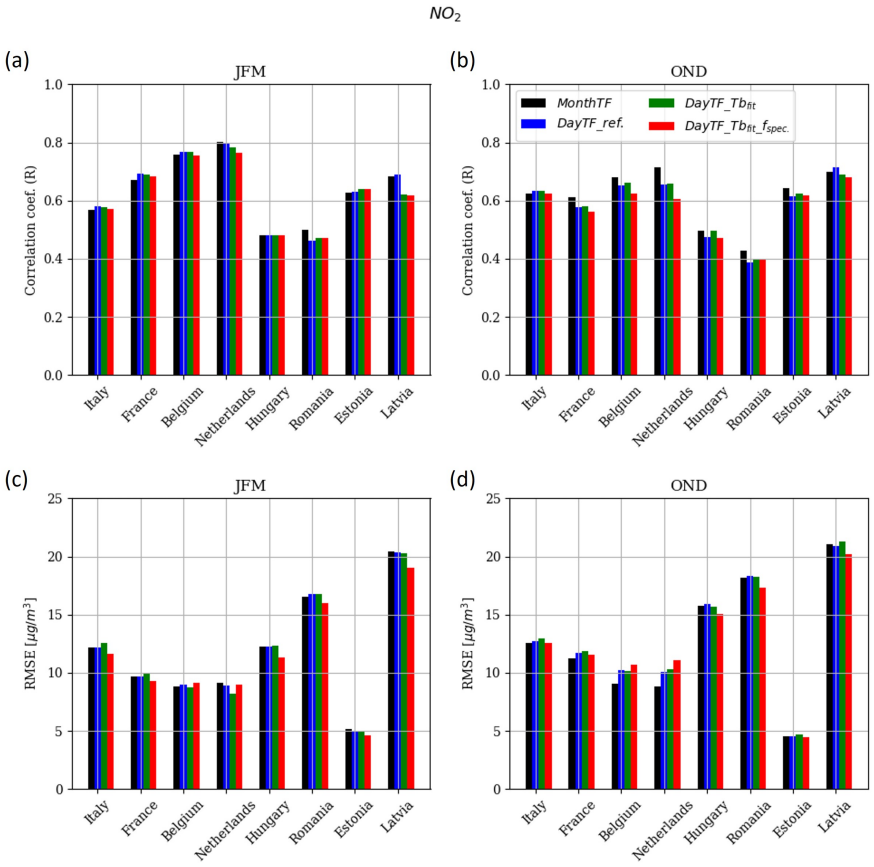

Figure 7. Same as Figure 5 for the NO_2 species.

4.1.2 Threshold level exceedance during the cold spells at the beginning of 2018

February and March 2018 were considerably colder than seasonal normals (up to -4°C on average across Europe). Several periods of intense cold have been identified and documented by the Copernicus Climate Change Service (C3S) (see Fig. S12 and S13 in SM). Three cold spells were reported to have affected most of the European region: from 7 to 10 February, from 26 February to 4 March and from 18 to 27 March. These periods are particularly relevant for analysing the use of HDDs when residential heating increases considerably and therefore to assess the model capabilities to capture threshold exceedance (e.g. Chen et al., 2017).

Figure 8 shows the temporal evolution of simulated concentrations of $PM_{2.5}$, PM_{10} and NO_2 during JFM, averaged over Europe. The example of France is also shown, with an adjusted identification of cold periods throughout the country from Météo-France (MF, 2019). Time series for other countries can be found in SM (see Fig. S7, S8 and S9). The simulation without HDD ("$MonthTF$") fails to correctly reproduce the variability driven by cold spells and relatively warmer winter periods, especially from February to March. Compared to $PM_{2.5}$ and PM_{10} observations averaged over Europe, the "$MonthTF$"

simulation completely misses the concentration peaks of 3 March (underestimation of around 20$\mu g/m^3$) and, in the case of $PM_{2.5}$, overestimates concentrations over the periods 15-20 February and 07-17 March (by +2 to +5$\mu g/m^3$).

The various peak concentrations of $PM_{2.5}$ and PM_{10} observed during cold periods were relatively well simulated by the HDD-based simulations. The different parametererizations can lead to differences of several $\mu g/m^3$ during PM peaks. The "$DayTF\_Tb_{fit}\_f_{spec.}$" simulation is the experiment that best reproduces the PM peaks during periods of intense cold at the European scale, with the exception of the last $PM_{2.5}$ peak around 25 March, when concentrations were slightly overestimated (up to 2$\mu g/m^3$).

For NO_2 species, concentrations between HDD-based experiments can vary slightly (up to 2$\mu g/m^3$) and more widely compared with "$MonthTF$" (up to 6$\mu g/m^3$). Compared to the European observations, all experiments tend to underestimate daily concentrations by -5 to -10$\mu g/m^3$. This does not apply to all countries, such as Estonia, Belgium and the Netherlands (see Fig. S7 and S9 in SM). Except for the peak on 8 February which was well modelled, NO_2 peaks are less well represented than PM. However, NO_2 peaks do not always correspond to cold periods. It should be pointed out that observations of NO_2 concentration levels during cold spells are not significantly higher than during other periods. This suggests that background NO_2 concentrations are less sensitive to abrupt variations in emissions linked to residential heating, such as during cold weather events. Recent literature (e.g. Grange et al., 2019; Wærsted et al., 2022) highlights a possible significant sensitivity of NO_2 concentrations to changes in road transport emissions linked to temperature changes.

The same analysis was carried out with the maximum daily concentration (see Fig. S6 in SM) and the findings are similar. To complete this spatially averaged analysis of JFM 2018, a specific study of threshold exceedance calculated at each measurement station is presented below, detailing the differences in performance in simulating high concentrations between HDD configurations.

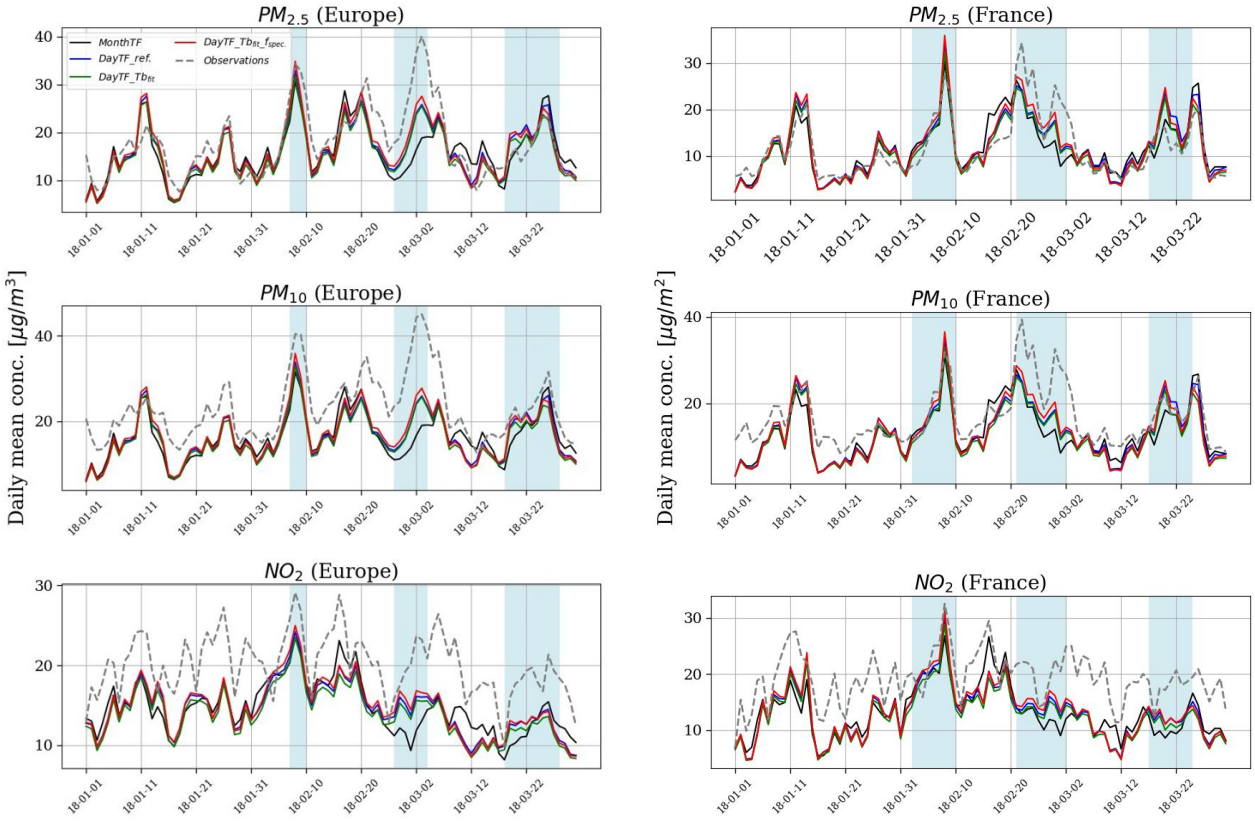

Figure 8. Average daily concentrations [$\mu g/m^3$] of $PM_{2.5}$, PM_{10} and NO_2 in Europe (left) and in France (right) between January and March from observations (~~AQ eReporting~~AQ e-Reporting) and different CHIMERE simulations. The blue areas indicate periods of intense cold as signaled by the C3S and MF for Europe and France respectively.

The number of exceedance of daily concentration thresholds of 25$\mu g/m^3$ for $PM_{2.5}$, 50$\mu g/m^3$ for PM_{10} and 40$\mu g/m^3$ for NO_2 (based on the Directive 2008/50/EC of the European Parliament and of the Council on ambient air quality and cleaner air for Europe) have been calculated at each monitoring station over Europe for JFM months of 2018. The table 4 shows the average number (per station) of good detection, false alarms and missed alarms, in regards to the observed concentrations.

Compared to the "$MonthTF$" experiment, the number of good detection increases and the number of missed alarm decreases when HDDs are included, for all combinations of parameters for the species $PM_{2.5}$ (+6% on average for good detection and -4% for missed alarms), PM_{10} (+41% and -7%) and NO_2 (+6% and -2%). The countries most concerned by these changes, according to the network of measuring stations available, are Belgium, the Netherlands, Poland and Slovakia. The number of false alarms decreases with "$DayTF\_ref.$" (-12% for PM_{10} and -2% for NO_2) and "$DayTF\_Tb_{fit}$" (-12% for $PM_{2.5}$, -20% for $PM_{2.5}$ and -11% for NO_2), and increases with "$DayTF\_Tb_{fit}\_f_{spec.}$" (+3% for $PM_{2.5}$, +1% for $PM_{2.5}$ and +19% for NO_2). However, this percentage increase in false alarms for "$DayTF\_Tb_{fit}\_f_{spec.}$" remains lower than its increase in

good detection for $PM_{2.5}$ and PM_{10}.

Among the different HDD configurations, the experiment "$DayTF\_Tb_{fit}\_f_{spec.}$" has the largest increase of good detection (+11% for $PM_{2.5}$, +57% for PM_{10} and -15% for NO_2) and the largest decrease of missed alarms (-7% for $PM_{2.5}$, -9% for PM_{10} and -5% for NO_2). Finally, "$DayTF\_Tb_{fit}\_f_{spec.}$" presents the best probability of detection for $PM_{2.5}$ (0.45), PM_{10} (0.21) and NO_2 (0.31), as illustrated on the performance diagrams (designed by Roebber (2009)) in SM (see Fig. S14).

Table 4. Average number per station (over Europe) of good detection, false alarms and missed alarms simulated by CHIMERE for the "$MonthTF$" experiment and variation in % with the other HDD-based experiments for JFM. ~~The~~ Based on the Directive 2008/50/EC, the threshold values for average daily concentrations not to be exceeded were set at 25$\mu g/m^3$ for $PM_{2.5}$, 50$\mu g/m^3$ for PM_{10} and 40$\mu g/m^3$ for NO_2.

Experiment	$MonthTF$ [average occurence per station over Europe]	$DayTF\_ref.$ [ave. occ. (% change)]	$DayTF\_Tb_{fit}$ [ave. occ. (% change)]	$DayTF\_Tb_{fit}\_f_{spec.}$ [ave. occ. (% change)]
		$PM_{2.5}$		
Good detection	7.03	7.32 (+4%)	7.18 (+2%)	7.80 (+11%)
False alarms	6.92	6.92 (±0%)	6.06 (-12%)	7.14 (+3%)
Missed alarms	10.80	10.51 (-3%)	10.65 (-1%)	10.03 (-7%)
		PM_{10}		
Good detection	0.89	1.17 (+32%)	1.19 (+34%)	1.39 (+57%)
False alarms	2.19	1.93 (-12%)	1.76 (-20%)	2.22 (+1%)
Missed alarms	5.66	5.37 (-5%)	5.36 (-5%)	5.15 (-9%)
		NO_2		
Good detection	1.95	2.02 (+4%)	1.95 (±0%)	2.19 (+12%)
False alarms	1.67	1.65 (-2%)	1.49 (-11%)	1.99 (+19%)
Missed alarms	5.21	5.14 (-1%)	5.21 (±0%)	4.97 (-5%)

4.2 Multi-year emission projections

~~This second part is devoted to the use of HDDs to estimate the total annual emissions of~~ Based on temperature, HDDs can be used to project emissions from the $GNFR\_C$ sector for a given year as an annual total (see Eq. 6), but also in near-real time on a daily time step (by normalising the HDDs of the current day with the daily average HDDs of the base year). The use of HDDs to simulate the day-to-day variability of residential emissions does not carry a specific risk in the context of simulations for a past year (reanalyses), as the annual heatsum of the corresponding year is known and can be used to normalised total emissions. On the contrary, when used in a forecast setup, using HDDs can induce a deviation from the input emissions, as the heatsum of the running year is unknown. This deviation is legitimate as in a colder (resp. milder) than expected winter, emissions should

rightfully be larger (resp. lower) than originally prescribed. There is no reason ~~than~~ that using emission from a past year (such an approach, referred to as persistence, is routinely used for the operational forecast) would be more legitimate. It is however important to document that risk of deviation. This second part is therefore devoted to the use of HDDs to estimate the total annual emissions of the $GNFR\_C$ sector.

In this section we compare the annual total emission modelled with HDDs to the uncertainty related to emission reporting mechanism (see Sect. 3.3.2), and we also compare the deviation when using a persistence approach. Because of the spatialisation of the $Tb_{fit}(c)$ and $f(c)$ parameters and the best modelling results in terms of concentration peaks and threshold exceedance, the HDD-based experiment "$DayTF\_Tb_{fit}\_f_{spec.}$" is used in this section. As the analyses results on $PM_{2.5}$ emissions are very similar to those for PM_{10}, the results in this section will be presented only for $PM_{2.5}$ and NO_x emissions. Figure 9 shows the average reported emissions of $PM_{2.5}$ and the average relative difference using "$DayTF\_Tb_{fit}\_f_{spec.}$" and the persistence for projected years n_{+2} and n_{+3} (over the 2009-2018 period). The same figure for NO_x emissions is available in SM (see Fig. S15). On a European scale (grid average) for $PM_{2.5}$ (NO_x) emissions, the relative difference in n_{+2} is 4.2% (5.0%) for the HDD-based projections and 4.6% (5.3%) for the persistence. The deviation from reported emissions varies spatially, depending on the country, both for HDD-based projections and for persistence. For France, Germany, Norway, Ireland and Latvia, the HDD method leads to significantly smaller differences with the reported $PM_{2.5}$ emissions (by about -10%) than the persistence method.

As expected ~~higher deviation are~~ a higher deviation is obtained for n_{+3} than for n_{+2} (with both the HDD and persistence approach) at the European level, but the HDD method again shows a lower average deviation (5.2% for $PM_{2.5}$ and 6.7% for NO_x) than persistence (6.2% for $PM_{2.5}$ and 7.8% for NO_x). Norway, Latvia and Lithuania benefited from a considerable reduction of relative difference with the HDD method for $PM_{2.5}$ emissions (by about -10%). Algeria shows a larger relative difference (compared to the persistence), but its total emissions remain low compared with other European countries. Finally, an interesting feature of the HDD method is that it ~~can~~ can provide more detailed spatial information within the same country (e.g. southeastern France, northern Italy), since HDDs depend on gridded temperature fields.

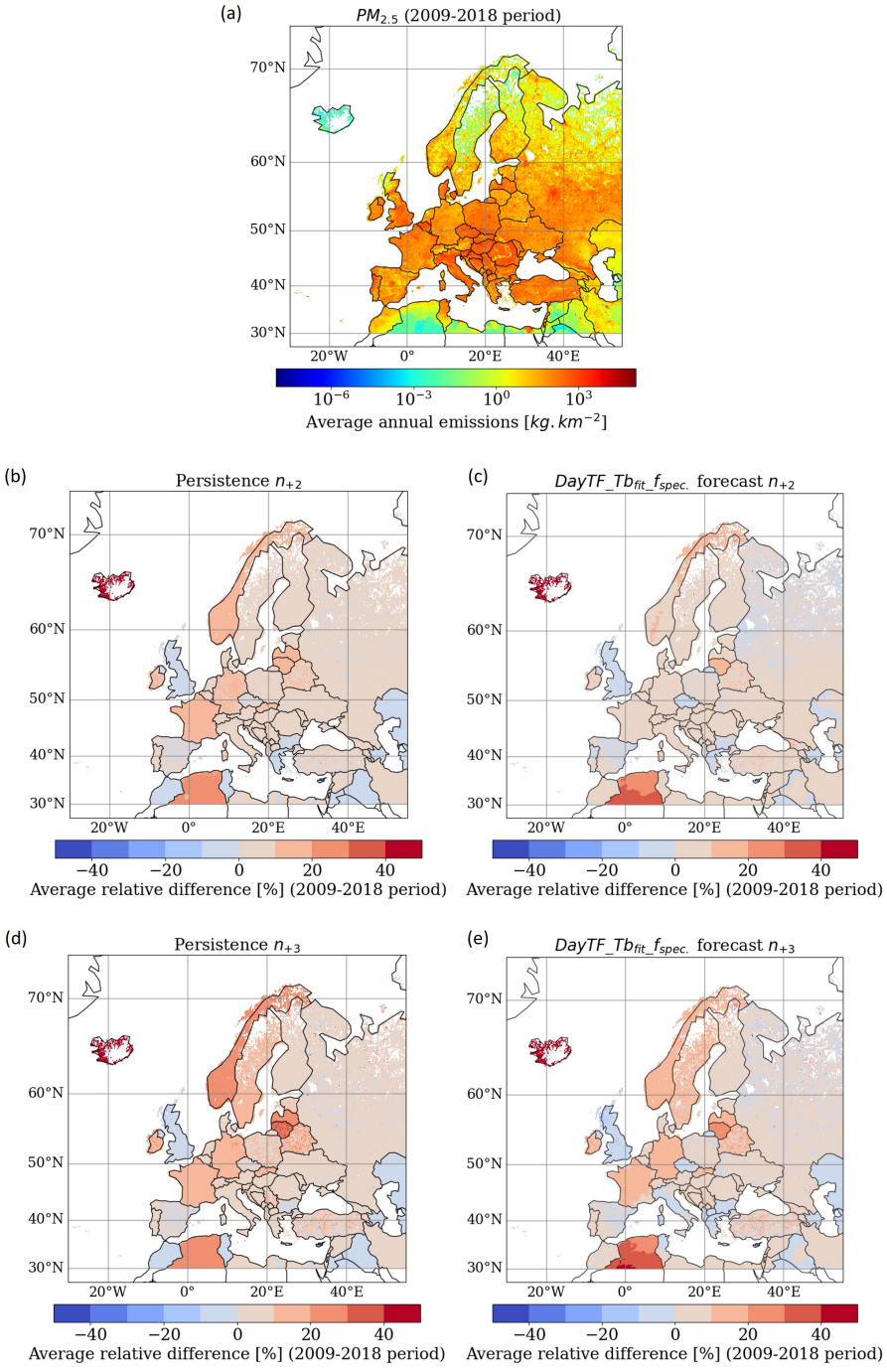

Figure 9. Spatial distribution of average annual $PM_{2.5}$ emissions $[kg/km^2]$ from $GNFR\_C$ (2009-2018 period), based on the CAMS-REG-AP-v5.1 inventory (a). Average relative difference using "$DayTF\_Tb_{fit}\_f_{spec.}$" to project in n_{+2} (b) in n_{+3} (d) each year between 2009 and 2019, compared to the reported emissions. Average relative difference using the persistence method for n_{+2} (c) and n_{+3} (e) over 2009-2018.

The average relative deviations induced by the "$DayTF\_Tb_{fit}\_f_{spec.}$" projection and the persistence method are compared to the reporting uncertainty estimated from the variability of emission reporting (see Sect. 3.3.2) for each EU-27 country. An example is given with France in Figure ~~11 and 12~~ 10 over the 2016-2018 period. Compared with 2016, 2017 and 2018 were warmer (+0.3°C and 0.8°C respectively on annual average) and the number of HDDs therefore decreased. Based on this method, a decrease in $PM_{2.5}$ emissions has been projected, in line with what has been officially reported. A deviation of only 1.49% from the official reporting has been calculated, which is still less than the estimated uncertainty of 3.0% (based on recalculations of emissions up to two years later). In the absence of inter-annual variability, the persistence method misses this decrease, and its deviation from the official report reaches 11.61%.

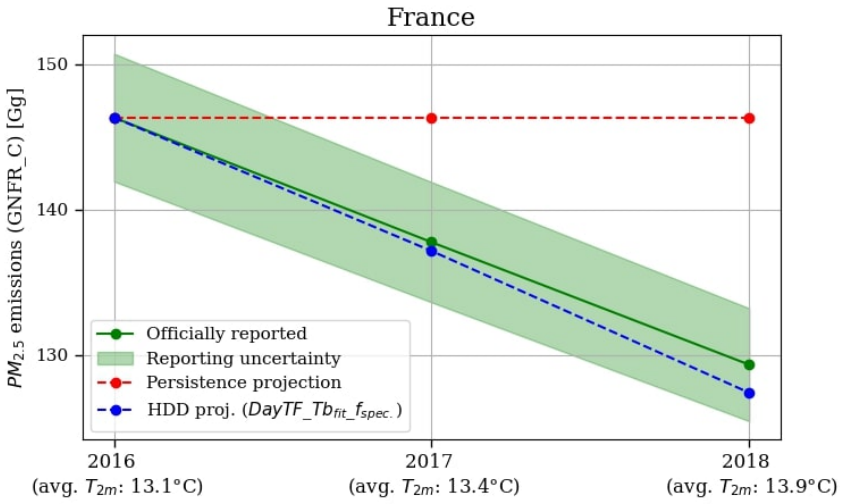

Figure 10. Example of a projection of total nation emissions ($GNFR\_C$) [Gg] for France up to years n_{+2} (2017 and 2018) from year n 2016. The blue curve represents the projection using the HDD method ("$DayTF\_Tb_{fit}\_f_{spec.}$"), the red curve the persistence projection, the green curve the official reported emissions (CEIP, 2023) and the green shaded area the reporting uncertainty estimated up to n_{+2} years.

The same analyses was carried out over the 2009-2018 period for $PM_{2.5}$ (Fig. 11) and NO_x emissions ~~respectively.~~(Fig. 12) respectively, by averaging the different projections to n_{+2} and n_{+3} over this period. For the $PM_{2.5}$ emissions projections to n_{+2} and n_{+3}, HDD-based projections and persistence fall within the same range of values; ranging from a few percents up to $\sim$30% depending on the country. However, more than half of the countries (18 out of 27) have a smaller average difference in reported $PM_{2.5}$ emissions with HDD-based projections than with persistence (both for n_{+2} and n_{+3}). The average reporting uncertainty (over the period 2018-2019) after 2 years (U_{+2}) is 17.8% for the EU-27 and rises to 23.4% after 3 years (U_{+3}). This estimated uncertainty remains greater than the deviation induced by the HDD method (for the EU-27 average but also for 16 EU countries individually in n_{+2} and 18 countries in n_{+3}). For the other countries, the uncertainty is admittedly lower, but remains close to the calculated difference (rarely more than 10%).

The average reporting uncertainty for NO_x emissions is 14.1% up to 2 years (U_{+2}) and 15.7% up to 3 years (U_{+3}) for the EU-27 average, being lower than for $PM_{2.5}$. However, it remains higher than the average difference induced by the HDD-based projections: 10.4% for n_{+2} and 11.6% for n_{+3} for the EU-27 average. At national level, the relative difference induced by "$DayTF\_Tb_{fit}\_f_{spec.}$" is lower than the respective uncertainty for 22 (20) countries in n_{+2} (n_{+3}).

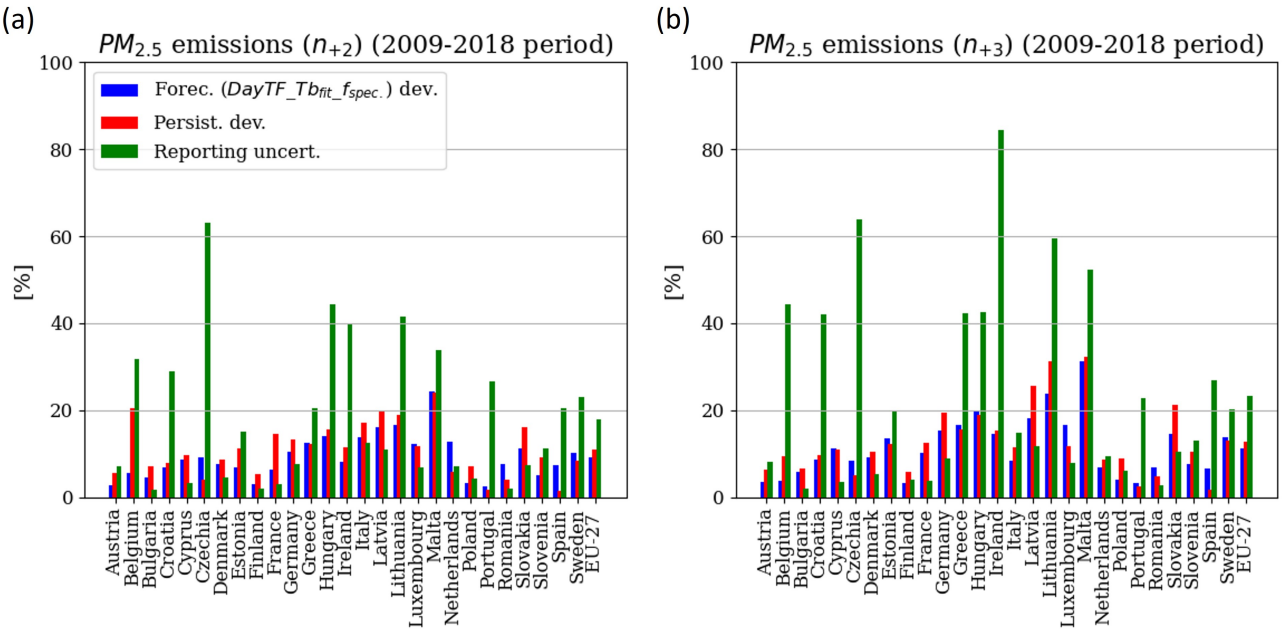

Figure 11. Average relative difference between reported emissions of $PM_{2.5}$ (CAMS-REG-AP-v5.1, $GNFR\_C$) and projected emissions using the "$DayTF\_Tb_{fit}\_f_{spec.}$" method (blue bars), and projected emissions by persistence (red bars) for n_{+2} (a) and n_{+3} (b) by country over the 2009-2018 period. HDD-based projections are also compared to the estimated reporting uncertainty (green bars) between 2009 and 2018 for emission recalculations up to 2 years (U_{+2}) and 3 years (U_{+3}) (from CEIP (2023)).

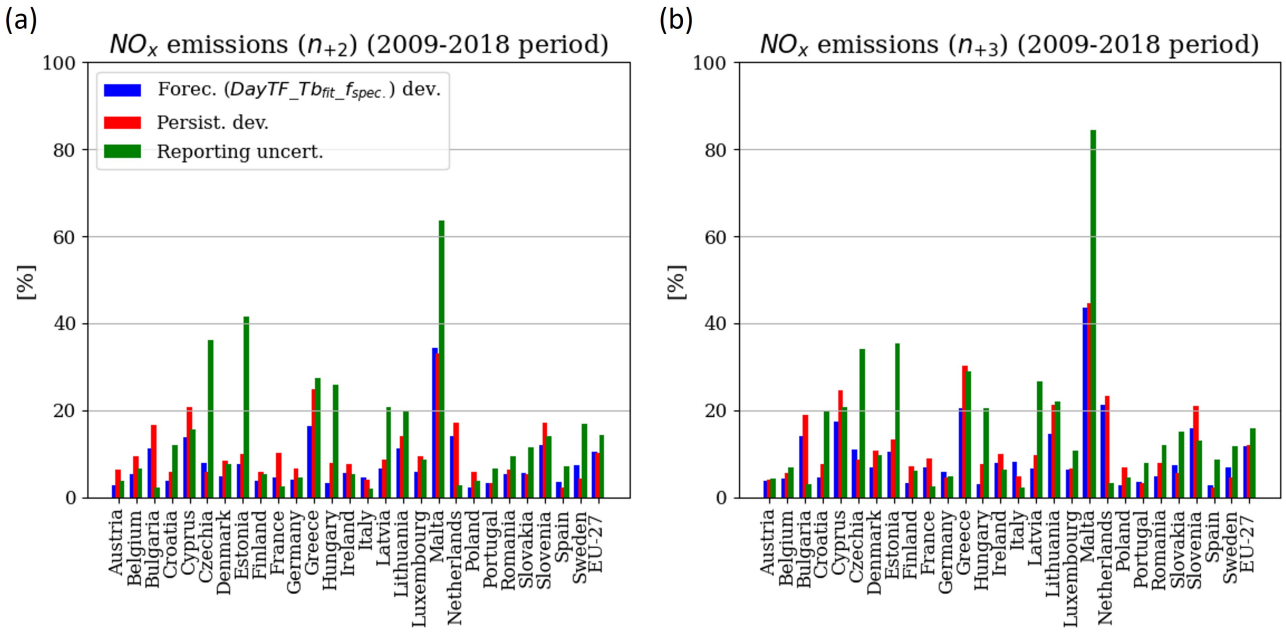

Figure 12. Same as Figure 11 for the NO_X emissions from $GNFR\_C$.

5 Discussion and conclusions

Based on statistical information on household energy consumption, this paper provides country- and species-dependent parameters ($Tb_{fit}(c)$ and $f_{spec.}(c)$) to derive HDD-based daily temporal factors to distribute PM and NO_x emissions from residential and commercial heating for EU-27 countries. $Tb_{fit}(c)$, the threshold ambient temperature at which building heating is activated and HDDs accumulated, is fitted with daily national domestic gas consumption data available over the period 2016-2021 for 8 countries (ENTSOG, 2023) and interpolated to the other EU countries using the IDW approach. $f_{spec.}(c)$, the non-temperature dependent fraction of residential emissions, is calculated for each EU-27 country based on the fraction of energy consumed by households for space heating (Eurostat, 2023); with the natural gas energy type for $f_{NOx}(c)$ and the solid fuels for $f_{PM}(c)$. It allows capturing well the specific behavioral characteristics of these countries (residential heating being used at different ambient temperature comfort values and with different energy mixes, and for different home insulation standards).

Several experiments were designed; the first with a monthly temporal factor ("$MonthTF$"), the second based on daily variations using HDDs with parameters from the literature ("$DayTF\_ref.$"), the third using HDDs with $Tb_{fit}(c)$ ("$DayTF\_Tb_{fit}$"), and the last using HDDs with $Tb_{fit}(c)$ and $f_{spec.}(c)$ ("$DayTF\_Tb_{fit}\_f_{spec.}$"). This work was designed to meet two objectives. We first aimed to assess the sensitivity of simulated $PM_{2.5}$, PM_{10} and NO_2 surface concentration to the different experiments (as described above) and to identify the best parameterization compared to in situ observations. Simulations car-

ried out with the CHIMERE model for 2018 have shown that the implementation of HDDs in the calculation of anthropogenic emissions from residential and commercial heating has an effect on $PM_{2.5}$, PM_{10} and NO_2 surface concentrations mainly from March to October. The performance scores were calculated at observation stations from the AQ e-Reporting database for the different experiments. "$DayTF\_Tb_{fit}\_f_{spec.}$" significantly improves the temporal correlation of the daily average concentration of $PM_{2.5}$ and PM_{10} during the winter season (JFM), with an average increase in the R coefficient of +0.11 compared to the simulation without HDD ("$MonthTF$") and by +0.05 compared to the simulation with HDDs using parameters from the literature ("$DayTF\_ref.$").

HDDs appear to have a fairly neutral effect on daily average concentration scores for the autumn season (OND). It should be noted that autumn 2018 has been relatively warmer than normal across Europe, with no significant cold spells. The scores obtained over this period do not allow us to highlight the benefits of HDDs. Furthermore, autumn, and especially its first few months, is a period of transition between the non-heating and heating seasons, which can vary considerably from one European country to another. Several studies have highlighted the difficulty of reproducing the beginning of the heating period (and its end in early spring), and this also applies to the HDD method (e.g. Grythe et al., 2019; Ciais et al., 2022).

Analyses carried out on the February and March 2018 pollution episodes related to cold spells showed that HDDs were necessary for a more correct simulation of $PM_{2.5}$ and PM_{10} concentration peaks. The "$DayTF\_Tb_{fit}\_f_{spec.}$" experiment increases the number of good detection of threshold exceedance by +11% and +57% for $PM_{2.5}$ and PM_{10} respectively compared to "$MonthTF$" on European stations, while decreasing the number of missed alarms (-7% and -9%). The increase in number of false alarms remains limited (+3% and +1%). This shows that "$DayTF\_Tb_{fit}\_f_{spec.}$" induces a better temporal distribution of concentration levels both below and above the exceedance thresholds, with the best probability of detection among the other parameterizations tested.

Concerning NO_2 species, the effect of HDDs on the simulated concentrations is low. As the contribution of $GNFR\_C$ emissions to total anthropogenic emissions remains limited (9.1% over the EU-27 (CEIP, 2023)), residential heating does not appear to be the main driver of NO_2 background concentrations, even during cold periods. Nevertheless, a significant decrease of the RMSE for the simulated concentrations in Italy, France, Hungary, Romania, Estonia and Latvia in winter was found, as well as a clear increase in good detection of threshold exceedance at European level (+12% with "$DayTF\_Tb_{fit}\_f_{spec.}$").

Finally, it should be stressed that the country- and species-specific parameters proposed in this work for the HDD method are based on national statistics which may be characterised by a degree of uncertainty. For instance, Eurostata data may underestimate wood consumption during summer recreational activities (e.g. barbecues). Compiling data on wood consumption can be challenging, as it cannot be monitored in the same way as natural gas. Furthermore, not everyone within a country lives in equally well-insulated houses, or has the same socio-economic level that allows them to turn on the heating appliances when they want to.

The second aim of this article was to use HDDs as a method to model national emission totals from $GNFR\_C$ and compare them with persistence and uncertainty (estimated from the magnitude of emission recalculations in subsequent years). The "$DayTF\_Tb_{fit}\_f_{spec.}$" parameterization was chosen. The analyses presented in this document have shown that this method performs better in regards to emission uncertainties and therefore can be used to estimate the effect of multi-year variability

in weather conditions to be taken into account, for both NO_x and PM emissions. The deviation obtained by the HDD projections is lower than the respective reporting uncertainty at European level for both n_{+2} and n_{+3}. The difference induced by HDD projections (compared with reported emissions) is even lower than with the persistence method (for more than half of the EU-27 countries).

The uncertainty estimated here is linked to changes in officially reported data at different years. The quality and accuracy of reported emissions data vary considerably from one country to another (EMEP, 2022). There are other ways of quantifying uncertainty in emissions inventories, for instance by including additional factors linked to the calculated emission factors or activity indicators. Kuenen et al. (2022) estimated that the uncertainty of $GNFR\_C$ in Europe is within the range 50%-200% for NO_x and 100%-300% for PM. The spatialization of emissions can be also associated with considerable uncertainties. As investigated in López-Aparicio et al. (2017) and Navarro-Barboza et al. (2024), wood combustion emissions may be over-allocated in urban areas compared to local inventories.

In conclusion, the HDD method shows positive results for the seasonal distribution and multi-year projection of emissions from commercial and residential heating. Based on the results of this paper, the use of HDDs and the spatialization of the parameters following the "$DayTF\_Tb_{fit}\_f_{spec.}$" experiment are recommended for the calculations of the emissions used as inputs in CTMs to improve the simulation of winter pollution episodes, including concentration peaks, but also for emission projections up to n_{+3}. In addition, different emission scenarios can be designed based on the use of HDDs. The HDD approach could be used to assess and isolate the meteorological effects of other major societal events that may have an impact on air quality, such as an energy crisis or a lockdown due to a pandemic situation (e.g. Guevara et al., 2021a). Finally, future work could focus on the dynamical modelling of emissions from other sectors that also appear to incorporate a weather-dependent component in order to improve the modelling of pollutant concentrations, such as for the combustion of diesel engine or the spreading of salt/sand on roads.

Data availability. Data on the natural gas consumption by country managed by ENSTOG are freely available at https://transparency.entsog.eu/ (last access: 20 July 2023; ENTSOG (2023)). The Eurostat data-set on disaggregated final energy consumption in households can be found at https://ec.europa.eu/eurostat/databrowser/view/NRG_D_HHQ__custom_2920041/bookmark/table?lang=en&bookmarkId=36e7b119-c46a-47b3-9c3d-aac3d44470d4 (last access: 20 July 2023; Eurostat (2023)). The CAMS-REG-AP-v5.1 inventory is available in the Emission of atmospheric Compounds and Compilation of Ancillary Data catalog at https://eccad.sedoo.fr (last access: 20 July 2023; Kuenen et al. (2022)). Reported multi-annual emissions by sector and country can be found at https://www.ceip.at/webdab-emission-database (last access: 20 July 2023; CEIP (2023)).

Author contributions. AG, FC, MG and AC developed the method. AG and FC conceptualized the paper. AG carried out the analysis and designed the figures. AG and FC wrote the original draft. AG, FC, MG and AC reviewed and edited the paper.

Competing interests. All the authors declare that there are no competing interests.

Acknowledgements. This research has been supported by the "Real-LIFE Emissions" project (n°LIFE 20 PRE-FI-000006) funded by the European Union, and by CAMS (implemented by ECMWF on behalf of the European Commission) under the service contracts $CAMS2\_61$ and $CAMS2\_40$. This work was granted access to the HPC resources of the "Centre de Calcul Recherche et Technologie" of "Commissariat à l'Energie Atomique". Finally, we acknowledge ENTSOG for the data on natural gas use, Eurostat for the data-set on disaggregated final energy consumption in households, ECCAD for the CAMS-REG-AP-v5.1 inventory and finally, CEIP for the multi-year data on emissions.

References

Alfaro, S. C. and Gomes, L.: Modeling mineral aerosol production by wind erosion: Emission intensities and aerosol size distributions in source areas, Journal of Geophysical Research: Atmospheres, 106, 18 075–18 084, https://doi.org/10.1029/2000JD900339, 2001.

Baykara, M., Im, U., and Unal, A.: Evaluation of impact of residential heating on air quality of megacity Istanbul by CMAQ, Science of The Total Environment, 651, 1688–1697, https://doi.org/10.1016/j.scitotenv.2018.10.091, 2019.

Bessagnet, B., Pirovano, G., Mircea, M., Cuvelier, C., Aulinger, A., Calori, G., Ciarelli, G., Manders, A., Stern, R., Tsyro, S., García Vivanco, M., Thunis, P., Pay, M.-T., Colette, A., Couvidat, F., Meleux, F., Rouïl, L., Ung, A., Aksoyoglu, S., Baldasano, J. M., Bieser, J., Briganti, G., Cappelletti, A., D'Isidoro, M., Finardi, S., Kranenburg, R., Silibello, C., Carnevale, C., Aas, W., Dupont, J.-C., Fagerli, H., Gonzalez, L., Menut, L., Prévôt, A. S. H., Roberts, P., and White, L.: Presentation of the EURODELTA III intercomparison exercise – evaluation ofthe chemistry transport models' performance on criteria pollutants and joint analysis with meteorology, Atmospheric Chemistry and Physics, 16, 12 667–12 701, https://doi.org/10.5194/acp-16-12667-2016, 2016.

Bian, H. and Prather, M. J.: Fast-J2: Accurate Simulation of Stratospheric Photolysis in Global Chemical Models, Journal of Atmospheric Chemistry, p. 16, https://link.springer.com/content/pdf/10.1023/A:1014980619462.pdf, 2002.

CEIP: Officially reported emission data, https://www.ceip.at/webdab-emission-database/emissions-as-used-in-emep-models, 2023.

Chen, G., Canonaco, F., Tobler, A., Aas, W., Alastuey, A., Allan, J., Atabakhsh, S., Aurela, M., Baltensperger, U., Bougiatioti, A., De Brito, J. F., Ceburnis, D., Chazeau, B., Chebaicheb, H., Daellenbach, K. R., Ehn, M., El Haddad, I., Eleftheriadis, K., Favez, O., Flentje, H., Font, A., Fossum, K., Freney, E., Gini, M., Green, D. C., Heikkinen, L., Herrmann, H., Kalogridis, A.-C., Keernik, H., Lhotka, R., Lin, C., Lunder, C., Maasikmets, M., Manousakas, M. I., Marchand, N., Marin, C., Marmureanu, L., Mihalopoulos, N., Močnik, G., Nęcki, J., O'Dowd, C., Ovadnevaite, J., Peter, T., Petit, J.-E., Pikridas, M., Matthew Platt, S., Pokorná, P., Poulain, L., Priestman, M., Riffault, V., Rinaldi, M., Różański, K., Schwarz, J., Sciare, J., Simon, L., Skiba, A., Slowik, J. G., Sosedova, Y., Stavroulas, I., Styszko, K., Teinemaa, E., Timonen, H., Tremper, A., Vasilescu, J., Via, M., Vodička, P., Wiedensohler, A., Zografou, O., Cruz Minguillón, M., and Prévôt, A. S.: European aerosol phenomenology - 8: Harmonised source apportionment of organic aerosol using 22 Year-long ACSM/AMS datasets, Environment International, 166, 107 325, https://doi.org/10.1016/j.envint.2022.107325, 2022.

Chen, J., Li, C., Ristovski, Z., Milic, A., Gu, Y., Islam, M. S., Wang, S., Hao, J., Zhang, H., He, C., Guo, H., Fu, H., Miljevic, B., Morawska, L., Thai, P., Lam, Y. F., Pereira, G., Ding, A., Huang, X., and Dumka, U. C.: A review of biomass burning: Emissions and impacts on air quality, health and climate in China, Science of The Total Environment, 579, 1000–1034, https://doi.org/10.1016/j.scitotenv.2016.11.025, 2017.

Ciais, P., Bréon, F.-M., Dellaert, S., Wang, Y., Tanaka, K., Gurriaran, L., Françoise, Y., Davis, S. J., Hong, C., Penuelas, J., Janssens, I., Obersteiner, M., Deng, Z., and Liu, Z.: Impact of Lockdowns and Winter Temperatures on Natural Gas Consumption in Europe, Earth's Future, 10, e2021EF002 250, https://doi.org/10.1029/2021EF002250, 2022.

Cincinelli, A., Guerranti, C., Martellini, T., and Scodellini, R.: Residential wood combustion and its impact on urban air quality in Europe, Current Opinion in Environmental Science & Health, 8, 10–14, https://doi.org/10.1016/j.coesh.2018.12.007, 2019.

Considine, T. J.: The impacts of weather variations on energy demand and carbon emissions, Resource and Energy Economics, 22, 295–314, https://doi.org/10.1016/S0928-7655(00)00027-0, 2000.

Couvidat, F., Debry, É., Sartelet, K., and Seigneur, C.: A hydrophilic/hydrophobic organic (H2O) aerosol model: Development, evaluation and sensitivity analysis, Journal of Geophysical Research: Atmospheres, 117, 2011JD017 214, https://doi.org/10.1029/2011JD017214, 2012.

Couvidat, F., Bessagnet, B., Garcia-Vivanco, M., Real, E., Menut, L., and Colette, A.: Development of an inorganic and organic aerosol model (CHIMERE 2017 β v1.0): seasonal and spatial evaluation over Europe, Geoscientific Model Development, 11, 165–194, https://doi.org/10.5194/gmd-11-165-2018, 2018.

Crippa, M., Solazzo, E., Huang, G., Guizzardi, D., Koffi, E., Muntean, M., Schieberle, C., Friedrich, R., and Janssens-Maenhout, G.: High resolution temporal profiles in the Emissions Database for Global Atmospheric Research, Scientific Data, 7, 121, https://doi.org/10.1038/s41597-020-0462-2, 2020.

Crippa, M., Guizzardi, D., Pisoni, E., Solazzo, E., Guion, A., Muntean, M., Florczyk, A., Schiavina, M., Melchiorri, M., and Hutfilter, A. F.: Global anthropogenic emissions in urban areas: patterns, trends, and challenges, Environmental Research Letters, 16, 074 033, https://doi.org/10.1088/1748-9326/ac00e2, 2021.

Denier van der Gon, H., Hendriks, C., Kuenen, J., Segers, A., and Visschedijk, A.: Description of current temporal emission patterns and sensitivity of predicted AQ for temporal emission patterns, Tech. Rep. EU FP7 MACC deliverable report D_D-EMIS_1.3, TNO, https://atmosphere.copernicus.eu/sites/default/files/2019-07/MACC_TNO_del_1_3_v2.pdf, 2011.

Denier Van Der Gon, H. A. C., Bergström, R., Fountoukis, C., Johansson, C., Pandis, S. N., Simpson, D., and Visschedijk, A. J. H.: Particulate emissions from residential wood combustion in Europe – revised estimates and an evaluation, Atmospheric Chemistry and Physics, 15, 6503–6519, https://doi.org/10.5194/acp-15-6503-2015, 2015.

Derognat, C.: Effect of biogenic volatile organic compound emissions on tropospheric chemistry during the Atmospheric Pollution Over the Paris Area (ESQUIF) campaign in the Ile-de-France region, Journal of Geophysical Research, 108, 8560, https://doi.org/10.1029/2001JD001421, 2003.

Ebel, A., Friedrich, R., and Rodhe, H.: GENEMIS: Assessment, Improvement, and Temporal and Spatial Disaggregation of European Emission Data, in: Tropospheric Modelling and Emission Estimation, edited by Ebel, A., Friedrich, R., and Rodhe, H., pp. 181–214, Springer Berlin Heidelberg, Berlin, Heidelberg, https://doi.org/10.1007/978-3-662-03470-5_6, 1997.

EEA: The European air quality observation network AQ e-Reporting, https://www.eea.europa.eu/en/datahub/datahubitem-view/3b390c9c-f321-490a-b25a-ae93b2ed80c1, 2023.

EMEP: EMEP Status Report 2022 : Transboundary particulate matter, photo-oxidants, acidifying and eutrophying components, https://emep.int/publ/reports/2022/EMEP_Status_Report_1_2022.pdf, 2022.

ENTSOG: European Gas Flow dashboard by ENTSOG, https://transparency.entsog.eu/, 2023.

Eurostat: Disaggregated final energy consumption in households, https://ec.europa.eu/eurostat/databrowser/view/NRG_D_HHQ__custom_4178907/default/table?lang=en, 2023.

Flemming, J., Huijnen, V., Arteta, J., Bechtold, P., Beljaars, A., Blechschmidt, A.-M., Diamantakis, M., Engelen, R. J., Gaudel, A., Inness, A., Jones, L., Josse, B., Katragkou, E., Marecal, V., Peuch, V.-H., Richter, A., Schultz, M. G., Stein, O., and Tsikerdekis, A.: Tropospheric chemistry in the Integrated Forecasting System of ECMWF, Geoscientific Model Development, 8, 975–1003, https://doi.org/10.5194/gmd-8-975-2015, 2015.

Gao, F. and Han, L.: Implementing the Nelder-Mead simplex algorithm with adaptive parameters, Computational Optimization and Applications, 51, 259–277, https://doi.org/10.1007/s10589-010-9329-3, 2012.

Grange, S. K., Farren, N. J., Vaughan, A. R., Rose, R. A., and Carslaw, D. C.: Strong Temperature Dependence for Light-Duty Diesel Vehicle NOx Emissions, Environmental Science & Technology, 53, 6587–6596, https://doi.org/10.1021/acs.est.9b01024, 2019.

Grythe, H., Lopez-Aparicio, S., Vogt, M., Vo Thanh, D., Hak, C., Halse, A. K., Hamer, P., and Sousa Santos, G.: The MetVed model: development and evaluation of emissions from residential wood combustion at high spatio-temporal resolution in Norway, Atmospheric Chemistry and Physics, 19, 10 217–10 237, https://doi.org/10.5194/acp-19-10217-2019, 2019.

Guenther, A., Karl, T., Harley, P., Wiedinmyer, C., Palmer, P. I., and Geron, C.: Estimates of global terrestrial isoprene emissions using MEGAN (Model of Emissions of Gases and Aerosols from Nature), Atmos. Chem. Phys., p. 31, 2006.

Guenther, A. B., Jiang, X., Heald, C. L., Sakulyanontvittaya, T., Duhl, T., Emmons, L. K., and Wang, X.: The Model of Emissions of Gases and Aerosols from Nature version 2.1 (MEGAN2.1): an extended and updated framework for modeling biogenic emissions, Geoscientific Model Development, 5, 1471–1492, https://doi.org/10.5194/gmd-5-1471-2012, 2012.

Guevara, M., Jorba, O., Soret, A., Petetin, H., Bowdalo, D., Serradell, K., Tena, C., Denier Van Der Gon, H., Kuenen, J., Peuch, V.-H., and Pérez García-Pando, C.: Time-resolved emission reductions for atmospheric chemistry modelling in Europe during the COVID-19 lockdowns, Atmospheric Chemistry and Physics, 21, 773–797, https://doi.org/10.5194/acp-21-773-2021, 2021a.

Guevara, M., Jorba, O., Tena, C., Denier van der Gon, H., Kuenen, J., Elguindi, N., Darras, S., Granier, C., and Pérez García-Pando, C.: Copernicus Atmosphere Monitoring Service TEMPOral profiles (CAMS-TEMPO): global and European emission temporal profile maps for atmospheric chemistry modelling, Earth System Science Data, 13, 367–404, https://doi.org/10.5194/essd-13-367-2021, 2021b.

Guevara, M., Petetin, H., Jorba, O., Denier Van Der Gon, H., Kuenen, J., Super, I., Jalkanen, J.-P., Majamäki, E., Johansson, L., Peuch, V.-H., and Pérez García-Pando, C.: European primary emissions of criteria pollutants and greenhouse gases in 2020 modulated by the COVID-19 pandemic disruptions, Earth System Science Data, 14, 2521–2552, https://doi.org/10.5194/essd-14-2521-2022, 2022.

Guion, A., Turquety, S., Cholakian, A., Polcher, J., Ehret, A., and Lathière, J.: Biogenic isoprene emissions, dry deposition velocity, and surface ozone concentration during summer droughts, heatwaves, and normal conditions in southwestern Europe, Atmospheric Chemistry and Physics, 23, 1043–1071, https://doi.org/10.5194/acp-23-1043-2023, 2023.

Juda-Rezler, K., Reizer, M., and Oudinet, J.-P.: Determination and analysis of PM10 source apportionment during episodes of air pollution in Central Eastern European urban areas: The case of wintertime 2006, Atmospheric Environment, 45, 6557–6566, https://doi.org/10.1016/j.atmosenv.2011.08.020, 2011.

Khomenko, S., Pisoni, E., Thunis, P., Bessagnet, B., Cirach, M., Iungman, T., Barboza, E. P., Khreis, H., Mueller, N., Tonne, C., De Hoogh, K., Hoek, G., Chowdhury, S., Lelieveld, J., and Nieuwenhuijsen, M.: Spatial and sector-specific contributions of emissions to ambient air pollution and mortality in European cities: a health impact assessment, The Lancet Public Health, 8, e546–e558, https://doi.org/10.1016/S2468-2667(23)00106-8, 2023.

Kuenen, J., Dellaert, S., Visschedijk, A., Jalkanen, J.-P., Super, I., and Denier van der Gon, H.: CAMS-REG-v4: a state-of-the-art high-resolution European emission inventory for air quality modelling, Earth System Science Data, 14, 491–515, https://doi.org/10.5194/essd-14-491-2022, 2022.

Liss, P. S. and Mervilat, L.: Air-Sea Gas Exchange Rates: Introduction and Synthesis, Springer Netherlands, Dordrecht, pp. 113–127, https://doi.org/10.1007/978-94-009-4738-2_5, 1986.

López-Aparicio, S., Guevara, M., Thunis, P., Cuvelier, K., and Tarrasón, L.: Assessment of discrepancies between bottom-up and regional emission inventories in Norwegian urban areas, Atmospheric Environment, 154, 285–296, https://doi.org/10.1016/j.atmosenv.2017.02.004, 2017.

López-Aparicio, S., Grythe, H., and Markelj, M.: High-Resolution Emissions from Wood Burning in Norway—The Effect of Cabin Emissions, Energies, 15, 9332, https://doi.org/10.3390/en15249332, 2022.

Mårtensson, E. M., Nilsson, E. D., De Leeuw, G., Cohen, L. H., and Hansson, H.-C.: Laboratory simulations and parameterization of the primary marine aerosol production: THE PRIMARY MARINE AEROSOL SOURCE, Journal of Geophysical Research: Atmospheres, 108, n/a–n/a, https://doi.org/10.1029/2002JD002263, 2003.

Marticorena, B. and Bergametti, G.: Modeling the atmospheric dust cycle: 1. Design of a soil-derived dust emission scheme, Journal of Geophysical Research, 100, 16 415, https://doi.org/10.1029/95JD00690, 1995.

Mbengue, S., Serfozo, N., Schwarz, J., Ziková, N., Šmejkalová, A. H., and Holoubek, I.: Characterization of Equivalent Black Carbon at a regional background site in Central Europe: Variability and source apportionment*, Environmental Pollution, 260, 113 771, https://doi.org/10.1016/j.envpol.2019.113771, 2020.

Menut, L., Bessagnet, B., Briant, R., Cholakian, A., Couvidat, F., Mailler, S., Pennel, R., Siour, G., Tuccella, P., Turquety, S., and Valari, M.: The CHIMERE v2020r1 online chemistry-transport model, Geoscientific Model Development, 14, 6781–6811, https://doi.org/10.5194/gmd-14-6781-2021, 2021.

MF: Bilan climatique de l'année 2018, https://meteofrance.fr/actualite/publications/les-publications-de-meteo-france/bilans-climatiques-annuels-de-2014-2018, 2019.

Mues, A., Kuenen, J., Hendriks, C., Manders, A., Segers, A., Scholz, Y., Hueglin, C., Builtjes, P., and Schaap, M.: Sensitivity of air pollution simulations with LOTOS-EUROS to the temporal distribution of anthropogenic emissions, Atmospheric Chemistry and Physics, 14, 939–955, https://doi.org/10.5194/acp-14-939-2014, 2014.

Muñoz-Sabater, J., Dutra, E., Agustí-Panareda, A., Albergel, C., Arduini, G., Balsamo, G., Boussetta, S., Choulga, M., Harrigan, S., Hersbach, H., Martens, B., Miralles, D. G., Piles, M., Rodríguez-Fernández, N. J., Zsoter, E., Buontempo, C., and Thépaut, J.-N.: ERA5-Land: a state-of-the-art global reanalysis dataset for land applications, Earth System Science Data, 13, 4349–4383, https://doi.org/10.5194/essd-13-4349-2021, 2021.

Navarro-Barboza, H., Pandolfi, M., Guevara, M., Enciso, S., Tena, C., Via, M., Yus-Díez, J., Reche, C., Pérez, N., Alastuey, A., Querol, X., and Jorba, O.: Uncertainties in source allocation of carbonaceous aerosols in a Mediterranean region, Environment International, 183, 108 252, https://doi.org/10.1016/j.envint.2023.108252, 2024.

Peuch, V.-H., Engelen, R., Rixen, M., Dee, D., Flemming, J., Suttie, M., Ades, M., Agustí-Panareda, A., Ananasso, C., Andersson, E., Armstrong, D., Barré, J., Bousserez, N., Dominguez, J. J., Garrigues, S., Inness, A., Jones, L., Kipling, Z., Letertre-Danczak, J., Parrington, M., Razinger, M., Ribas, R., Vermoote, S., Yang, X., Simmons, A., Garcés De Marcilla, J., and Thépaut, J.-N.: The Copernicus Atmosphere Monitoring Service: From Research to Operations, Bulletin of the American Meteorological Society, 103, E2650–E2668, https://doi.org/10.1175/BAMS-D-21-0314.1, 2022.

Pokorná, P., Schwarz, J., Krejci, R., Swietlicki, E., Havránek, V., and Ždímal, V.: Comparison of PM2.5 chemical composition and sources at a rural background site in Central Europe between 1993/1994/1995 and 2009/2010: Effect of legislative regulations and economic transformation on the air quality, Environmental Pollution, 241, 841–851, https://doi.org/10.1016/j.envpol.2018.06.015, 2018.

Quayle, R. G. and Diaz, H. F.: Heating Degree Day Data Applied to Residential Heating Energy Consumption, Journal of Applied Meteorology, 19, 241–246, https://doi.org/https://doi.org/10.1175/1520-0450(1980)019<0241:HDDDAT>2.0.CO;2, 1980.

Rémy, S., Kipling, Z., Huijnen, V., Flemming, J., Nabat, P., Michou, M., Ades, M., Engelen, R., and Peuch, V.-H.: Description and evaluation of the tropospheric aerosol scheme in the Integrated Forecasting System (IFS-AER, cycle 47R1) of ECMWF, Geoscientific Model Development, 15, 4881–4912, https://doi.org/10.5194/gmd-15-4881-2022, 2022.

Roebber, P. J.: Visualizing Multiple Measures of Forecast Quality, Weather and Forecasting, 24, 601–608, https://doi.org/10.1175/2008WAF2222159.1, 2009.

Rudziński, K. J., Sarang, K., Nestorowicz, K., Asztemborska, M., Żyfka-Zagrodzińska, E., Skotak, K., and Szmigielski, R.: Winter sources of PM2.5 pollution in Podkowa Leśna, a Central-European garden town (Mazovia, Poland), Environmental Science and Pollution Research, 29, 84 504–84 520, https://doi.org/10.1007/s11356-022-21673-1, 2022.

Schindlbacher, S., Matthews, B., and Ullrich, B.: Uncertainties and recalculations of emission inventories submitted under CLRTAP, https://www.ceip.at/fileadmin/inhalte/ceip/00_pdf_other/2021/uncertainties_and_recalculations_of_emission_inventories_submitted_under_clrtap.pdf, 2021.

Sicard, P., Agathokleous, E., De Marco, A., Paoletti, E., and Calatayud, V.: Urban population exposure to air pollution in Europe over the last decades, Environmental Sciences Europe, 33, 28, https://doi.org/10.1186/s12302-020-00450-2, 2021.

Spinoni, J., Vogt, J. V., and Barbosa, P.: European degree-day climatologies and trends for the period 1951–2011, International Journal of Climatology, 35, 25–36, https://doi.org/https://doi.org/10.1002/joc.3959, 2015.

StatFin: Heating degree day (HDD), https://stat.fi/meta/kas/lammitystarvelu_en.html), 2023.

Stirnberg, R., Cermak, J., Kotthaus, S., Haeffelin, M., Andersen, H., Fuchs, J., Kim, M., Petit, J.-E., and Favez, O.: Meteorology-driven variability of air pollution (PM1) revealed with explainable machine learning, Atmospheric Chemistry and Physics, 21, 3919–3948, https://doi.org/10.5194/acp-21-3919-2021, 2021.

Stohl, A., Klimont, Z., Eckhardt, S., Kupiainen, K., Shevchenko, V. P., Kopeikin, V. M., and Novigatsky, A. N.: Black carbon in the Arctic: the underestimated role of gas flaring and residential combustion emissions, Atmospheric Chemistry and Physics, 13, 8833–8855, https://doi.org/10.5194/acp-13-8833-2013, 2013.

Tammekivi, T., Kaasik, M., Hamer, P., Santos, G. S., and Šteinberga, I.: Air pollution situation in small towns, including winter resorts: a comparative study of three cases in Northern Europe, Air Quality, Atmosphere & Health, 16, 945–961, https://doi.org/10.1007/s11869-023-01315-2, 2023.

Thom, H. C. S.: The rational relationship between heating degree days and temperature, Monthly Weather Review, 82, 1–6, https://doi.org/10.1175/1520-0493(1954)082<0001:TRRBHD>2.0.CO;2, 1954.

Van Leer, B.: Towards the ultimate conservative difference scheme. IV. A new approach to numerical convection, Journal of Computational Physics, 23, 276–299, https://doi.org/10.1016/0021-9991(77)90095-X, 1977.

Velikou, K., Lazoglou, G., Tolika, K., and Anagnostopoulou, C.: Reliability of the ERA5 in Replicating Mean and Extreme Temperatures across Europe, Water, 14, 543, https://doi.org/10.3390/w14040543, 2022.

Wærsted, E. G., Sundvor, I., Denby, B. R., and Mu, Q.: Quantification of temperature dependence of NO emissions from road traffic in Norway using air quality modelling and monitoring data, Atmospheric Environment: X, 13, 100 160, https://doi.org/10.1016/j.aeaoa.2022.100160, 2022.

Supplement

Table S1. Value of Tb_{fit}, f_{NOx} and f_{PM} for each EU-27 country. The missing data for the calculation of $f_{spec.}$ have been replaced by the European average available (numbers underlined).

Country	Tb_{fit}	f_{NOx}	f_{PM}
Austria	15.92	0.12	0.07
Belgium	15.92	0.16	0.01
Bulgaria	15.65	0.19	0.04
Croatia	15.82	0.27	0.05
Cyprus	15.35	<u>0.25</u>	0.02
Czechia	15.58	0.35	0.04
Denmark	15.03	0.28	0.06
Estonia	12.49	0.27	0.12
Finland	11.99	0.22	0.24
France	15.96	0.17	0.04
Germany	15.45	0.21	0.07
Greece	15.60	0.13	0.01
Hungary	17.69	0.17	0.01
Ireland	15.39	0.29	0.12
Italy	15.68	0.25	0.04
Latvia	11.49	0.48	0.15
Lithuania	13.52	0.34	0.07
Luxembourg	15.67	0.12	0.02
Malta	15.57	<u>0.25</u>	0.01
Netherlands	15.43	0.23	0.01
Poland	15.10	0.47	0.10
Portugal	15.52	0.94	0.29
Romania	16.63	0.41	0.23
Slovakia	16.14	0.20	0.25
Slovenia	15.85	0.27	0.22
Spain	15.62	0.54	0.11
Sweden	11.99	0.24	0.11

Table S2. Characteristics of domestic gas data-sets from the ENTSOG Transparency platform. For each country, the gas supplier used and the period covered by the data are detailed.

Country	Gas supplier	Time cover
Hungary	FGSZ	2019-2021
Roumania	Transgas	2018-2021
Italy	Snam Rete Gas	2016-2021
France	GRTgas	2016-2021
Belgium	Fluxys Belgium	2016-2021
Netherlands	GTS	2016-2021
Latvia	Conexus	2020-2021
Estonia	Elering Gaas	2016-2021

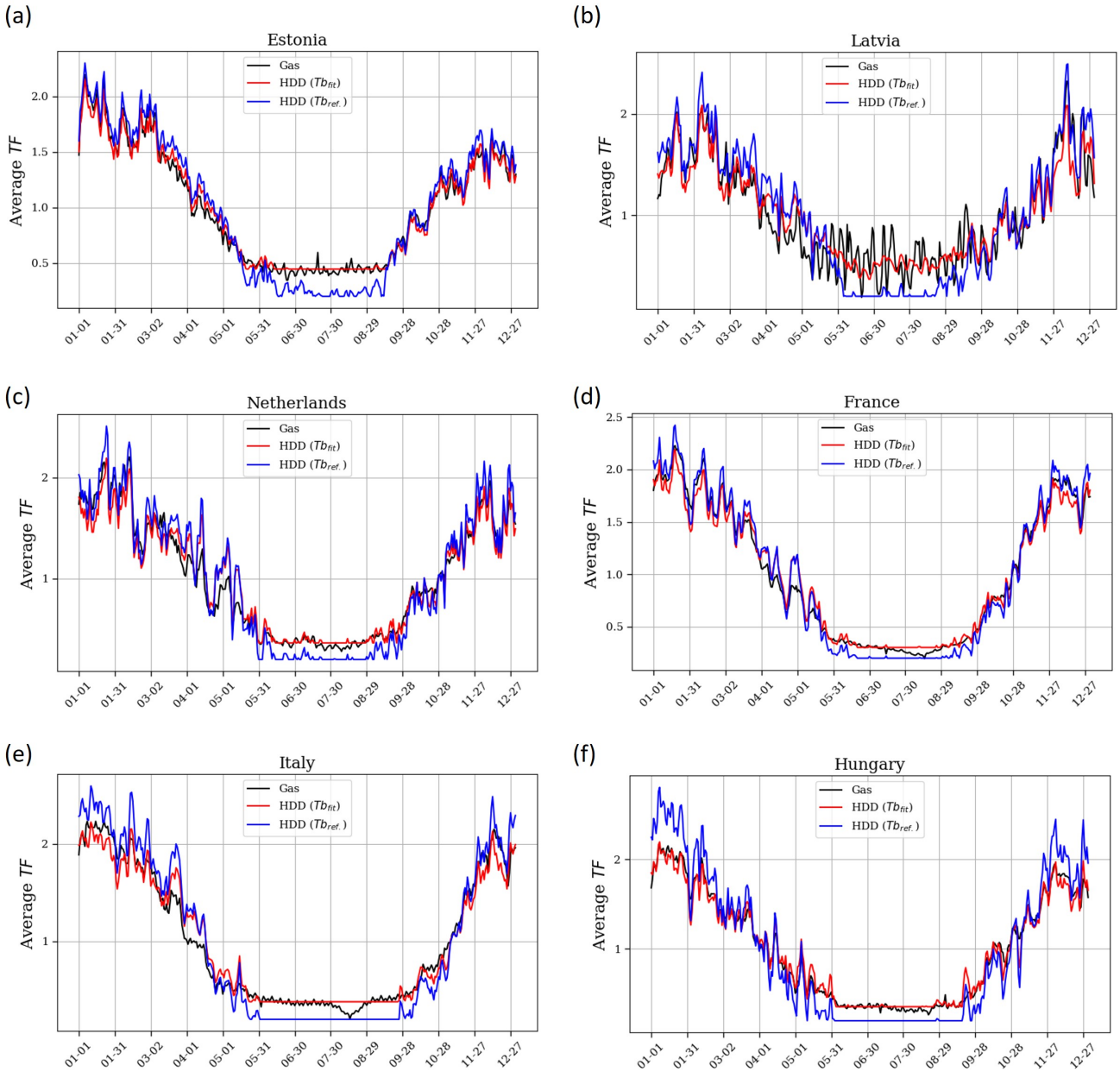

Figure S1. Daily evolution of $TF(c)$ (unitless) for gas consumption (in black), for HDDs with Tb_{fit} (in red) and for HDDs with $Tb_{ref.}$ (in blue). $TF(c)$ is averaged over the period 2018-2021 for Estonia (a), over 2020-2021 for Latvia (b), over 2016-2021 for Netherlands (c), over 2016-2021 for France (d), over 2016-2021 for Italy (e) and over 2019-2021 for Hungary (f).

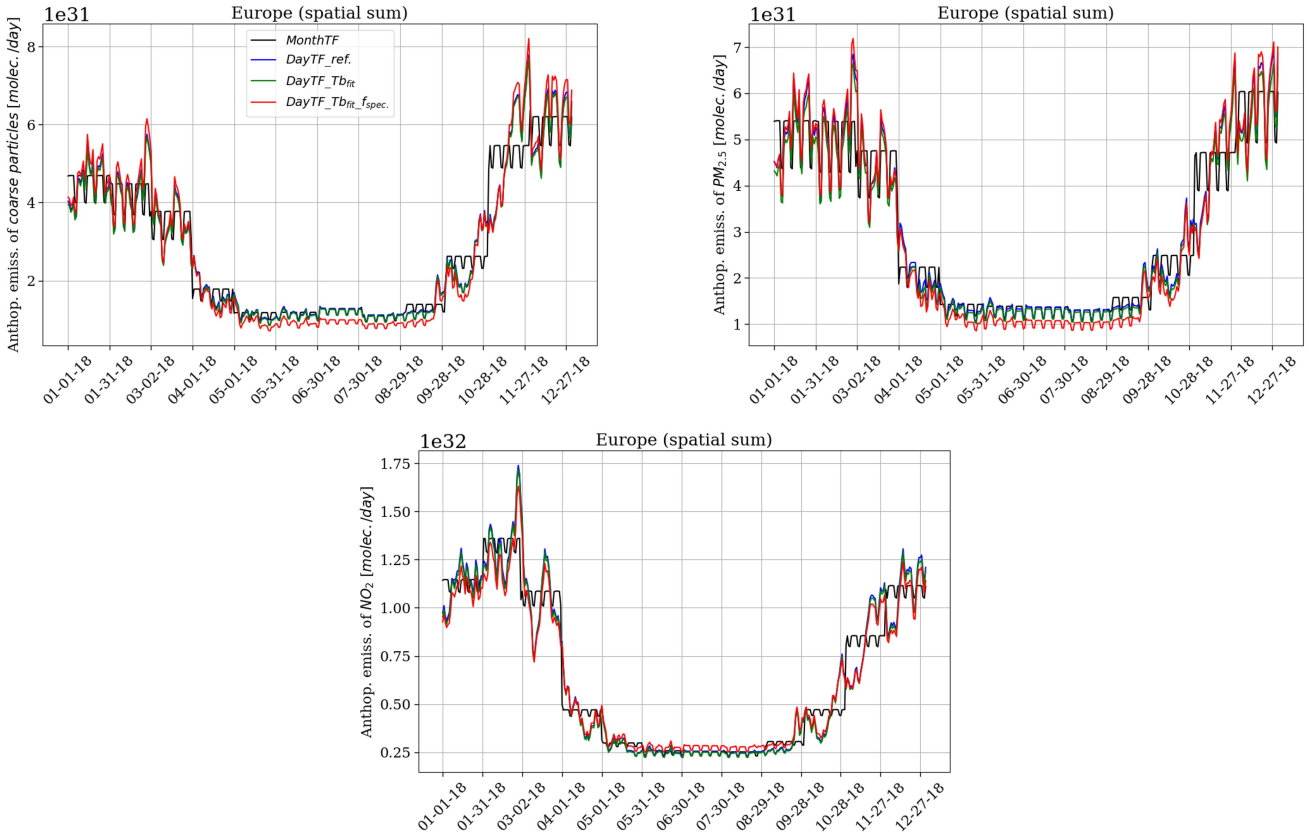

Figure S2. Time series at daily time step of total anthropogenic emissions of coarse particles (top left), fine particles (top right) and NO_2 (bottom) in Europe for the different experiments detailed in Table 2.

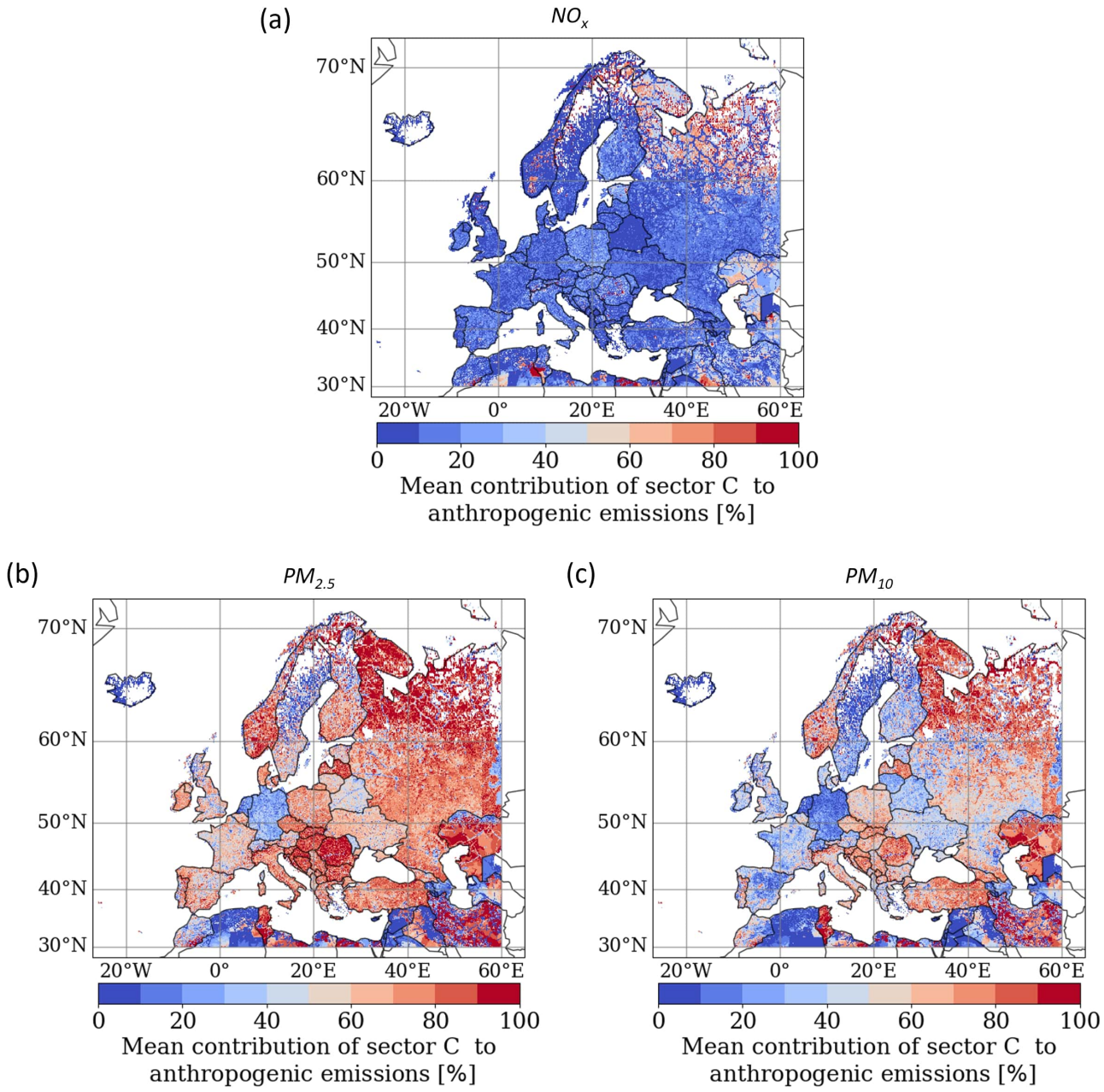

Figure S3. Spatial distribution of the average contribution of GNFR sector "Other stationary combustions" (C) to total anthropogenic emissions for NO_x (a), $PM_{2.5}$ (b) and PM_{10} (c) species over the period 2009-2018, based on the CAMS-REG-AP-v5.1 inventory (Kuenen et al., 2022).

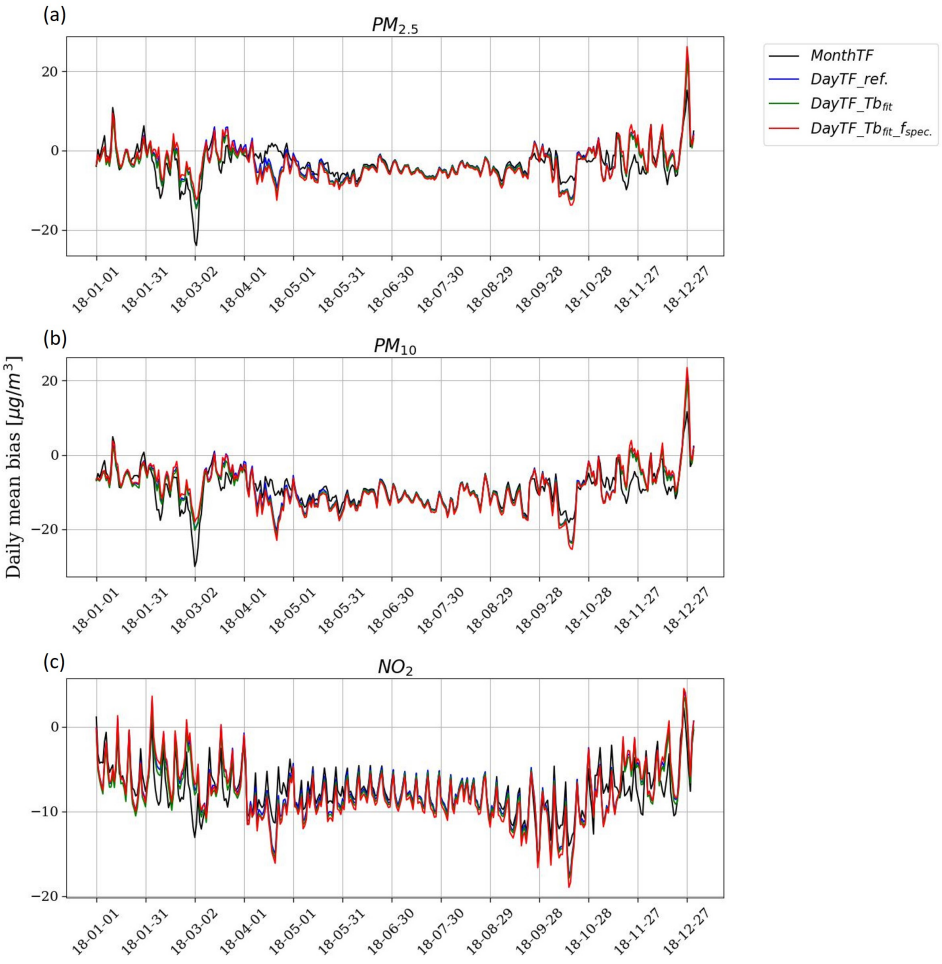

Figure S4. Evolution in 2018 of the daily average bias [$\mu g/m^3$] over the European domain between the CHIMERE simulations and the AQ e-Reporting observation stations of $PM_{2.5}$ (a), PM_{10} (b) and NO_2 (c).

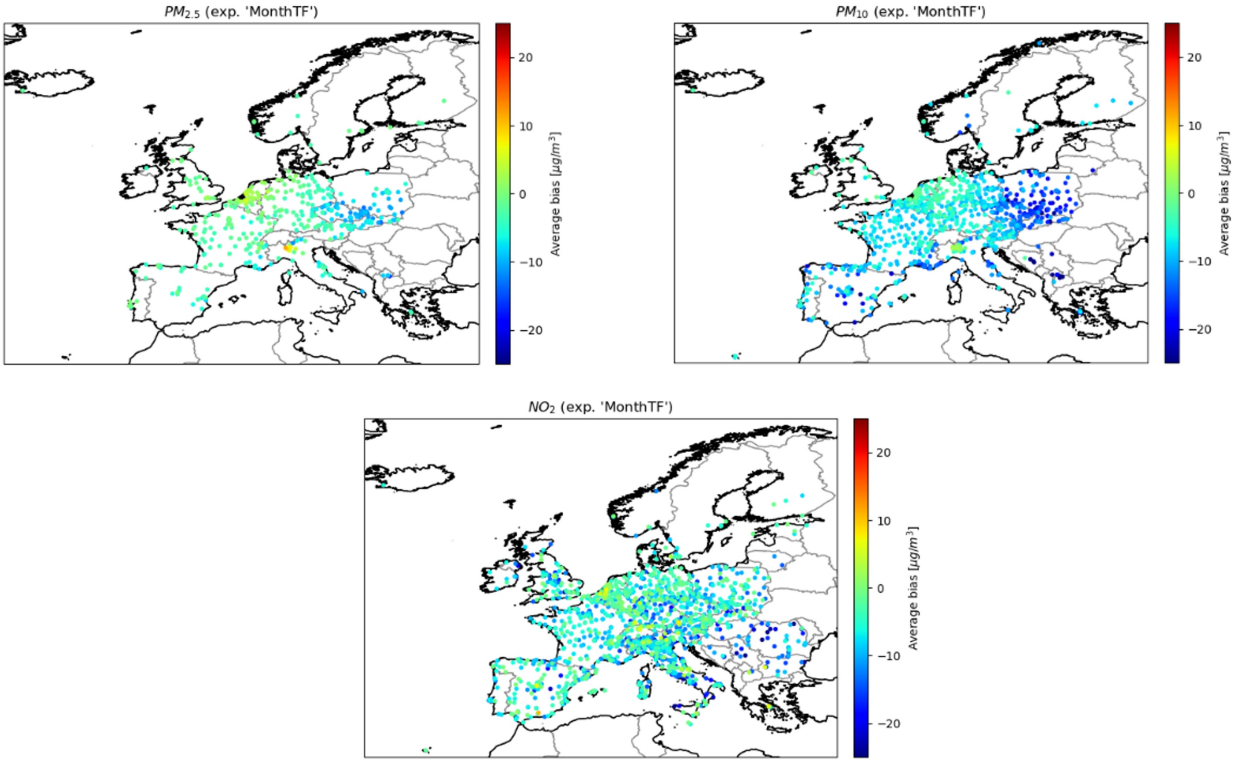

Figure S5. Average bias (mod. - obs.) [$\mu g/m^3$] for 2018 for PM_{10} (top left), $PM_{2.5}$ (top right) and NO_2 (bottom).

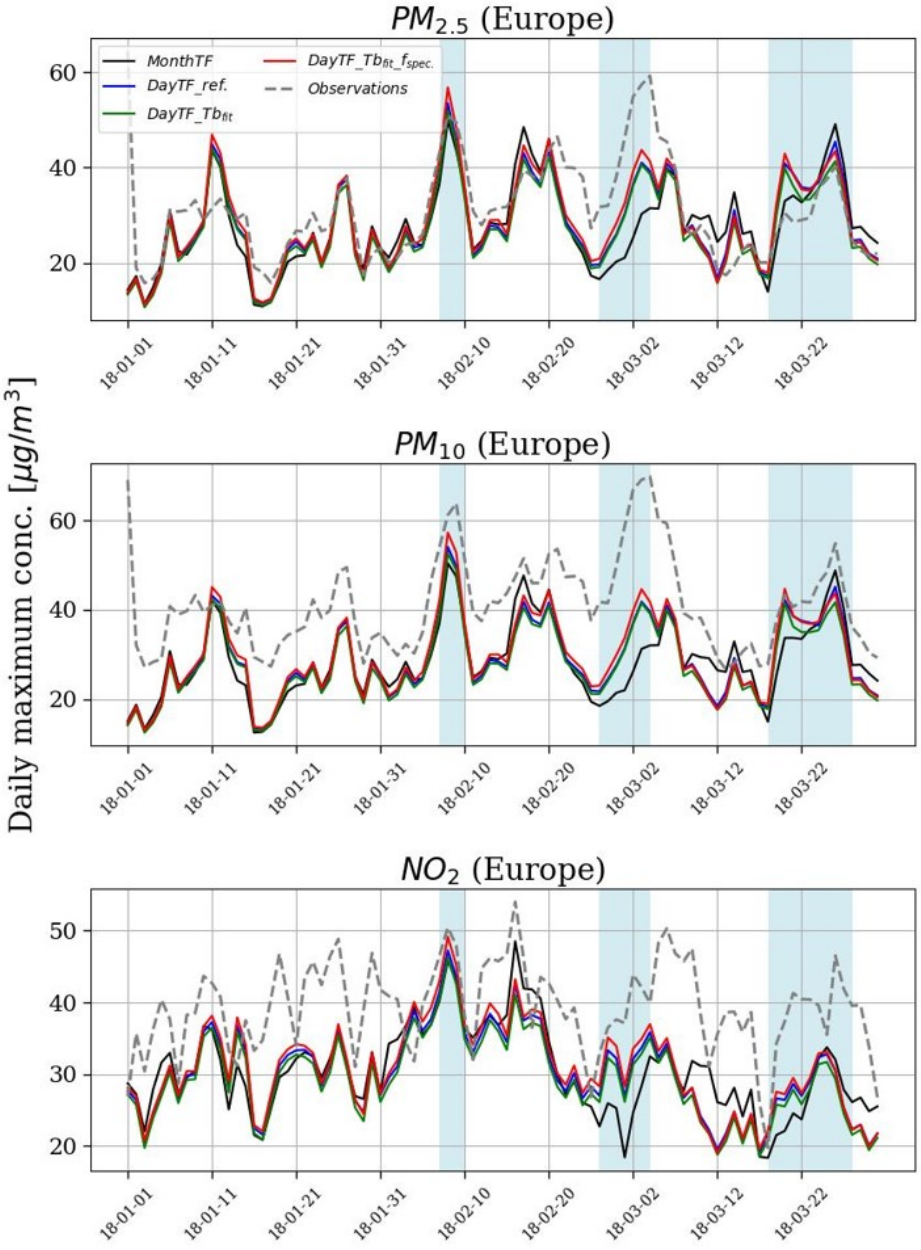

Figure S6. Daily maximum concentrations of $PM_{2.5}$ (top), PM_{10} (middle) and NO_2 (bottom) in Europe between January and March from observations (~~AQ eReporting~~AQ e-Reporting) and different CHIMERE experiments. The blue areas indicate periods of intense cold as signaled by the Climate Copernicus service.

.

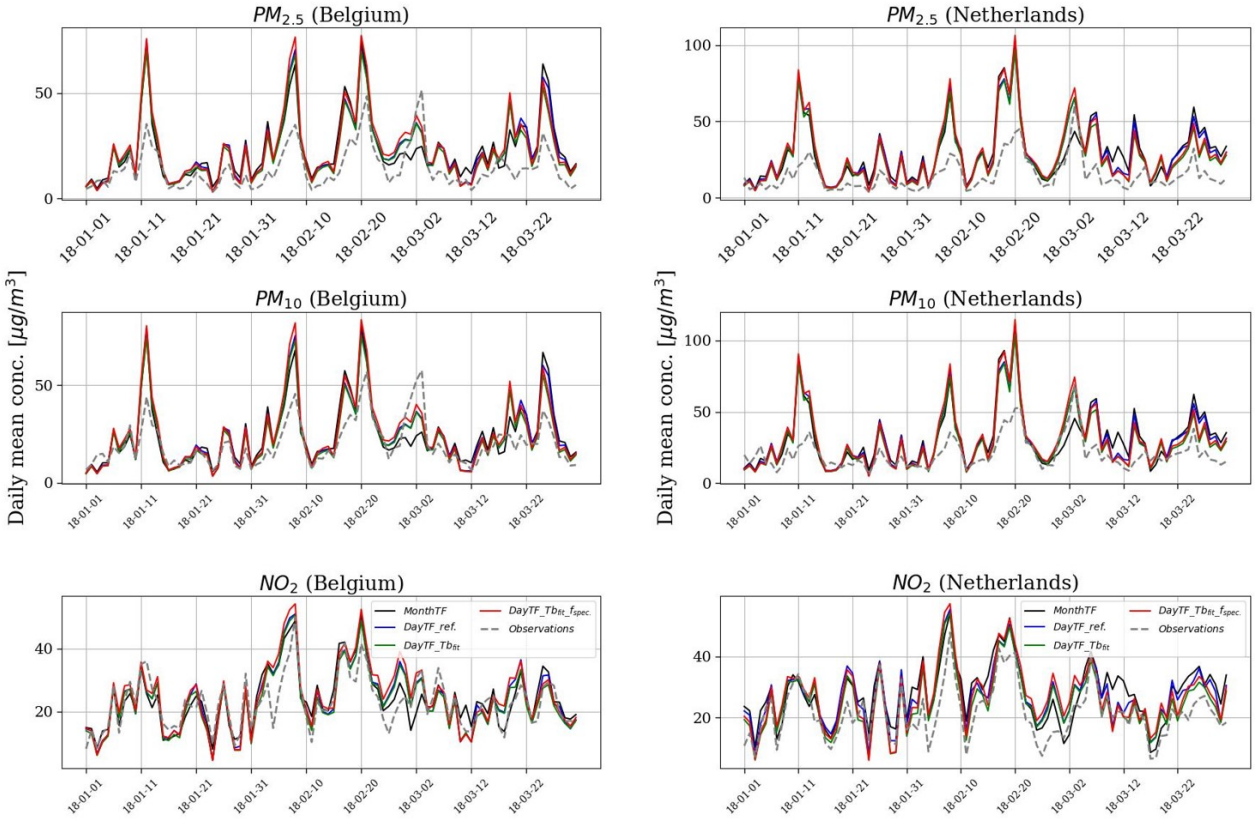

Figure S7. Average daily concentrations of $PM_{2.5}$ (top), PM_{10} (middle) and NO_2 (bottom) in Belgium (left) and Netherlands (right) between January and March from observations (~~AQ eReporting~~AQ e-Reporting) and different CHIMERE simulations.

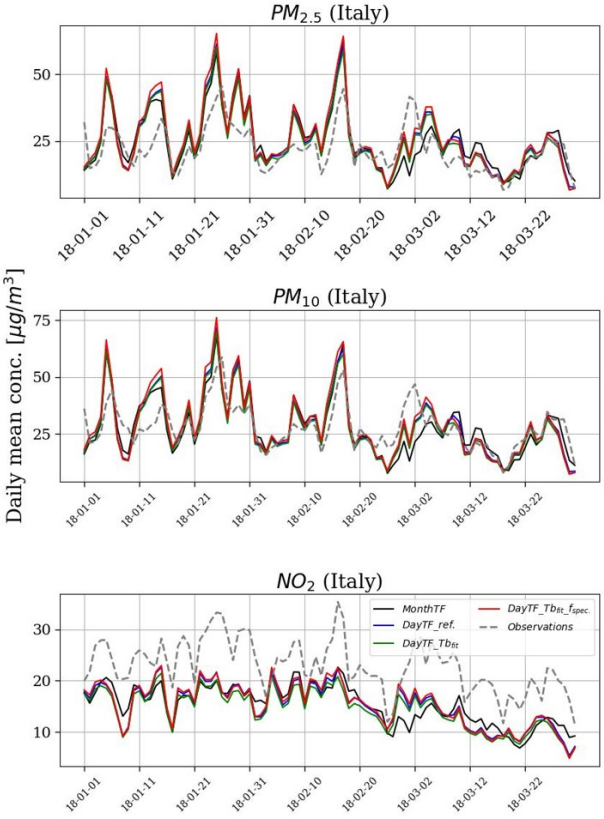

Figure S8. Same as Figure S7 for Italy.

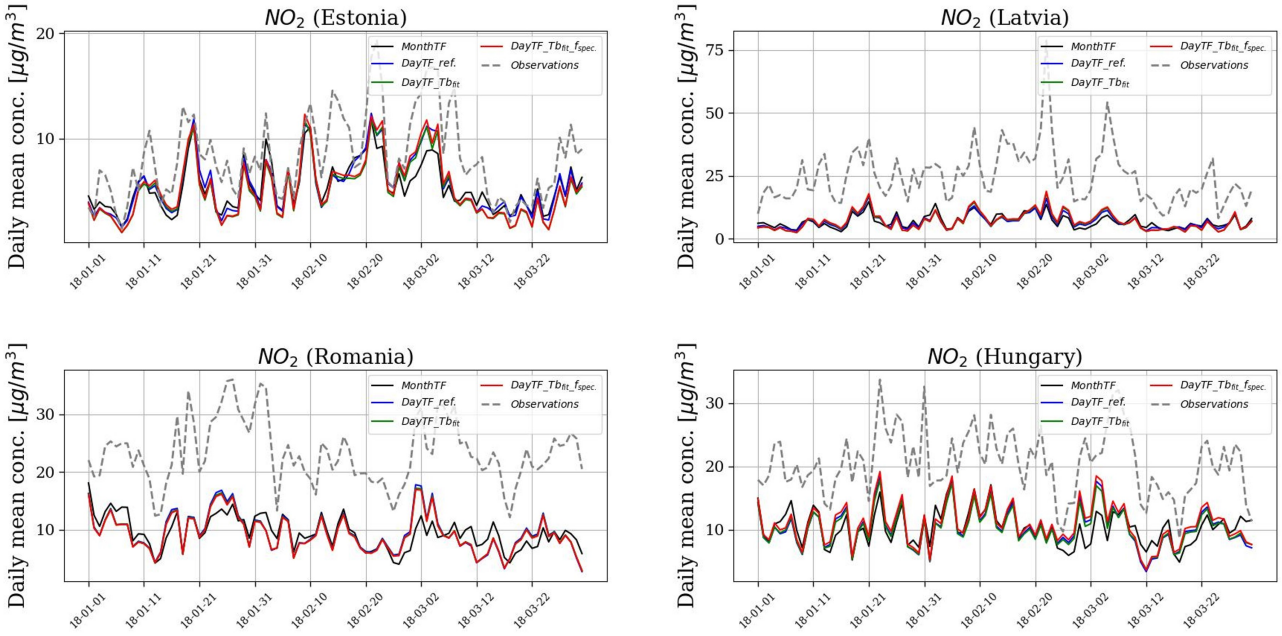

Figure S9. Average daily concentrations of NO_2 in Estonia (upper left), Latvia (upper right), Romania (lower left) and Hungary (lower right) between January and March from observations (~~AQ eReporting~~AQ e-Reporting) and different CHIMERE simulations.

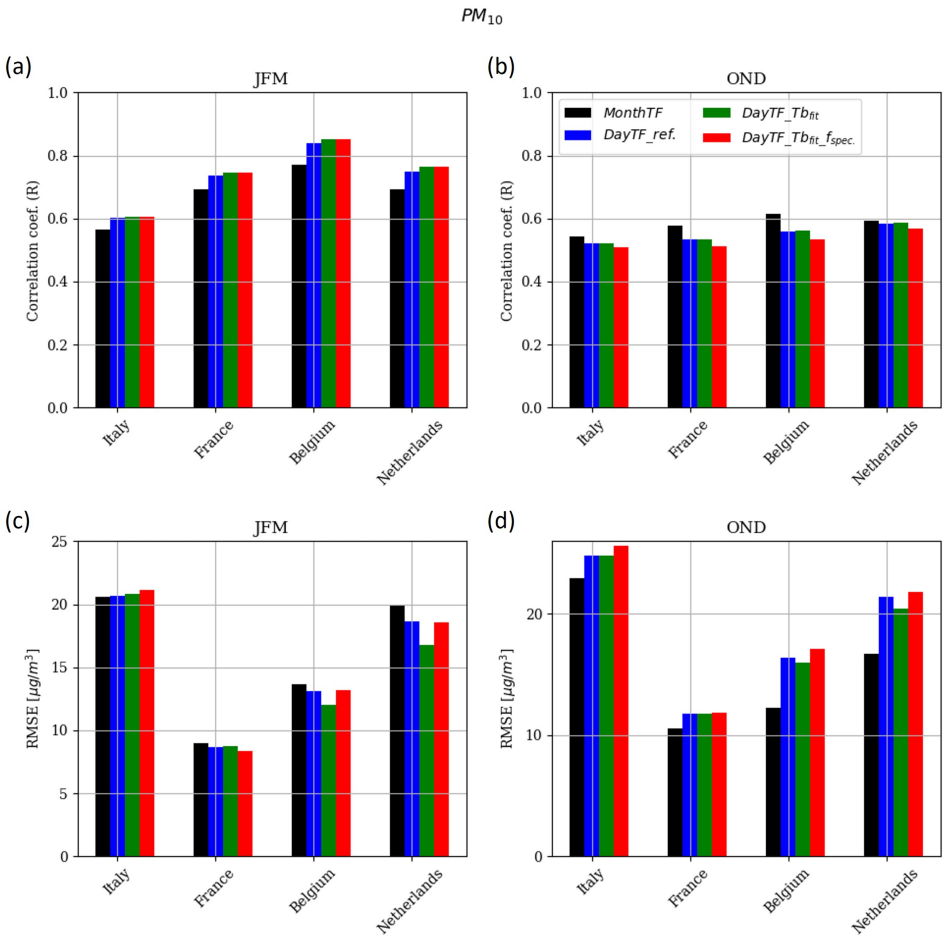

Figure S10. Spearman correlation (R coefficient) and RMSE ($\mu g/m^3$) of hourly PM_{10} concentrations for the months JFM (panel (a) and (c) respectively) and OND (panel (b) and (d) respectively), averaged over stations in countries that have been fitted with gas consumption data and for which concentration measurements are available.

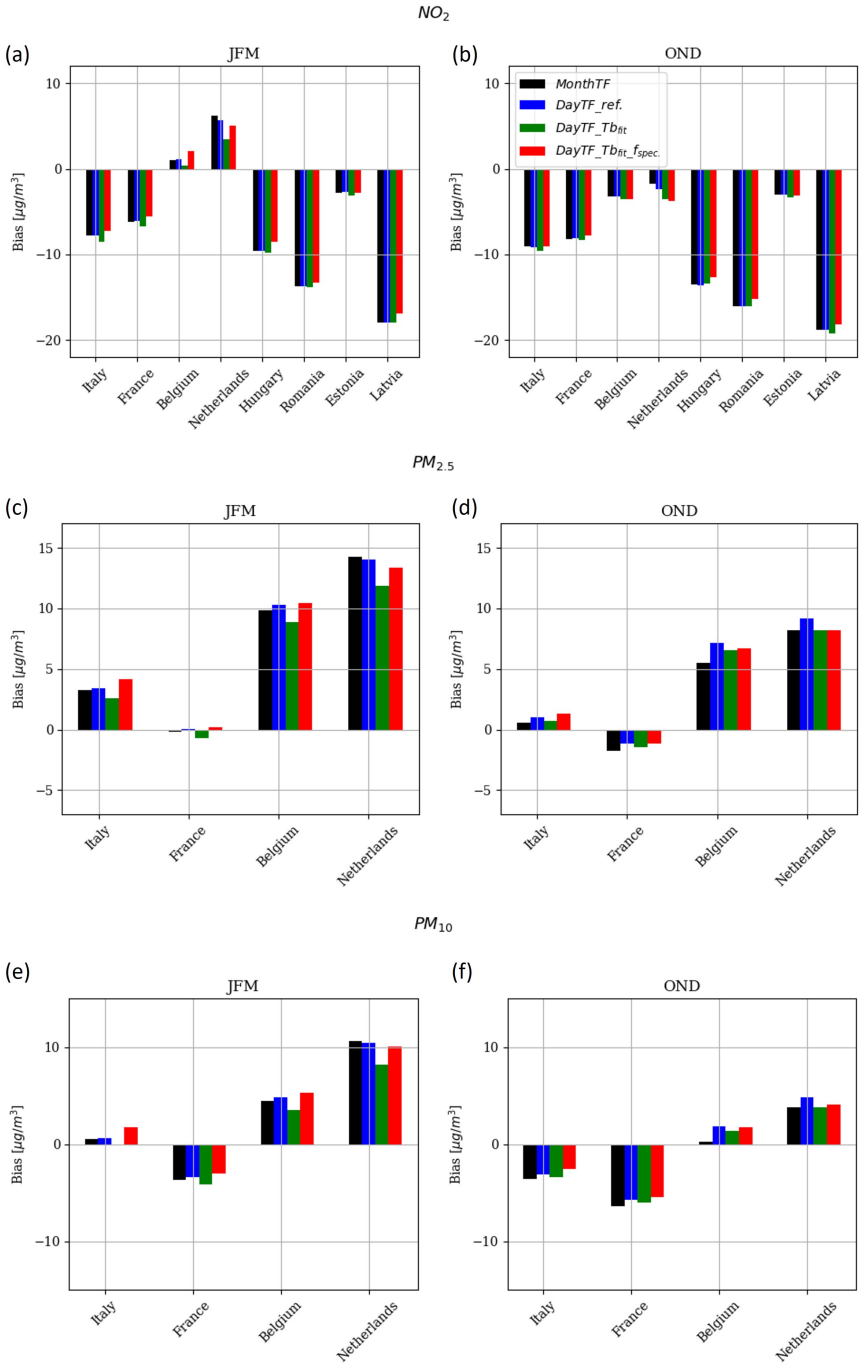

Figure S11. Average bias ($\mu g/m^3$) of hourly NO_2, $PM_{2.5}$, PM_{10} concentrations for the months JFM (panel (a), (c), (e) respectively) and OND (panel (b), (d), (f) respectively), averaged over stations in countries that have been fitted with gas consumption data and for which concentration measurements are available.

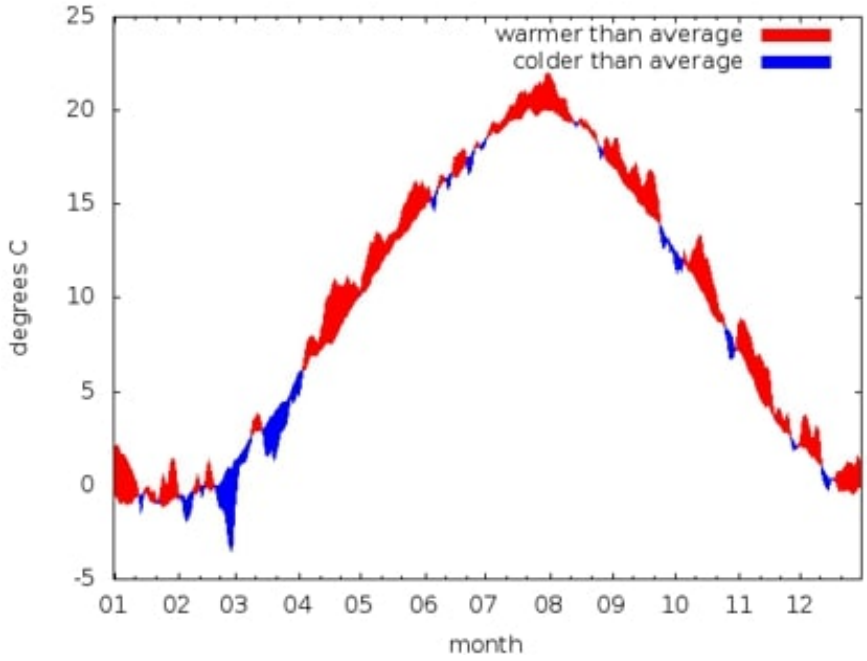

Figure S12. Daily anomalies of the European mean surface temperature for 2018, relative to 1981-2010. Data source: E-OBS, Credit: Copernicus Climate Change Service (C3S)/KNMI (https://climate.copernicus.eu/cold-start-year).

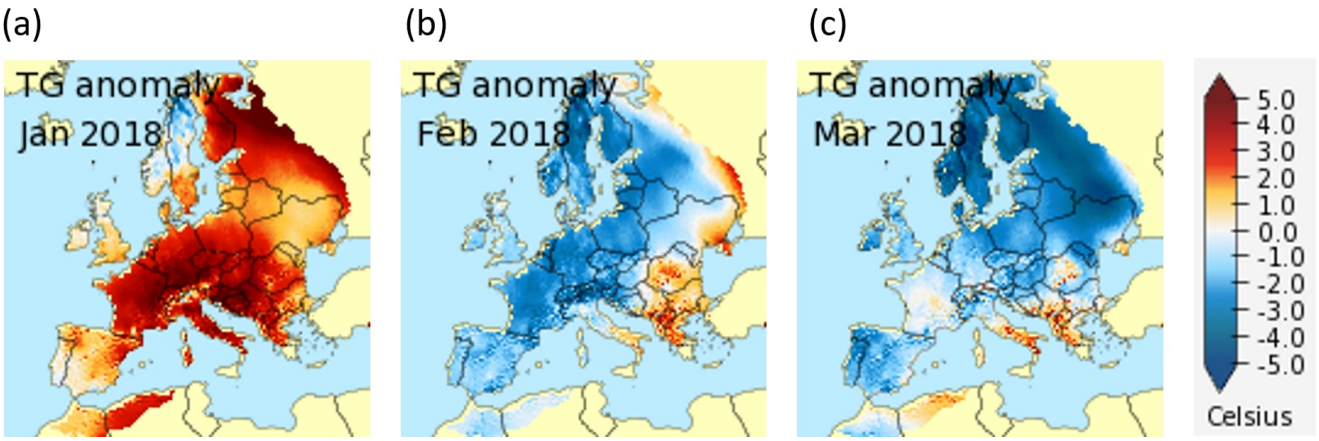

Figure S13. Differences of average temperature between January 2018 (a), February 2018 (b) and March 2018 (c), and the reference period 1981-2010. Data source: E-OBS, Credit: Copernicus Climate Change Service (C3S)/KNMI (https://surfobs.climate.copernicus.eu/stateofthecliimate/).

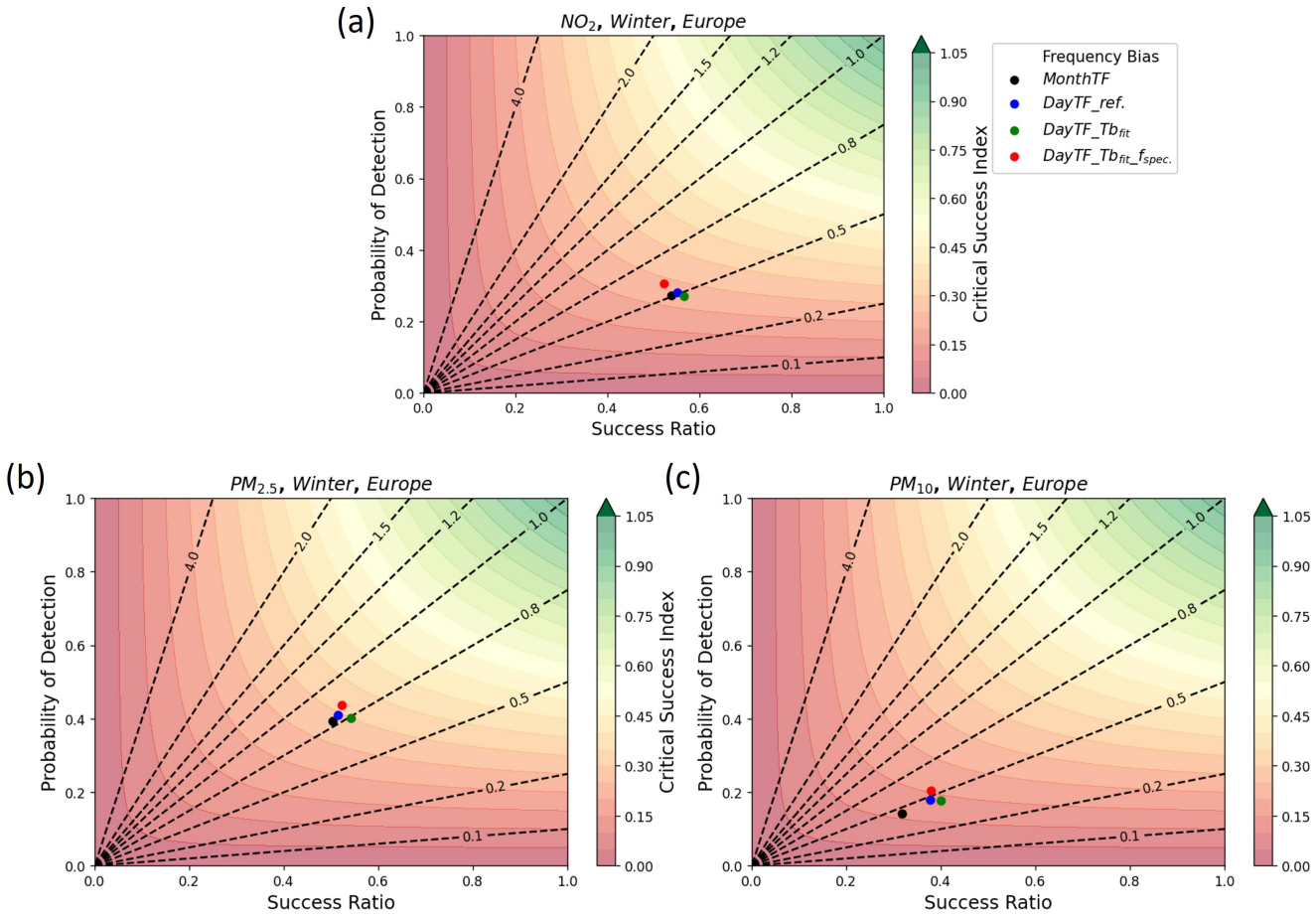

Figure S14. Performance diagram, as designed by Roebber (2009), comparing the ability to simulate the threshold exceedance (as a daily average) for the different experiments for NO_2 (a), $PM_{2.5}$ (b) and PM_{10} (c) averaged over Europe during the JFM months. The probability of detection is shown on the vertical axis, the success ratio on the horizontal axis, the frequency bias by the dashed line and the critical success index by the curved lines.

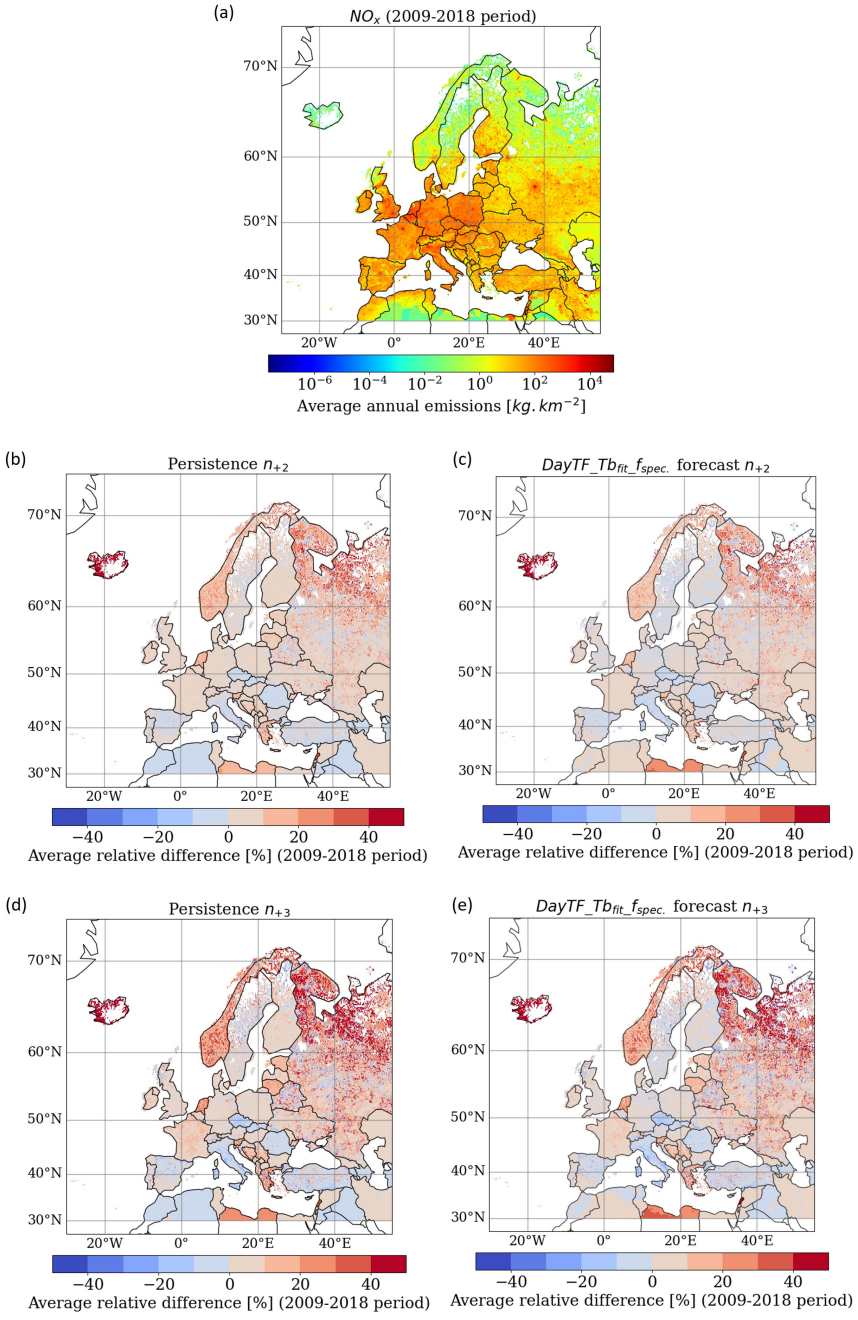

Figure S15. Spatial distribution of average annual NO_x emissions [kg/km^2] from $GNFR\_C$ (2009-2018 period), based on the CAMS-REG-AP-v5.1 inventory (a). Average relative difference using "$DayTF\_Tb_{fit}\_f_{spec.}$" to project in n_{+2} (b) in n_{+3} (d) for each year between 2009 and 2019, compared to the reported emissions. Average relative difference using the persistence method for n_{+2} (c) and n_{+3} (e) over 2009-2018.