Article

Machine Learning to Predict Area Fugitive Emission Fluxes of GHGs from Open-Pit Mines

Seyedahmad Kia [1], Manoj K. Nambiar [1] , Jesse Thé [2], Bahram Gharabaghi [1] and Amir A. Aliabadi [1],*

[1] School of Engineering, University of Guelph, Guelph, ON N1G 2W1, Canada; skia@uoguelph.ca (S.K.); manojunbc@gmail.com (M.K.N.); bgharaba@uoguelph.ca (B.G.)

[2] Lakes Environmental, Waterloo, ON N2L 3L3, Canada; jesse.the@weblakes.com

* Correspondence: aliabadi@uoguelph.ca

Abstract: Greenhouse gas (GHG) emissions from open-pit mines pose a global climate challenge, which necessitates appropriate quantification to support effective mitigation measures. This study considers the area-fugitive methane advective flux (as a proxy for emission flux) released from a tailings pond and two open-pit mines, denominated "old" and "new", within a facility in northern Canada. To estimate the emission fluxes of methane from these sources, this research employed near-surface observations and modeling using the weather research and forecasting (WRF) passive tracer dispersion method. Various machine learning (ML) methods were trained and tested on these data for the operational forecasting of emissions. Predicted emission fluxes and meteorological variables from the WRF model were used as training and input datasets for ML algorithms. A series of 10 ML algorithms were evaluated. The four models that generated the most accurate forecasts were selected. These ML models are the multi-layer perception (MLP) artificial neural network, the gradient boosting (GBR), XGBOOST (XGB), and support vector machines (SVM). Overall, the simulations predicted the emission fluxes with R^2 (-) values higher than 0.8 (-). Considering the bias (Tonnes h^{-1}), the ML predicted the emission fluxes within 6.3%, 3.3%, and 0.3% of WRF predictions for the old mine, new mine, and the pond, respectively.

Keywords: emission flux; machine learning (ML) method; open-pit mines; weather research and forecasting (WRF)

1. Introduction

In developed and developing countries, energy is a critical component supporting the development of the economy and society [1]. Demand for energy is forecasted to expand globally by 33% from the 2010s to the 2030s, and the demand will lead to an increase in the production of more greenhouse gas (GHG) emissions [2]. In the United States, more than 25% of the methane emitted into the atmosphere is a result of the oil and gas industry. Abating such emissions from the oil and gas sector is a critical component of mitigating climate change [3–5]. The warming trend can be slowed down by reducing the amount of methane emissions [6,7]. Methane has a global warming potential (GWP) 28 times greater than carbon dioxide for a duration of 100 years. Approximately one quarter of total global warming is related to the methane emissions from natural and anthropogenic sources. The short-term warming trends are affected by the methane lifetime, which is approximately 10 years [8].

The oil sands in Canada are proven to be the third largest reservoir of oil in the world [9]. The exploitation of such a reserve demands rigorous environmental management, given the energy intensity required to extract the crude oil from the oil sands [10,11]. In 2015, Alberta oil sands accounted for the largest share of GHG emissions from the oil and gas sector, 37% and 10%, respectively, of GHG emissions in Canada [9]. The unconventional oil resources industry, such as oil sands, is more energy-intensive than the conventional oil, with higher GHG emissions associated with resource extraction, which shows the need

for more attention to the emissions from these industries [12]. Oil sands operations emit GHGs due to various activities such as drilling, land development, extraction, chemical and physical processing, venting, flaring, and the release from large area sources (also known as fugitive emissions) [2].

For the development of adequate mitigation measures, a reliable quantification of GHG emission fluxes from open-pit mines is needed [13]. Large estimation differences of methane emission fluxes from such facilities are the result of ignoring the meteorological effects that complicate pollutant transport in complex topographies [14]. While orographic complexities confound atmospheric boundary layer (ABL) flows compared to ABL flows over flat and homogeneous terrains [15–23], only a handful of investigations are dedicated to ABL flows over open-pit mines [24,25], primarily due to operational, risk, access, economic, and other concerns [25,26]. Open-pit mines represent an unusual topographic feature. The meteorological patterns within and near these mines can be quite different from flat terrain [20–22,27,28], and understanding these patterns is important given the environmental impact of mining activities. The release of GHGs from such environments is associated with multiple steps of the resource extraction process, starting from land modification to the exportation of the mined goods [4,14]. Mine activities often create large amounts of fugitive dust, odorous compounds, and GHGs [29], and the flow circulations, shear layers, and plume meandering created by mine pits [30] impact the transport of materials to downwind environments. Approved emissions calculation methods for regulatory reporting fail in quantifying area fugitive emission fluxes from complex terrains. For instance, concerning open-pit mines, Liggio et al. (2019) [31] investigated that area-fugitive emissions of carbon dioxide probed in situ by an aircraft differed by 13–123% of values reported by emission inventory datasets. To address the common challenges of quantifying GHG emission fluxes from open-pit mines, measurement, modelling, and machine learning techniques of emission quantification are briefly reviewed in the subsequent sections.

1.1. Measurement Techniques to Quantify Emission Fluxes

The extrapolation of observations and point measurements is the basis for most of the open-pit mine GHG emission flux predictions by industrial reports. Flux chambers (FCs) are a simple but intrusive means of measuring gas emission fluxes. One of the difficulties with FC measurements is their spatially small measurement footprint, as well as their impracticality for measuring emission fluxes from vertical surfaces (e.g., mine faces). A spatially inhomogeneous and topographically complex source like an open-pit mine would require an extensive and difficult FC sampling survey [32,33]. Micro-meteorological measurement approaches, such as eddy-covariance or flux gradient, overcome some of the weaknesses of FCs (e.g., no interference with the emitting surface, larger measurement footprint, and capability for long-term monitoring), but they are fundamentally challenged by complex wind conditions, and the difficulty of interpreting their measurements for spatially inhomogeneous emission sources [34–36]. Another measurement paradigm is based on the concept of mass balance, which attempts to enclose the emitting facility in a hypothetical box in the atmosphere, while the emission flux is measured on the boundary of this box using aircraft or drone sampling [31,37–39]. While overcoming the limitations of the previous techniques, this approach has its own shortcomings. For instance, aircraft sampling misses the first few tens of meters near the ground, the measurement of which is crucial for the accurate estimation of the emission fluxes [37]. In addition, mass balance measurements are conducted for a facility only a few times a year, while a more continuous measurement technique is ideal to capture diurnal and seasonal variations of emission fluxes [22]. Because of the large footprint of open-pit mining facilities, precise measurements of their emission fluxes are not possible. Therefore, past research has attempted to offer modeling tools for the estimation of emission fluxes for operational use.

1.2. Modeling Techniques to Quantify Emission Fluxes

Access to high computational power has propelled methods for the calculation of area-fugitive emission fluxes by atmospheric models under two categories of modeling

approaches. The first category groups models into diagnostic versus prognostic paradigms. In the diagnostic paradigm, wind field and pollutant dispersion modeling is achieved using empirical formulations, while in the prognostic paradigm, fundamental transport equations of atmospheric flow and dispersion are solved. Diagnostic models are computationally efficient, while they lack accuracy. On the other hand, prognostic models are computationally expensive, while they provide better accuracy. Using either of these paradigms, the second category groups models under either forward or backward dispersion modes. In the forward dispersion approach, the input pollutant sources are resolved both in space and time, while the model calculates the emission flux downwind. However, in the backward dispersion approach, the pollutant concentration is measured downwind, while by relying on a dispersion model, the emission source strength upwind is inferred. The backward dispersion modelling paradigm is also known as inverse dispersion modelling (IDM) [40,41].

The CALifornia PUFF Model (CALPUFF) is an example of a diagnostic model [42]. CALPUFF has been utilized to analyze emission fluxes from industrial and urban sites, as well as the associated air quality and health implications [43–48]. When applied to complex topography, CALPUFF has been shown to be less accurate when modeling atmospheric dispersion under the thermally stable condition [49]. In addition, CALPUFF is shown to overpredict the turbulent diffusion, and as a result, underpredict the concentration of pollutants when compared to observations or more accurate models [50]. CALPUFF is also shown to suffer from lack of adequate meteorological forcing for its wind field and dispersion predictions. For instance, when near-surface and upper-air grid points of the model are not sufficiently forced by appropriate meteorological conditions, the wind field predictions lose accuracy [51]. One remedy in operational CALPUFF modeling is to force it with meteorological fields obtained from a prognostic mesoscale weather model [52]. As far as emission flux estimation is concerned, CALPUFF is usually used in an IDM configuration, where a measurement of concentration of a pollutant downwind is carried out, while the emission flux at the source is estimated by relying on CALPUFF's dispersion calculations. This paradigm in using CALPUFF can be seriously challenged, given its deficiencies when applied to non-conventional topographies, such as a mine pit.

The Weather Research and Forecasting (WRF) model is an example of a prognostic model. Such mesoscale models are employed to provide weather, air quality, and dispersion analyses. These models require physical parameterizations to account for the exchanges of momentum, heat, and atmospheric species (e.g., water) between the earth surface and the atmosphere [53]. Present-day mesoscale models allow for the selection of a rich spectrum of parameterizations, and choosing the effective parameterizations among the alternatives is a very complex task, informed by the desired outcome of a modeling investigation [54]. The accuracy of the actual terrain representation, which is under question, is the primary limitation of WRF at refined grid spacings [55]. In order to make accurate projections in weather, land-use and topographic data are needed, which are counted as significant inputs used for weather forecasting models. In addition, high-resolution terrain data are required for the elucidation of complex meteorological phenomena of different areas [56]. The complex terrains impact weather through changing different atmospheric variables, e.g., wind velocity vector components, radiative properties, and surface sensible and latent heat fluxes [57]. It has been demonstrated that an increase in input topographic and land-use resolutions, combined with refining grid spacing, plays a more significant role than simply using more refined grid spacings in the model toward simulations with more accuracy [58]. The weather patterns that are influenced by terrain features at the regional level play a crucial role in the numerical model by changing the distribution of rainfall, wind field, and atmospheric state variables [59,60]. The land-use information plays an integral part in weather alteration across the scales by influencing the exchanges of momentum, heat, and atmospheric species between the earth surface and the atmosphere [61]. Atmospheric state variables can vary, caused by the influences of changing aerodynamic roughness length scale when higher terrain and land-use resolution information are used [62]. The thermodynamic variables show improved agreement with measurements when detailed terrain information is utilized [20,63].

In regards to emission flux calculations, the WRF model is furnished with a built-in passive tracer dispersion method [22,64,65], which is accompanied by other chemistry modules that can be utilized to compute the temporal and spatial evolution of the mixing ratio of tracers, and consequently, emission fluxes in the forward dispersion paradigm of modeling [66–69]. The functionality of WRF for dispersion simulations in complex terrain and the land use environment of a mine pit has been investigated comprehensively [22]. The boundary condition for gases at the lowest model layer in mesoscale models is implemented either as fixed-value (fixed-mixing ratio) [22,67] or constant gradient (fixed-flux informed from inventory datasets) [68,70–72] (e.g., in accordance to calculations or emission records) representing spatial and temporal distributions. As proposed by Karion et al. (2019) [71], adopting the fixed-flux approach on the basis of emission inventories has been criticized, due to the model's inability to conserve the emission flux near the surface and model's exit boundaries, along with the spatio-temporal inaccuracies regarding the inventory datasets [68]. Based on the given observations, the use of the fixed-value (i.e., fixed-mixing ratio) boundary condition may be preferred. Despite the fact that fixed-value approaches do not provide an estimate for the discharged pollutants at the source, they aid the estimation of the emission flux at locations far away from the source, which is affected by weather modulations in the wind field and other state variables. This is named the advective flux, a proxy for the emission flux in the present study.

Motivated by the limitations of the diagnostic and prognostic models, machine learning (ML) may be employed to estimate emissions at a mining facility, given limited observed and simulated data. As opposed to regression techniques, which are linear models, the goal of neural networks is to elucidate non-linear patterns in information by constructing layers of nodes (neurons) in the model. The more popular and accurate machine learning methods, like support vector machines (SVM) [73], which engineer hyperplanes to separate instances, can be used as the main method for emission flux calculation from open-pits. SVM is optimized to create separating hyperplanes in the data among classes to increase the margin among classes. Another ML method is the multi-layer perceptron (MLP) artificial neural network, which has mostly been applied in hydrology [74], although the model is not preferred for present-day machine learning tasks. The gradient boosting (GBR) technique is an approach for classification, regression, and similar problems, offering a prediction among an ensemble of weak prediction models. These weak models are composed of decision trees [75]. XGBOOST (XGB) is a branch of the GBR model, which utilizes the Newton–Raphson method in a function space not similar to the gradient boosting that utilizes the gradient descent method. In the XGB model, a second order Taylor's expansion is used in the loss function to construct the connection to the Newton–Raphson method. The long short-term memory (LSTM) technique [76] is used in deep learning problems. Not similar to standard feed-forward neural networks, LSTM includes feedback connections. Implementations of procedures such as LSTM allow network training to take place without having long-term parameters "explode" or "vanish" as a result of multiple learning updates [77,78]. ML models based on SVM, deep learning, LSTM, and more have been used in various facets of energy engineering predictions, such as power plant heat transfer rate, power plant emission reduction, fluidized adsorption bed processes, and generator power curves [79–81]. However, despite the importance of quantifying emissions from open-pits, the authors of this article have not been informed of any studies attempting to estimate methane emission fluxes from open-pit mining facilities during different diurnal times and seasons using ML models.

1.3. Research Gaps and Objectives

The above review uncovers deficiencies and strengths in all emission flux quantification methods applicable to open-pit mines. Techniques purely based on measurement can be accurate, but suffer from limited spatio-temporal coverage. Diagnostic models are computationally fast but struggle to provide accurate estimates of the emission flux. On the other hand, prognostic models are complicated to set up, and computationally expensive for operational use, but they offer more accurate estimates of the emission flux. Finally,

machine learning techniques appear to require extensive training datasets and be specific for each emission source, without the possibility of extrapolating the use of one technique developed for one emission source to another emission source under different conditions.

This study is motivated by the deficiencies and strengths of each method to propose a surrogate method with operational capability to estimate the emission flux of GHGs from an open-pit mine fairly accurately and computationally fast. In the proposed method, meteorological variables and emission flux from a weather research and forecasting (WRF) mesoscale model were used as an input and training dataset for a few machine learning techniques to predict the emission flux. Two sets of models are used. The first set is composed of the GBR, XGBOOST, SVM, and MLP algorithms. The ML model learns in general. The second set includes LSTM, which is only considered as a research model, since in the operational sense, it is difficult to measure the emission flux at any given time (e.g., each hour) on a continuous basis to predict the emission flux in the next time interval (e.g., next hour).

The study is organized as follows. Section 2 describes the methodology. Section 2.1 introduces the open-pit mine site. The details of the WRF model are presented in Section 2.2, and the details of the machine learning models and the statistical analysis are presented in Section 2.3. In Section 3, the results of the WRF and machine learning model simulations for two mines (old and new) and a tailings pond are presented. Section 4 includes the main conclusions and recommendations.

2. Methodology

2.1. Open-Pit Mining Site

Figure 1a shows the site location of the present study, which is an open-pit mine in Canada. The mining site was composed of two open-pits, an old mine and a new mine, and a tailings pond. The old mine, with a depth of roughly 100-m and a span of roughly 2000-m, was mainly the place where the mining excavations were conducted. Another mine, the new mine, which was newly exploited, was smaller than the old mine.

The present study takes a forward dispersion modeling paradigm employing WRF 4.0 [82]. The WRF 4.0 source code is developed in Fortran (https://www2.mmm.ucar.edu/wrf/users/ (accessed 17 January 2020)). A tracer dispersion method was used and the field-measured methane (tracer) mixing ratio was forced at the model's lowest level. This represented the spatio-temporal variations of methane (tracer) mixing ratio over the two mines and the pond. The advective methane flux (a proxy for emission flux in this study) was simulated using WRF and presented as it experienced changes diurnally and seasonally. At the model's inner domain boundary, the dynamics of the surface-atmosphere interactions were not accounted for (e.g., biological processes within the pond, the intensity of excavations, or methane build-up inside the pit cavity at night); as an alternative, the meteorologically modulated emission flux downstream of the emission source was computed. Due to the presence of uncertainties regarding this method, it was not possible to accurately quantify the absolute flux of methane emissions from the entire facility. However, the quantification of diurnal and seasonal change in the emission flux could be performed with enough accuracy to motivate future research [22].

The supporting observations and modeling occurred in four campaigns. The first campaign lasted for twenty days in May 2018 (Summer'18 or S18). The second campaign lasted for twenty days in February–March 2019 (Winter'19 or W19). The third campaign lasted for 30 days in July–August 2019 (Summer'19 or S19). The fourth campaign lasted for 30 days in October–November 2019 (Fall'19 or F19). Overall, 100 (days) (2400 (hours)) of observations and modeling were conducted, which are summarized in Table 1. The training and testing data for the ML models are provided at an hourly resolution, so the entire dataset includes 2400 records for each input or output variable. There are 960, 960, and 480 data records for each of the tailings pond, old mine, and new mine, respectively.
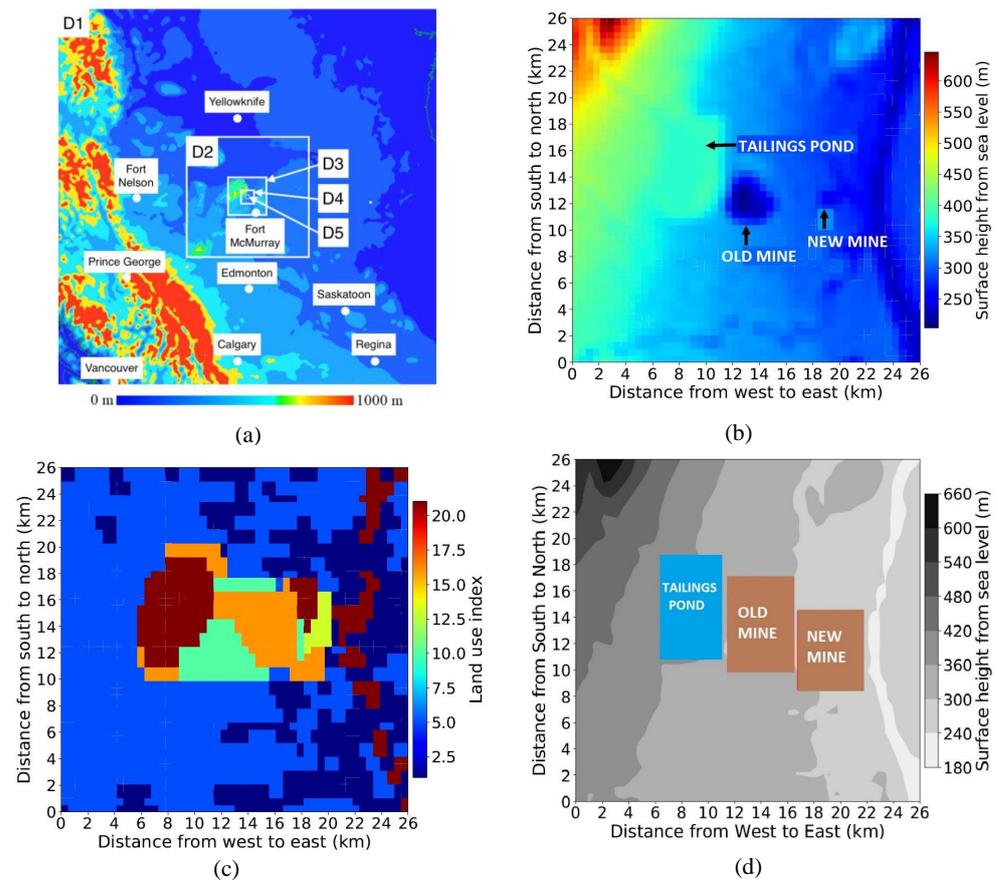
Figure 1. (**a**) WRF simulation domains (D1 is the largest domain and D5 is the smallest domain); (**b**) smallest domain WRF surface topography; (**c**) smallest domain WRF land-use configuration; (**d**) smallest domain WRF area patches for forcing the tracer boundary condition; measurement sites are approximately uniformly distributed within each rectangle.

Table 1. Supporting observations and modeling in four emissions monitoring campaigns; the date durations for each of the emissions sources are stated.

Campaign	Year	Tailings Pond	Old Mine	New Mine
S18	2018	1 May–11 May	18 May–28 May	-
W19	2019	14 February–24 February	16 March–26 March	-
S19	2019	9 July–19 July	29 July–8 August	12 August–22 August
F19	2019	25 October–4 November	10 November–20 November	8 October–18 October

2.2. Weather Research and Forecasting Model

2.2.1. Model Configurations

Five nested domains were set based on one-way nesting (i.e., no feedback interaction between two adjacent domains), with the outermost domain (D1) including a large area of north-western America, and the innermost domain (D5) mainly encompassing the mining site, including the tailings pond and the two mines. Detailed information about horizontal and vertical grid spacings, time step, grid spacing and time step ratios between adjacent grids, and the time advance scheme are provided by Nambiar et al. (2020) [22]. Most notably, the inner most domain had a horizontal grid spacing of 0.51 (km), and the model included 12 pressure levels from 0.025 to 2 (km) above the surface. The map illustrating the domains is presented in Figure 1a. The WRF settings and physical parameterizations are summarized in Table 2.

Table 2. WRF physical parameterizations.

Physics	Parameterization
Planetary Boundary Layer (PBL)	Mellor-Yamada-Janić [83]
Microphysics	Thompson Parameterization [84] (D1, D2, and D3 only)
Longwave Radiation	Rapid Radiative Transfer Model [85]
Shortwave Radiation	Rapid Radiative Transfer Model [85]
Cumulus Parameterization	Tiedtke Parameterization [86] (D1, D2, and D3 only)
Surface Layer (SL)	Monin-Obukhov Eta Similarity Parameterization [83]
Land Surface (LS)	Noah Land Surface Parameterization

The standard Global 30 Arc-Second (GTOPO 30s) database was employed to define the latest topography of the site in the three largest domains with a horizontal resolution of 0.9 (km). For the two smallest domains, the Shuttle Radar Topography Mission (SRTM) 1s information was employed with a horizontal resolution of 0.030 (km) [62,87]. For the smallest domain, the extracted topographical resolution (SRTM 1s) was further overwritten with a LiDAR dataset with a horizontal and vertical resolution of 0.001 (km). Figure 1b presents the topography of the site updated with SRTM 1s and LiDAR datasets.

Nahian et al. (2020) [20] investigated that, by means of utilizing the latest modified land-use configurations and topography, the computations of atmospheric state variables at the same facility would agree better with the measurements. The representation of land-use according to the Moderate Resolution Imaging Spectroradiometer (MODIS) 30s dataset (updated 21-class IGBP-MODIS land use dataset) by Friedl et al. (2010) [88] was utilized with categorizations for the pond (21), the two mines (16), mine processing facilities and housing (13), and cleared forest (10). The lake models in WRF as parameterized by Subin et al. (2012) [89] and Gu et al. (2015) [90] could simulate water–atmospheric exchanges of momentum, heat, and humidity. A depth of 0.050 (km) was implemented for the water bodies at the site, and so the lake model included 25 layers of water, soil, and snow combined. Figure 1c presents the settings of the land use implemented for WRF calculations. The initial and boundary conditions in WRF were implemented, resorting to the Global Data Assimilation System (GDAS) dataset (from the National Centers for Environmental Prediction (NCEP)), with grid spacing of 0.25° and a time stamp updated every 6 (hours).

2.2.2. Methane Transport and Flux Calculation

The campaigns provided field observations of methane mixing ratio near the surface at various locations of the mining facility. These data were acquired at heights from 2 to 10 (m) above the surface in four spots uniformly distributed in each of the areas of the tailings pond and the two mines, as depicted in Figure 1d. The instrument used was the Los Gatos Research Ultra-Portable Greenhouse Gas Analyzers (LGRs). Allen et al. (2019) [91] report that this instrument operates based on the principle of off-axis integrated cavity output spectroscopy and two near-infrared tunable diode lasers, which promptly scan a single strong (and isolated) absorption line of the target gas (here methane). The overall release areas, tiled by the rectangular patches, were equalized for the mines and the pond. This area, approximately 30 (km^2), was the same for all simulations and release sources. Methane measurements were acquired every 15 (min). The measurements were averaged every 4 (hours) and forced in the model as a boundary condition at the model's lowest layer [64]; this was necessitated by the WRF model, which required the recompilation of the program every 4 (hours). In summary, the program was customized to allow tracer release from a desired spatial area.

As proposed by Bhimireddy et al. (2018) [65], the passive tracer dispersion method was employed in WRF without accounting for any chemical reactions, and the tracer moved vertically and horizontally by the computed mean wind field and turbulent transport by means of the surface layer (SL) and planetary boundary layer (PBL) parameterizations. The Monin–Obukhov SL parameterization accounts for the vertical transport of the tracer by relying on similarity theory [83]. The Mellor–Yamada–Janić PBL parameterization accounts

for turbulence kinetic energy (*TKE*) employing a 2.5 closure technique [92]. Both horizontal and vertical turbulent motions within PBL are modelled as diffusion using eddy diffusion functions that depend on *TKE* and other model variables [93]. Twelve (hours) were counted as the spin-up time [22]. To discover the diurnal and seasonal variations in the emission flux, simulations were performed for 4-hour time intervals.

The tracer (methane) emission flux *F* was computed through a box placed around the emission sources by considering Gauss' theorem. *F* can be given by [37]

$$F = F_a + F_t + F_s + F_m + F_c, \tag{1}$$

where F_a is the advective flux and F_t is the turbulent flux through the box surfaces, F_s is the surface flux because of deposition or biogenic release of methane due to surface activity, F_m is the flux due to density variations in the box, and F_c is the flux because of chemical reactions encompassing methane as it interacts with other species in air. Large-eddy simulations (LES) predict a horizontal turbulent flux of atmospheric species upwind, which accounts for 10–20% of the advective flux [38,70]. However, this flux could not be computed in the present simulations due to the computational cost associated with overall simulations, extending a time span of 100 (days). F_s was disregarded considering the fact that the site land was primarily barren or developed for industrial operations and housing [72]. In accordance with Oertel et al. (2016) [94], the boreal forest is not recognized as a source of methane either (on the contrary, this forest is suggested to absorb methane). Thus, the basic assumption is that the measured methane mixing ratio is only generated from mine pits and ponds, and is not affected by other processes. So, overall, only the advective flux F_a and flux due to density variations in the box F_m are considered to be the major contributors to the total flux, i.e., $F \approx F_a + F_m$. We further disregard F_m; after all, any such flux contribution would average to zero over many days. Therefore, F_a was considered as the single major operational contributor to the total flux *F*. Considering these assumptions and former evidence for the same mining facility, it was suggested that the advective flux solely accounted for more than 95% of the overall emission flux [37]

$$F \approx F_a = \sum_{north} \overline{V}\,\overline{S}\Delta A + \sum_{east} \overline{U}\,\overline{S}\Delta A - \sum_{south} \overline{V}\,\overline{S}\Delta A - \sum_{west} \overline{U}\,\overline{S}\Delta A + \sum_{top} \overline{W}\,\overline{S}\Delta A, \tag{2}$$

where $\overline{U}$, $\overline{V}$, and $\overline{W}$ are the wind velocity vector components (m s^{-1}) simulated at each grid point on the box enclosing the emission source, ΔA is area element (m^2), and $\overline{S}$ is the concentration of the passive tracer (here methane) (μg m^{-3}) for a specified grid point. It is worth mentioning that when the box height stretches to altitudes well beyond the PBL height, then the last term in the above equation can be neglected (equal to zero). By its formulation WRF simulates the tracer transport in units of (ppm); however to calculate the emission flux in (Tonne h^{-1}), the tracer unit must first be converted to (μg m^{-3}) for methane. Using thermodynamic considerations, the units for the tracer can be converted from (ppm) to (μg m^{-3}). This is possible since all the required thermodynamic variables (e.g., temperature, potential temperature, and pressure) are available in WRF and can be extracted. Nambiar et al. (2020) [22] provide the detailed calculations. Overall, this procedure allows the operational calculation of the advective flux as a proxy for emission flux.

2.2.3. Uncertainty in WRF Estimate of the Methane Emission Flux

The ML model predictions are based on the WRF model estimates, so the success of the ML model predictions depends on the success of the WRF model in predicting the emissions flux of methane. If the accuracy of the WRF predictions of the methane emission flux can be quantified, then ML predictions will have practical relevance in the field.

For quantifying the WRF uncertainty of methane emission flux from the open-pit mine, the relevant uncertainties are bias in wind speed and bias in methane concentrations (or alternatively, mixing ratio). The bias (rather than root mean square error) is relevant

because any emission quantification method attempts to report the cumulative emission flux accurately over an entire year, as opposed to a snapshot prediction.

In an earlier study, we quantified the WRF uncertainty in wind speed using airborne measurements in the actual mine field [20]. In that study, the WRF uncertainty and average wind speed from field observations were $\Delta U = 0.0819$ and $\overline{U} = 2.614$ (m s^{-1}). In another study, we quantified the WRF uncertainty in the mixing ratio of methane using aircraft measurements in the same mine field [22]. In that study, the WRF uncertainty and average mixing ratio from field observations were $\Delta S = 0.0543$ and $\overline{S} = 0.417$ (ppm). Using these uncertainties and average values, we can estimate the uncertainty in WRF's advective flux (proxy for emission flux) using

$$\frac{\Delta F}{F} \approx \sqrt{\left(\frac{\Delta U}{\overline{U}}\right)^2 + \left(\frac{\Delta S}{\overline{S}}\right)^2} = \sqrt{\left(\frac{0.0819}{2.614}\right)^2 + \left(\frac{0.0543}{0.417}\right)^2} = 0.13, \tag{3}$$

which suggests that the uncertainty in WRF's prediction of the methane emission flux is about 13%. Given the lack of alternative methods for prolonged measurements of the emission flux, this uncertainty justifies the use of the WRF dataset as training data for the ML model.

2.3. Machine Learning Model and Statistical Analysis

Under practical considerations, up to four measurement periods (fall, winter, spring, and summer) with no more than 10 days per season are required to quantify emission fluxes every year. This measurement period is sufficient because of the high cost of accurate observation methods, such as deploying instrumented aircraft or another technologically sophisticated system. The GBR, XGBOOST (XGB), SVM, and MLP are candidate ML models that can be used operationally, while predicting the emission flux in the next hour with LSTM, based on measuring the emission flux at the current hour, will remain as a research tool. All the available data for the four campaigns include 10 days of observations and WRF modelling for a given emission source (i.e., old mine, new mine, or pond).

Model tuning was achieved and optimal hyper-parameters were set by minimizing the error statistics with respect to the reference dataset (emission flux predicted by WRF). The settings can be provided as: GBR: number of estimators = 1000, minimum samples per leaf = 4, minimum weight fraction per leaf = 0.01; XGBOOST: gamma = 1×10^{-4}, max tree depth = 20, number of estimators = 100, learning rate = 5×10^{-2}, minimum child weight = 2, sub-samples = 0.5; SVM: gamma = 1×10^{-2}, epsilon = 1×10^{-3}, $C = 1 \times 10^4$, kernel = rbf (radius basis function), cache = 1000 Mbytes; MLP: activation function = ReLu, loss function = MSE, learning rate = 1×10^{-3}, solver = ADAM, epsilon (for the ADA solver) = 1×10^{-8}, number of hidden layers = 1 with 100 nodes, alpha (L2 penalty to control over-fitting) = 1×10^{-4}, batch size = 100. The ML model source code is developed in Python 3.8 (https://www.python.org/ (accessed 10 August 2021)), and it utilized python's basic data analysis and visualization packages [95,96]. Specific ML algorithms were imported from the Scikit-Learn (sklearn) version 1.0 [97,98] and xgboost version 1.6 [99] libraries.

Figure 2 shows the methodology diagram for the present analysis. Overall, up to 10 features are chosen from WRF as input to the ML models. These are meteorological variables: diurnal time (h), season, wind speed (m s^{-1}) at 10 (m) above surface, temperature (K) at 2 (m) above surface, precipitation (mm h^{-1}), PBL height (m), surface pressure (Pa), relative humidity (%), surface sensible heat flux (W m^{-2}), and surface latent heat flux (W m^{-2}). These variables are averaged over the footprint of each emission source. The WRF-generated emission flux is chosen as the training data for the ML models. The ML models generate the emission flux as output. All the data are shuffled and 80% of data are separated, as the training set from the remaining 20% of data, as the test set. Ten ML algorithms are tested to check for the accuracy. Since the models are fairly simple and run reasonably fast, it is needed to choose the models that best fit the purpose, ignoring the "complexity" of the models. The inclusion of the deep neural network (DNN), in the form of the MLP algorithm, indicated that the benefits of DNNs cannot be fully exploited here,

as the training dataset is limited, and therefore, traditional ML algorithms perform better, with a reduced chance for over-fitting. In the present work the analysis is performed by adding one input variable (called feature in the ML method) at a time to check the accuracy of the prediction for a greater number of input variables.
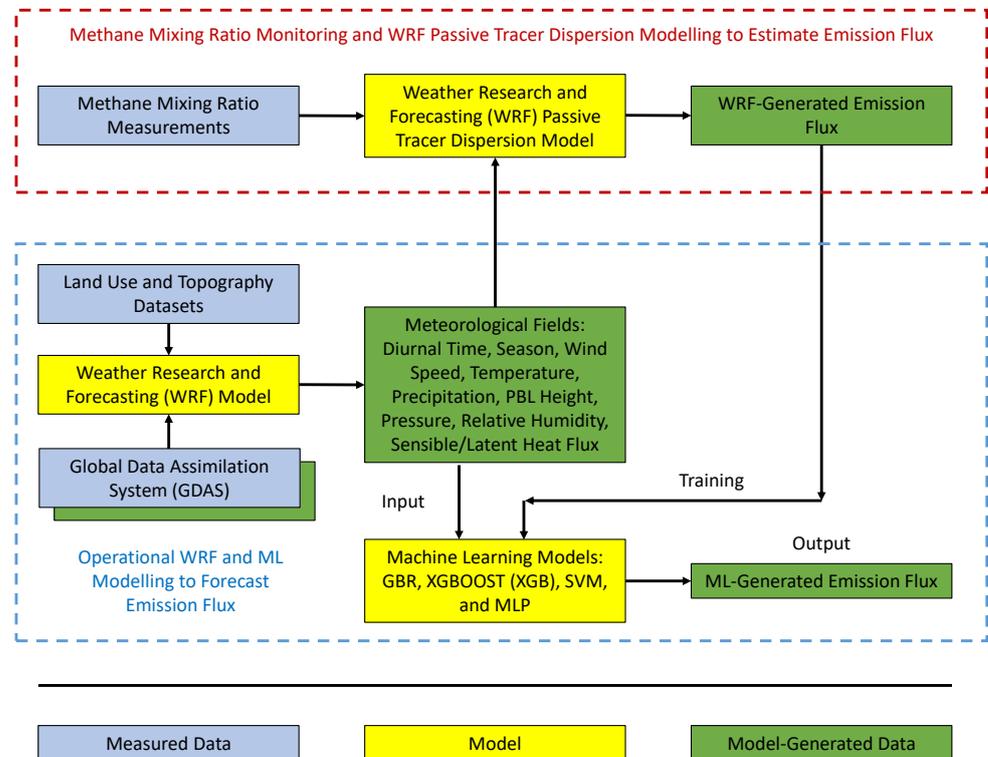
Figure 2. Research methodology: the weather research and forecasting (WRF) model relies on measured methane mixing ratio data, land use data, topography data, and the Global Data Assimilation System (GDAS) to generate the emission flux; the machine learning (ML) models use the WRF-generated meteorological fields as input data and the WRF-generated emission flux as training data to generate the emission flux as output data.

Quantitative comparisons between the model prediction (M_i) and a reference dataset (O_i) (here methane emission flux from WRF) are performed by determining the bias and root mean square error (RMSE) (Tonnes h^{-1}) defined by

$$Bias = \frac{\sum_{i=1}^{n}(M_i - O_i)}{n}, \tag{4}$$

$$RMSE = \sqrt{\frac{\sum_{i=1}^{n}(M_i - O_i)^2}{n}}, \tag{5}$$

where n is the number of data points accounted for in the error statistic calculation. The coefficient of determination R-squared (R^2 (-)) is the statistical measure of the correlation between a model's performance and a reference dataset defined by

$$R^2 = \frac{[n\sum_{i=1}^{n}(M_i O_i) - (\sum_{i=1}^{n} M_i)(\sum_{i=1}^{n} O_i)]^2}{[n\sum_{i=1}^{n}(O_i)^2 - (\sum_{i=1}^{n} O_i)^2][n\sum_{i=1}^{n}(M_i)^2 - (\sum_{i=1}^{n} M_i)^2]}. \tag{6}$$

The Pearson's correlation coefficient (*PCC* (-)) between M_i and O_i is often used to assess the ability of a model to simulate temporal variation and is given by

$$PCC = \frac{\sum_{i=1}^{n}(M_i - \overline{M})(O_i - \overline{O})}{\sqrt{\sum_{i=1}^{n}(M_i - \overline{M})^2}\sqrt{\sum_{i=1}^{n}(O_i - \overline{O})^2}}, \tag{7}$$

where overbars signify averages and *n* is the sample size.

3. Results and Discussion

Time series of wind speed at 10 (m) above surface, potential temperature at 2 (m) above surface, surface heat flux, and emission flux of methane during 10 days of F19 campaign for the old/new mines and the pond are presented in Figure 3. All data shown are extracted from WRF simulations. In the time period from 48 to 96 (hours), higher wind speed and surface heat flux are predicted over the pond compared to the mines (Figure 3a,c). The higher wind speed and surface heat flux caused a higher emission flux from the pond (Figure 3d). Overall, the pond shows greater variability in the wind speed, surface heat flux, and emission flux noticed by the peaks from Figure 3a,c,d. The reason for such variations could be related to meteorological effects over water bodies. For the mines, the emission flux drops during the night as the wind speed decreases over night in the mine cavity. The change in the diurnal variation of temperature for the mines are higher than the pond as the heat capacity of water bodies are higher than soil, which could damp the amplitude of temperature oscillations in the atmosphere above the pond. Figure 3d shows that as the new mine is more active, the emission flux from it is higher. The new mine emission flux exhibits a notable diurnal pattern, possibly due to the regular schedule of the mining activity. Meanwhile, such diurnal pattern is not as notable for the pond, which emits methane given the natural variation in the meteorological conditions. Likewise, such diurnal variability in the emission flux is weaker for the old mine.

Figures 4–6 show the *PCC* matrix computed for measured tracer (methane) mixing ratio, normalized emission flux (computed by WRF), and other atmospheric state variables (computed by WRF) for the old mine, new mine, and pond, respectively. For this analysis all the data is used from the WRF simulations, i.e., the entire 2400 records, with 960, 960, and 480 records related to the tailings pond, old mine, and new mine, respectively. As can be seen for the old mine in Figure 4 for S18, W19, S19, and F19 cases, the emission flux was greatly correlated with wind speed at 10 [m] above surface ($S10$) in all seasons but highly correlated with Surface Heat Flux (SHF) (W m^{-2}) only during the warm conditions (S18 and S19 campaigns). For example, in the S18 case, the *PCCs* for $S10$ and SHF are 0.82 (-) and 0.74 (-), respectively, for the old mine (Figure 4a). The same correlation for the new mine and pond can be seen for $S10$ and emission flux in warm conditions (S18 and S19 campaigns) (Figures 5a and 6a,c). The emission flux of the pond in S18 and S19 campaigns was highly correlated with $S10$ ($PCC = 0.81, 0.62$ (-)) but anti-correlated with SHF ($PCC = -0.43, -0.37$(-)) (Figure 6a,c). The anti-correlation was foreseen having the knowledge that the water temperature would be different than that of the nearby earth surfaces. This is itself due to larger thermal capacity of water compared to soil [20].

In S19 and F19 campaigns, the emission fluxes from the old and new mines were highly correlated with $S10$ for the old mine ($PCC = 0.74, 0.84$ (-)) and for the new mine ($PCC = 0.89, 0.85$ (-)). In S19 and F19 campaigns, the emission fluxes were highly correlated with SHF for the old mine ($PCC = 0.35, 0$ (-)) and for the new mine ($PCC = 0.4, 0.68$ (-)). However, in S19 and F19 campaigns, the emission flux from the pond was highly correlated with $S10$ ($PCC = 0.62, 0.77$ (-)), but either anti-correlated or correlated with SHF ($PCC = -0.37, 0.73$ (-)), possibly due to transitioning from the warm to the cold conditions. The greatest correlation with the emission flux is with the wind speed. The higher the wind speed, the greater the emission flux. For the case of the mines, surface sensible heat flux is positively correlated with the emission flux. On the other hand, for the case of the pond, surface sensible heat flux is negatively correlated with the emission flux in the warm conditions ($PCC = -0.43, -0.37$ (-) for S18 and S19, respectively) but positively correlated with the emission flux in the cold conditions ($PCC = 0.12, 0.73$ (-) for W19 and F19, respectively). This difference can be explained given the meteorological processes. In the warm conditions (S18 and S19), the high emission flux is accompanied with high wind speed during the daytime, where the pond surface is cooling the atmosphere due to being at a lower temperature (negative heat flux). While in the cold conditions (W19 and F19), the high emission flux is accompanied with high wind speed during synoptic events, where

the pond surface is warming the atmosphere due to being at a higher temperature (positive heat flux).
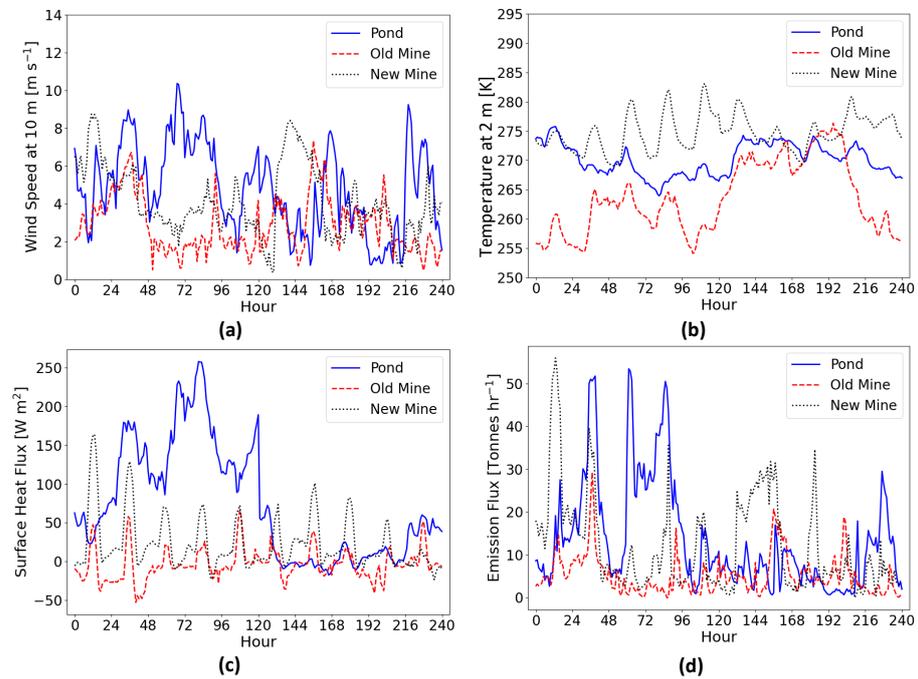
Figure 3. Time series of WRF predictions of (**a**) wind speed at 10 (m) above surface, (**b**) potential temperature at 2 (m) above surface, (**c**) surface heat flux, and (**d**) emission flux of methane during 10 (days) of F19 campaign for emissions monitoring from the old/new mines and the pond.
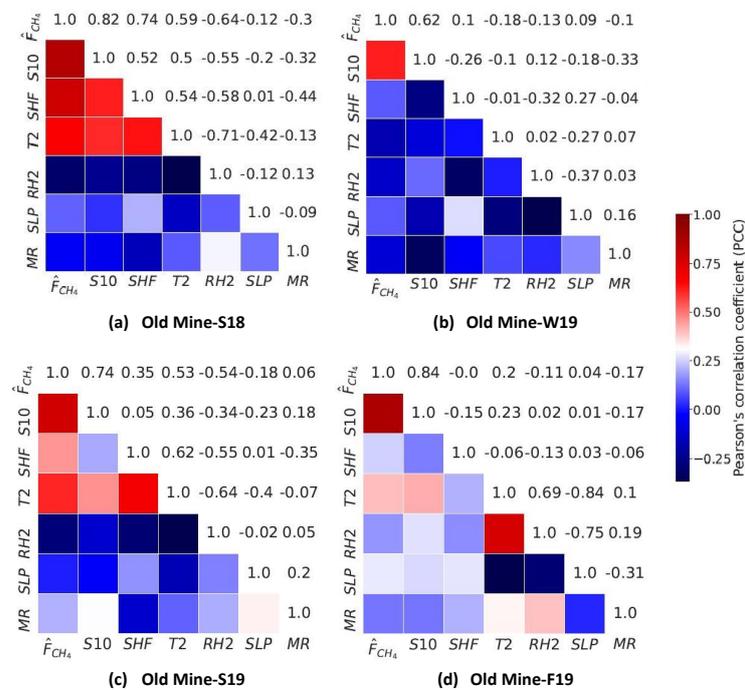
Figure 4. Pearson's correlation coefficient (*PCC*) of the old mine (**a**–**d**); tracer (methane) mixing ratio (*MR*) measured near surface; WRF normalized emission flux ($\hat{F}_{CH_4}$); WRF wind speed at 10 (m) (*S*10); WRF upward surface sensible heat flux (*SHF*); WRF temperature at 2 (m) (*T*2); WRF relative humidity at 2 (m) (*RH*2); and WRF sea level pressure (*SLP*).
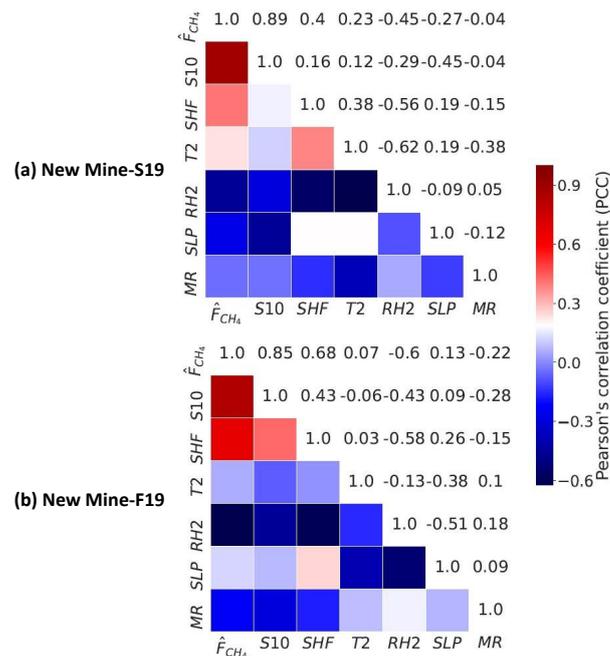
Figure 5. Pearson's correlation coefficient (*PCC*) of the new mine (**a**,**b**); variables the same as in Figure 4.
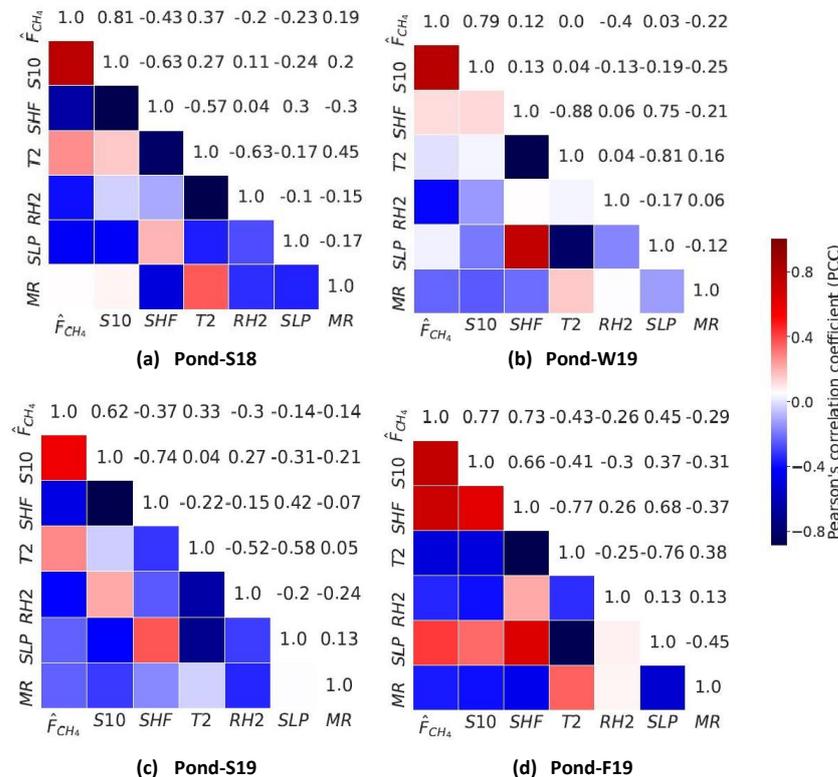
Figure 6. Pearson's correlation coefficient (*PCC*) of the pond (**a**–**d**); variables the same as in Figure 4.

Figures 7–9 show the scatter plots of emission fluxes predicted by the MLP, GBR, XGB, and SVM models versus the WRF predictions. This data is associated with the 20% testing data, for which the ML models were not trained. On the same figures some statistics for the comparison are calculated and reported, which include the Bias (Tonnes h^{-1}), RMSE (Tonnes h^{-1}), and R^2 (-) for the old mine, new mine, and the pond.

Table 3 shows the Bias (Tonnes h^{-1}), RMSE (Tonnes h^{-1}), and R^2 (-) values for the four machine learning methods (MLP, GBR, XGB, SVM) applied to the old and new mines and the pond during the individual campaigns. The error statistics are associated with the 20% testing data, for which the ML models were not trained. The averaged Bias (Tonnes h^{-1}) shows lower values for the pond and new mine, while the old mine has the highest Bias (Tonnes h^{-1}) and RMSE (Tonnes h^{-1}). This may suggest a more successful prediction of emission flux by the ML methods for the pond and the new mine compared to the old mine. The average R^2 (-) of prediction by the deployed ML methods shows better forecast for the new mine although the two other predictions have R^2 (-) higher than 0.8 (-). This may be due to more regular and scheduled mining activity in the new mine, compared to the old mine and the pond. Comparing the Bias (Tonnes h^{-1}), RMSE (Tonnes h^{-1}), and R^2 (-) shows that the GBR and XGB models performed better predictions for all the different emission sources and seasons. The average WRF emission flux for the old mine, new mine, and the pond are 9.90, 11.15, and 9.19 (Tonnes h^{-1}), respectively. Considering the Bias (Tonnes h^{-1}), the ML-predicted emission fluxes are on average within 6.3%, 3.3%, and 0.3% of WRF prediction for the old mine, new mine, and the pond, respectively. The same numbers considering RMSE (Tonnes h^{-1}) would be around 30% of WRF predictions.
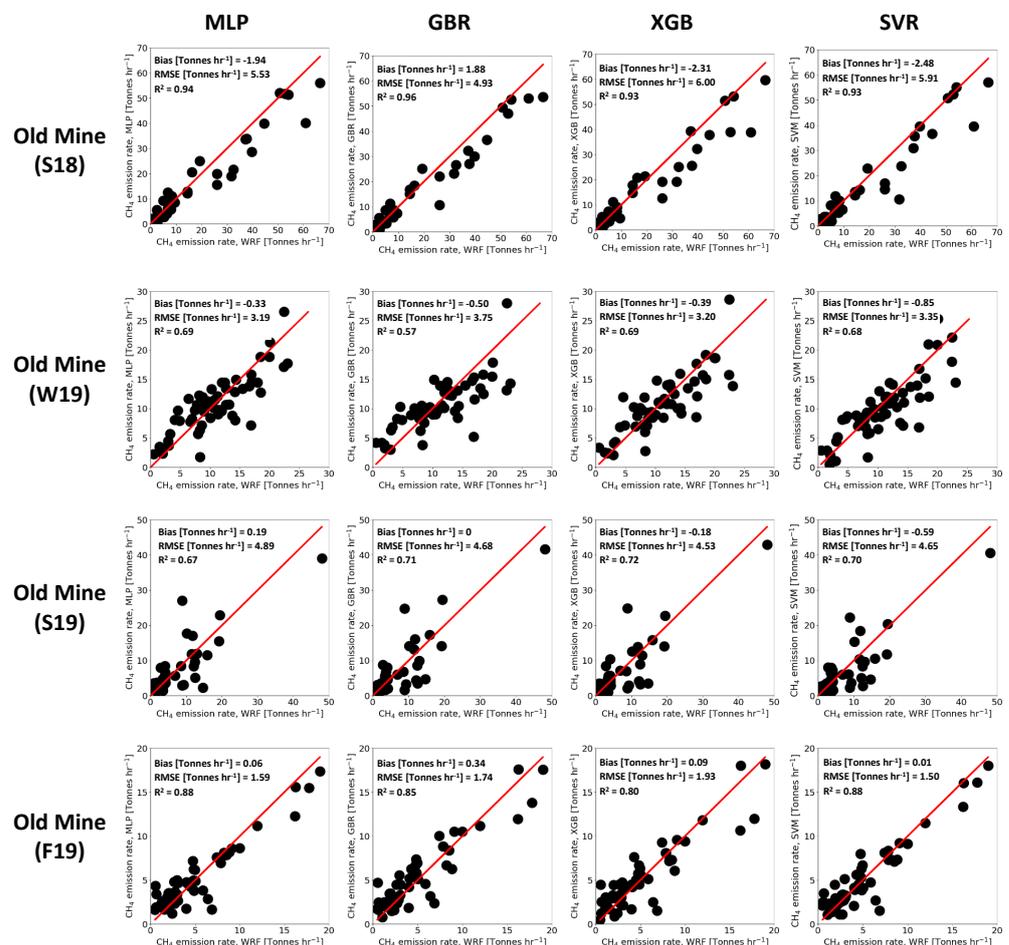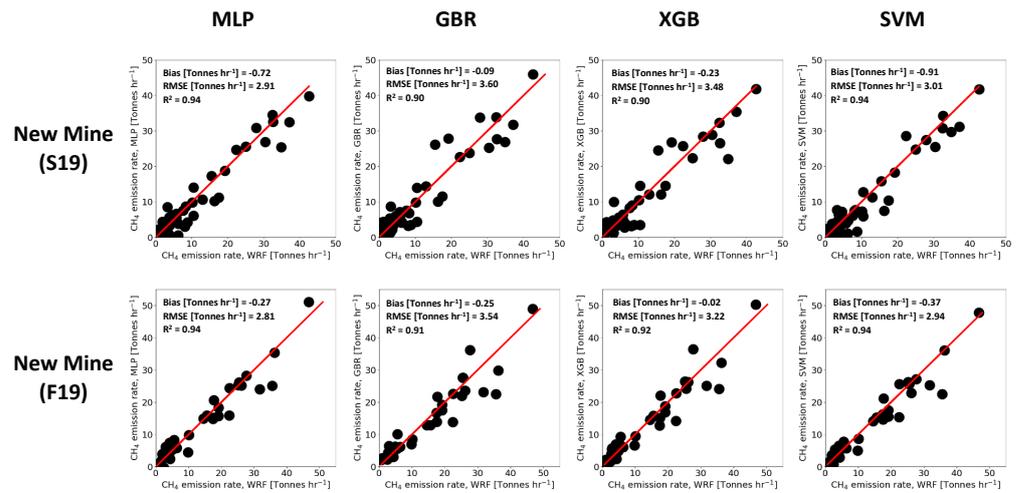

Figure 7. Scatter plots of emission rates predicted by the MLP, GBR, XGB, and SVM models versus the WRF predicted emission rates and the calculated Bias (Tonnes h^{-1}), RMSE (Tonnes h^{-1}), and R^2 (-) for the old mine for S18, W19, S19, and F19 campaigns.
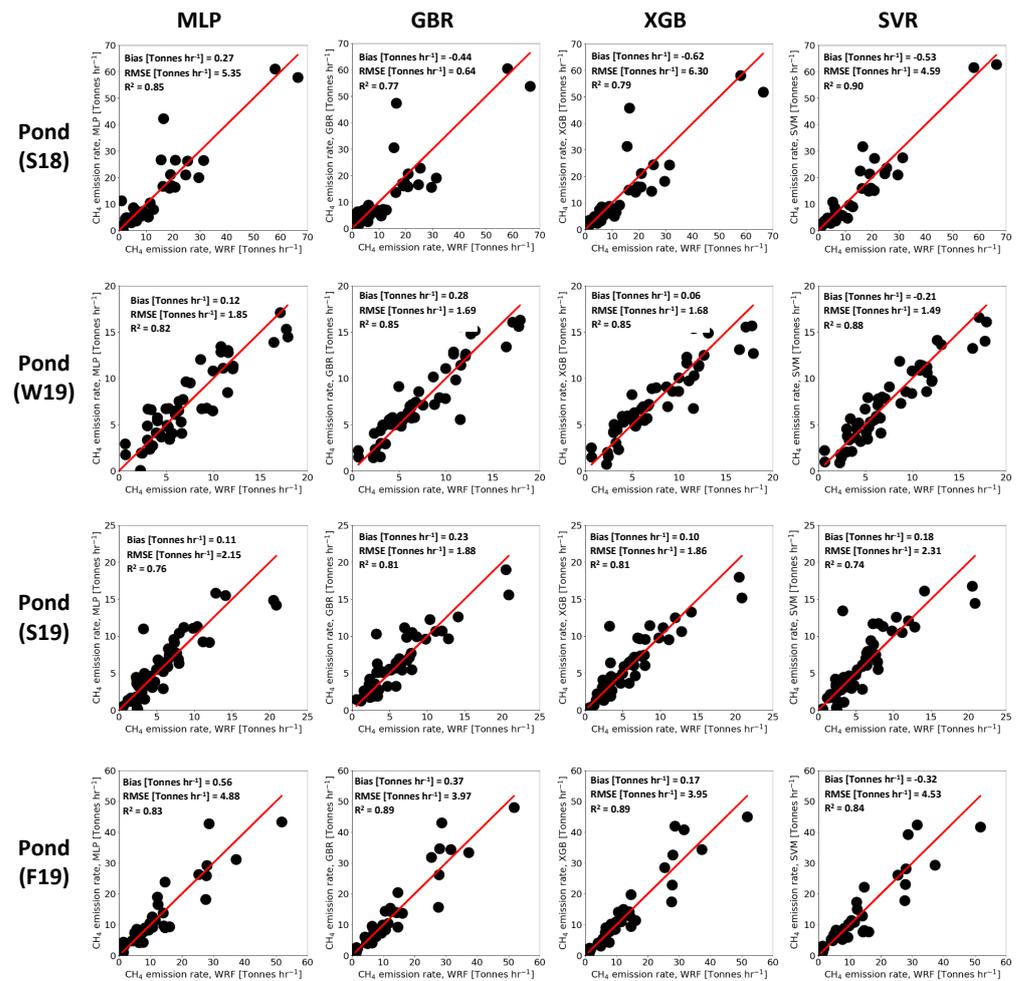
Figure 8. Scatter plots of emission rates predicted by the MLP, GBR, XGB, and SVM models versus the WRF predicted emission rates and the calculated Bias (Tonnes h^{-1}), RMSE (Tonnes h^{-1}), and R^2 (-) for the new mine for S19 and F19 campaigns.

Figure 9. Scatter plots of emission rates predicted by the MLP, GBR, XGB, and SVM models versus the WRF predicted emission rates and the calculated Bias (Tonnes h^{-1}), RMSE (Tonnes h^{-1}), and R^2 (-) for the pond for S18, W19, S19, and F19 campaigns.

Table 3. Bias (Tonnes h^{-1}), RMSE (Tonnes h^{-1}), and R^2 (-) values for the four machine learning methods (MLP, GBR, XGB, SVM) of old and new mines and the pond for S18, W19, S19, and F19.

		Old Mine (S18, W19, S19, F19)		New Mine (S19, F19)		Pond (S18, W19, S19, F19)	
Bias (Tonnes h^{-1})	MLP	−0.47		−0.51		0.25	
	GBR	−0.46	Average = −0.62	−0.17	Average = −0.37	0.11	Average = 0.03
	XGB	−0.63		−0.13		−0.06	
	SVM	−0.91		−0.65		−0.20	
RMSE (Tonnes h^{-1})	MLP	3.93		2.87		3.75	
	GBR	3.85	Average = 3.96	3.57	Average = 3.20	3.96	Average = 3.73
	XGB	4.04		3.36		3.81	
	SVM	4.02		2.98		3.38	
R^2 (-)	MLP	0.89		0.94		0.85	
	GBR	0.89	Average = 0.89	0.90	Average = 0.92	0.83	Average = 0.85
	XGB	0.88		0.91		0.84	
	SVM	0.89		0.94		0.88	

4. Conclusions and Recommendations

Open-pit mines represent an unusual topographic feature. The meteorological patterns within and near these mines can be quite different from those of flat terrain, and understanding these patterns is important given the environmental impact of mining activities. Mine activities often create large amounts of fugitive dust, odorous compounds, and greenhouse gases (GHGs), the quantification of which is confounded by flow circulations, shear layers, and plume meandering created by mine pits. Open-pit mine emissions measurement techniques have limitation of spatio-temporal coverage. Although the diagnostic models are computationally fast, the models have difficulty to estimates the emission flux accurately. On the other hand, more accurate prognostic models are complicated to set up and computationally expensive for operational use. The machine learning techniques are fast, but require extensive training datasets to be accurate.

This work proposed a surrogate method with operational capability to estimate the emission flux of GHGs from a mine pit in Canada fairly accurately and computationally fast, identifying diurnal and seasonal variations of the flux. The study used Weather Research and Forecasting (WRF) and machine learning (ML) methods. The in situ field measurements of methane mixing ratio were conducted in various spots of two mine pits, denominated old and new, as well as a tailings pond in May 2018 (S18), February–March 2019 (W19), July–August 2019 (S19), and October–November 2019 (F19) to provide tracer boundary conditions for WRF. Methane transport in WRF was simulated using a passive tracer dispersion method. The meteorological variables and emission flux from the WRF model were used as training and test datasets for various ML models. Four ML techniques (multi-layer perceptron (MLP) artificial neural network, gradient boosting (GBR), XGBOOST (XGB), and support vector machines (SVM)) were used as the main modeling methods. The long short-term memory (LSTM) technique was considered as a research model since, in the operational sense, it is difficult to measure the emission flux at any given time interval on a continuous basis to predict the emissions in the next time interval.

As the new mine was more active than the old mine, the emission flux out of the new mine was higher than the old mine and showed significant diurnal variations. The emission flux was greatly correlated with wind speed at 10 (m) above the surface ($S10$), but it exhibited lower or anti-correlations with other meteorological variables. The ML methods were more successful at the prediction of emission flux for the pond and the new mine compared to the old mine. The average bias and RMSE (Tonnes h^{-1}) were lower for the pond and new mine, while the old mine had the highest bias (Tonnes h^{-1}) and RMSE (Tonnes h^{-1}). Because of the more scheduled mining activity in the new mine, compared to the old mine and the pond, the average R^2 (-) of emission flux prediction by the deployed ML methods was higher than the two other predictions; however, the old mine and pond still exhibit an R^2 (-) higher than 0.8 (-). Considering the bias (Tonnes h^{-1}),

RMSE (Tonnes h^{-1}), and R^2 (-), the GBR and XGB algorithms performed better predictions for all the different emission sources and seasons. Combining all emission sources (mines and pond), seasons, and the four ML methods, the new WRF-ML modeling paradigm can predict the emission flux with an average Bias, RMSE, R^2 of -0.32 (Tonnes h^{-1}), 3.63 (Tonnes h^{-1}), and 0.88 (-), which reveals a potential interest in use of ML methods for predicting emission fluxes from the complex terrains of open-pit mines.

Some future work is possible. More accurate observations of methane mixing ratio with higher spatial resolution near the surface and over prolonged periods can potentially improve the WRF-ML prediction results. One of the main tasks which remains is whether WRF can conserve the mass flux of methane near the surface and at the exit boundary of its inner domain. If such conservation can be established, the proposed method can offer an accurate proxy for the emission flux from open-pit mines.

Author Contributions: Conceptualization, A.A.A., J.T. and B.G.; data curation, A.A.A., M.K.N. and J.T.; formal analysis, A.A.A., S.K. and J.T.; funding acquisition, A.A.A., B.G. and J.T.; methodology, A.A.A., M.K.N. and J.T.; visualization, S.K. and M.K.N.; writing—original draft, A.A.A., S.K. and J.T.; writing—review and editing, A.A.A., S.K., M.K.N., B.G. and J.T. All authors have read and agreed to the published version of the manuscript.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The Atmospheric Innovations Research (AIR) Laboratory at the University of Guelph may provide the datasets for this article pending the authorization of the data owners. For access, please visit http://www.aaa-scientists.com/ (accessed on 10 August 2021) or contact Principal Investigator Amir A. Aliabadi (aliabadi@uoguelph.ca).

Conflicts of Interest: The authors declare no conflict of interest.

Nomenclature

ΔA:	Area element (m^2)
Bias:	Bias of emission flux (Tonne h^{-1})
F:	Emission flux (Tonne h^{-1})
M:	Modeled emission flux (Tonne h^{-1})
MR:	Methane mixing ratio (ppm)
n:	Sample size (-)
O:	Reference emission flux (Tonne h^{-1})
PCC:	Pearson's correlation coefficient (-)
R^2:	Coefficient of determination (-)
$RH2$:	Relative humidity at 2 m (%)

RMSE:　　　Root mean square error of emission flux (Tonne h^{-1})
$\overline{S}$:　　　　Concentration of passive tracer (µg m^{-3})
SHF:　　　Surface Heat Flux (W m^{-2})
*S*10:　　　Wind speed at 10 m (m s^{-1})
SLP:　　　Sea level pressure (Pa)
*T*2:　　　Temperature at 2 m (K)
$\overline{U}$:　　　Average wind velocity along x direction (m s^{-1})
$\overline{V}$:　　　Average wind velocity along y direction (m s^{-1})
$\overline{W}$:　　　Average wind velocity along z direction (m s^{-1})

Abbreviations

ABL:　　　　Atmospheric Boundary Layer
CALPUFF:　　CALifornia PUFF Model
FCs:　　　　Flux Chambers
GTOPO 30s:　Global 30 Arc-Second
GDAS:　　　Global Data Assimilation System
GBR:　　　　Gradient Boosting
GHG:　　　　Greenhouse Gas
IDM:　　　　Inverse Dispersion Modelling
LS:　　　　　Land Surface
LSTM:　　　Long Short-Term Memory
LGR:　　　　Los Gatos Research
ML:　　　　Machine Learning
MODIS:　　　Moderate Resolution Imaging Spectroradiometer
MLP:　　　　Multi-Layer Perceptron
NCEP:　　　National Centers for Environmental Prediction
PBL:　　　　Planetary Boundary Layer
SRTM:　　　Shuttle Radar Topography Mission
SVM:　　　　Support Vector Machines
SL:　　　　　Surface Layer
TKE:　　　　Turbulence Kinetic Energy
WRF:　　　　Weather Research and Forecasting
XGB:　　　　XGBOOST

References

1. Betancourt-Torcat, A.; Elkamel, A.; Ricardez-Sandoval, L. A modeling study of the effect of carbon dioxide mitigation strategies, natural gas prices and steam consumption on the Canadian Oil Sands operations. *Energy* **2012**, *45*, 1018–1033. [CrossRef]
2. Rahman, M.M.; Canter, C.; Kumar, A. Greenhouse gas emissions from recovery of various North American conventional crudes. *Energy* **2014**, *74*, 607–617. [CrossRef]
3. Lan, X.; Talbot, R.; Laine, P.; Torres, A. Characterizing Fugitive Methane Emissions in the Barnett Shale Area Using a Mobile Laboratory. *Environ. Sci. Technol.* **2015**, *49*, 8139–8146. [CrossRef] [PubMed]
4. Hendrick, M.F.; Ackley, R.; Sanaie-Movahed, B.; Tang, X.; Phillips, N.G. Fugitive methane emissions from leak-prone natural gas distribution infrastructure in urban environments. *Environ. Pollut.* **2016**, *213*, 710–716. [CrossRef] [PubMed]
5. Ocko, I.B.; Hamburg, S.P.; Jacob, D.J.; Keith, D.W.; Keohane, N.O.; Oppenheimer, M.; Roy-Mayhew, J.D.; Schrag, D.P.; Pacala, S.W. Unmask temporal trade-offs in climate policy debates. *Science* **2017**, *356*, 492–493. [CrossRef]
6. Höglund-Isaksson, L. Global anthropogenic methane emissions 2005-2030: Technical mitigation potentials and costs. *Atmos. Chem. Phys.* **2012**, *12*, 9079–9096. [CrossRef]
7. Shindell, D.; Kuylenstierna, J.C.I.; Vignati, E.; van Dingenen, R.; Amann, M.; Klimont, Z.; Anenberg, S.C.; Muller, N.; Janssens-Maenhout, G.; Raes, F.; et al. Simultaneously Mitigating Near-Term Climate Change and Improving Human Health and Food Security. *Science* **2012**, *335*, 183–189. [CrossRef]
8. Zhang, Y.; Gautam, R.; Pandey, S.; Omara, M.; Maasakkers, J.D.; Sadavarte, P.; Lyon, D.; Nesser, H.; Sulprizio, M.P.; Varon, D.J.; et al. Quantifying methane emissions from the largest oil-producing basin in the United States from space. *Sci. Adv.* **2020**, *6*. [CrossRef]
9. Janzen, R.; Davis, M.; Kumar, A. Evaluating long-term greenhouse gas mitigation opportunities through carbon capture, utilization, and storage in the oil sands. *Energy* **2020**, *209*, 118364. [CrossRef]
10. Nimana, B.; Canter, C.; Kumar, A. Life cycle assessment of greenhouse gas emissions from Canada's oil sands-derived transportation fuels. *Energy* **2015**, *88*, 544–554. [CrossRef]

11. Di Lullo, G.; Zhang, H.; Kumar, A. Uncertainty in well-to-tank with combustion greenhouse gas emissions of transportation fuels derived from North American crudes. *Energy* **2017**, *128*, 475–486. [CrossRef]

12. Hmiel, B.; Petrenko, V.V.; Dyonisius, M.N.; Buizert, C.; Smith, A.M.; Place, P.F.; Harth, C.; Beaudette, R.; Hua, Q.; Yang, B.; et al. Preindustrial 14CH4 indicates greater anthropogenic fossil CH4 emissions. *Nature* **2020**, *578*, 409–412. [CrossRef] [PubMed]

13. Hempel, S.; Adolphs, J.; Landwehr, N.; Willink, D.; Janke, D.; Amon, T. Supervised Machine Learning to Assess Methane Emissions of a Dairy Building with Natural Ventilation. *Appl. Sci.* **2020**, *10*, 6938. [CrossRef]

14. Alvarez, R.A.; Zavala-Araiza, D.; Lyon, D.R.; Allen, D.T.; Barkley, Z.R.; Brandt, A.R.; Davis, K.J.; Herndon, S.C.; Jacob, D.J.; Karion, A.; et al. Assessment of methane emissions from the U.S. oil and gas supply chain. *Science* **2018**, *361*, 186–188. [CrossRef] [PubMed]

15. Rotach, M.W.; Zardi, D. On the boundary-layer structure over highly complex terrain: Key findings from MAP. *Q. J. R. Meteor. Soc.* **2007**, *133*, 937–948. [CrossRef]

16. Medeiros, L.E.; Fitzjarrald, D.R. Stable boundary layer in complex Terrain. Part I: Linking fluxes and intermittency to an average stability index. *J. Appl. Meteorol. Clim.* **2014**, *53*, 2196–2215. [CrossRef]

17. Medeiros, L.E.; Fitzjarrald, D.R. Stable boundary layer in complex terrain. Part II: Geometrical and sheltering effects on mixing. *J. Appl. Meteorol. Clim.* **2015**, *54*, 170–188. [CrossRef]

18. Soiket, M.I.; Oni, A.; Kumar, A. The development of a process simulation model for energy consumption and greenhouse gas emissions of a vapor solvent-based oil sands extraction and recovery process. *Energy* **2019**, *173*, 799–808. [CrossRef]

19. Byerlay, R.A.E.; Nambiar, M.K.; Nazem, A.; Nahian, M.R.; Biglarbegian, M.; Aliabadi, A.A. Measurement of land surface temperature from oblique angle airborne thermal camera observations. *Int. J. Remote Sens.* **2020**, *41*, 3119–3146. [CrossRef]

20. Nahian, M.R.; Nazem, A.; Nambiar, M.K.; Byerlay, R.; Mahmud, S.; Seguin, A.M.; Robe, F.R.; Ravenhill, J.; Aliabadi, A.A. Complex meteorology over a complex mining facility: Assessment of topography, land use, and grid spacing modifications in WRF. *J. Appl. Meteorol. Clim.* **2020**, *59*, 769–789. [CrossRef]

21. Nambiar, M.K.; Byerlay, R.A.E.; Nazem, A.; Nahian, M.R.; Moradi, M.; Aliabadi, A.A. A Tethered Air Blimp (TAB) for observing the microclimate over a complex terrain. *Geosci. Instrum. Meth.* **2020**, *9*, 193–211. [CrossRef]

22. Nambiar, M.K.; Robe, F.R.; Seguin, A.M.; Endsin, M.; Aliabadi, A.A. Diurnal and Seasonal Variation of Area-Fugitive Methane Advective Flux from an Open-Pit Mining Facility in Northern Canada Using WRF. *Atmosphere* **2020**, *11*, 1227. [CrossRef]

23. Kia, S.; Flesch, T.K.; Freeman, B.S.; Aliabadi, A.A. Atmospheric transport over open-pit mines: The effects of thermal stability and mine depth. *J. Wind Eng. Ind. Aerod.* **2021**, *214*, 104677. [CrossRef]

24. Kelly, E.N.; Schindler, D.W.; Hodson, P.V.; Short, J.W.; Radmanovich, R.; Nielsen, C.C. Oil sands development contributes elements toxic at low concentrations to the Athabasca River and its tributaries. *Proc. Natl. Acad. Sci. USA* **2010**, *107*, 16178–16183. [CrossRef] [PubMed]

25. Small, C.C.; Cho, S.; Hashisho, Z.; Ulrich, A.C. Emissions from oil sands tailings ponds: Review of tailings pond parameters and emission estimates. *J. Petrol. Sci. Eng.* **2015**, *127*, 490–501. [CrossRef]

26. Simpson, I.J.; Blake, N.J.; Barletta, B.; Diskin, G.S.; Fuelberg, H.E.; Gorham, K.; Huey, L.G.; Meinardi, S.; Rowland, F.S.; Vay, S.A.; et al. Characterization of trace gases measured over Alberta oil sands mining operations: 76 speciated C_2–C_{10} volatile organic compounds (VOCs), CO_2, CH_4, CO, NO, NO_2, NO_y, O_3 and SO_2. *Atmos. Chem. Phys.* **2010**, *10*, 11931–11954. [CrossRef]

27. Clements, C.B.; Whiteman, C.D.; Horel, J.D. Cold-Air-Pool Structure and Evolution in a Mountain Basin: Peter Sinks, Utah. *J. Appl. Meteorol.* **2003**, *42*, 752–768. [CrossRef]

28. Whiteman, C.D.; Haiden, T.; Pospichal, B.; Eisenbach, S.; Steinacker, R. Minimum Temperatures, Diurnal Temperature Ranges, and Temperature Inversions in Limestone Sinkholes of Different Sizes and Shapes. *J. Appl. Meteorol.* **2004**, *43*, 1224–1236. [CrossRef]

29. Bhowmick, T.; Bandopadhyay, S.; Ghosh, T. Three-dimensional CFD modeling approach to approximate air pollution conditions in high-latitude open-pit mines. *WIT Trans. Built. Environ.* **2015**, *168*, 741–753. [CrossRef]

30. Silvester, S.A.; Lowndes, I.S.; Hargreaves, D.M. A computational study of particulate emissions from an open pit quarry under neutral atmospheric conditions. *Atmos. Environ.* **2009**, *43*, 6415–6424. [CrossRef]

31. Liggio, J.; Li, S.M.; Staebler, R.M.; Hayden, K.; Darlington, A.; Mittermeier, R.L.; O'Brien, J.; McLaren, R.; Wolde, M.; Worthy, D.; et al. Measured Canadian oil sands CO_2 emissions are higher than estimates made using internationally recommended methods. *Nat. Commun.* **2019**, *10*, 1863. [CrossRef] [PubMed]

32. Rochette, P.; Eriksen-Hamel, N.S. Chamber measurements of soil Nitrous Oxide flux: Are absolute values reliable? *Soil Sci. Soc. Am. J.* **2008**, *72*, 331–342. [CrossRef]

33. Bai, M.; Suter, H.; Lam, S.K.; Flesch, T.K.; Chen, D. Comparison of slant open-path flux gradient and static closed chamber techniques to measure soil N_2O emissions. *Atmos. Meas. Tech.* **2019**, *12*, 1095–1102. [CrossRef]

34. Meyers, T.P.; Hall, M.E.; Lindberg, S.E.; Kim, K. Use of the modified Bowen-ratio technique to measure fluxes of trace gases. *Atmos. Environ.* **1996**, *30*, 3321–3329. [CrossRef]

35. Foken, T.; Leuning, R.; Oncley, S.R.; Mauder, M.; Aubinet, M. Eddy Covariance. In *Eddy Covariance*; Aubinet, M., Vesala, T., Papale, D., Eds.; Springer: Dordrecht, The Netherlands, 2012. [CrossRef]

36. You, Y.; Staebler, R.M.; Moussa, S.G.; Beck, J.; Mittermeier, R.L. Methane emissions from an oil sands tailings pond: A quantitative comparison of fluxes derived by different methods. *Atmos. Meas. Tech.* **2021**, *14*, 1879–1892. [CrossRef]

37. Gordon, M.; Li, S.M.; Staebler, R.; Darlington, A.; Hayden, K.; O'Brien, J.; Wolde, M. Determining air pollutant emission rates based on mass balance using airborne measurement data over the Alberta oil sands operations. *Atmos. Meas. Tech.* **2015**, *8*, 3745–3765. [CrossRef]

38. Conley, S.; Faloona, I.; Mehrotra, S.; Suard, M.; Lenschow, D.H.; Sweeney, C.; Herndon, S.; Schwietzke, S.; Pétron, G.; Pifer, J.; et al. Application of Gauss's theorem to quantify localized surface emissions from airborne measurements of wind and trace gases. *Atmos. Meas. Tech.* **2017**, *10*, 3345–3358. [CrossRef]

39. Baray, S.; Darlington, A.; Gordon, M.; Hayden, K.L.; Leithead, A.; Li, S.M.; Liu, P.S.K.; Mittermeier, R.L.; Moussa, S.G.; O'Brien, J.; Staebler, R.; et al. Quantification of methane sources in the Athabasca Oil Sands Region of Alberta by aircraft mass balance. *Atmos. Chem. Phys.* **2018**, *18*, 7361–7378. [CrossRef]

40. Flesch, T.K.; Wilson, J.D.; Harper, L.A.; Crenna, B.P. Estimating gas emissions from a farm with an inverse-dispersion technique. *Atmos. Environ.* **2005**, *39*, 4863–4874. [CrossRef]

41. Pernini, T.G.; Zaccheo, T.S.; Dobler, J.; Blume, N. Estimating oil sands emissions using horizontal path-integrated column measurements. *Atmos. Meas. Tech.* **2022**, *15*, 225–240. [CrossRef]

42. Scire, J.S.; Strimaitis, D.G.; Yamartino, R.J. *A User's Guide for the CALPUFF Dispersion Model (Version 5)*; Technical Report; Earth Tech, Inc.: Concord, MA, USA, 2000.

43. Jittra, N.; Pinthong, N.; Thepanondh, S. Performance evaluation of AERMOD and CALPUFF air dispersion models in industrial complex area. *Air Soil Water Res.* **2015**, *8*, ASWR–S32781. [CrossRef]

44. Arregocés, H.; Rojano, R.; Restrepo, G.; Angulo, L. Using CALPUFF to determine the environmental impact of a coal mine open pit. *WIT Trans. Ecol. Environ.* **2016**, *207*, 55–66. [CrossRef]

45. Holnicki, P.; Kałuszko, A.; Nahorski, Z.; Stankiewicz, K.; Trapp, W. Air quality modeling for Warsaw agglomeration. *Arch. Environ. Prot.* **2017**, *43*, 48–64. [CrossRef]

46. Oleniacz, R.; Rzeszutek, M. Intercomparison of the CALMET/CALPUFF modeling system for selected horizontal grid resolutions at a local scale: A case study of the MSWI Plant in Krakow, Poland. *Appl. Sci.* **2018**, *8*, 2301. [CrossRef]

47. Rzeszutek, M. Parameterization and evaluation of the CALMET/CALPUFF model system in near-field and complex terrain-Terrain data, grid resolution and terrain adjustment method. *Sci. Total Environ.* **2019**, *689*, 31–46. [CrossRef]

48. Sówka, I.; Paciorek, M.; Skotak, K.; Kobus, D.; Zathey, M.; Klejnowski, K. The Analysis of the Effectiveness of Implementing Emission Reduction Measures in Improving Air Quality and Health of the Residents of a Selected Area of the Lower Silesian Voivodship. *Energies* **2020**, *13*, 4001. [CrossRef]

49. Cox, R.M.; Sontowski, J.; 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.* **2005**, *12*, 329–341. [CrossRef]

50. Cui, H.; Yao, R.; Chen, L.; Lv, M.; Xin, C.; Wu, Q. Field study of atmospheric boundary layer observation in a hilly Gobi Desert region and comparison with the CALMET/CALPUFF model. *Atmos. Environ.* **2020**, *235*, 117576. [CrossRef]

51. Wang, W.; Shaw, W.J.; Seiple, T.E.; Rishel, J.P.; Xie, Y. An evaluation of a diagnostic wind model (CALMET). *J. Appl. Meteorol. Clim.* **2008**, *47*, 1739–1756. [CrossRef]

52. Qiu, X.; Cheng, I.; Yang, F.; Horb, E.; Zhang, L.; Harner, T. Emissions databases for polycyclic aromatic compounds in the Canadian Athabasca oil sands region—Development using current knowledge and evaluation with passive sampling and air dispersion modelling data. *Atmos. Chem. Phys.* **2018**, *18*, 3457–3467. [CrossRef]

53. Gilliam, R.C.; Pleim, J.E. Performance assessment of new land surface and planetary boundary layer physics in the WRF-ARW. *J. Appl. Meteorol. Clim.* **2010**, *49*, 760–774. [CrossRef]

54. Nossent, J.; Elsen, P.; Bauwens, W. Sobol' sensitivity analysis of a complex environmental model. *Environ. Modell. Softw.* **2011**, *26*, 1515–1525. [CrossRef]

55. Jiménez, P.A.; Dudhia, J. Improving the representation of resolved and unresolved topographic effects on surface wind in the WRF model. *J. Appl. Meteorol. Clim.* **2012**, *51*, 300–316. [CrossRef]

56. Lorente-Plazas, R.; Jiménez, P.A.; Dudhia, J.; Montávez, J.P. Evaluating and improving the impact of the atmospheric stability and orography on surface winds in the WRF model. *Mon. Weather Rev.* **2016**, *144*, 2685–2693. [CrossRef]

57. Cerlini, P.B.; Emanuel, K.A.; Todini, E. Orographic effects on convective precipitation and space-time rainfall variability: Preliminary results. *Hydrol. Earth Syst. Sci.* **2005**, *9*, 285–299. [CrossRef]

58. Rong, Z.; Dahai, X. Multi-scale turbulent planetary boundary layer parameterization in mesoscale numerical simulation. *J. Appl. Meteorol. Sci.* **2004**, *15*, 543–555.

59. Houze Jr., R.A. Orographic effects on precipitating clouds. *Rev. Geophys.* **2012**, *50*, RG1001. [CrossRef]

60. Yáñez-Morroni, G.; Gironás, J.; Caneo, M.; Delgado, R.; Garreaud, R. Using the Weather Research and Forecasting (WRF) model for precipitation forecasting in an Andean region with complex topography. *Atmosphere* **2018**, *9*, 304. [CrossRef]

61. Taylor, C.M.; Lambin, E.F.; Stephenne, N.; Harding, R.J.; Essery, R.L. The influence of land use change on climate in the Sahel. *J. Clim.* **2002**, *15*, 3615–3629. [CrossRef]

62. De Meij, A.; Vinuesa, J. Impact of SRTM and Corine Land Cover data on meteorological parameters using WRF. *Atmos. Res.* **2014**, *143*, 351–370. [CrossRef]

63. Zhang, C.; Lin, H.; Chen, M.; Yang, L. Scale matching of multiscale digital elevation model (DEM) data and the Weather Research and Forecasting (WRF) model: A case study of meteorological simulation in Hong Kong. *Arab. J. Geosci.* **2014**, *7*, 2215–2223. [CrossRef]

64. Blaylock, B.K.; Horel, J.D.; Crosman, E.T. Impact of Lake Breezes on Summer Ozone Concentrations in the Salt Lake Valley. *J. Appl. Meteorol. Climatol.* **2017**, *56*, 353–370. [CrossRef]

65. Bhimireddy, S.R.; Bhaganagar, K. Short-term passive tracer plume dispersion in convective boundary layer using a high-resolution WRF-ARW model. *Atmos. Pollut. Res.* **2018**, *9*, 901–911. [CrossRef]

66. Grell, G.A.; Peckham, S.E.; Schmitz, R.; McKeen, S.A.; Frost, G.; Skamarock, W.C.; Eder, B. Fully coupled online chemistry within the WRF model. *Atmos. Environ.* **2005**, *39*, 6957–6975. [CrossRef]

67. Beck, V.; Gerbig, C.; Koch, T.; Bela, M.M.; Longo, K.M.; Freitas, S.R.; Kaplan, J.O.; Prigent, C.; Bergamaschi, P.; Heimann, M. WRF-Chem simulations in the Amazon region during wet and dry season transitions: Evaluation of methane models and wetland inundation maps. *Atmos. Chem. Phys.* **2013**, *13*, 7961–7982. [CrossRef]

68. Ahmadov, R.; McKeen, S.; Trainer, M.; Banta, R.; Brewer, A.; Brown, S.; Edwards, P.M.; de Gouw, J.A.; Frost, G.J.; Gilman, J.; et al. Understanding high wintertime ozone pollution events in an oil- and natural gas-producing region of the western US. *Atmos. Chem. Phys.* **2015**, *15*, 411–429. [CrossRef]

69. Barkley, Z.R.; Lauvaux, T.; Davis, K.J.; Deng, A.; Miles, N.L.; Richardson, S.J.; Cao, Y.; Sweeney, C.; Karion, A.; Smith, M.; et al. Quantifying methane emissions from natural gas production in north-eastern Pennsylvania. *Atmos. Chem. Phys.* **2017**, *17*, 13941–13966. [CrossRef]

70. Leukauf, D.; Gohm, A.; Rotach, M.W. Quantifying horizontal and vertical tracer mass fluxes in an idealized valley during daytime. *Atmos. Chem. Phys.* **2016**, *16*, 13049–13066. [CrossRef]

71. Karion, A.; Lauvaux, T.; Lopez Coto, I.; Sweeney, C.; Mueller, K.; Gourdji, S.; Angevine, W.; Barkley, Z.; Deng, A.; Andrews, A.; et al. Intercomparison of atmospheric trace gas dispersion models: Barnett Shale case study. *Atmos. Chem. Phys.* **2019**, *19*, 2561–2576. [CrossRef]

72. Zhao, X.; Marshall, J.; Hachinger, S.; Gerbig, C.; Frey, M.; Hase, F.; Chen, J. Analysis of total column CO_2 and CH_4 measurements in Berlin with WRF-GHG. *Atmos. Chem. Phys.* **2019**, *19*, 11279–11302. [CrossRef]

73. Cortes, C.; Vapnik, V. Support-vector networks. *Mach. Learn.* **1995**, *20*, 273–297. [CrossRef]

74. Maier, H.R.; Dandy, G.C. The use of artificial neural networks for the prediction of water quality parameters. *Water Resour. Res.* **1996**, *32*, 1013–1022. [CrossRef]

75. Piryonesi, S.M.; El-Diraby, T.E. Data analytics in asset management: Cost-effective prediction of the pavement condition index. *J. Infrastruct. Syst.* **2020**, *26*, 04019036. [CrossRef]

76. Hochreiter, S.; Schmidhuber, J. Long short-term memory. *Neural Comput.* **1997**, *9*, 1735–1780. [CrossRef] [PubMed]

77. Freeman, B.S.; Taylor, G.; Gharabaghi, B.; Thé, J. Forecasting air quality time series using deep learning. *J. Air Waste Manag.* **2018**, *68*, 866–886. [CrossRef]

78. Tabrizi, S.E.; Xiao, K.; Van Griensven, J.; Saad, M.; Farghaly, H.; Yang, S.X.; Gharabaghi, B. Hourly Road Pavement Surface Temperature Forecasting Using Deep Learning Modelsa. *J. Hydrol.* **2021**, *603*, 126877. [CrossRef]

79. Skrobek, D.; Krzywanski, J.; Sosnowski, M.; Kulakowska, A.; Zylka, A.; Grabowska, K.; Ciesielska, K.; Nowak, W. Prediction of Sorption Processes Using the Deep Learning Methods (Long Short-Term Memory). *Energies* **2020**, *13*, 6601. [CrossRef]

80. Ashraf, W.M.; Uddin, G.M.; Arafat, S.M.; Krzywanski, J.; Xiaonan, W. Strategic-level performance enhancement of a 660 MWe supercritical power plant and emissions reduction by AI approach. *Energy Convers. Manag.* **2021**, *250*, 114913. [CrossRef]

81. Ashraf, W.M.; Uddin, G.M.; Farooq, M.; Riaz, F.; Ahmad, H.A.; Kamal, A.H.; Anwar, S.; El-Sherbeeny, A.M.; Khan, M.H.; Hafeez, N.; et al. Construction of Operational Data-Driven Power Curve of a Generator by Industry 4.0 Data Analytics. *Energies* **2021**, *14*, 1227. [CrossRef]

82. Powers, J.G.; Klemp, J.B.; Skamarock, W.C.; Davis, C.A.; Dudhia, J.; Gill, D.O.; Coen, J.L.; Gochis, D.J.; Ahmadov, R.; Peckham, S.E.; et al. The weather research and forecasting model: Overview, system efforts, and future directions. *B. Am. Meteorol. Soc.* **2017**, *98*, 1717–1737. [CrossRef]

83. Janjić, Z.I. The Step-Mountain Eta Coordinate Model: Further Developments of the Convection, Viscous Sublayer, and Turbulence Closure Schemes. *Mon. Weather Rev.* **1994**, *122*, 927–945. [CrossRef]

84. Thompson, G.; Field, P.R.; Rasmussen, R.M.; Hall, W.D. Explicit Forecasts of Winter Precipitation Using an Improved Bulk Microphysics Scheme. Part II: Implementation of a New Snow Parameterization. *Mon. Weather Rev.* **2008**, *136*, 5095–5115. [CrossRef]

85. Iacono, M.J.; Delamere, J.S.; Mlawer, E.J.; Shephard, M.W.; Clough, S.A.; Collins, W.D. Radiative forcing by long-lived greenhouse gases: Calculations with the AER radiative transfer models. *J. Geophys. Res. Atmos.* **2008**, *113*, D13103. [CrossRef]

86. Zhang, C.; Wang, Y.; Hamilton, K. Improved Representation of Boundary Layer Clouds over the Southeast Pacific in ARW-WRF Using a Modified Tiedtke Cumulus Parameterization Scheme. *Mon. Weather Rev.* **2011**, *139*, 3489–3513. [CrossRef]

87. Lupaşcu, A.; Iriza, A.; Dumitrache, R. Using a high resolution topographic data set and analysis of the impact on the forecast of meteorological parameters. *Rom. Rep. Phys.* **2015**, *67*, 653–664.

88. Friedl, M.A.; Sulla-Menashe, D.; Tan, B.; Schneider, A.; Ramankutty, N.; Sibley, A.; Huanga, X. MODIS Collection 5 global land cover: Algorithm refinements and characterization of new datasets. *Remote Sens. Environ.* **2010**, *114*, 168–182. [CrossRef]

89. Subin, Z.M.; Riley, W.J.; Mironov, D. An improved lake model for climate simulations: Model structure, evaluation, and sensitivity analyses in CESM1. *J. Adv. Model. Earth Syst.* **2012**, *4*, M02001. [CrossRef]

90. Gu, H.; Jin, J.; Wu, Y.; Ek, M.B.; Subin, Z.M. Calibration and validation of lake surface temperature simulations with the coupled WRF-lake model. *Clim. Chang.* **2015**, *129*, 471–483. [CrossRef]

91. Allen, G.; Hollingsworth, P.; Kabbabe, K.; Pitt, J.R.; Mead, M.I.; Illingworth, S.; Roberts, G.; Bourn, M.; Shallcross, D.E.; Percival, C.J. The development and trial of an unmanned aerial system for the measurement of methane flux from landfill and greenhouse gas emission hotspots. *Waste Manag.* **2019**, *87*, 883–892. [CrossRef]

92. Gibbs, J.A.; Fedorovich, E.; van Eijk, A.M.J. Evaluating Weather Research and Forecasting (WRF) Model Predictions of Turbulent Flow Parameters in a Dry Convective Boundary Layer. *J. Appl. Meteorol. Climatol.* **2011**, *50*, 2429–2444. [CrossRef]

93. Xue, L.; Chu, X.; Rasmussen, R.; Breed, D.; Boe, B.; Geerts, B. The Dispersion of Silver Iodide Particles from Ground-Based Generators over Complex Terrain. Part II: WRF Large-Eddy Simulations versus Observations. *J. Appl. Meteorol. Climatol.* **2014**, *53*, 1342–1361. [CrossRef]

94. Oertel, C.; Matschullat, J.; Zurba, K.; Zimmermann, F.; Erasmi, S. Greenhouse gas emissions from soils—A review. *Geochemistry* **2016**, *76*, 327–352. [CrossRef]

95. Hunter, J.D. Matplotlib: A 2D graphics environment. *Comput. Sci. Eng.* **2007**, *9*, 90–95. [CrossRef]

96. Waskom, M.L. seaborn: Statistical data visualization. *J. Open Source Softw.* **2021**, *6*, 3021. [CrossRef]

97. Pedregosa, F.; Varoquaux, G.; Gramfort, A.; Michel, V.; Thirion, B.; Grisel, O.; Blondel, M.; Prettenhofer, P.; Weiss, R.; Dubourg, V.; et al. Scikit-learn: Machine Learning in Python. *J. Mach. Learn. Res.* **2011**, *12*, 2825–2830.

98. Buitinck, L.; Louppe, G.; Blondel, M.; Pedregosa, F.; Mueller, A.; Grisel, O.; Niculae, V.; Prettenhofer, P.; Gramfort, A.; Grobler, J.; et al. API design for machine learning software: Experiences from the scikit-learn project. In Proceedings of the European Conference on Machine Learning and Principles and Practice of Knowledge Discovery In Databases: Languages for Data Mining and Machine Learning, Prague, Czech Republic, 23–27 September, 2013; pp. 108–122.

99. Chen, T.; Guestrin, C. XGBoost: A Scalable Tree Boosting System. In Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, San Francisco, CA, USA, 13–17 August 2016; pp. 785–794. [CrossRef]