Multitask Learning for Estimating Power Plant Greenhouse Gas Emissions from Satellite Imagery

Joëlle Hanna
AIML Lab, School of Computer Science
University of St.Gallen
joelle.hanna@unisg.ch

Michael Mommert
AIML Lab, School of Computer Science
University of St.Gallen
michael.mommert@unisg.ch

Linus Scheibenreif
AIML Lab, School of Computer Science
University of St.Gallen
linus.scheibenreif@unisg.ch

Damian Borth
AIML Lab, School of Computer Science
University of St.Gallen
damian.borth@unisg.ch

Abstract

The burning of fossil fuels produces large amounts of carbon dioxide (CO_2), a major Greenhouse Gas (GHG) and a main driver of Climate Change. Quantifying GHG emissions is crucial for accurate predictions of climate effects and to enforce emission trading schemes. The reporting of such emissions is only required in some countries, resulting in insufficient global coverage. In this work, we propose an end-to-end method to predict power generation rates for fossil fuel power plants from satellite images based on which we estimate GHG emission rates. We present a multitask deep learning approach able to simultaneously predict: (i) the pixel-area covered by plumes from a single satellite image of a power plant, (ii) the type of fired fuel, and (iii) the power generation rate. We then convert the predicted power generation rate into estimates for the rate at which CO_2 is being emitted. Experimental results show that our model approach allows us to estimate the power generation rate of a power plant to within 139 MW (MAE, for a mean sample power plant capacity of 1177 MW) from a single satellite image and CO_2 emission rates to within 311 t/h. This multitask learning approach improves the power generation estimation MAE by 39% compared to a baseline single-task network trained on the same dataset.

1 Introduction

Despite a recent decrease in their use in most European countries, fossil fuels like coal, oil and gas, still account for 71% of the continent's energy production [Eurostat, 2021]. This popularity comes at a dangerously high environmental cost that affects humanity in the long term: burning fossil fuels leads to air and water pollution, and constitutes the main driver of climate change.

Quantifying GHG emissions from individual industrial sites typically requires the use of dedicated measuring devices that report detailed emission information, as may be mandated by environmental protection guidelines. Such data may be used to enforce environmental protection regulations or pollutant certificate trading schemes. Unfortunately, reporting requirements differ between countries, resulting in a highly heterogeneous coverage of GHG emission reportings, globally.

In this work, we aim to estimate GHG emission rates for fossil fuel power plants at a given time through observations of the emitted plumes from Earth-observing satellites. GHG emission rates are correlated to power generation rates and as such can be related to the extent of plumes emerging from

Tackling Climate Change with Machine Learning workshop at NeurIPS 2021.

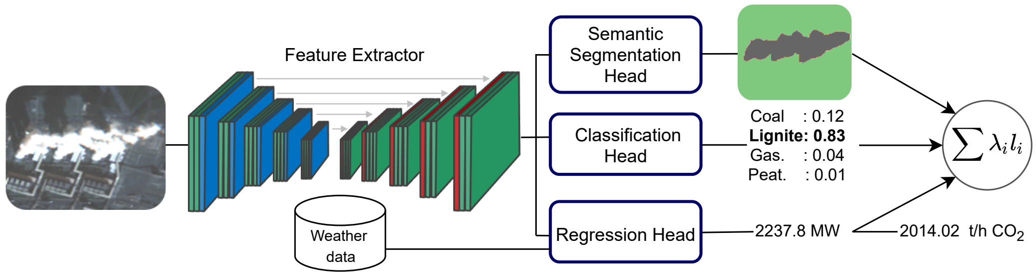

Figure 1: Diagram of the proposed multitask learning method. The model takes as input a multi-spectral satellite image and learns simultaneously three tasks utilizing a U-Net backbone: (i) semantic segmentation of plumes, (ii) classification of type of fired fuel, and (iii) regression with respect to power generation rate.

smokestacks or cooling towers of power plants. Our definition of plumes include plumes of steam released from cooling towers, and, to a smaller scale, plumes of smoke released from smokestacks. We propose a novel multi-task learning approach to primarily predict the power generation rate from a satellite image of a power plant (treated as a regression problem), together with the type of fired fuel (treated as a classification problem) and plume footprint (treated as a segmentation problem). We incorporate external weather data to consider environmental conditions. Using the predicted power generation rate, we can estimate the amount of emitted CO_2 at a given time based on an empirical relation. Both the predicted power generation rate and the CO_2 emission rate predicted in this work are of interest to various stakeholders and inform efforts to minimize the effects of Climate Change. Our contribution is threefold: (1) we compile a data set of active power plants in Europe together with their plume segmentation maps and the corresponding actual power generation rate, (2) we propose a multitask learning approach able to simultaneously segment plumes, predict the type of fired fuel as well as the power generation rate – out-performing single-task approaches for all these tasks – and (3) we estimate CO_2 emission rates using the predicted power generation rates and derived emission factors.

2 Related Work

To the best of our knowledge, little work has been done concerning the estimation of GHG emissions from fossil fuel power plants based on observational data. The estimation of the power plant generation has mostly been restricted to extracting statistics from available annual data. Yin et al. [2020] estimate the deviation of each plant from the average generation of other similar plants based on detailed information on plant-level (such as its size, fuel-type, or installed capacity) and environmental factors using machine learning algorithms such as gradient boosting trees.

Gray et al. [2020] estimate the utilisation of fossil fuel power plants by identifying plumes in satellite image data then counting the number of hours it is actually generating power over a year. Finally, they compute the capacity factor by dividing the annual generation by an empirical relationship between the annual mean plume area and mean hourly generation rate. Couture et al. [2020] showed that it was possible to achieve high quality estimates of power plant behavior, i.e. whether it was on or off, using satellite images. Going one step further, Mommert et al. [2020] successfully segment smoke plumes over industrial sites and distinguish them from natural clouds and show that classifying different types of power plants is possible with high confidence [Mommert et al., 2021].

In contrast to previous works, we estimate the power generation rate (and from those CO_2 emission rate estimates) directly from satellite images without any prior assumptions, allowing our pipeline to be applied on a global scale. We choose a multitask approach that has proven successful in other remote sensing applications [Bischke et al., 2019].

3 Dataset

In this work, we use remote imaging data of fossil fuel power plants taken by ESA's Sentinel-2 Earth-observing satellites. Extending the dataset from Mommert et al. [2020], we acquired geographic

Table 1: Test set multitask vs. singletask baseline performance for plume segmentation (seg.), power generation regression (reg.) and fuel type classification (cls.).

Loss	Task Weights (λ_i) Seg.	Reg.	Cls.	Seg. IoU	Reg. MAE / R^2	Cls. accuracy
Seg. only	1	0	0	0.640	-	-
Reg. only	0	1	0	-	225 / 0.66	-
Cls. only	0	0	1	-	-	0.775
Reg. + Seg.	0.4	0.6	0	0.643	145 / 0.81	-
Reg. + Cls.	0	0.6	0.4	-	151 / 0.81	0.779
All tasks	0.15	0.7	0.15	**0.668**	**139 / 0.83**	**0.853**

Table 2: Ablation study on impact of weather variables on the regression task performance on the test set and using the multitask model.

Temperature	Humidity	Wind	MAE	R^2
✗	✗	✗	169	0.74
✓	✗	✗	148	0.77
✗	✓	✗	142	0.81
✗	✗	✓	147	0.79
✗	✓	✓	145	0.81
✓	✓	✓	**139**	**0.83**

coordinates of 300 European power plants based on the *Joint Research Centre (JRC) Open Power Plants Database* [Kanellopoulos et al., 2019] for which power generation data [Entsoe, 2021] is available. For each site we retrieved Sentinel-2 images taken during 2020 that we manually annotated.

Our final dataset [1] contains 1639 satellite observations of 146 different fossil fuel power plants with their segmentation label, that we carefully divide into train (80%) and test (20%) sets making sure not to include the same site in more than one set. We supplement our processed satellite image data with contemporaneous weather data (temperature, relative humidity and wind speed) from the *ERA-5 data set* [Hersbach et al., 2020]. The installed capacity of these power plants ranges from 29 MW to 5230 MW with a mean of 1177 MW. Our sample includes power plants that use 4 different types of fuel: hard coal (41%), gas (29%), lignite (29%), and peat ($\leq$ 1 %) .

For each observation, our dataset contains a Sentinel-2 satellite image with its segmentation map, the type of fired fuel, the actual power generation rate at the corresponding timestamp, and weather data.

4 Approach

Our approach consists of a multi-task deep learning model (see Figure 1) that takes as input a multi-spectral satellite image centered on a power station to predict primarily its actual power generation output. We added two other tasks: semantic segmentation of plumes and classification of type of fired fuel, to boost the performance of the desired primary task. This is achieved by pushing the network towards learning a robust representation that generalizes well to different related tasks.

Feature Extractor The first part of the proposed method is a U-Net [Ronneberger et al., 2015] used as a feature extractor. Its architecture consists of a contracting and an expanding path. We take advantage of hard parameter sharing through the entire U-Net backbone leading to a representation that is shared between all the tasks, reducing the risk of overfitting.

Specific Tasks The three tasks' branches receive the shared representation and specialise on one task. The *first task* segments plumes. It consists of a convolutional layer which outputs a segmentation map of the same height and width as the input image. The *second task* uses a convolutional layer followed by a fully connected one and a softmax function to detect the type of fired fuel: hard coal, lignite, gas, or peat (see Section 3). *Estimating actual power generation*, the third and most important task, is done using a convolutional layer followed by 3 blocks of batch normalization, fully connected layer and ReLu for the activation function. Our targets for the regression task are power generation rates reported through Entsoe [2021] that are contemporaneous with the Sentinel-2 observations and available for each power plant in our sample.

Losses and Metrics We perform single task versus multitask training. Each of the three tasks is learned using a specific loss function. For the plume segmentation we use the binary cross entropy loss, for the fuel classification task we use the cross entropy loss, and, for the regression task, we use the L1 loss. In the multitask learning setup, we scale each loss and weight them (λ_i); weights were optimized as hyper-parameters with a focus on the regression task. We measure the performance of the segmentation task with IoU, fuel classification task with accuracy, and the regression task with Mean Absolute Error (MAE), Root Mean Square Error (RMSE) and Mean Absolute Percentage Error (MAPE)

[1] The code base for this work is available at github.com/HSG-AIML/RemoteSensingCO2Estimation; the complete data set is available at zenodo.org

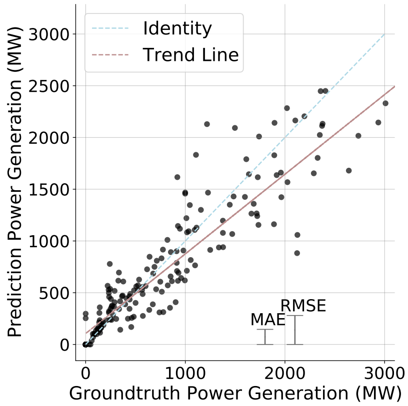

Figure 2: Test set power generation rates.

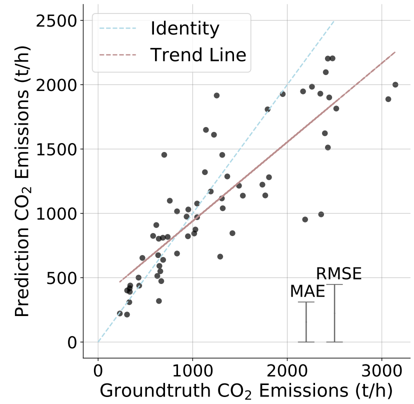

Figure 3: Test set CO_2 emission rates.

CO_2 Emission Estimation CO_2 emission rates are directly correlated to power generation rates and depend on other plant-specific properties, like the fuel type. Based on reported annual emissions and power generation for a selection of hard coal and lignite power plants [Gutmann et al., 2014], we derive emission factors (0.9 ± 0.1 t CO_2/MWh for lignite and 0.6 ± 0.1 t CO_2/MWh for hard coal) to convert our predicted power generation rates to CO_2 emission rates. Figure 3 compares the predicted and ground-truth emission rates for 65 observations of 6 different hard coal and lignite power plants for which we know plant-specific emission factors [Gutmann et al., 2014]. Ground-truth emission rates are derived for these 6 plants based on the plant-specific emission factors and their ground-truth power generation rates.

5 Experimental Results

We perform a hyper-parameter search for the multitask architecture outlined in Section 4 and Figure 1, leading to the optimized hyper-parameters and resulting in the evaluation metrics shown in Table 1. In Figure 2 we display the predicted power generation rate versus the ground-truth for our test sample based on the multi-task approach (”All tasks“ in Table 1), resulting in an R^2 of 0.83, an MAE of 139 MW, an RMSE of 261 MW and a MAPE of 19 %. CO_2 emission rates are estimated from predicted power generation rates utilizing fuel-specific emission factors (see Section 4 for 6 plants in our test set for which such ground-truth information is available; the derived MAE for CO_2 is 311 t/h and the MAPE is 34 %.r We furthermore perform an ablation study with respect to weather variables that supplement the regression task (Table 2).

6 Discussion

Our multitask approach enables us to predict power generation rates within 139 MW (MAE) and estimate CO_2 emission rates within 311 t/h (MAE) for our test set power plants. While the regression performance is likely to improve with more training data, we found that the segmentation task performance is already on-par with human labeling efforts. Based on Table 1 we note that the multi-task model outperforms the single-task models trained on the same data: we observe a relative improvement of 5% on the segmentation IoU, 39 % on the regression MAE, 26 % on the regression R^2, and 10% on the classification accuracy.

We discuss some limitations of our approach. Our plume definition, on which the power generation rate regression is based, does not distinguish between smoke plumes resulting from the combustion of fossil fuels and steam plumes from cooling devices. The effect of this simplification on our results is negligible since we find that most plumes in our data set emerge from cooling towers. Furthermore, emission factors utilized in the conversion from power generation to CO_2 emission are empirically based on fuel-specific statistical considerations. While uncertainties introduced by this simplification are propagated into our CO_2 emission rates, power plant-specific emission factors might lead to better results, but are only available for select power plants.

7 Conclusion

This work shows that it is possible to predict power generation rates with high confidence and CO_2 emission rates with some confidence from single Sentinel-2 satellite images. Our proposed deep multitask architecture was trained on a combination of three tasks and experiments confirmed that auxiliary tasks can indeed boost the network performance. Our model is able to predict power generation rates from individual images with R^2=0.83 or within 139 MW (MAE) and CO_2 emission rates within 311 t/h. For the average power plant in our sample (1177 MW capacity), our generation predictions are of high confidence. Our method is thus able to contribute to the estimation of CO_2 emission rates from power plants on a global scale.

8 Acknowledgment

This work utilizes satellite imagery provided by the European Copernicus Programme, power generation data provided by ENTSO-E and weather data provided by ECMWF Integrated Forecasting System. The authors would like to thank Entsorgung St. Gallen and Kehrichtheizkraftwerk St. Gallen for their collaboration on this project.

References

B. Bischke, P. Helber, J. Folz, D. Borth, and A. Dengel. Multi-task learning for segmentation of building footprints with deep neural networks. *2019 IEEE International Conference on Image Processing (ICIP)*, pages 1480–1484, 2019.

H. D. Couture, J. O'Connor, G. Mitchell, I. Söldner-Rembold, D. D'souza, K. Karra, K. Zhang, A. R. Kargar, T. Kassel, B. Goldman, D. Tyrrell, W. Czerwinski, A. Talekar, and C. McCormick. Towards tracking the emissions of every power plant on the planet. *ArXiv*, 2020. Tackling Climate Change with Machine Learning Workshop at NeurIPS 2020.

Entsoe. Actual generation per generation unit table, 2021. `https://transparency.entsoe.eu/generation/r2/actualGenerationPerGenerationUnit/show`.

Eurostat. Share of fossil fuels in gross available energy. `https://ec.europa.eu/eurostat/fr/web/products-eurostat-news/-/ddn-20210204-1`, 2021. (accessed: 2021-03-20).

M. Gray, L. Watson, S. Ljungwaldh, and E. Morris. Using satellite imagery to estimate the utilisation of fossil fuel power plants, 2020. `http://www.indiaenvironmentportal.org.in/files/file/Statellite_report.pdf`.

K. Gutmann, J. Huscher, D. Urbaniak, A. White, and C. Schaible. Europe's dirty 30: How the eu's coal-fired power plants are undermining its climate efforts. Published by CAN Europe, WWF European Policy Office, HEAL, the EEB and Climate Alliance Germany, in Brussels, Belgium; `http://awsassets.panda.org/downloads/dirty_30_report_finale.pdf`, 2014.

H. Hersbach, B. Bell, P. Berrisford, G. Biavati, A. Horányi, M. S. J, J. Nicolas, C. Peubey, R. Radu, I. Rozum, D. Schepers, A. Simmons, C. Soci, D. Dee, and J.-N. Thépaut. Era5 hourly data on pressure levels from 1979 to present. copernicus climate change service (c3s) climate data store (cds), 2020. `10.24381/cds.bd0915c6`.

K. Kanellopoulos, M. De Felice, I. Hidalgo Gonzalez, and A. Bocin. Jrc open power plants database (jrc-ppdb-open), Dec. 2019. URL `https://doi.org/10.5281/zenodo.3574566`.

M. Mommert, M. Sigel, M. Neuhausler, L. Scheibenreif, and D. Borth. Characterization of industrial smoke plumes from remote sensing data, 2020. Tackling Climate Change with Machine Learning Workshop at NeurIPS 2020.

M. Mommert, L. Scheibenreif, J. Hanna, and D. Borth. Power plant classification from remote imaging with deep learning. *ArXiv*, abs/2107.10894, 2021.

O. Ronneberger, P. Fischer, and T. Brox. U-Net: Convolutional networks for biomedical image segmentation. *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015*, May 2015.

T. Yin, L. Byers, L. Valeri, and J. Friedrich. Estimating power plant generation in the global power plant database. 2020.