The Lagrangian Atmospheric Radionuclide Transport Model (ARTM) - Sensitivity studies and evaluation using airborne measurements of power plant emissions

Robert Hanfland[1,2], Dominik Brunner[3], Christiane Voigt[2,4], Alina Fiehn[4], Anke Roiger[4], and Margit Pattantyús-Ábrahám[1]

[1]Bundesamt für Strahlenschutz, Oberschleißheim, Germany
[2]Johannes Gutenberg-Universität Mainz, Institute of Atmospheric Physics, Mainz, Germany
[3]Empa Swiss Federal Laboratories for Material Science and Technology, Dübendorf, Switzerland
[4]Deutsches Zentrum für Luft- und Raumfahrt (DLR), Institute of Atmospheric Physics, Oberpfaffenhofen, Germany
Correspondence: Robert Hanfland (rhanfland@bfs.de)

Abstract. The Atmospheric Radionuclide Transport Model (ARTM) operates at the meso-γ-scale and simulates the dispersion of radionuclides originating from nuclear facilities under routine operation within the planetary boundary layer. This study presents the extension and validation of this Lagrangian particle dispersion model and consists of three parts: i) a sensitivity study that aims to assess the impact of key input parameters on the simulation results; ii) the evaluation of the mixing prop-
5 erties of five different turbulence models using the well-mixed criterion; and iii) a comparison of model results to airborne observations of carbon dioxide (CO_2) emissions from a power plant and the evaluation of related uncertainties. In the sensitivity study, we analyse the effects of stability class, roughness length, zero-plane displacement factor and source height on the three-dimensional plume extent as well as the distance between source and maximum concentration at the ground. The results show that the stability class is the most sensitive input parameter as expected. The five turbulence models are the default turbu-
10 lence models of ARTM 2.8.0 and ARTM 3.0.0, one alternative built-in turbulence model of ARTM and two further turbulence models implemented for this study. The well-mixed condition tests showed that all five turbulence models are able to preserve an initially well-mixed atmospheric boundary layer reasonably well. The models deviate only 6% from the expected uniform concentration below 80% of the mixing layer height except for the default turbulence model of ARTM 3.0.0 with deviations by up to 18%, respectively. CO_2 observations along a flight path in the vicinity of the lignite power plant Bełchatów, Poland
15 measured by the DLR Cessna aircraft during the CoMet campaign in 2018 allow ~~to evaluate the model performance~~ evaluation of model performance (Reviewer (R) 1; Comment (C) 1) for the different turbulence models under unstable boundary layer conditions. All simulated mixing ratios are in the same order of magnitude as the airborne in situ data. An extensive uncertainty analysis using probability distribution functions, statistical tests and direct spatio-temporal comparisons of measurements and model results help to quantify the model uncertainties. With the default turbulence setups of ARTM version 2.8.0 and 3.0.0, the
20 plume widths are underestimated by up to 50% resulting in a strong overestimation of the maximum plume CO_2 mixing ratios. The comparison of the three alternative turbulence models shows a good agreement of the peak plume CO_2 concentrations, the CO_2 distribution within the plumes and the plume width with 30% deviations in peak CO_2 concentration and less than 25%

deviation of the measured CO_2 plume width. Uncertainties of the simulations may arise from the different spatial and temporal resolution of simulations and measurements in addition to the turbulence parametrisation and boundary conditions. The results of this work may help to improve the accurate representation of real plumes in very unstable atmospheric conditions by the selection of distinct turbulence models. Further comparisons at different stability regimes are required for a final assessment of model uncertainties.

Copyright statement. TEXT

1 Introduction

Atmospheric dispersion models (ADMs) are widely used by the scientific community and authorities. They are applied to a variety of problems, such as the study of the impact of pollutant emissions on air quality (Gariazzo et al., 2007; Stohl et al., 2007; Berchet et al., 2017; Lonati et al., 2022; Shupe et al., 2022) or the dispersion of radioactive discharges to the air (Chino et al., 2011; Connan et al., 2013; Draxler et al., 2015; Arnold et al., 2015) and they can operate at the full meteorological scale ranging from the micro-scale (shorter than km) to the meso-scale (km to thousand km) up to the synoptic or global scale (larger than thousand km).

The Atmospheric Radionuclide Transport Model (ARTM), analysed in this study, belongs to the class of models operating at the micro-β to meso-γ-scale (approx. $0.5\,km$ to $20\,km$). It is a Lagrangian particle dispersion model (LPDM) designed for the dispersion of radionuclides from nuclear facilities under routine operation in the planetary boundary layer (PBL).

However, any ADM has to demonstrate the applicability to the system of study. The most important method to confirm this is validation (Kleijnen, 1995; Schlesinger et al., 1979). This includes i) sensitivity analysis (SAR2; C7), which relates the model response to variations of the model's input parameters as well as ii) the comparison with observations revealing whether a model is an accurate representation of the system and whether simulation results are, to a certain degree, in agreement with observations (Kleijnen, 1995; Rao, 2005).

Concerning ARTM, Hettrich (2017) performed a sensitivity study to analyse the effect of input parameters (e.g. emission rate, source geometry, stability class and particle number) on concentrations at selected locations near the ground. Hanfland et al. (2022) provided an overview of the physical basis and mathematical formulations of the model and presented a qualitative description of the influences of different input parameters on three-dimensional plume characteristics for a general simulation setup. However, both lack in quantifying sensitivities.

Here we expand the former studies with a more systematic and quantitative sensitivity analysis. Different sensitivity coefficients are calculated, which describe the dependence of the simulation output on the input parameters stability class (SC), roughness length (z_0), zero-plane displacement factor (d) and source height (h_s) within the whole simulated PBL and rank them according to their effects on the model output.

Publications presenting comparisons of ARTM's mixing ratio simulation results with measurements are rare. Hettrich (2017) compared ARTM simulation results with measurements at a few selected locations near the surface showing discrepancies which are related to complex orography or local thermal induced winds that are not covered by ARTM's wind field model TALdia. Martens et al. (2012) studied the influence of a single large building close to the source on the dispersion showing that at distances larger than $4\,\mathrm{km}$ the influence decreases. These comparisons covered only a few atmospheric conditions and were limited to near surface concentration measurements.

Brunner et al. (2023) presented an intercomparison of six different atmospheric transport models, including ARTM, with airborne in situ and remote sensing carbon dioxide (CO_2) measurements sampling the exhaust plume of the Bełchatów lignite power plant in Poland under very unstable atmospheric conditions (Fix et al., 2018). The data set comprises a considerable number of plume transects at different distances from the source and heights within the PBL, is characterised by a strong contrast between background and plume CO_2 mixing ratio and provides a three-dimensional description of the mixing ratio field. The spatial extent of the area covered by measurements is around the maximum domain size ARTM can tackle. In that study ARTM simulations were performed using the default turbulence model of version 2.8.0. including a workaround for meandering plumes because the simulated plume appeared to be too narrow at very unstable atmospheric conditions.

ARTM's dispersion depends on the used turbulence model (Hanfland et al., 2022) and since the three-dimensional data set facilitates further analysis of the model we investigate whether the modelled plume could become more realistic by using different turbulence models. Two new turbulence models were implemented based on the ideas of Hanna (1982) and Degrazia et al. (2000) additionally to three built-in models of ARTM. The same workaround for meandering plumes based on the default turbulence model of ARTM 2.8.0, as primarily presented in Brunner et al. (2023), is also included to this study. Furthermore, all the five turbulence models are evaluated concerning turbulence model characteristics and mixing efficiency as well as compared with measurements.

The structure of this study is as follows: Section 2 gives a short description of the atmospheric dispersion model ARTM. Section 3 introduces several different sensitivity analysis methods and presents the sensitivity of typical model simulation outputs to key input parameters. Section 4 presents the five turbulence models and assesses their performance with respect to the well-mixed condition. Section 5 shows the comparison and evaluation of ARTM simulation results for the five turbulence models with the three-dimensional airborne data. Section 6 concludes the results of this paper.

2 The model ARTM

The Atmospheric Radionuclide Transport Model is a LPDM developed specifically for the dispersion of radioactive emissions from nuclear facilities in an area of typically $10\,\mathrm{km} \times 10\,\mathrm{km}$. Its purpose is to provide annual activity concentration fields in the area around nuclear facilities under routine operation in slightly structured non-urban terrain which are used in a follow-up step to calculate the additional exposure of the population (Hanfland et al., 2022).

The dispersion model propagates numerical particles representing radioactive tracers in space and time according to wind and turbulence fields obtained by a diagnostic approach. Meteorological data from measurements in the vicinity of the nuclear

facilities are used to calculate a mass-conserving diagnostic wind field. The turbulence is obtained by a Markov process which uses wind speed fluctuations and Lagrangian correlation times as input parameters, both depending on the Obukhov length as turbulence parameter. This static diagnostic approach employed by ARTM is similar to those of e.g. SWIFT/micro-SWIFT or CALMET (Cox et al., 2005; Moussafir et al., 2004; Scire et al., 1998) but (R2; C1) differs from other larger scale LPDMs such as FLEXPART, STILT, NAME or HYSPLIT, which use prognostic meteorological fields from a numerical weather prediction model to drive the particle propagation (Hanfland et al., 2022; Lin et al., 2003; Stohl et al., 2005; Ryall and Maryon, 1998; Draxler and Hess, 1998). The advantage of this static diagnostic (R2; C6) approach is its high computational efficiency but its application is limited to the area in the vicinity of the meteorological measurement (Ratto et al., 1994; Eichhorn and Kniffka, 2010; Wan et al. . with typical applications from $10\,\mathrm{km}$ to a few hundred km for horizontal extent depending on the number of meteorological measurement locations and the terrain (Ratto et al., 1994). Since ARTM works with one single meteorological measurement location its simulation domain can horizontally extend up to a few tens of km depending on the terrain. (R2; C6) The simulation domain is divided into grid cells for which the average activity concentration for the whole simulation period is calculated. A detailed model description is given in Hanfland et al. (2022).

3   Sensitivity study

Sensitivity analysis (SA) (R1; C2 and R2; C7) is a method to study the model's response to variations of input parameters in a systematic way and it may answer the following questions: how does the uncertainty of input parameters influence the model output; which parameters require additional research in order to reduce output uncertainty; which parameters are most significant or insignificant for the model's output; does the model behave as expected when varying a certain input parameter (Hamby, 1994; Frey and Patil, 2002; Rao, 2005; Saltelli et al., 2008). SA methods are either local or global depending on the sampled input parameter space (Saltelli et al., 2008; Morio, 2011; Zagayevskiy and Deutsch, 2015).

The results of the methods may differ depending on the shape of the input parameter space. Thus, the application of several methods is recommended (Iman and Helton, 1988; Hamby, 1995). In this work, several different local and global SA methods are therefore applied to provide ARTM 2.8.0 using its default turbulence model (Hanfland et al., 2022). A description for unstable atmospheric conditions is given in Sec. 4.1. The application of several SA methods provides a comprehensive assessment of the response of ARTM to different input parameters. (R2;General Comment (GC) 1)

3.1   Local sensitivity analysis methods

Local SA focuses on one single point in the input parameter space. The output of a model is represented by $Y = g(X_1, \ldots X_k)$ where the random variables X_i with $i = 1, \ldots, k$ denote the different input parameters. The representations (or values) of X_i are denoted with x_i. The input parameters X_i are varied one at a time while all the others are held constant at their reference values x_i^{ref}. This local SA approach is similar to estimating the partial derivative $\frac{\partial Y}{\partial X_i}$ and characterises the effect of the input parameter X_i on Y at one reference point $\boldsymbol{X}^{\mathrm{ref}} = (x_1^{\mathrm{ref}}, \ldots, x_k^{\mathrm{ref}})$ (Morio, 2011).

3.1.1 Sensitivity index

The sensitivity index described by Hoffman and Gardner (1983) uses the parameters at the reference point where each parameter is varied one at a time by their full range. The sensitivity index is calculated as

$$120 \quad SI_i = \frac{Y_{i,\mathrm{max}} - Y_{i,\mathrm{min}}}{Y_{i,\mathrm{max}}} \tag{1}$$

where $Y_{i,\mathrm{max(min)}}$ indicates the maximum (minimum) output value, respectively. The sensitivity index is a value between $0 \leq SI_i \leq 1$ and gives the fraction of output variation caused by the varied input parameter (Hamby, 1994).

3.1.2 One-at-a-time sensitivity measure

The one-at-a-time sensitivity measure calculates the variation of the model output normalised to the largest output variation ΔY_{max} that had been observed for the different input parameters. Starting from the default parameter set, the parameters are varied one at a time by a percentage α. The sensitivity coefficient for the input parameter X_i are calculated as

$$SM_i^\alpha = \frac{|Y_{i,+\alpha} - Y_{i,-\alpha}|}{\Delta Y_{\mathrm{max}}} \tag{2}$$

where $\Delta Y_{\mathrm{max}} = \max\left(|Y_{l,+\alpha} - Y_{l,-\alpha}|\right) \ \forall l \in i$ (Link et al., 2018). In this work the percentages $\pm 25\,\%$ and $\pm 50\,\%$ are used for α. SM_i^α is a value between $0 \leq SM_i^\alpha \leq 1$ where unity identifies the input parameter with the biggest effect on the model output Y.

3.2 Global sensitivity analyses

Global SAs sample the whole input parameter space, which leads to a broader representation of the sensitivity compared to local methods but also increase computation time. A general discussion about global SA can be found in Saltelli et al. (2008).

3.2.1 Sobol' indices

The variance-based Sobol' indices use variance decomposition to calculate indices of different orders (Sobol', 1993). Usually, only two key Sobol' indices are determined: the first-order index S_i; and the total effect index $S_{\mathrm{T}i}$.

For the first one, the conditional expected value of the model output $E_{\boldsymbol{X}_{\sim i}}(Y|X_i)$ with a constant value of X_i and varying values for all other input parameters $\boldsymbol{X}_{\sim i}$ is computed. For different realisations of X_i, $V_{X_i}[E_{\boldsymbol{X}_{\sim i}}(Y|X_i)]$ reflects the variance of the model output Y originating from a variation of the input parameter X_i. The first Sobol' index is then given by

$$140 \quad S_i = \frac{V_{X_i}[E_{\boldsymbol{X}_{\sim i}}(Y|X_i)]}{V(Y)} \tag{3}$$

where $V(Y)$ is the unconditional variance of the output where all X_i are varied. $V_{X_i}[E_{\boldsymbol{X}_{\sim i}}(Y|X_i)]$ cannot be larger than $V(Y)$ and thus for the sensitivity coefficient $0 \leq S_i \leq 1$ is valid. This index is called first-order sensitivity index as it does not take higher-order effects (i.e. interactions between different input parameters) into account (Saltelli et al., 2008).

The second index considered here is the total effect index. It takes higher order terms into account, which might be important depending on the model. The total effect is calculated as

$$S_{\mathrm{T}i} = 1 - \frac{V_{\boldsymbol{X}_{\sim i}}[E_{X_i}(Y|\boldsymbol{X}_{\sim i})]}{V(Y)} \tag{4}$$

where $V_{\boldsymbol{X}_{\sim i}}[E_{X_i}(Y|\boldsymbol{X}_{\sim i})] = V_{\boldsymbol{X}_{\sim i}}[E_{X_i}(Y|X_1, X_2, \ldots, X_{i-1}, X_{i+1}, \ldots, X_k)] \leq V(Y)$ is the total variance of all input parameters except X_i. As the first order index, this total effect index is a value between zero and unity where a value of zero indicates no influence of X_i on the output Y while unity indicates a strong influence (Saltelli et al., 2008). A comprehensive description of the method is given by Saltelli et al. (2008).

For the analysis presented here, the python library SALib (Herman and Usher, 2017) is used for the quasi-random sampling with low discrepancy after Joe and Kuo (2008) of the input parameter space as well as for the calculation of the Sobol' indices. It furthermore allows the estimation of the 95% confidence intervals (Herman and Usher, 2021).

3.2.2 δ-method

In comparison with the Sobol' indices, the δ-method takes the complete density distribution of the model output into account, which ensures the conservation of the whole information of the output density distribution (Borgonovo, 2007). The probability density function of X_i is denoted $f_{X_i}(x_i)$. The sensitivity coefficient δ_i for the input parameter X_i is calculated using the marginal density distribution of the input parameter $f_{X_i}(x_i)$ and the difference between the unconditional and the conditional density functions $f_Y(y)$ and $f_{Y|X_i}(y)$ of the model output with fixed representation $X_i = x_i$ as

$$\delta_i = \frac{1}{2} \int f_{X_i}(x_i) \left[\int |f_Y(y) - f_{Y|X_i}(y)| \, \mathrm{d}y \right] \, \mathrm{d}x_i \tag{5}$$

(Borgonovo, 2007). δ_i represents the total effect of an input parameter X_i on Y. It can take a value between zero and unity $(0 \leq \delta_i \leq 1)$ where zero means that the output is independent of X_i (Plischke et al., 2013). The same library SALib (Herman and Usher, 2017) was used to apply the δ-method including the estimation of the 95% confidence interval.

3.3 Model setup for sensitivity analyses

For the SA a simple model setup with a single source and constant wind was chosen. Figure 1 illustrates the simulation domain with the plume sampled at a height of $20.5\,\mathrm{m}$ above ground level (agl). The simulation domain has an extent of $10\,\mathrm{km} \times 1.5\,\mathrm{km} \times 1.5\,\mathrm{km}$ in x-, y-, and z-direction. The x-direction is defined in west-east orientation, the y-direction in north-south orientation. The domain is divided into grid cells with a horizontal resolution of $50\,\mathrm{m}$. Vertically, the 1.5 km high simulation (R2; C8) domain is divided into 19 levels of varying thickness gradually increasing from the lowest layer ($3\,\mathrm{m}$ thick) to the top simulation layer ($300\,\mathrm{m}$ thick). A table of the level thicknesses is shown in the Table S1. The point source is located at the coordinates ($x = 25\,\mathrm{m}$, $y = 25\,\mathrm{m}$). The vertical position is varied during the SAs. A constant westerly wind ($270°$) was used for the entire simulation period of 24 hours with a velocity of $1\,\mathrm{m\,s^{-1}}$ at $10\,\mathrm{m}$ height. In order to focus on the evolving dispersion pattern, the topography is assumed as flat surface. The gaseous krypton isotope $^{85}\mathrm{Kr}$ with a decay constant of
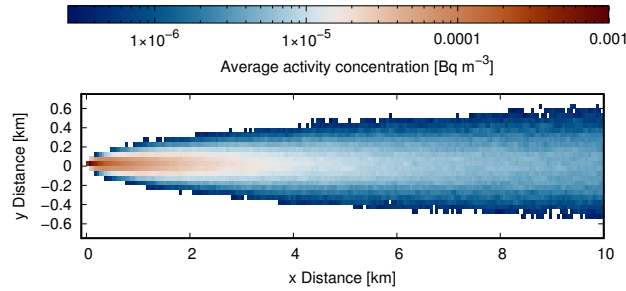
Figure 1. X-y-plane of the simulation domain for the sensitivity analyses with activity concentration at a height of $20.5\,\mathrm{m}$. The generic plume is simulated for SC = neutral, $z_0 = 0.5\,\mathrm{m}$, $d = 6$ and $h_s = 20\,\mathrm{m}$ with a wind speed of $1\,\mathrm{m\,s^{-1}}$ from the west at $10\,\mathrm{m}$ height and an emission rate of $1\,\mathrm{Bq\,s^{-1}}$. The activity concentration distribution in $\mathrm{Bq\,m^{-3}}$ is given in a logarithmic scale.

$\lambda_{\mathrm{decay}} = 2.05 \times 10^{-9}\,\mathrm{s^{-1}}$ was used as tracer. This results in a decay of less than $0.02\,\%$ within the simulation period and can
175 therefore be neglected. The emission source is represented as a source with a constant activity rate of $1\,\mathrm{Bq\,s^{-1}}$.

The input parameters stability class SC, roughness length z_0, zero-plane displacement factor d and source height h_s (R2; C9) are assumed to be the key parameters of ARTM. The zero-plane displacement d_0 depends on d as $d_0 = d \cdot z_0$. When using d instead of d_0 the input parameters for the SA are independent. This allows the analysis of the unbiased effects of input parameter variations. (R2; C9) These parameters and their ranges are summarised in Table 1. Only a discrete set of six SCs is
180 allowed in ARTM. The range of z_0 values allowed by ARTM is limited to the roughness lengths that correspond ~~the typical~~ to typical (R1; C3) land covers in the vicinity of German nuclear facilities. German authorities recommend a value of the zero-plane displacement factor $d = 6$ (TA Luft, 2002; VDI 3783 part 8, 2017). The range for the variation of d is centred on this value and limited to forest canopy heights typical for mixed forest (Lang et al., 2022). In Table 1, the parameter values for the reference point for the local SAs are marked with * symbol. For global sensitivity analyses, the whole parameter ranges are
185 sampled.

The target quantities of the SAs are two important characteristics of the plume, i) the plume volume, which is a measure for the tracer dispersion and is closely linked to the maximum mixing ratio and ii) the distance between the location of the maximum activity concentration and the source position at ground level ,which is of special interest for radiation exposure assessment.

190 **3.4 Results of the sensitivity analyses**

The results of the calculations of the local and global sensitivity coefficients are summarised in Table 2. Concerning the plume volume, all SA methods compute the highest SA coefficients for the stability class. Although less prominent, this is also observable for the distance between source and maximum concentration at the ground level except for sensitivity indices SI_i.

For $SM_{z_0}^{25}$ no value can be calculated because a variation of $\pm 25\,\%$ from the reference roughness length value does not lead
195 to a change of the categorial z_0 value. For the other parameters, the two different ranges of variation ($\alpha = 25$ and $\alpha = 50$)

Table 1. Input parameters and their values and ranges. The default parameters for local SAs are marked with (*).

Parameter	Values/Range
Stability class (SC) [1]	very stable, stable, neutral*, indifferent, unstable, very unstable
Roughness length (z_0) [1]	$0.10\,\mathrm{m}, 0.20\,\mathrm{m}, 0.50\,\mathrm{m}^*, 1.00\,\mathrm{m}, 1.50\,\mathrm{m}, 2.00\,\mathrm{m}$
Zero-plane displacement factor (d) [2]	$3 \ldots 6^* \ldots 15$
Source height (h_s) [3]	$10\,\mathrm{m} \ldots 20\,\mathrm{m}^* \ldots 120\,\mathrm{m}$

[1] For the global SAs the given values are sampled.

[2] For the global SAs the values are sampled continuously within the range.

[3] For the global SAs the values are sampled with a resolution of $1\,\mathrm{m}$ within the range.

Table 2. Sensitivity coefficients of local and global sensitivity analyses for the plume volume and for the distance between the source and the maximum concentration at the ground level. For the Sobol' indices (S_i and $S_{\mathrm{T}i}$) and the δ-method (δ_i) $95\,\%$ confidence intervals are given as well. Coefficients with very large relative confidence intervals are marked with (*), coefficients of one method, which cannot be distinguished within their confidence intervals are marked with (†).

Plume volume						
Parameter	SI_i	SM_i^{25}	SM_i^{50}	S_i	$S_{\mathrm{T}i}$	δ_i
SC	0.987	1	1	0.981 ± 0.032	0.99 ± 0.04	0.666 ± 0.009
z_0	0.718	-	0.202	$0.005 \pm 0.006^*$	0.017 ± 0.002	0.130 ± 0.001
d	0.291	0.022	0.027	$(0.3 \pm 8) \cdot 10^{-4*}$	$(28 \pm 5) \cdot 10^{-5}$	$0.126 \pm 0.002^\dagger$
h_s	0.119	0.004	0.004	$(0.9 \pm 9) \cdot 10^{-4*}$	$(39.5 \pm 2.3) \cdot 10^{-5}$	$0.126 \pm 0.002^\dagger$

Distance between source and maximum concentration						
Parameter	SI_i	SM_i^{25}	SM_i^{50}	S_i	$S_{\mathrm{T}i}$	δ_i
SC	0.884	1	1	0.16 ± 0.06	0.90 ± 0.06	0.328 ± 0.005
z_0	0.769	-	0.158	$0.00 \pm 0.04^*$	$0.75 \pm 0.07^\dagger$	$0.118 \pm 0.003^\dagger$
d	0.222	0.250	0.053	$0.00 \pm 0.01^*$	0.06 ± 0.02	$0.115 \pm 0.003^\dagger$
h_s	0.971	0.750	0.211	$0.02 \pm 0.04^*$	$0.74 \pm 0.07^\dagger$	0.129 ± 0.004

for SM_i^α provide valuable additional information. For example, it can be seen from Table 2 that the deviations between the coefficients of SM_i^{25} and SM_i^{50} are small for the plume volume while they are large for the distance between the source and the maximum concentration. The influence of the input parameters seems to be rather linear for the plume volume but highly non-linear for the distance between source and maximum concentration at the ground.

For the global SA methods, both target quantities show a distinct importance not only of the first order (direct influence of one single input parameter) but also of higher order (including interactions of two or more input parameters) effects. A small difference between S_i and $S_{\mathrm{T}i}$ shows a large first order effect as it can be seen for the plume volume. In contrary to this, a large difference reveals a small first order effect compared to a higher order effect as it can be seen for the distance between source and maximum concentration at ground level. This agrees with the conclusions that can be drawn from the δ_i coefficients. The sum of the sensitivity coefficients for the plume volume $\sum_i \delta_i = 1.05 \pm 0.01 \cong 1$ indicates that the effects of variation in the input parameters on variation in the plume volume are separable, i.e. interactions between input parameters play a minor role. For the distance between source and maximum concentration, the sum of the sensitivity coefficients $\sum_i \delta_i = 0.690 \pm 0.008 \ncong 1$ indicates the important role of cross interactions between the input parameters (Borgonovo, 2007). In contrary to the findings of the Sobol' indices that some input parameters having negligible influences, the δ-method suggests that the output characteristics are sensitive to all parameters. This difference could be due to the different amounts of information processed by the two methods. While the Sobol' indices compare conditional and unconditional variances of the output distribution the δ-method takes the entire output distribution into account.

Some of the global SA coefficients have very large relative confidence intervals and cannot be distinguished from zero (marked with *). Others cannot be distinguished from each other within their confidence intervals (marked with †). Increasing the sample size of 24,576 further would be necessary to get smaller confidence intervals but this would also increase the computation time (Herman and Usher, 2021).

Based on the coefficients from Table 2, the input parameters were ranked according to their importance as summarised in Table 3. The rankings obtained for the individual SA methods differ not only for the two target quantities but also between different methods. The overall ranking, which is simply computed as the sum (Σ) over the different methods, is provided in the last column.

The most unambiguous result is that all SA methods show the plume volume to be most sensitive to the SC. The ranks for the other input parameters, in contrast, are not the same. At this point we want to mention that the ranking of SM_i^{25} given in Table 3 is the average of all possible rankings for this method when taking into account that there is no coefficient for $SM_{z_0}^{25}$. The rankings of the remaining local SA methods SI_i and SM_i^{50} are in agreement with each other for the plume volume, while the rankings of the global SA methods disagree. Compared to the rankings for the plume volume, those for the distance between source and maximum concentration at the ground level are less uniform.

The overall rankings for both target quantities differ from each other, which emphasises that different target quantities are not necessarily sensitive to the same input parameters. Both target quantities are most sensitive to the SC, which is thus a potential source of high uncertainty. The source height h_{s} has little influence on the plume volume but it is the second most important parameter for the distance between source and maximum concentration at the ground level. ~~The strong influence of h_{s} on this target quantity is intuitively understandable, but interestingly SC is still more important.~~ (R2; C10)

Table 3. Ranking of the influence of the input parameters on the plume volume and on the distance between the source and the maximum concentration at ground level for local and global sensitivity analysis methods.

	Plume volume							
Parameter	SI_i	SM_i^{25}	SM_i^{50}	S_i	$S_{\mathrm{T}i}$	δ_i	Σ	Rank
SC	1	1	1	1	1	1	6	1
z_0	2	3	2	3	2	2	14	2
d	3	2	3	3	4	3.5	18.5	3
h_s	4	4	4	3	3	3.5	21.5	4
	Distance between source and maximum concentration							
Parameter	SI_i	SM_i^{25}	SM_i^{50}	S_i	$S_{\mathrm{T}i}$	δ_i	Σ	Rank
SC	2	1	1	1	1	1	7	1
z_0	3	3	3	3	2.5	3.5	18	3
d	4	4	4	3	4	3.5	22.5	4
h_s	1	2	2	3	2.5	2	12.5	2

4 Turbulence models for unstable conditions (R2; C11) and their performance on the well-mixed test

In LPDMs the turbulent motion is described via a Markov chain approach in the form of a Langevin equation (Lin and Gerbig, 2013). ARTM uses the wind speed fluctuation σ and Lagrangian correlation time T_L as input parameters for this Markov chain approach (Hanfland et al., 2022). The turbulence model specifies these two parameters and thus, influences the tracer dispersion and hence the simulated mixing ratio field (Katharopoulos et al., 2022).

4.1 Description of turbulence models

The turbulence model implemented in ARTM 2.8.0 as the default model is not widely used in the scientific community. Besides this, it has been reported by Janicke and Janicke (2011) that it sometimes underestimates plume dispersion. Therefore, they introduced a modified turbulence model leading to stronger dispersion, which can optionally be activated in ARTM. In 2022, the new version 3.0.0 of ARTM was released. It implements a new turbulence model according to the Association of German Engineers (VDI) guideline VDI 3783 part 8 (2017). All three models deviate from the model suggested by Hanna (1982), which is quite widely used and thoroughly tested against tracer release experiments. However, in this model the turbulence may abruptly change between SCs. To overcome this issue of discontinuity, Degrazia et al. (2000) proposed a continuous description of the turbulence throughout all atmospheric conditions. Since the used measurement data set for the comparison of simulations and observations was collected under unstable atmospheric conditions the following analyses focus on unstable stratification. (R2; C11) The wind speed fluctuation σ and the Lagrangian correlation time scales T_L of the five turbulence

models are presented in the following Eqs. 6 - 26 for unstable stratification and their profiles are displayed in Fig. 2. For the following quantities we define the x-components along the average horizontal wind direction, the y-components perpendicular

250 to it in the horizontal plane and the z-components in the vertical direction. ~~For simplicity~~ Although the zero-plane displacement is ~~not taken into account in the~~ used in ARTM (GRS, 2015) to displace the wind profile vertically to account for the influence of obstacles, in the sake of simplicity they are not included in the following equations. (R2; C12)

The first model, the default boundary layer model (BLM) of ARTM 2.8.0, was initially suggested by Kerschgens et al. (2000) and is based on the works of Lenschow et al. (1980); Panofsky et al. (1977); Hicks (1985) and Gryning et al. (1987). It (R2;

255 C13) describes profiles for the wind speed fluctuations as

$$\sigma_{\mathrm{x}} = 2.4 \cdot u_* \left(1 + 0.01486 \frac{-h_{\mathrm{m}}}{\kappa L} \right)^{\frac{1}{3}} \cdot \exp\left(\frac{-z}{h_{\mathrm{m}}} \right), \tag{6}$$

$$\sigma_{\mathrm{y}} = 1.8 \cdot u_* \left(1 + 0.03522 \frac{-h_{\mathrm{m}}}{\kappa L} \right)^{\frac{1}{3}} \cdot \exp\left(\frac{-z}{h_{\mathrm{m}}} \right) \tag{7}$$

and

260 $$\sigma_{\mathrm{z}} = 1.3 \cdot u_* \left[\left(1 - 0.8 \frac{z}{h_{\mathrm{m}}} \right)^3 \cdot \frac{-z}{\kappa L} + \exp\left(\frac{-z}{h_{\mathrm{m}}} \right)^3 \right]^{\frac{1}{3}} \tag{8}$$

where u_* is the friction velocity, h_{m} is the mixing layer height, $\kappa = 0.4$ is the von Kármán constant, L is the Obukhov length and z is the height agl ~~(Kerschgens et al., 2000; VDI 3783 part 8, 2002; Hanfland et al., 2022)~~(VDI 3783 part 8, 2002; Hanfland et al., 2022). This model is called ARTM2 in the following.

The second turbulence model available in ARTM is based on ARTM2 with a modification in the exponents as well as in the

265 prefactor of the crosswind component as

$$\sigma_{\mathrm{x}} = 2.4 \cdot u_* \left(1 + 0.01486 \frac{-h_{\mathrm{m}}}{\kappa L} \right)^{\frac{1}{3}} \cdot \exp\left(\frac{-0.3 \cdot z}{h_{\mathrm{m}}} \right), \tag{9}$$

$$\sigma_{\mathrm{y}} = 2.0 \cdot u_* \left(1 + 0.03522 \frac{-h_{\mathrm{m}}}{\kappa L} \right)^{\frac{1}{3}} \cdot \exp\left(\frac{-0.3 \cdot z}{h_{\mathrm{m}}} \right) \tag{10}$$

and

270 $$\sigma_{\mathrm{z}} = 1.3 \cdot u_* \left[\left(1 - 0.8 \frac{z}{h_{\mathrm{m}}} \right)^3 \cdot \frac{-z}{\kappa L} + \exp\left(\frac{-0.3 \cdot z}{h_{\mathrm{m}}} \right)^3 \right]^{\frac{1}{3}} \tag{11}$$

(Janicke and Janicke, 2011). This model leads to wider plumes and is called PRFMOD in the following.

In addition to the two previous models we added a turbulence model to ARTM based on ~~the~~ ARTM2 modified with (R2;

C14) formulations used in other ADMs (Stohl et al., 2005). This model uses σ_z from ARTM2 given in Eq. 8 but the horizontal wind speed fluctuations

275 $$\sigma_{\mathrm{x}} = \sigma_{\mathrm{y}} = u_* \left(12 + \frac{h_{\mathrm{m}}}{2|L|} \right)^{\frac{1}{3}} \tag{12}$$
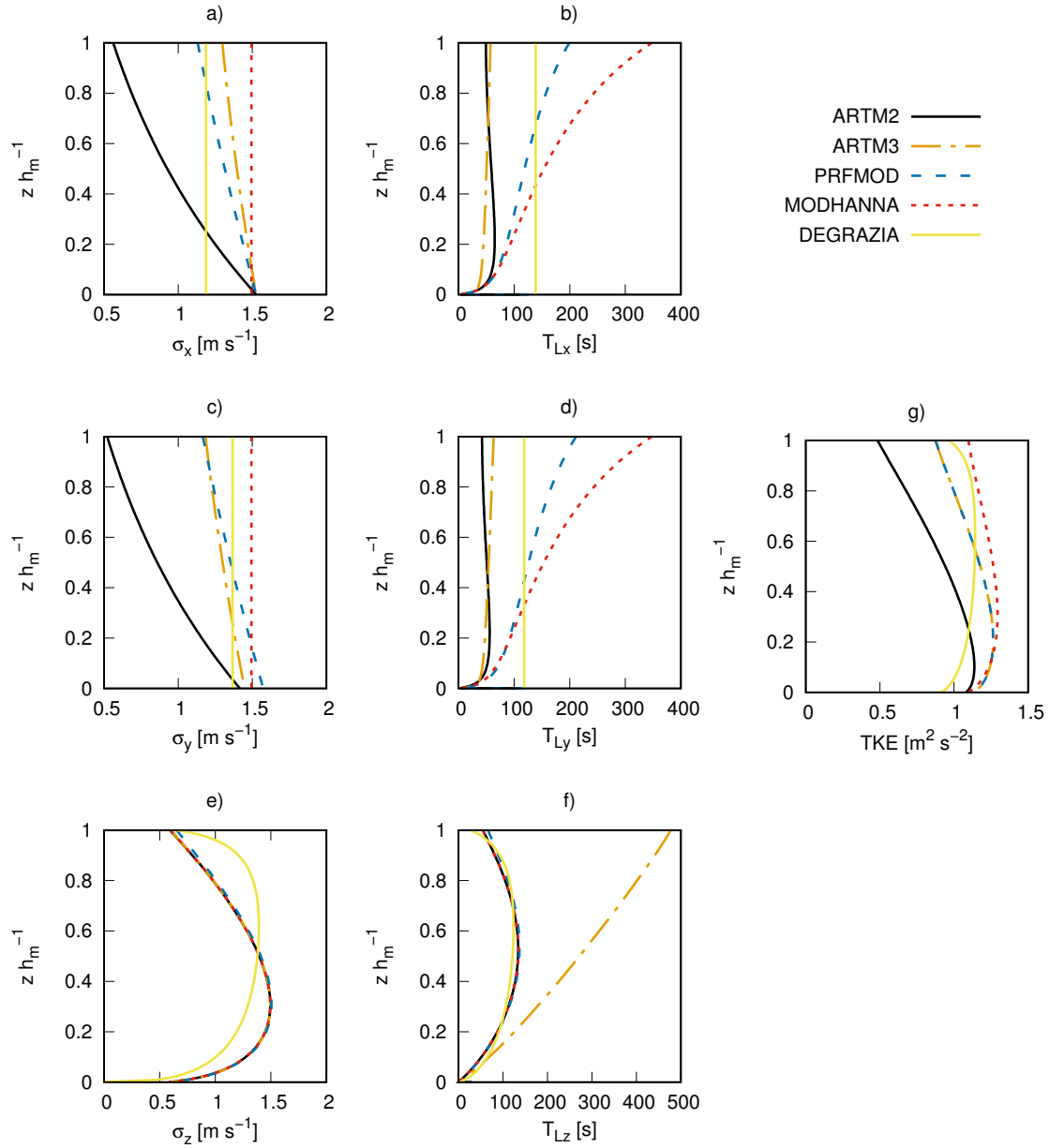
Figure 2. Vertical profiles of model characteristics for the five turbulence models ARTM2, ARTM3, PRFMOD, MODHANNA and DE-GRAZIA for very unstable atmospheric conditions. Wind speed fluctuations σ: a) along wind direction σ_x; c) in the crosswind direction σ_y; and e) in the vertical direction σ_z against the normalized height (normalized to the boundary layer height). The corresponding Lagrangian correlation times T_L are shown in b), d) and f), respectively. The turbulent kinetic energy per unit mass TKE is shown in g).

are equal to the equations suggested by Hanna (1982). The aim of this modification is to increase the turbulent kinetic energy and to analyse the effect of the horizontal wind speed fluctuations on the dispersion. (R2; C14) In the following, this model is called MODHANNA.

The Lagrangian correlation times of the three models above are given according to Kolmogorov's theory as

$$T_{\mathrm{L}i} = \frac{2 \cdot \sigma_i^2}{\mathrm{C}_0 \cdot \eta} \tag{13}$$

(Luhar and Britter, 1989) with the Kolmogorov constant $\mathrm{C}_0 = 5.7$ and the dissipation rate of the turbulent kinetic energy

$$\eta = \max \left\{ \frac{u_*^3}{\kappa z} \left[\left(1 - \frac{z}{h_{\mathrm{m}}} \right)^2 + \frac{z}{h_{\mathrm{m}}} \right] + \frac{-u_*^3}{\kappa L} \left[1.5 - 1.3 \left(\frac{z}{h_{\mathrm{m}}} \right)^{\frac{1}{3}} \right], \frac{u_*^3}{\kappa z} \right\}. \tag{14}$$

The fourth model is the default model of the new version 3.0.0 of ARTM with the wind speed fluctuations given as

$$\sigma_{\mathrm{x}} = 2.4 \cdot u_* \left[1 + 0.01486 \frac{-h_{\mathrm{m}}}{\kappa L} \cdot \exp \left(-0.9 \frac{z}{h_{\mathrm{m}}} \right) \right]^{\frac{1}{3}}, \tag{15}$$

$$\sigma_{\mathrm{x}} = 2.0 \cdot u_* \left[1 + 0.02568 \frac{-h_{\mathrm{m}}}{\kappa L} \cdot \exp \left(-0.9 \frac{z}{h_{\mathrm{m}}} \right) \right]^{\frac{1}{3}} \tag{16}$$

and

$$\sigma_{\mathrm{z}} = 1.3 \cdot u_* \left[\left(1 - 0.8 \frac{z}{h_{\mathrm{m}}} \right)^3 \cdot \frac{-z}{\kappa L} + \exp \left(-0.9 \frac{z}{h_{\mathrm{m}}} \right)^3 \right]^{\frac{1}{3}} \tag{17}$$

(VDI 3783 part 8, 2017). The Lagrangian correlation time scales are calculated via the turbulent diffusion coefficients K_i as

$$T_{\mathrm{L}i} = \frac{K_i}{\sigma_i^2} \tag{18}$$

with

$$K_j = 0.9 \frac{u(z) \cdot h_{\mathrm{m}}}{100 \cdot u_*} \sigma_j \tag{19}$$

for the horizontal components j and

$$K_z = \kappa u_* z \left[\left(1 - 0.8 \frac{z}{h_{\mathrm{m}}} \right)^4 \frac{9z}{-L} + \exp \left(-3.6 \frac{z}{h_{\mathrm{m}}} \right) \right]^{\frac{1}{2}} \tag{20}$$

for the vertical component, respectively (VDI 3783 part 8, 2017). This model is called ARTM3 in the following.

We implemented a fifth model, which in contrast to the previous four turbulence models that are based on similarity theory, is based on the spectral distribution of the turbulent kinetic energy of the boundary layer and was presented by Degrazia et al. (2000). For very unstable boundary conditions the wind speed fluctuations are given as

$$\sigma_{\mathrm{x}} = 0.53 \cdot u_* \left(\frac{-h_{\mathrm{m}}}{\kappa L} \right)^{\frac{1}{3}}, \tag{21}$$

$$\sigma_{\mathrm{y}} = 0.61 \cdot u_* \left(\frac{-h_{\mathrm{m}}}{\kappa L} \right)^{\frac{1}{3}} \tag{22}$$

and

$$\sigma_{\mathrm{z}} = 0.54 \cdot u_* \left(\frac{-h_{\mathrm{m}}}{\kappa L} \right)^{\frac{1}{3}} \cdot \left\{ 1.8 \left[1 - \exp\left(\frac{-4z}{h_{\mathrm{m}}} \right) - 0.0003 \cdot \exp\left(\frac{8z}{h_{\mathrm{m}}} \right) \right] \right\}^{\frac{1}{3}} \tag{23}$$

with the Lagrangian correlation times

$$T_{\mathrm{L}i} = \frac{l_i}{\sigma_i} \tag{24}$$

where l_i is the Lagrangian correlation length given as

$$l_{\mathrm{x}} = l_{\mathrm{y}} = 0.21 \cdot h_{\mathrm{m}} \left(0.01 \frac{h_{\mathrm{m}}}{-L} \right)^{\frac{1}{2}} \tag{25}$$

and

$$l_{\mathrm{z}} = 0.14 \cdot h_{\mathrm{m}} \left(0.01 \frac{h_{\mathrm{m}}}{-L} \right)^{\frac{1}{2}} \cdot \left\{ 1.8 \left[1 - \exp\left(\frac{-4z}{h_{\mathrm{m}}} \right) - 0.0003 \cdot \exp\left(\frac{8z}{h_{\mathrm{m}}} \right) \right] \right\}. \tag{26}$$

In this work this turbulence model is denoted as DEGRAZIA.

The turbulent kinetic energy per unit mass is determined as

$$TKE = \frac{1}{2} \left(\sigma_{\mathrm{x}}^2 + \sigma_{\mathrm{y}}^2 + \sigma_{\mathrm{z}}^2 \right) \tag{27}$$

(Stull, 1988).

4.2 Evaluation of turbulent mixing

The degree of preservation of well-mixed conditions is a key quality indicator for any LPDM, similar to the preservation of mass in a Eulerian model. It tests whether an initially uniform distribution of a tracer in an incompressible flow remains uniform as postulated by the second law of thermodynamics (Sawford, 1986; Thomson, 1987; Lin and Gerbig, 2013; Bahlali et al., 2020). Exactly fulfilling this criterion is challenging, but it is important to quantify the degree of deviation from this ideal behaviour to judge the magnitude of systematic model biases and whether these biases are acceptable. In this work the tests of the turbulence models are limited to the case of very unstable atmospheric conditions because the observations that are used for the comparison were collected at very unstable atmospheric conditions, too.

The well-mixed condition test can characterise the vertical mixing homogeneity of a model. For these tests simulation domains with periodic horizontal boundaries and reflecting vertical boundaries are used. (R2; C15) This virtually expands the simulation domain to infinite extent and prevents the simulation from losing tracer mass. The whole simulation domain serves as a volume source where 115,200 numerical particles are inserted uniformly within the first simulation hour because ARTM does not take into account the changing density with height. The domain size is $2000\,\mathrm{m} \times 2000\,\mathrm{m} \times 1100\,\mathrm{m}$ in x-,

y-, and z-direction with a horizontal (vertical) resolution of $200\,\mathrm{m}$ ($25\,\mathrm{m}$), respectively. The vertical extent of the domain is equal to the assumed mixing depth. A temporally constant wind profile for unstable atmospheric conditions, as described in Hanfland et al. (2022) (R2; C16) with a wind speed of ~~$1\,\mathrm{m\,s^{-1}}$~~ $2.3\,\mathrm{m\,s^{-1}}$ (R2; C16) at $10\,\mathrm{m}$ height and a direction of $270°$
(westerly) is used. The selected wind speed originates from measurement sites at very unstable stratification in Germany. (R2; C16) For the evaluation, the hourly mean concentration and its standard deviation was derived for each vertical level.

The concentration profiles of the different turbulence models for very unstable PBL conditions are shown in Fig. 3. The concentrations of the state of mixing after one hour (red line) and after two hours (blue dashed line) are shown. Concentration values are normalised to the mean concentration ($c\,\overline{c}^{-1}$) and the height is normalised to the mixing depth ($z\,h_{\mathrm{m}}^{-1}$). We used the same initial numerical particle distribution for all turbulence models to eliminate possible differences arising from different initial distributions.

The concentration profiles after one hour differ from the uniform distribution $c\,\overline{c}^{-1} = 1$. This indicates a certain degree of segregation of the numerical particles but most deviations are less than 5% (vertical dashed lines). The largest deviations can be found at the top of the PBL for the ARTM3 model (~~$> 30\%$~~ $> 35\%$) (R2; C16) and the DEGRAZIA model ($> 15\%$). The profiles of the ARTM2 and the MODHANNA turbulence models are very similar since they both contain the same vertical turbulence parametrisation. The PRFMOD turbulence model differs slightly from the ARTM2 model due to modifications described in Eq. 11. The profile of ARTM3 shows trends of dilution and accumulation similar to ARTM2, PRFMOD and MODHANNA but magnified in its extent. The profile of the DEGRAZIA turbulence model shows a different shape because of its different formulation of the turbulence parameters (see Eqs. 23, 24 and 26).

By $t = 2\,\mathrm{h}$, the dilution of concentration has further increased at the bottom and the top of the PBL and the accumulation at $z\,h_{\mathrm{m}}^{-1} \approx 0.3$ (horizontal dashed black line) has further increased partly beyond 5% but well below 10% deviation for ARTM2, PRFMOD, MODHANNA and DEGRAZIA. For ARTM3, the dilution at the ground almost vanishes while the dilution above $z\,h_{\mathrm{m}}^{-1} = 0.8$ increases to 40% and the accumulation in the middle of the PBL increases to 18%. After the second hour, no further changes are observed as it can be seen in Fig. 3 f) for the ARTM2 turbulence model at $z\,h_{\mathrm{m}}^{-1} \approx 0.3$. Time series for other heights and other turbulence models show similar behaviour and are given in Supplement S2.

This well-mixed condition test shows that the simulation result systematically overestimates the concentration values at $z\,h_{\mathrm{m}}^{-1} \approx 0.3$ for the ARTM2, PRFMOD and the MODHANNA model after the second hour. Near the surface, which is important for estimation of exposure to the population, the concentration values are underestimated. In both cases, the errors are only 5 - 6%. At the top of the PBL, the models underestimate the expected concentration significantly. The ARTM3 turbulence model shows the smallest deviation from the mean domain concentration near the ground but it overestimates the concentration in the middle of the PBL before turning into a substantial underestimation towards the mixing layer top. Below $z\,h_{\mathrm{m}}^{-1} = 0.8$ the turbulence model DEGRAZIA performs best. At the top of the PBL the model decreases well below the expected concentration. All the tested turbulence models can be assumed as acceptable for simulations at very unstable atmospheric conditions but the partly large deviations of the concentration from the expected values at certain heights have to be taken into account when interpreting model results. At low-wind conditions ($1\,\mathrm{m\,s^{-1}}$ at $10\,\mathrm{m}$ height) the deviation from the uniform concentration is similar (see Supplement S2). (R2; C16)
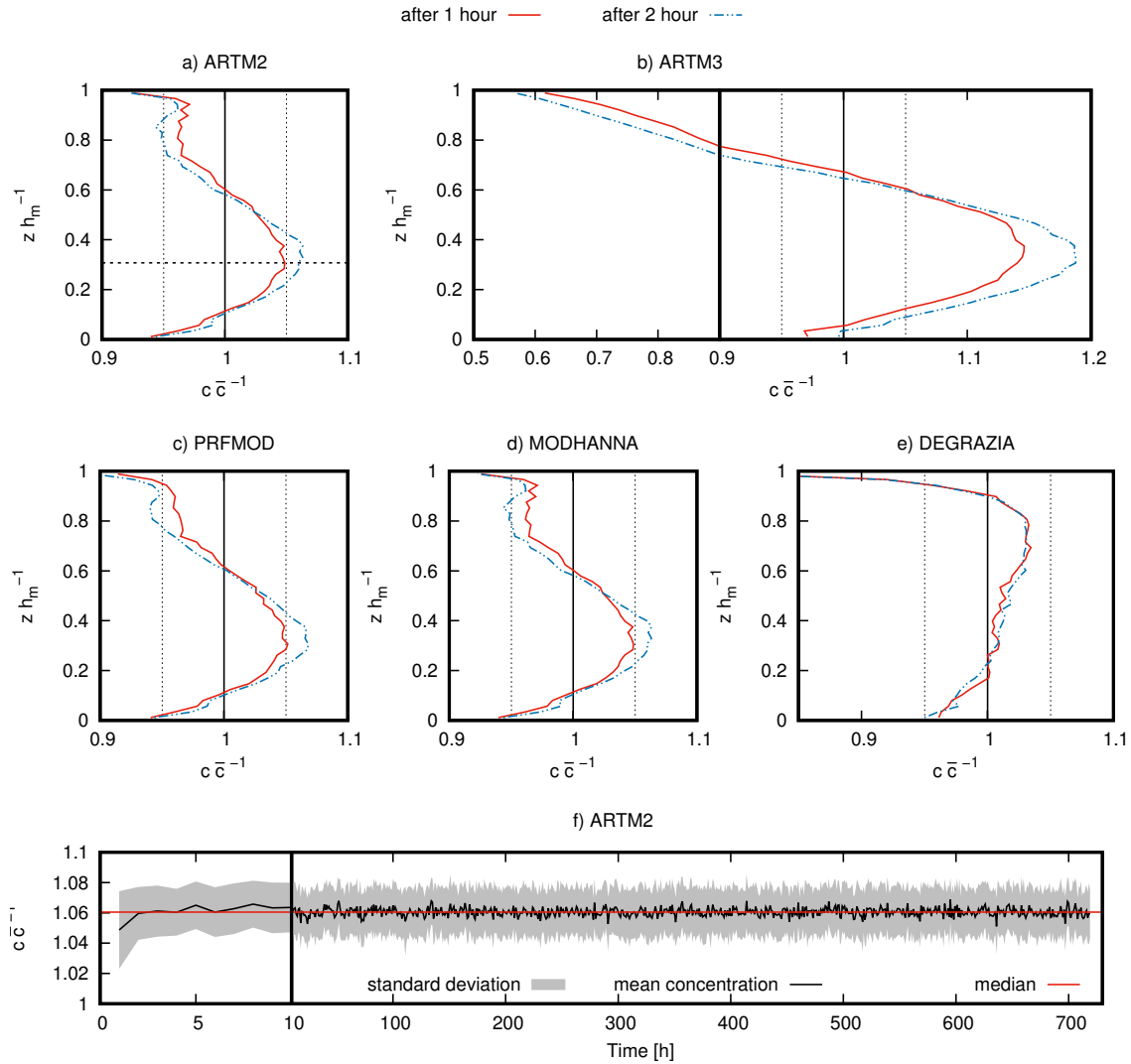
Figure 3. Profiles of the concentration normalised to the mean concentration $c\bar{c}^{-1}$ (a, b, c, d, e) of the different turbulence models ARTM2, ARTM3, PRFMOD, MODHANNA and DEGRAZIA after one hour (red lines) and two hours (blue dashed dotted lines) for periodic lateral simulation domain boundaries and reflecting bottom and top boundaries. In b) the x-axis scale changes at $c\bar{c}^{-1} = 0.9$. f) Time series of the normalised concentration at normalised height $z\,h_{\mathrm{m}}^{-1} \approx 0.3$ for the ARTM2 model, which is indicated by the dashed horizontal line in a. The x-axis scale changes at 10 hours.

5   Comparison of ARTM simulation with airborne observations

The comparison of atmospheric dispersion simulation results with measurements near the ground is not sufficient to derive any conclusions about the three-dimensional structure of simulated emission plumes. To study the agreement of simulated

365   and observed plume dispersion it is inevitable to use observations that resolve the structure of the real plume. Since ARTM

simulates the emissions of nuclear facilities ~~from quite high stack~~with source heights of mainly 100 m to 200 m (R1; C4), it is useful to choose observational data originating from similar height levels. In this work, we present a comparison of ARTM simulations with airborne CO_2 observations within the PBL. In such a case, the comparison is challenging because of the turbulent character of the PBL. As pointed out by ?Brunner et al. (2023), observations only provide snap-shots of the real world while simulations provide one realisation of a multitude of stochastic representations of the real world. Simulations with slightly perturbed initial conditions could result in different dispersion patterns of the plume. Furthermore, simulation results and observations may have different spatial and temporal resolutions and uncertainties, which complicate the comparison of simulations with observations (Farchi et al., 2016). Thus, in this work, the comparison of simulation results with observations for five turbulence models are given using rather general plume characteristics such as the plume width per transect and maximum mixing ratios.

5.1 Observational data

The aircraft observations used for this investigation originate from the Carbon Dioxide and Methane Mission (CoMet 1.0) ~~(Fix et al., 2018; Luther et al., 2019; Fiehn et al., 2020; Gałkowski et al., 2021; Wolff et al., 2021; Krautwurst et al., 2021; Andersen et al.~~ (Fix et al., 2018; Luther et al., 2019; Fiehn et al., 2020; Gałkowski et al., 2021; Wolff et al., 2021; Krautwurst et al., 2021; Andersen et al. . The campaign took place in May and June 2018 and involved three aircraft performing in situ and remote sensing measurements. The objective was to study CO_2 and methane (CH_4) emissions from different sources in Europe including power plants, as well as to compare the different observational methods.

For the evaluation of ARTM, airborne in situ CO_2 measurements in the vicinity of the Bełchatów lignite power plant in Poland were used (Fiehn et al., 2020; Kostinek et al., 2021). An overview map with the CO_2 mixing ratios along the flight path is shown in Fig. 4. The in situ measurements had been performed on 7 June 2018 between 13:00 and 15:00 UTC aboard the DLR Cessna Grand Caravan 208B. One transect on the upwind side of the emitter was performed at the beginning in order to derive the mean background CO_2 mixing ratio $c_{CO_2} = 401.2$ ppmv. The exhaust plume of the power plant was probed during several transects on the downwind side at heights between 500 m and 1.7 km agl. They form two wall patterns at meridional distances of approx. 13 km (wall 1) and 23 km (wall 2) and a single transect at approx. 6 km from the source.

~~The~~ CO_2 ~~had been~~ was (R2; GC3) measured with a cavity ring-down spectroscopy analyser (G1301-m, Piccaro) specifically modified for the airborne deployment ~~and with an uncertainty of~~ operating at 0.5 Hz. the CO_2 measurement uncertainty is ± 0.15 ppmv and the temporal resolution was increased to 1 s by interpolation to make the data comparable with other data collected during the campaign (R2; GC3). Details of the measurement equipment and related uncertainties are described by Klausner et al. (2020). The sampling repetition and the velocity of the aircraft results in a spacial distance of about 140 m between the 0.5 Hz data points. The Picarro instrument measures CO_2, methane and water vapor sequentially and , thus, the values are representative for the last third of the measurement interval. (R2; GC3) Observational data for wind direction, wind speed and flight height are shown in Supplement S3.
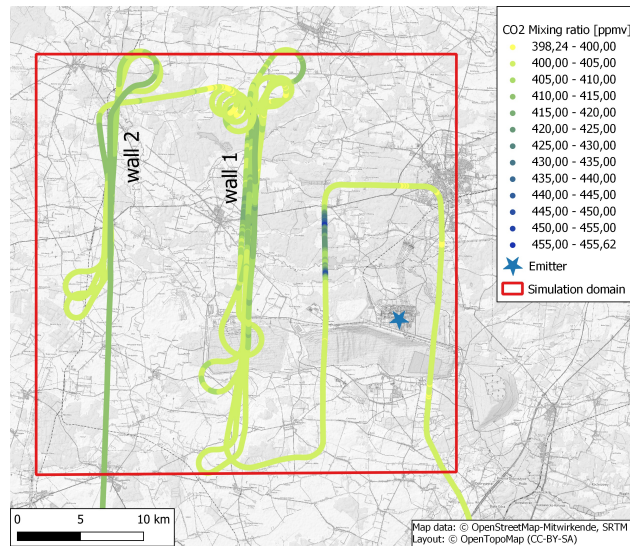
Figure 4. Map showing the flight path of the DLR Cessna aircraft in the vicinity of the Bełchatów lignite power plant (blue star), color-coded by the in situ measured CO_2 values. Transects were performed both east (upwind side) and west (downwind side) of the emitting power plant. The red box indicates the simulation domain.

5.2 Model setup

We chose a simulation domain of $33.3 \, \text{km} \times 33.3 \, \text{km} \times 1.9 \, \text{km}$ that covers the horizontal extent of the flight trajectory and vertically extends beyond the mixing layer depth by four simulation level. The horizontal resolution was $150 \, \text{m}$. The extent of the simulation domain with the location of the emission source (two stacks in a distance of $300 \, \text{m}$) is shown in Fig. 4. Vertically, the grid spacing gradually increases from $3 \, \text{m}$ to $35 \, \text{m}$ until $100 \, \text{m}$ height is reached. Above, $50 \, \text{m}$ level thickness was used. All level thicknesses are given in Table S2.

ARTM requires several input parameters: SC, z_0, d_0, orography, several source specific parameters as well as wind speed and direction at one location in the simulation domain. Since there were no stationary ground based wind measurements available, wind direction and wind velocity as well as SC and mixing layer height were derived from the airborne measurements. The actual emission rates are unknown. However, ~~?~~ Brunner et al. (2023) estimated the overall CO_2 emission rate according to the generated electrical power of the power plant resulting in $1503.0 \, \text{kg} \, \text{s}^{-1}$ during the measurement flight. This corresponds to 123% of the annual mean emission rate of $38.4 \, \text{Mt} \, CO_2$ reported by the power plant to the European Pollutant Release and Transfer Register (E-PRTR) for the year 2018. The description of the derivation of SC, z_0, d_0, and the plume rise as well as the orography data are given in Supplement S4. The parameter values and the origin of the orography data are summarised in Table 4.

ARTM requires radionuclide emission rates in $\text{Bq} \, \text{s}^{-1}$. As tracer ~~in its bounded form as~~ CO_2 with the radioactive isotope ^{14}C is used. (R1; C5) Its decay constant $\lambda = 5730 \pm 40$ years leads to a decay of $5.5 \times 10^{-6} \, \%$ which is negligible within the

Table 4. Input parameters needed by ARTM that are constant during the simulation run.

Parameter	Value	Reference
Stability class	very unstable	(KTA 1508, 2017)
Roughness length	$0.50\,\mathrm{m}$	(TA Luft, 2002)
~~Zero plane~~ Zero-plane displacement	$6 \cdot 0.50\,\mathrm{m}$	(TA Luft, 2002)
Mixing layer height	$1650\,\mathrm{m}$	-
Stack heights	$300\,\mathrm{m}$	~~(?)~~(SkyscraperPage, 2023)
Plume rise (western stack)	$202\,\mathrm{m}$	-
Plume rise (eastern stack)	$74\,\mathrm{m}$	-
Emission rate (western source)	$1002.0\,\mathrm{kg\,s^{-1}}$	~~(?)~~(Brunner et al., 2023)
Emission rate (eastern source)	$501.0\,\mathrm{kg\,s^{-1}}$	~~(?)~~(Brunner et al., 2023)
Orography	SRTM3 data	(Farr et al., 2007)

415   simulation period. Thus, ARTM's internal emission rates in $\mathrm{Bq\,s^{-1}}$ can be used as an equivalent for a mass rate in $\mathrm{kg\,s^{-1}}$ and to convert activity concentration into mixing ratio.

The wind speed ($4.4\,\mathrm{m\,s^{-1}}$) and directions driving the simulation were derived from one flight transect (13:28:03 UTC to 13:33:14 UTC) at a distance of $\approx 13\,\mathrm{km}$ to the west of the power plant at a height of $\approx 600\,\mathrm{m}$ agl. This transect is located close to the middle of the simulation domain and therefore assumed to be representative. The histogram of the wind directions of the
420   transect is shown in Fig. 5. Based on this histogram, two different setups of the model were selected:

i) A single wind direction of $120°$ (mean of the distribution) was selected, assuming that the wind fluctuations are part of the turbulence spectrum and should therefore be represented by the turbulence parametrisation of ARTM.

ii) Two different wind directions were used alternatingly to drive ARTM, a direction of $106°$ (mean of all directions $< 120°$) and a direction of $134°$ (mean of all directions $> 120°$). This assumes that part of the wind variation is due to meso-scale
425   variability that cannot be resolved by ARTM's turbulence scheme.

The first setup was applied for all turbulence models while the second setup was only tested for ARTM2. The hourly sequence of wind inputs for the model is summarised in Supplement S5.

5.3 Horizontal dispersion

The mixing ratio maps simulated with the five turbulence models at a height of $750\,\mathrm{m}$ to $800\,\mathrm{m}$ are shown in Fig. 6 together with
430   the observations between $700\,\mathrm{m}$ to $800\,\mathrm{m}$. We subtracted the background CO_2 mixing ratio of $401\,\mathrm{ppmv}$ from the observation to make them comparable to the simulation results.

The simulated and observed mixing ratios of the plumes are in the same order of magnitude. The simulated plumes show the mean wind direction to be in agreement with the observed one, however, the meandering behaviour of the real plume can be
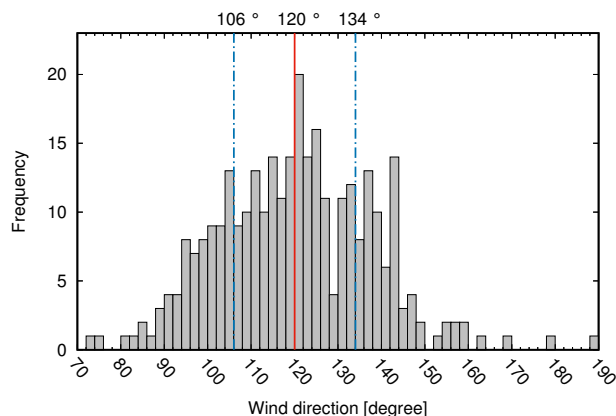
Figure 5. Histogram of the wind directions of the transect chosen for the determination of the wind direction and wind velocity. The transect covers a duration from 13:28:03 UTC to 13:33:14 UTC with the mean position $53.31°$ N, $19.15°$ E. The mean measurement height is $599\,\mathrm{m}$ agl. The mean value of the wind direction is $120°$ (red line). The mean value for the wind directions below $120°$ is $106°$, above $120°$ is $134°$ (blue dashed lines), respectively.

observed at transect 1, 2 and 3 in Fig. 6 revealing that this behaviour is not covered by all turbulence models. The mixing ratio
435 profile in lateral (crosswind) direction simulated by ARTM resembles a Gaussian distribution. This is expected for a constant wind direction and wind speed (Thykier-Nielsen et al., 1999).

The different turbulence models clearly affect the simulated plume widths. The ARTM2 turbulence model simulates the narrowest plume. The ARTM3 model results in a slightly wider plume but compared to the observations both are too narrow. The PRFMOD and DEGRAZIA turbulence models show much broader plumes that cover the observed one to a large extent.
440 The widest plume is simulated by the MODHANNA turbulence model and is in good agreement with the observed plume width. The width of the plumes of the turbulence models is mainly attributed to the horizontal wind speed fluctuations and Lagrangian correlation times displayed in Fig. 2. The highest values for σ_y and $T_{\mathrm{L}y}$ are simulated by MODHANNA, PRFMOD and DEGRAZIA followed by ARTM3 and ARTM2 in the upper half of the PBL in agreement with the simulated plume width.

Fig. 7 shows the simulated and observed plumes of the different turbulence models together with the flight height agl along
445 the flight path. Transects 1, 2 and 3 are shaded grey. Data above the simulated boundary layer top is excluded from the figures. In Fig. 7, the simulated maximum CO_2 mixing ratios of ARTM2 are at all transects larger compared to the observations. Within the simulated boundary layer this deviation reaches 300% at 14:08 UTC and is attributed to the too narrow simulated plume. With increasing plume width of the different turbulence models the maximum mixing ratios decrease (see Figs. 7 a – e). The turbulence models ARTM3, PRFMOD, MODHANNA and DEGRAZIA simulate mixing ratio peaks similar to or below
450 the observed values. Due to dispersion the mixing ratio maximums decrease with increasing distance from the source for all models in agreement with the observation. It is important to point out that simulated mixing ratio values are highly dependent on the emission rates.
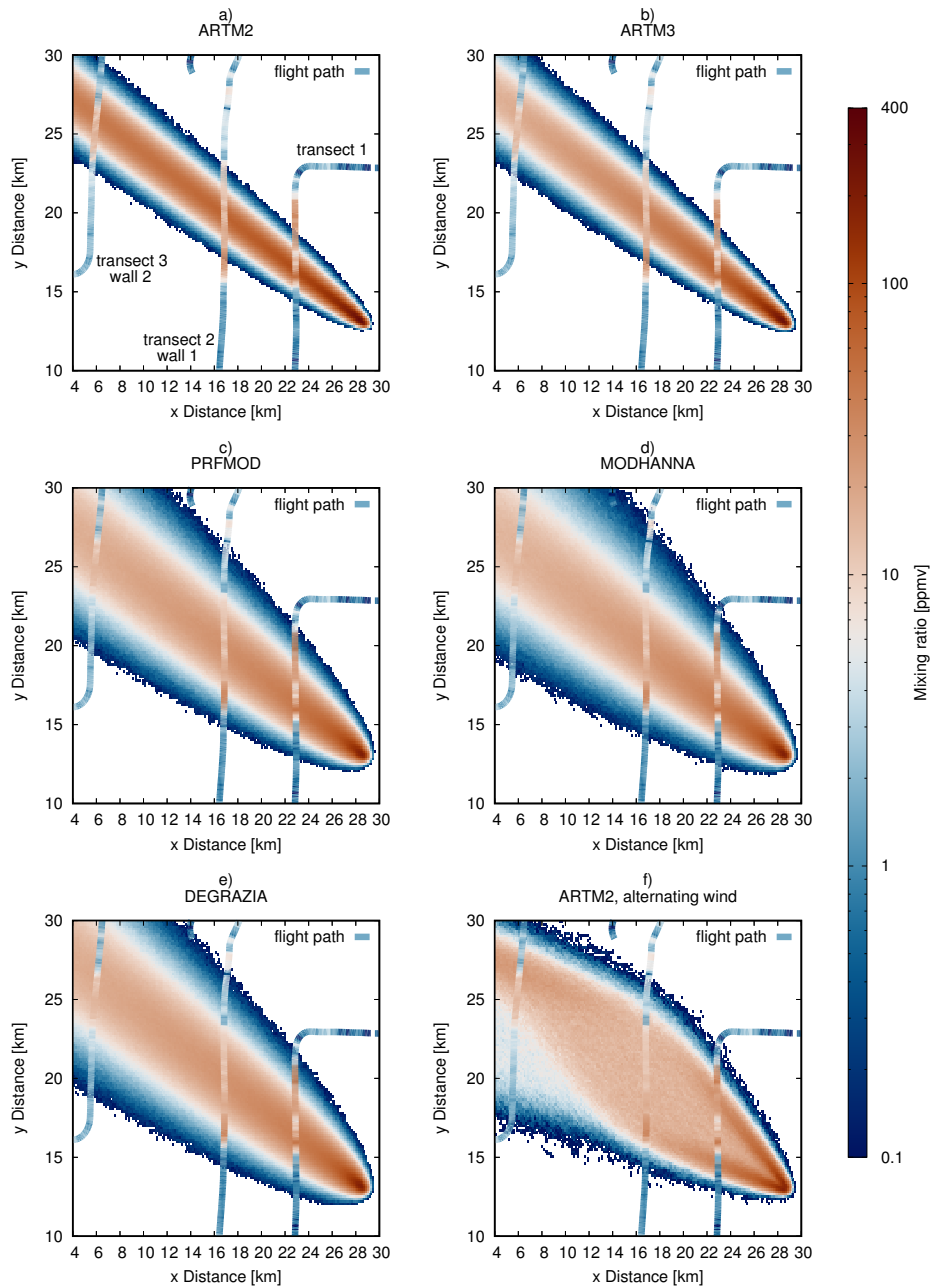
Figure 6. Modelled CO_2 mixing ratio for the case of one wind direction in a) ARTM2, b) ARTM3, c) PRFMOD, d) MODHANNA and e) DEGRAZIA; and two wind directions in f) ARTM2. The wind directions and speeds are given in Table S3; the input parameters in Table 4. Mixing ratios of the simulated plume (averaged over the simulation time) at heights between $750\,m$ to $800\,m$ and the in situ data along the flight path between $700\,m$ to $800\,m$ is shown in logarithmic scale in ppmv. The two wind direction case in f) shows the mean CO_2 mixing ratio of two subsequent hours for the duration of the measurement flight from 13:00 UTC to 15:00 UTC. The background CO_2 mixing ratio of $401\,ppmv$ is subtracted from the observation.
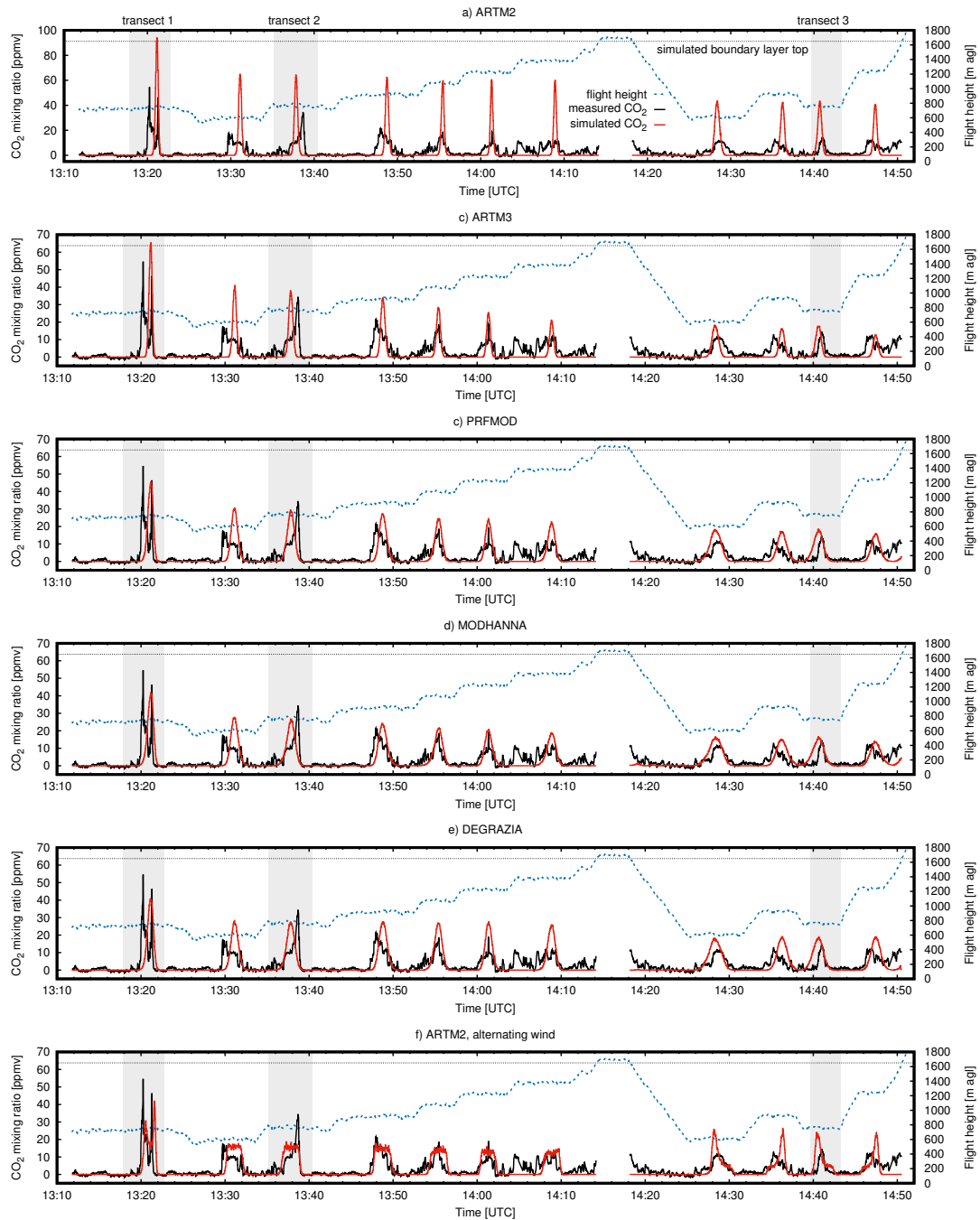
Figure 7. Simulated (red line) and measured (black line) CO_2 data along the flight path together with the flight height (blue dotted line) within the simulation domain. a) ARTM2, b) ARTM3, c) PRFMOD, d) MODHANNA, e) DEGRAZIA turbulence models and f) ARTM2 turbulence model with two alternating wind directions. The transects shown in Fig. 6 are shaded grey.

The simulation gives one hour averages of the exhaust plume, which is expected as the mean of several realisations of meandering plumes. It is not expected that simulated values are much larger than the observed ones but can occur if the width of the simulated plume or the mixing layer depth are underestimated or the emission rate is overestimated.

An alternative to model the meandering behaviour via the turbulence is the usage of alternating wind directions for subsequent simulation hours for the ARTM2 turbulence model to explicitly simulate the meandering plume (Figs. 6 f and 7 f). Simulation results from subsequent hours are combined by calculating the average concentrations. The wind direction derivation is explained in Sec. 5.2. This method generates the widest plume covering the observations and mimics the structure of two maxima at transect 1. However, these two observed maxima originate from snap-shots of the meandering plume and is not expected to be reproduced by the time-averaged simulation. Moreover, physically unrealistic plateaus of mixing ratios are simulated in wall 1 and a single narrow mixing ratio peak in wall 2 which is a result of the alternating wind directions. Mixing ratio maps of simulations and observations at other selected heights are given in Supplement S6.

5.4 Vertical dispersion

For the analysis of the vertical plume behaviour, the cross sections of the simulated plumes at wall 1 are presented in Fig. 8. The narrowest simulated plume is obtained by the ARTM2 model and underestimates the width of the observed plume at heights from $600\,\mathrm{m}$ to $1400\,\mathrm{m}$ agl. The plume of ARTM3 model is slightly wider throughout the PBL. In both, the ARTM3 and PRFMOD model, the values of σ_y (T_{Ly}) decrease (increase) with height, respectively (see Fig.2). While these opposing trends cancel out each other for the ARTM3 model they lead to a slightly increase of lateral dispersion with height for the PRFMOD model. The vertical profiles of σ_y and T_{Ly} of the MODHANNA model shown in Figs. 2 c) and d) appear to lead to a slightly increasing dispersion, too. The DEGRAZIA model in contrast, shows a constant behaviour for both, σ_y and T_{Ly}. Below $200\,\mathrm{m}$ the width of all five simulated plumes decrease towards the surface.

All turbulence models show a slight decrease of the mixing ratio with increasing height at a constant distance from the source (see Fig. 7), which agrees with observations. From the cross sections at wall 1 (Fig. 8) the average horizontal mixing ratio profiles are derived and shown in Fig 9. Except for the DEGRAZIA model, the decreasing mixing ratio with increasing height above $600\,\mathrm{m}$ can be recognised here as well. However, in Fig. 7 the highest maximum mixing ratios at wall 1 ~~occur at different transects for the simulations and the observation. This indicates that the effective source height of the simulations was assumed too low compared to the~~ (13:25 UTC - 14:10 UTC) occur at transect 2 for the measurements. The simulations instead show very similar peaks at the two lowest transects in wall 1 having the highest concentration values. (R2; C18)

In contrast to the Gaussian lateral mixing ratio distribution of the plume of Fig. 8 a), the ARTM2 turbulence model with two alternating wind directions (Fig. 8 f) shows the uniform mixing ratio distribution in the plume's core region (mixing ratio > $1\,\mathrm{ppmv}$) as it was already shown in Fig. 7 f.

The cross sections of wall 2 in Fig. ~~S11~~ S12 of the Supplement show a similar behaviour of the plumes. The measured data show a large variation of the plume width on the different transects emphasising the meandering and turbulent character of the real plume. Furthermore, it can be recognised that the real plume is not entirely recorded; the transects are too short at this wall.
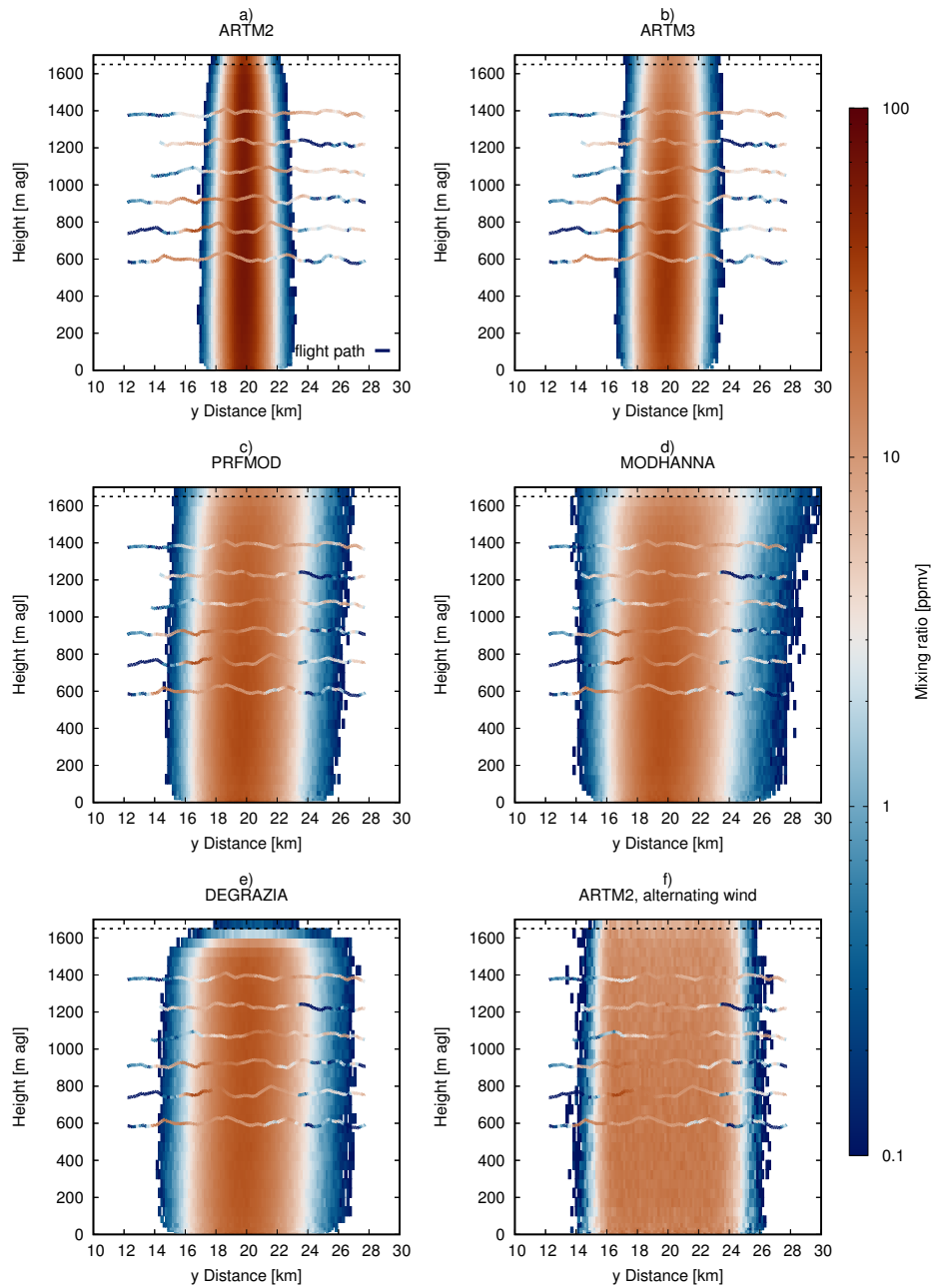
Figure 8. Cross section of the simulated plumes at wall 1 of the observations for the different turbulence models a) ARTM2, b) ARTM3, c)PRFMOD, d)MODHANNA and e) DEGRAZIA. f) the two wind directions case for ARTM2. The x-axis "y Distance" is in south-north orientation. The dashed line at 1650 m agl marks the simulated mixing layer top.
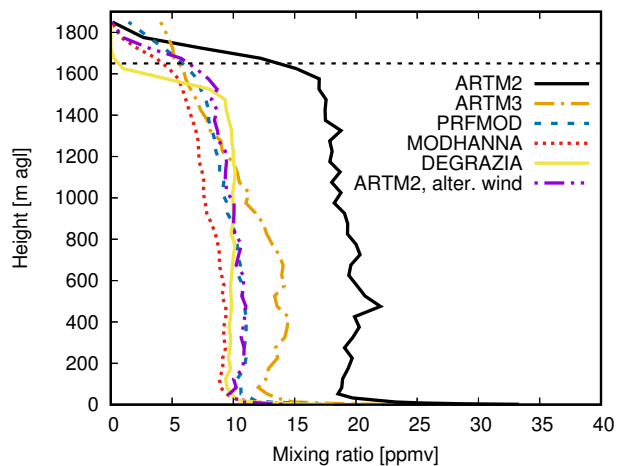
Figure 9. Profile of the average horizontal mixing ratio of the six simulation cases at wall 1 (see Fig. 8). The dashed line at $1650\,\text{m}$ agl marks the simulated mixing layer top.

5.5 Validation and uncertainty evaluation

In order to quantify the simulations~~uncertainty~~' uncertainty (R1; C6), we investigate the deviations of the simulated and the observed CO_2 mixing ratios in the plume by probability distributions (PDs), comparisons of integrated plume mass and point-
490   to-point mixing ratio comparisons.

The deviation of model results and measurements in a plume can be accessed by the comparison of the PDs and the cumulative probability distributions (CDPs) of simulated and observed CO_2 mixing ratios in the plume. The PDs of simulation and measurement are normalised to the maximum mixing ratio of the measurements with the integrals of simulated and measured distributions being equal. To get rid of the mixing ratio fluctuation of the excess mixing ratio of the measurement, mixing ratio
495   values below $1\,\text{ppmv}$ are not taken into account. The PDs and CPDs of the five different turbulence models and the observation for all transects below the simulated boundary layer top are given in Fig. 10. The PDs of the simulated and measured plume show the occurrence of mixing ratio values relative to the maximum mixing ratio of the measurements. There is an overestimation of simulated maximum mixing ratios for the ARTM2 and ARTM3 turbulence model. The high number of data points at approx. 20% of the maximum mixing ratio of the measurements is due to the fine structure, shoulders beside peaks and broad
500   indistinct peaks of the plume not represented in the simulations. MODHANNA can be identified as the turbulence model that shows the best agreement with the observations concerning the PDs and CPDs, i.e. the occurrence of the mixing ratio values is most similar to those of the measurement. To quantify the similarity and to decide whether simulations and measurements are significantly different three statistical test were applied: The Z-test, the Kolmogorov-Smirnov (KS) test and the Cramér-von Mises (CvM) test (Conover, 1980; Wilks, 2006; University of Oregon, 2020). The Z statistic represents the distance between
505   the means of two PDs normalised to the standard error. Statistics below 2 indicate no significant difference between the distributions. Additional interpretation limits are given in Supplement S7. The KS statistic represents the supremum of the distance
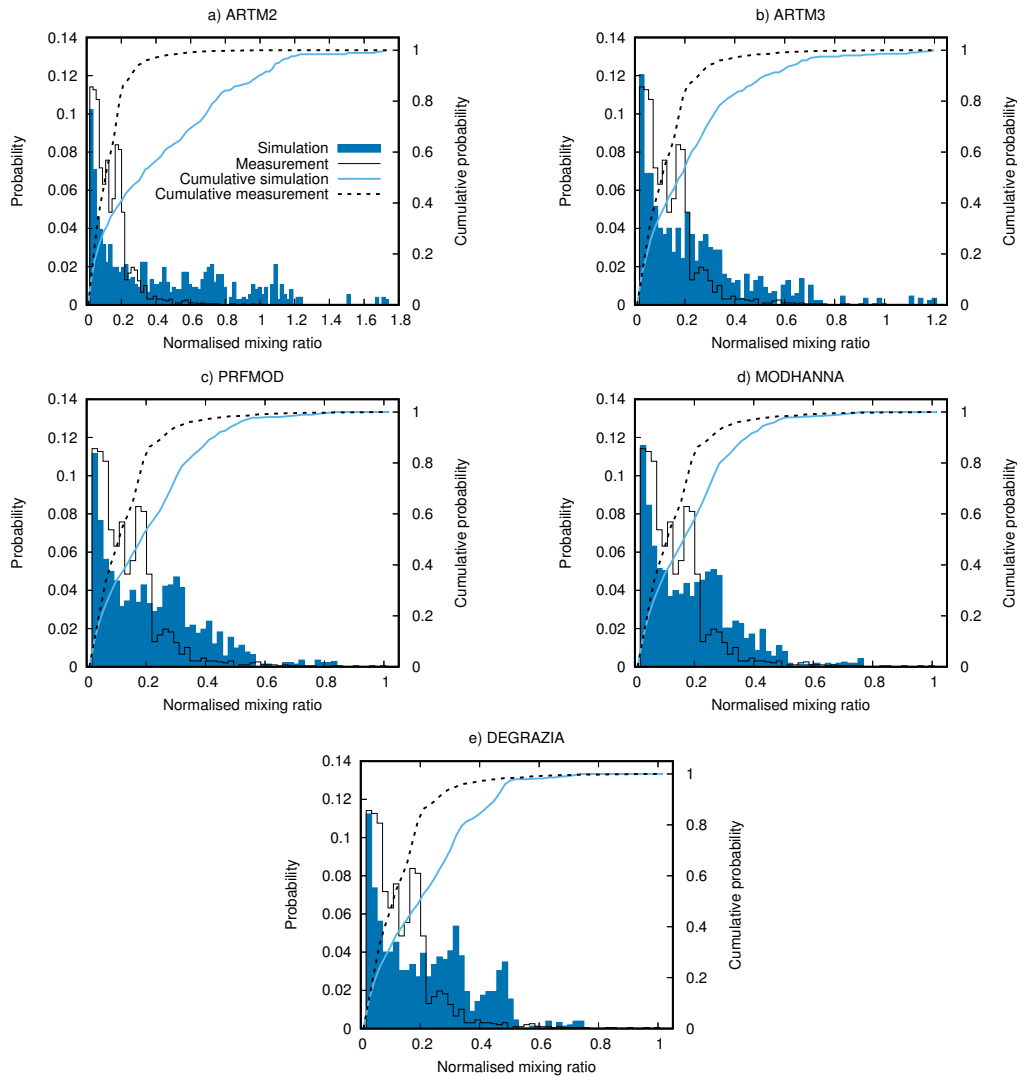
Figure 10. Probability distribution (bars) and cumulative probability distribution (lines) of simulated and measured mixing ratios of the five turbulence models. The PDs are normalised according to the maximum mixing ratio of the measurements and the integral of simulated and measured PDs are equal. Mixing ratio values below $1\,\mathrm{ppmv}$ are not considered in the PDs and CPDs.

between two CPDs while the CvM statistic is proportional to the integral of the distances between two CPDs. For both we assumed a significance level of 0.05. The statistics and their p-values (in brackets) are summarised in Table 5. The three statistical tests show that all simulated mixing ratio distributions differ significantly from the observed one. Nevertheless, the statistics

510   can be used to rank the models. MODHANNA shows the best agreement with the observations i.e. the distribution of mixing ratio values in the transects is most similar to those of the observations compared to the other turbulence models. The statistical

Table 5. Z statistics, Kolmogorov-Smirnov (KS) statistics and Cramér-von Mises (CvM) statistics of the mixing ratio distributions of the five turbulence models. The p-values are given in brackets, respectively. The significance level is 0.05.

Turbulence model	Z statistic	KS statistic	CvM statistic
ARTM2	16.8	0.45 (0.00)	25.1 (0.00)
ARTM3	11.5	0.23 (0.00)	12.8 (0.00)
PRFMOD	13.5	0.24 (0.00)	15.5 (0.00)
MODHANNA	10.2	0.20 (0.00)	9.8 (0.00)
DEGRAZIA	15.5	0.27 (0.00)	19.8 (0.00)

tests rank ARTM3 second but this may be biased by the statistical tests being very sensitive to deviations in the regions of the PDs with high numbers of low mixing ratio values. We want to point out that the results do not mean that the MODHANNA model produces mixing ratio peaks that are structured like the observed ones but the relative occurrence of mixing ratio values is most similar.

To compare the simulation results, the integral of the mixing ratio values along the flight path below the simulated boundary layer top (see Fig. 7) within the plume is shown in Table 6. We used the method above to get rid of the baseline fluctuations of the excess mixing ratios to calculate the integrals. This procedure is also applied to the simulations. Except for ARTM3, there is a good agreement between the modelled and the measured data: The deviation is less than 13%. Concerning ARTM3, there is a strong vertical gradient in the simulated mixing ratios of the plume above $700\,\mathrm{m}$ as it is illustrated in the Figs. 8 b and 9. Tracers are stronger diluted (accumulated) in the upper (lower) half of the PBL than for other turbulence models. This corresponds to the findings of Sec. 4. Since the flight path is mainly located in the upper half of the PBL, the integral along the flight path results in a lower value for ARTM3. The higher mixing ratios in the lower half of the PBL might become important

Table 6. Integrals of mixing ratio values (values below 1 ppmv are not considered) along the flight path for simulations A_{sim} and observations A_{obs} within the simulated PBL (see Fig. 7) given in $\mathrm{ppmv}\cdot\mathrm{km}$ as well as their ratio.

Turbulence model	A_{obs}	A_{sim}	$A_{\mathrm{sim}}A_{\mathrm{obs}}^{-1}$
ARTM2	1094	1186	108.4%
ARTM3	1094	742	67.8%
PRFMOD	1094	1194	109.2%
MODHANNA	1094	1114	101.8%
DEGRAZIA	1094	1236	112.9%

when simulations are used for radiation exposure assessment. The results suggest, that the original assumption of the emission

525 rate may not deviate much from the actual value. However, observations below $600\,\mathrm{m}$ are necessary to get a more complete

comparison of simulated and actual plume.

The deviation between simulations of the five turbulence models and observation at a specific position can be assessed using

density scatter-plots as given in Fig. 11. All mixing ratio values larger than $1\,\mathrm{ppm}$ along the flight path below the simulated

boundary layer top are considered. As a guide to the eye the regression with slope $m = 1$ is given as a dashed line and represents
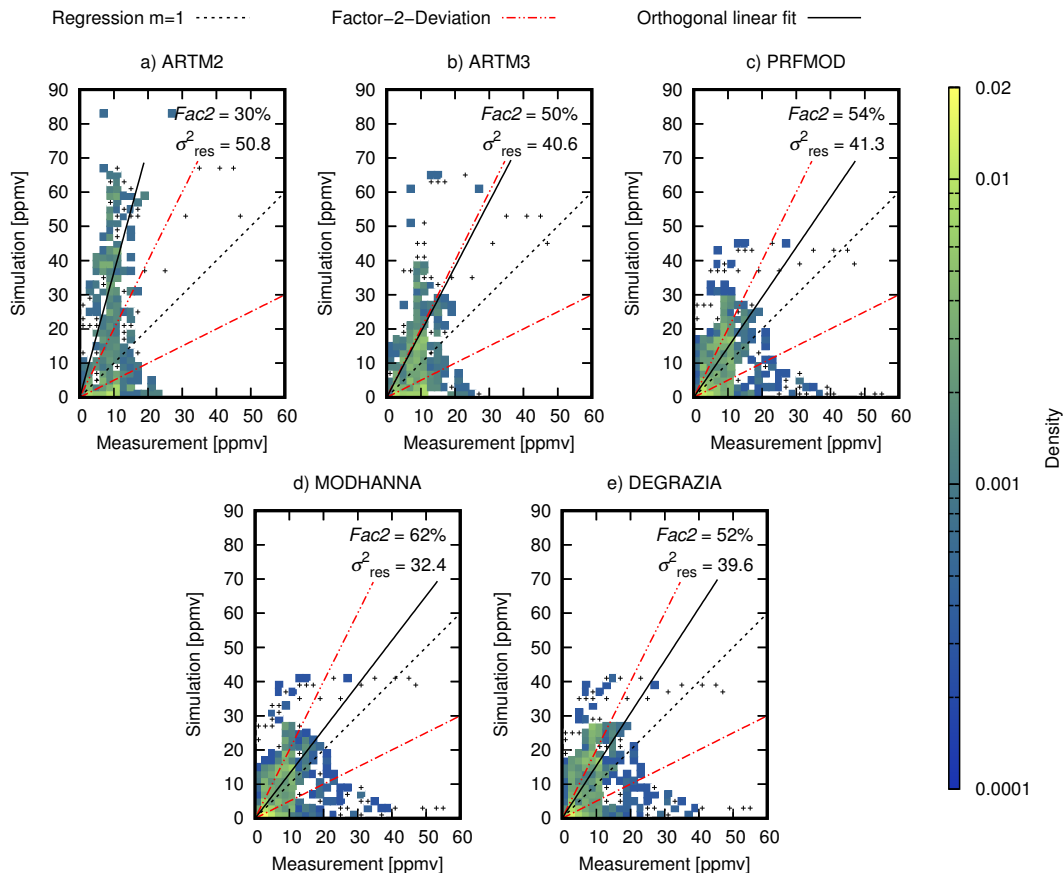
Figure 11. Density scatter-plots of the simulated mixing ratios of the five turbulence models ARTM2 (a), ARTM3 (b), PRFMOD (c),

MODHANNA (d) and DEGRAZIA (e) against the observations. Single data points in a bin are indicated with $(+)$, more data points in a bin

are colour-coded. The regression with slope $m = 1$ (dotted black line) represents the identity of simulation with measurement, the dotted

dashed red line represents a deviation from the regression with $m = 1$ by the factor two and the solid black line represents the orthogonal

linear fit to the data points. $Fac2$ gives the percentage of data points with deviations of not more than a factor two from the regression $m = 1$.

The residual variance of the orthogonal fit is given by σ_{res}^2.

530 the equality of simulated and observed mixing ratios. The deviation from this equality by the factor two is confined by the red

dashed dotted lines. It is not expected to find a lot of data points at the regression $m = 1$ due to the fundamental differences of

the data set properties of simulation and observation. However, a large amount of data points within a deviation of a factor two

decreases the uncertainty. The percentage of data points within these boarders is represented as $Fac2$ given in Fig. 11. Low $Fac2$ values can also be explained by the large number of measurement data points outside the simulated plume because they are too

535 narrow. The smallest $Fac2$ is derived for the ARTM2 model. This coincide with the unbalanced distribution of the data points around the regression $m = 1$ with the simulation overestimating the observed mixing ratios and simultaneously simulating too narrow plumes. This is represented by the orthogonal regression of data points (black line) given in the figure with a slope above three. The residual variance σ_{res}^2 quantifies the scattering of data points. The large value for ARTM2 indicates a less compact data point distribution. ARTM3 shows a more balanced and compacter data but still distinctly overestimates mixing ratios

540 and underestimates plume widths. PRFMOD, MODHANNA and DEGRAZIA show similar properties with $Fac2 > 50\,\%$, compacter, well balanced data and less overestimated mixing ratios and underestimated plume widths, with the MODHANNA model performed slightly best for the given measurement and turbulence conditions.

The uncertainty of the CO_2 measurement device of $\pm 0.15\,\mathrm{ppmv}$ is at least one order of magnitude smaller than the measured enhanced CO_2 concentrations. Thus, the measurement uncertainty has only a minor impact on the evaluation results. (R2; GC3)

545

6   Conclusions

In this work we presented an extensive evaluation of ARTM with three different elements, a sensitivity analysis, an analysis of turbulence models and a comparison with aircraft observations. Based on the sensitivity analysis, we identified the stability class to be the most important input parameter followed by the roughness length, the source height and the displacement height

550 factor. Therefore, special care has to be taken to determine the stability class for a simulation because uncertainties of this parameter cause large uncertainties in model results. This emphasises the general disadvantage of the rather coarse stability class concept being used to describe atmospheric turbulence. A finer classification or a continuous parameter such as the Obukhov length could be a better option but would generally require detailed measurements of turbulence parameters such as friction velocity and sensible heat flux.

555 In addition to the three turbulence models already implemented in ARTM 3.0.0, two further turbulence models, the MOD-HANNA model and the DEGRAZIA model, were implemented and tested. Evaluation of the models by applying the well-mixed condition test showed that the ARTM2, PRFMOD and MODHANNA models produced a moderate ~~segregation~~ deviation of an initially uniform concentration profile ~~in the~~ under unstable (R2; GC4) atmospheric boundary layer conditions. Underestimations from the uniform concentration occur primarily at the ground and at the top of the boundary layer with up to $10\,\%$

560 while overestimations occur in between with up to $7\,\%$ at one third of the PBL. ~~The~~ Under the same conditions the (R2; GC4) ARTM3 model produces the strongest deviations up to four times higher than the other models. However, near the ground ARTM3 performs best. The DEGRAZIA model showed a less inhomogeneous profile with deviations from the uniform concentration of $5\,\%$ or less below $z = 0.8 h_{\mathrm{m}}$. The discrepancies under 6% below 80% of the boundary layer height show good mixing properties of the unstable (R2; GC4) planetary boundary layer for the ARTM2, PRFMOD, MODHANNA and DE-

565   GRAZIA model. ~~The mixing properties of the ARTM3 model may bias simulation results when handling with γ-cloud-shine or wet deposition.~~ (R2; C19)

    Three dimensional airborne in situ observational data measuring a power plant emission plume were compared to ARTM simulation results. The time resolution of ARTM results is one hour, which is larger than the expected time scale of the observed meandering plume and therefore, ARTM is expected to capture the time integrated real plume and not the fine structures on
570 small scales. ARTM simulated the mean wind in agreement with the observations throughout the simulation domain. The different turbulence models simulate plume mixing ratios in the same order of magnitude as the measurements although the exact mixing ratio values depend on the emission rate. ARTM2 underestimates the plume spread under very unstable conditions and overestimates the maximum mixing ratio by a factor of two or more. The ARTM3 model produces only slightly wider plumes in lateral direction but lower maximum mixing ratio values at the upper half of the PBL. This is attributed to the
575 inhomogeneous vertical mixing and the horizontal turbulence parametrisation of the ARTM3 turbulence model under unstable conditions (R2; GC4). The other turbulence models PRFMOD, MODHANNA and DEGRAZIA simulate a wider plume spread in the range of the measurements. Maximum mixing ratios are close to the measurements, the integral of the mixing ratios of simulations and observations along the flight path are comparable. The models were evaluated with measurements at heights larger than $600\,\mathrm{m}$, hence do not cover the heights below. The differences in temporal and local resolution of simulation and
580 measurements lead to differences in the distributions of mixing ratio values. According to Fig. 10, under unstable conditions, (R2; GC4) all turbulence models underestimate the occurrence of mixing ratio values around 20% of the maximum mixing ratio of the measurements which are a result of the fine structure of the plumes. The smallest deviations in PDs and CPDs are found for MODHANNA. Using point-to-point comparisons, the ARTM2 model showed the largest deviations from the measured plume, ARTM3 shows better agreement and PRFMOD, MODHANNA and DEGRAZIA showed comparable good
585 performances, with MODHANNA slightly matches the measurements best.

    With the results of this study we showed that ARTM is able to simulate the extension and mixing ratios of a plume under unstable stratification (R2; GC4) when the proper turbulence model is used. The ARTM3 model showed to be a suitable turbulence model for radiation exposure assessment when conservative long-term simulations are requested. However, the PRFMOD, MODHANNA and DEGRAZIA models simulate the exhaust plume closer to real exhaust plumes under the given
590 conditions and under the limitations of temporal and spatial uncertainties. Within this validation using the in situ data from the Bełchatów power plant, the MODHANNA turbulence model performed slightly best. However, this ranking cannot be generalized to other stability conditions without further comparison studies. (R2; GC2) Further analyses with known emission terms at different atmospheric turbulence properties~~could~~, stable and neutral conditions, are necessary and can (R2; GC4) lead to better validation of ARTM. We encourage the comparison of other turbulence formulations such as the one suggested by
595 Hanna (1982). This would also reveal differences between the model of Hanna and the DEGRAZIA model and extends the analyses given by Carvalho et al. (2002) to larger heights. (R2; GC2 and C20) The collection of measurement data in the upper and lower half of the PBL as well as transects sampling the entire extent of a plume are beneficial. Also, the use of the Obukhov length as a measure for atmospheric stability is encouraged.

Code and data availability. The program ARTM is available on request from the Federal Office for Radiation Protection (BfS) of Germany via artm@bfs.de. The data from the CoMeT 1.0 campaign is available at https://halo-db.pa.op.dlr.de/mission/94.

Author contributions. Conceptualisation: Robert Hanfland, Margit Pattantyús-Ábrahám, Dominik Brunner, Christiane Voigt; Data Collection: Alina Fiehn, Anke Roiger; Methodology: Robert Hanfland, Margit Pattantyús-Ábrahám, Dominik Brunner, Christiane Voigt; Formal Analysis: Robert Hanfland; Writing - original draft preparation: Robert Hanfland; Writing - review and editing: Margit Pattantyús-Ábrahám, Robert Hanfland, Dominik Brunner, Christiane Voigt, Alina Fiehn, Anke Roiger; Supervision: Margit Pattantyús-Ábrahám, Dominik Brunner, Christiane Voigt

Competing interests. The authors have no competing interests to declare that are relevant to the content of this article.

Disclaimer. TEXT

Acknowledgements. We thank the BfS for funding and Christopher Strobl (BfS) for his support. Furthermore, the authors would like to thank Stephan Henne (Empa) for helpful discussions.

References

Andersen, T., de Vries, M., Necki, J., Swolkien, J., Menoud, M., Röckmann, T., Roiger, A., Fix, A., Peters, W., and Chen, H.: Local to regional methane emissions from the Upper Silesia Coal Basin (USCB) quantified using UAV-based atmospheric measurements, Atmos. Chem. Phys. Discuss., 2022, 1–35, https://doi.org/10.5194/acp-2021-1061, 2022.

Arnold, D., Maurer, C., Wotawa, G., Draxler, R., Saito, K., and Seibert, P.: Influence of the meteorological input on the atmospheric transport modelling with FLEXPART of radionuclides from the Fukushima Daiichi nuclear accident, J. Environ. Radioactiv., 139, 212–225, https://doi.org/10.1016/j.jenvrad.2014.02.013, 2015.

Bahlali, M. L., Henry, C., and Carissimo, B.: On the Well-Mixed Condition and Consistency Issues in Hybrid Eulerian/Lagrangian Stochastic Models of Dispersion, Bound-Lay. Meteorol., 174, 275–296, https://doi.org/10.1007/s10546-019-00486-9, 2020.

Berchet, A., Zink, K., Oettl, D., Brunner, J., Emmenegger, L., and Brunner, D.: Evaluation of high-resolution GRAMM–GRAL (v15.12/v14.8) NOx simulations over the city of Zürich, Switzerland, Geosci. Model Dev., 10, 3441–3459, https://doi.org/10.5194/gmd-10-3441-2017, 2017.

Borgonovo, E.: A new uncertainty importance measure, Reliab. Eng. Syst. Safe., 92, 771–784, https://doi.org/10.1016/j.ress.2006.04.015, 2007.

Brunner, D., Kuhlmann, G., Henne, S., Koene, E., Kern, B., Wolff, S., Voigt, C., Jöckel, P., Kiemle, C., Roiger, A., Fiehn, A., Krautwurst, S., Gerilowski, K., Bovensmann, H., Borchardt, J., Galkowski, M., Gerbig, C., Marshall, J., Klonecki, A., Prunet, P., Hanfland, R., Pattantyús-Ábrahám, M., Wyszogrodzki, A., and Fix, A.: Evaluation of simulated CO2 power plant plumes from six high-resolution atmospheric transport models, Atmos. Chem. Phys., 23, 2699–2728, https://doi.org/10.5194/acp-23-2699-2023, 2023.

Carvalho, J. C., Anfossi, D., Trini Castelli, S., and Degrazia, G. A.: Application of a model system for the study of transport and diffusion in complex terrain to the TRACT experiment, Atmos. Environ., 36, 1147–1161, https://doi.org/10.1016/S1352-2310(01)00559-3, 2002.

Chino, M., Nakayama, H., Nagai, H., Terada, H., Katata, G., and Yamazawa, H.: Preliminary Estimation of Release Amounts of 131I and 137Cs Accidentally Discharged from the Fukushima Daiichi Nuclear Power Plant into the Atmosphere, J. Nucl. Sci. and Technol., 48, 1129–1134, https://doi.org/10.1080/18811248.2011.9711799, 2011.

Connan, O., Smith, K., Organo, C., Solier, L., Maro, D., and Hébert, D.: Comparison of RIMPUFF, HYSPLIT, ADMS atmospheric dispersion model outputs, using emergency response procedures, with 85Kr measurements made in the vicinity of nuclear reprocessing plant, J. Environ. Radioactiv., 124, 266–277, https://doi.org/10.1016/j.jenvrad.2013.06.004, 2013.

Conover, W. J.: Practical Nonparametric Statistics, Wiley & Sons, New York, 2 edn., 1980.

Cox, R. M., Sontowski, J., and Dougherty, C. M.: An evaluation of three diagnostic wind models (CALMET, MCSCIPUF, and SWIFT) with wind data from the Dipole Pride 26 field experiments, Meteorol. Appl., 12, 329–341, https://doi.org/10.1017/S1350482705001908, 2005.

Degrazia, G. A., Anfossi, D., Carvalho, J. C., Mangia, C., Tirabassi, T., and Campos Velho, H. F.: Turbulence parameterisation for PBL dispersion models in all stability conditions, Atmos. Environ., 34, 3575–3583, https://doi.org/10.1016/S1352-2310(00)00116-3, 2000.

Draxler, R., Arnold, D., Chino, M., Galmarini, S., Hort, M., Jones, A., Leadbetter, S., Malo, A., Maurer, C., Rolph, G., Saito, K., Servranckx, R., Shimbori, T., Solazzo, E., and Wotawa, G.: World Meteorological Organization's model simulations of the radionuclide dispersion and deposition from the Fukushima Daiichi nuclear power plant accident, J. Environ. Radioactiv., 139, 172–184, https://doi.org/10.1016/j.jenvrad.2013.09.014, 2015.

Draxler, R. R. and Hess, G. D.: An overview of the HYSPLIT_4 modelling system for trajectories, dispersion and deposition, Aust. Meteorol. Mag., 47, 295–308, 1998.

Eichhorn, J. and Kniffka, A.: The numerical flow model MISKAM: State of development and evaluation of the basic version, Meteorologische Zeitschrift, 19, 81–90, https://doi.org/10.1127/0941-2948/2010/0425, 2010.

Farchi, A., Bocquet, M., Roustan, Y., Mathieu, A., and Quérel, A.: Using the Wasserstein distance to compare fields of pollutants: application to the radionuclide atmospheric dispersion of the Fukushima-Daiichi accident, Tellus B, 68, 31 682, https://doi.org/10.3402/tellusb.v68.31682, 2016.

Farr, T. G., Rosen, P. A., Caro, E., Crippen, R., Duren, R., Hensley, S., Kobrick, M., Paller, M., Rodriguez, E., Roth, L., Seal, D., Shaffer, S., Shimada, J., Umland, J., Werner, M., Oskin, M., Burbank, D., and Alsdorf, D.: The Shuttle Radar Topography Mission, Rev. Geophys., 45, https://doi.org/10.1029/2005RG000183, 2007.

Fiehn, A., Kostinek, J., Eckl, M., Klausner, T., Gałkowski, M., Chen, J., Gerbig, C., Röckmann, T., Maazallahi, H., Schmidt, M., Korbeń, P., Nęcki, J., Jagoda, P., Wildmann, N., Mallaun, C., Bun, R., Nickl, A. L., Jöckel, P., Fix, A., and Roiger, A.: Estimating CH4, CO2 and CO emissions from coal mining and industrial activities in the Upper Silesian Coal Basin using an aircraft-based mass balance approach, Atmos. Chem. Phys., 20, 12 675–12 695, https://doi.org/10.5194/acp-20-12675-2020, 2020.

Fix, A., Amediek, A., Bovensmann, H., Ehret, G., Gerbig, C., Gerilowski, K., Pfeilsticker, K., Roiger, A., and Zöger, M.: CoMet: An Airborne Mission to Simultaneously Measure CO2 and CH4 Using Lidar, Passive Remote Sensing, and In-Situ Techniques, EPJ Web Conf., 176, 02 003, https://doi.org/10.1051/epjconf/201817602003, 2018.

Frey, H. C. and Patil, S. R.: Identification and Review of Sensitivity Analysis Methods, Risk Anal., 22, 553–578, https://doi.org/10.1111/0272-4332.00039, 2002.

Gariazzo, C., Papaleo, V., Pelliccioni, A., Calori, G., Radice, P., and Tinarelli, G.: Application of a Lagrangian particle model to assess the impact of harbour, industrial and urban activities on air quality in the Taranto area, Italy, Atmos. Environ., 41, 6432–6444, https://doi.org/10.1016/j.atmosenv.2007.06.005, 2007.

Gałkowski, M., Jordan, A., Rothe, M., Marshall, J., Koch, F. T., Chen, J., Agusti-Panareda, A., Fix, A., and Gerbig, C.: In situ observations of greenhouse gases over Europe during the CoMet 1.0 campaign aboard the HALO aircraft, Atmos. Meas. Tech., 14, 1525–1544, https://doi.org/10.5194/amt-14-1525-2021, 2021.

GRS: ARTM Atmospheric Radionuclide-Transport-Model, source Code, 2015.

Gryning, S. E., Holtslag, A. A. M., Irwin, J. S., and Sivertsen, B.: Applied Dispersion Modelling Based on Meteorological Scaling Parameters, Atmos. Environ. (1967), 21, 79–89, https://doi.org/10.1016/0004-6981(87)90273-3, 1987.

Hamby, D. M.: A Review of Techniques for Parameter Sensitivity Analysis of Environmental Models, Environ. Monit. Assess., 32, 135–154, https://doi.org/10.1007/BF00547132, 1994.

Hamby, D. M.: A Comparison of Sensitivity Analysis Techniques, Health Phys., 68, 195–204, https://journals.lww.com/health-physics/Fulltext/1995/02000/A_Comparison_of_Sensitivity_Analysis_Techniques.5.aspx, 1995.

Hanfland, R., Pattantyús-Ábrahám, M., Richter, C., Brunner, D., and Voigt, C.: The Lagrangian Atmospheric Radionuclide Transport Model (ARTM) - development, description and sensitivity analysis, Air Qual. Atmos. Health., https://doi.org/10.1007/s11869-022-01188-x, 2022.

Hanna, S. R.: Applications in Air Pollution Modeling, pp. 275–310, Springer Netherlands, Dordrecht, https://doi.org/10.1007/978-94-010-9112-1_7, 1982.

Herman, J. and Usher, W.: SALib: An open-source Python library for Sensitivity Analysis, Journal of Open Source Software, 2, 97, https://doi.org/10.21105/joss.00097, 2017.

Herman, J. and Usher, W.: SALib: Basics, https://salib.readthedocs.io/en/latest/user_guide/basics.html, 2021.

685 Hettrich, S.: Validation and Verification of the Atmospheric Radionuclide Transport Model (ARTM), Thesis, 2017.

Hicks, B. B.: Behavior of Turbulence Statistics in the Convective Boundary Layer, J. Appl. Meteorol. Clim., 24, 607–614, https://doi.org/10.1175/1520-0450(1985)024<0607:BOTSIT>2.0.CO;2, 1985.

Hoffman, F. O. and Gardner, R. H.: Evaluation of Uncertainties in Environmental Radiological Assessment Models, book section 11, U.S. Nuclear Regulatory Commission, Washington, https://www.nrc.gov/docs/ML0917/ML091770419.pdf, 1983.

690 Iman, R. L. and Helton, J. C.: An Investigation of Uncertainty and Sensitivity Analysis Techniques for Computer Models, Risk Anal., 8, 71–90, https://doi.org/10.1111/j.1539-6924.1988.tb01155.x, 1988.

Janicke, U. and Janicke, L.: Some aspects of the definition of meteorological boundary layer profiles and comparisons with measurements, Report, Janicke Consulting,, https://www.janicke.de/data/bzu/bzu-007-01.pdf, 2011.

Joe, S. and Kuo, F. Y.: Constructing Sobol' Sequences with Better Two-Dimensional Projections, SIAM J. Sci. Comput., 30, 2635–2654,
695 https://doi.org/10.1137/070709359, 2008.

Katharopoulos, I., Brunner, D., Emmenegger, L., Leuenberger, M., and Henne, S.: Lagrangian Particle Dispersion Models in the Grey-Zone of Turbulence: Adaptions to FLEXPART-COSMO for Simulations at 1 km Grid Resolution, Bound-Lay. Meteorol., https://doi.org/10.1007/s10546-022-00728-3, 2022.

Kerschgens, M. J., Nölle, C., and Martens, R.: Comments on turbulence parameters for the calculation of dispersion in the atmospheric
700 boundary layer, Meteorologische Zeitschrift, 9, 155–163, https://doi.org/10.1127/metz/9/2000/155, 2000.

Klausner, T., Mertens, M., Huntrieser, H., Gałkowski, M., Kuhlmann, G., Baumann, R., Fiehn, A., Jöckel, P., Pühl, M., and Roiger, A.: Urban greenhouse gas emissions from the Berlin area: A case study using airborne CO2 and CH4 in situ observations in summer 2018, Elementa Science of the Anthropocene, 8, https://doi.org/10.1525/elementa.411, 2020.

Kleijnen, J. P. C.: Verification and validation of simulation models, Eur. J. Oper. Res., 82, 145–162, https://doi.org/10.1016/0377-
705 2217(94)00016-6, 1995.

Kostinek, J., Roiger, A., Eckl, M., Fiehn, A., Luther, A., Wildmann, N., Klausner, T., Fix, A., Knote, C., Stohl, A., and Butz, A.: Estimating Upper Silesian coal mine methane emissions from airborne in situ observations and dispersion modeling, Atmos. Chem. Phys., 21, 8791–8807, https://doi.org/10.5194/acp-21-8791-2021, 2021.

Krautwurst, S., Gerilowski, K., Borchardt, J., Wildmann, N., Gałkowski, M., Swolkień, J., Marshall, J., Fiehn, A., Roiger, A., Ruhtz, T.,
710 Gerbig, C., Necki, J., Burrows, J. P., Fix, A., and Bovensmann, H.: Quantification of CH4 coal mining emissions in Upper Silesia by passive airborne remote sensing observations with the Methane Airborne MAPper (MAMAP) instrument during the CO2 and Methane (CoMet) campaign, Atmos. Chem. Phys., 21, 17 345–17 371, https://doi.org/10.5194/acp-21-17345-2021, 2021.

KTA 1508: Instrumentation for Determining the Dispersion of Radioactive Substances in the Atmosphere, Report, Nuclear Safety Standards Commission (KTA), 2017.

715 Lang, N., Jetz, W., Schindler, K., and Wegner, J. D.: A high-resolution canopy height model of the Earth, arXiv preprint, https://doi.org/10.48550/arXiv.2204.08322, 2022.

Lenschow, D. H., Wyngaard, J. C., and Pennell, W. T.: Mean-Field and Second-Moment Budgets in a Baroclinic, Convective Boundary Layer, Journal of Atmospheric Sciences, 37, 1313–1326, https://doi.org/10.1175/1520-0469(1980)037<1313:Mfasmb>2.0.Co;2, 1980.

Lin, J. and Gerbig, C.: How can we satisfy the well-mixed criterion in highly inhomogeneous flows? A practical approach, pp. 59–70,
720 Geophysical Monograph Series, American Geophysical Union, Washington, DC, https://doi.org/10.1029/2012GM001232, 2013.

Lin, J. C., Gerbig, C., Wofsy, S. C., Andrews, A. E., Daube, B. C., Davis, K. J., and Grainger, C. A.: A near-field tool for simulating the upstream influence of atmospheric observations: The Stochastic Time-Inverted Lagrangian Transport (STILT) model, J. Geophys. Res-Atmos, 108, https://doi.org/10.1029/2002JD003161, 2003.

Link, K. G., Stobb, M. T., Di Paola, J., Neeves, K. B., Fogelson, A. L., Sindi, S. S., and Leiderman, K.: A local and global sensitivity analysis of a mathematical model of coagulation and platelet deposition under flow, PLOS ONE, 13, 1–38, https://doi.org/10.1371/journal.pone.0200917, 2018.

Lonati, G., Cernuschi, S., and Giani, P.: Air Quality Impact Assessment of a Waste-to-Energy Plant: Modelling Results vs. Monitored Data, Atmosphere-Basel, 13, 516, https://doi.org/10.3390/atmos13040516, 2022.

Luhar, A. K. and Britter, R. E.: A Random Walk Model for Dispersion in Inhomogeneous Turbulence in a Convective Boundary Layer, Atmos. Environ. (1967), 23, 1911–1924, https://doi.org/10.1016/0004-6981(89)90516-7, 1989.

Luther, A., Kleinschek, R., Scheidweiler, L., Defratyka, S., Stanisavljevic, M., Forstmaier, A., Dandocsi, A., Wolff, S., Dubravica, D., Wildmann, N., Kostinek, J., Jöckel, P., Nickl, A. L., Klausner, T., Hase, F., Frey, M., Chen, J., Dietrich, F., Nęcki, J., Swolkień, J., Fix, A., Roiger, A., and Butz, A.: Quantifying CH4 emissions from hard coal mines using mobile sun-viewing Fourier transform spectrometry, Atmos. Meas. Tech., 12, 5217–5230, https://doi.org/10.5194/amt-12-5217-2019, 2019.

Martens, R., Bruecher, W., Richter, C., Sentuc, F., Sogalla, M., and Thielen, H.: Extension and validation of ARTM (atmospheric radionuclide transportation model) for the application as dispersion calculation model in AVV (general administrative provision) and SBG (incident calculation bases); Erweiterung und Validierung von ARTM für den Einsatz als Ausbreitungsmodell in AVV und SBG, Report, Gesellschaft für Anlagen- und Reaktorsicherheit mbH (GRS), Köln (Germany), 2012.

Morio, J.: Global and local sensitivity analysis methods for a physical system, Eur. J. Phys., 32, 1577–1583, https://doi.org/10.1088/0143-0807/32/6/011, 2011.

Moussafir, J., Oldrini, O., Tinarelli, G., Sontowski, J., and Dougherty, C. M.: A New Operational Approach to Deal with Dispersion Around Obstacles: the MSS (Micro Swift Spray) Software Suite, in: 9th Int. Conf. on Harmonisation within Atmospheric Dispersion Modelling for Regulatory Purposes, https://harmo.org/Conferences/Proceedings/_Garmisch-Partenkirchen/publishedSections/5.26.pdf, 2004.

Panofsky, H. A., Tennekes, H., Lenschow, D. H., and Wyngaard, J. C.: The Characteristics of Turbulent Velocity Components in the Surface Layer Under Convective Conditions, Bound-Lay. Meteorol., 11, 355–361, https://doi.org/10.1007/BF02186086, 1977.

Plischke, E., Borgonovo, E., and Smith, C. L.: Global sensitivity measures from given data, Eur. J. Oper. Res., 226, 536–550, https://doi.org/10.1016/j.ejor.2012.11.047, 2013.

Rao, K. S.: Uncertainty Analysis in Atmospheric Dispersion Modeling, Pure Appl. Geophys., 162, 1893–1917, https://doi.org/10.1007/s00024-005-2697-4, 2005.

Ratto, C. F., Festa, R., Romeo, C., Frumento, O. A., and Galluzzi, M.: Mass-consistent models for wind fields over complex terrain: The state of the art, Environ. Softw, 9, 247–268, https://doi.org/10.1016/0266-9838(94)90023-X, 1994.

Ryall, D. B. and Maryon, R. H.: Validation of the UK Met. Office's NAME Model Against the ETEX Dataset, Atmos. Environ., 32, 4265–4276, https://doi.org/10.1016/S1352-2310(98)00177-0, 1998.

Saltelli, A., Ratto, M., Andres, T., Campolongo, F., Cariboni, J., Gatelli, D., Saisana, M., and Tarantola, S.: Global Sensitivity Analysis. The Primer, John Wiley & Sons, Ltd, Chichester, 2008.

Sawford, B. L.: Generalized random forcing in random-walk turbulent dispersion models, The Physics of Fluids, 29, 3582–3585, https://doi.org/10.1063/1.865784, 1986.

Schlesinger, S., Crosbie, R. E., Gagné, R. E., Innis, G. S., Lalwani, C. S., Loch, J., Sylvester, R. J., Wright, R. D., Kheir, N., and Bartos, D.: Terminology for model credibility, SIMULATION, 32, 103–104, https://doi.org/10.1177/003754977903200304, 1979.

760 Scire, J. S., Robe, F., Fernau, M., and Yamartino, R.: A user's guide for the CALMET meteorological model (Version 5), Report, Earth Tech, Inc., https://www.researchgate.net/publication/225089751_A_user%27s_guide_for_the_CALMET_meteorological_model_Version_5, 1998.

Shupe, M. D., Rex, M., Blomquist, B., Persson, P. O. G., Schmale, J., Uttal, T., Althausen, D., Angot, H., Archer, S., Bariteau, L., Beck, I., Bilberry, J., Bucci, S., Buck, C., Boyer, M., Brasseur, Z., Brooks, I. M., Calmer, R., Cassano, J., Castro, V., Chu, D., Costa, D., Cox, C. J.,
765 Creamean, J., Crewell, S., Dahlke, S., Damm, E., de Boer, G., Deckelmann, H., Dethloff, K., Dütsch, M., Ebell, K., Ehrlich, A., Ellis, J., Engelmann, R., Fong, A. A., Frey, M. M., Gallagher, M. R., Ganzeveld, L., Gradinger, R., Graeser, J., Greenamyer, V., Griesche, H., Griffiths, S., Hamilton, J., Heinemann, G., Helmig, D., Herber, A., Heuzé, C., Hofer, J., Houchens, T., Howard, D., Inoue, J., Jacobi, H.-W., Jaiser, R., Jokinen, T., Jourdan, O., Jozef, G., King, W., Kirchgaessner, A., Klingebiel, M., Krassovski, M., Krumpen, T., Lampert, A., Landing, W., Laurila, T., Lawrence, D., Lonardi, M., Loose, B., Lüpkes, C., Maahn, M., Macke, A., Maslowski, W., Marsay, C., Maturilli,
770 M., Mech, M., Morris, S., Moser, M., Nicolaus, M., Ortega, P., Osborn, J., Pätzold, F., Perovich, D. K., Petäjä, T., Pilz, C., Pirazzini, R., Posman, K., Powers, H., Pratt, K. A., Preußer, A., Quéléver, L., Radenz, M., Rabe, B., Rinke, A., Sachs, T., Schulz, A., Siebert, H., Silva, T., Solomon, A., Sommerfeld, A., et al.: Overview of the MOSAiC expedition: Atmosphere, Elementa: Science of the Anthropocene, 10, https://doi.org/10.1525/elementa.2021.00060, 2022.

SkyscraperPage: Drawing of Belchatow Power Station Units 1-6, https://skyscraperpage.com/diagrams/?buildingID=6131, 2023.

775 Sobol', I. M.: Sensitivity Estimates for Nonlinear Mathematical Models, Mathematical Modelling and Computational Experiments, 4, 407–414, 1993.

Stohl, A., Forster, C., Frank, A., Seibert, P., and Wotawa, G.: Technical note: The Lagrangian particle dispersion model FLEXPART version 6.2, Atmos. Chem. Phys., 5, 2461–2474, https://doi.org/10.5194/acp-5-2461-2005, 2005.

Stohl, A., Berg, T., Burkhart, J. F., Fjæraa, A. M., Forster, C., Herber, A., Hov, Ø., Lunder, C., McMillan, W. W., Oltmans, S., Shiobara, M.,
780 Simpson, D., Solberg, S., Stebel, K., Ström, J., Tørseth, K., Treffeisen, R., Virkkunen, K., and Yttri, K. E.: Arctic smoke - record high air pollution levels in the European Arctic due to agricultural fires in Eastern Europe in spring 2006, Atmos. Chem. Phys., 7, 511–534, https://doi.org/10.5194/acp-7-511-2007, 2007.

Stull, R. B.: An Introduction to Boundary Layer Meteorology, vol. 13 of *Atmospheric Sciences Library*, Kluwer, Dordrecht, https://doi.org/10.1007/978-94-009-3027-8, 1988.

785 TA Luft: Erste Allgemeine Verwaltungsvorschrift zum Bundes-Immissionsschutzgesetz (Technische Anleitung zur Reinhaltung der Luft–TA Luft), 2002.

Thomson, D. J.: Criteria for the selection of stochastic models of particle trajectories in turbulent flows, J. Fluid Mech., 180, 529–556, https://doi.org/10.1017/S0022112087001940, 1987.

Thykier-Nielsen, S., Deme, S., and Mikkelsen, T.: Description of the Atmospheric Dispersion Module RIMPUFF, Report, Risø National Lab-
790 oratory, https://www.researchgate.net/publication/228741276_Description_of_the_atmospheric_dispersion_module_RIMPUFF, 1999.

University of Oregon: Comparing Distributions: Z Test, http://homework.uoregon.edu/pub/class/es202/ztest.html, 2020.

VDI 3783 part 8: Turbulence parameters for dispersion models supported by measurement data, https://www.beuth.de/de/technische-regel/vdi-3783-blatt-8/59262605, VDI-Handbuch Reinhaltung der Luft, Band 1b, 2002.

VDI 3783 part 8: Turbulence parameters for dispersion models supported by measurement data, https://www.beuth.de/de/technische-regel/
795 vdi-3783-blatt-8/261571914?websource=vdin, VDI-Handbuch Reinhaltung der Luft, Band 1b, 2017.

Wan, H., Giorgetta, M. A., Zängl, G., Restelli, M., Majewski, D., Bonaventura, L., Fröhlich, K., Reinert, D., Rípodas, P., Kornblueh, L., and Förstner, J.: The ICON-1.2 hydrostatic atmospheric dynamical core on triangular grids – Part 1: Formulation and performance of the baseline version, Geosci. Model Dev., 6, 735–763, https://doi.org/10.5194/gmd-6-735-2013, 2013.

Wilks, D. S.: Statistical Methods in the Atmospheric Sciences, vol. 91 of *International Geophysics Series*, Elsevier, Amsterdam, 2 edn., 2006.

Wolff, S., Ehret, G., Kiemle, C., Amediek, A., Quatrevalet, M., Wirth, M., and Fix, A.: Determination of the emission rates of CO2 point sources with airborne lidar, Atmos. Meas. Tech., 14, 2717–2736, https://doi.org/10.5194/amt-14-2717-2021, 2021.

Zagayevskiy, Y. and Deutsch, C. V.: A Methodology for Sensitivity Analysis Based on Regression: Applications to Handle Uncertainty in Natural Resources Characterization, Natural Resources Research, 24, 239–274, https://doi.org/10.1007/s11053-014-9241-0, 2015.