

# Validating the Global Power Plant Database with Automatized Recognition of Wind Turbines on Google Maps

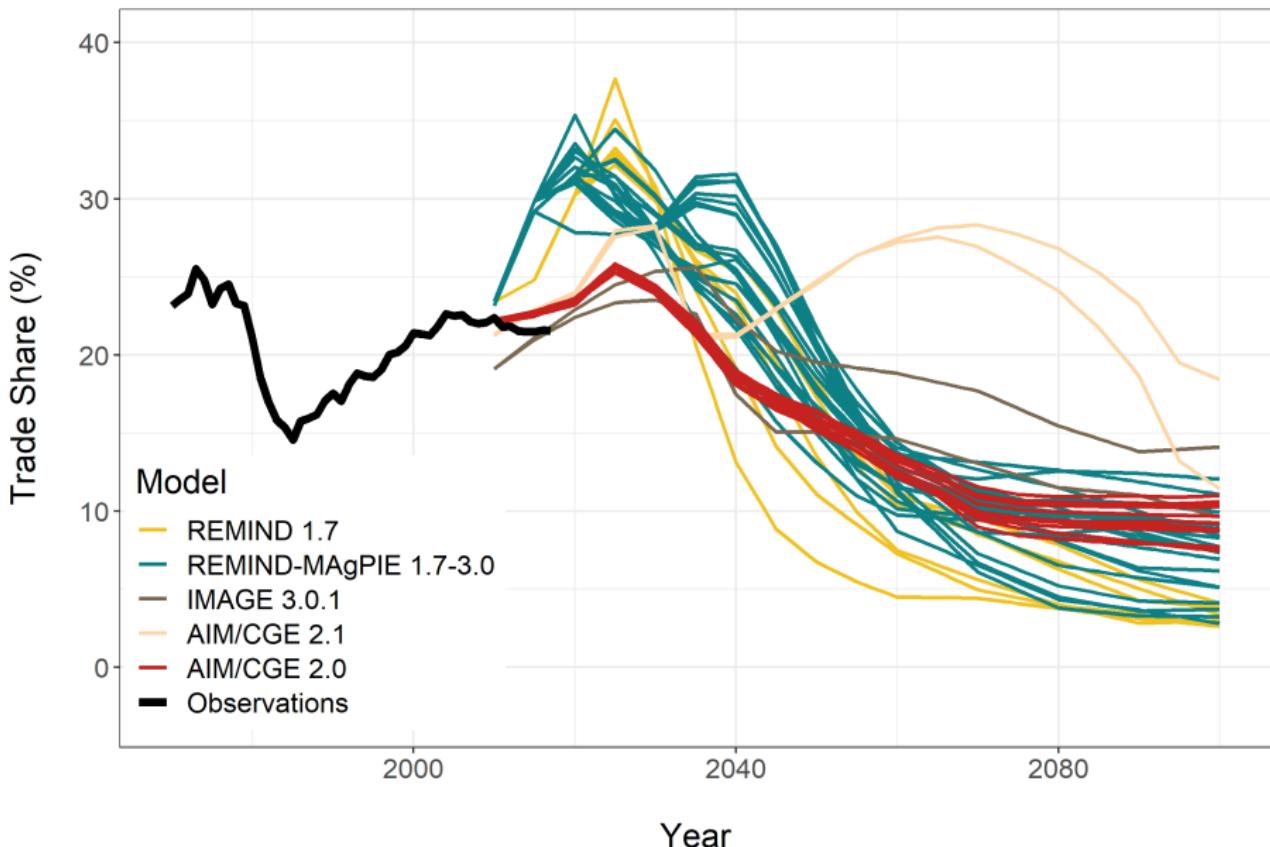
Johannes Schmidt

University of Natural Resources and Life Sciences, Vienna

20.03.2019



# Introduction to reFUEL (I)



## Introduction to reFUEL (II)

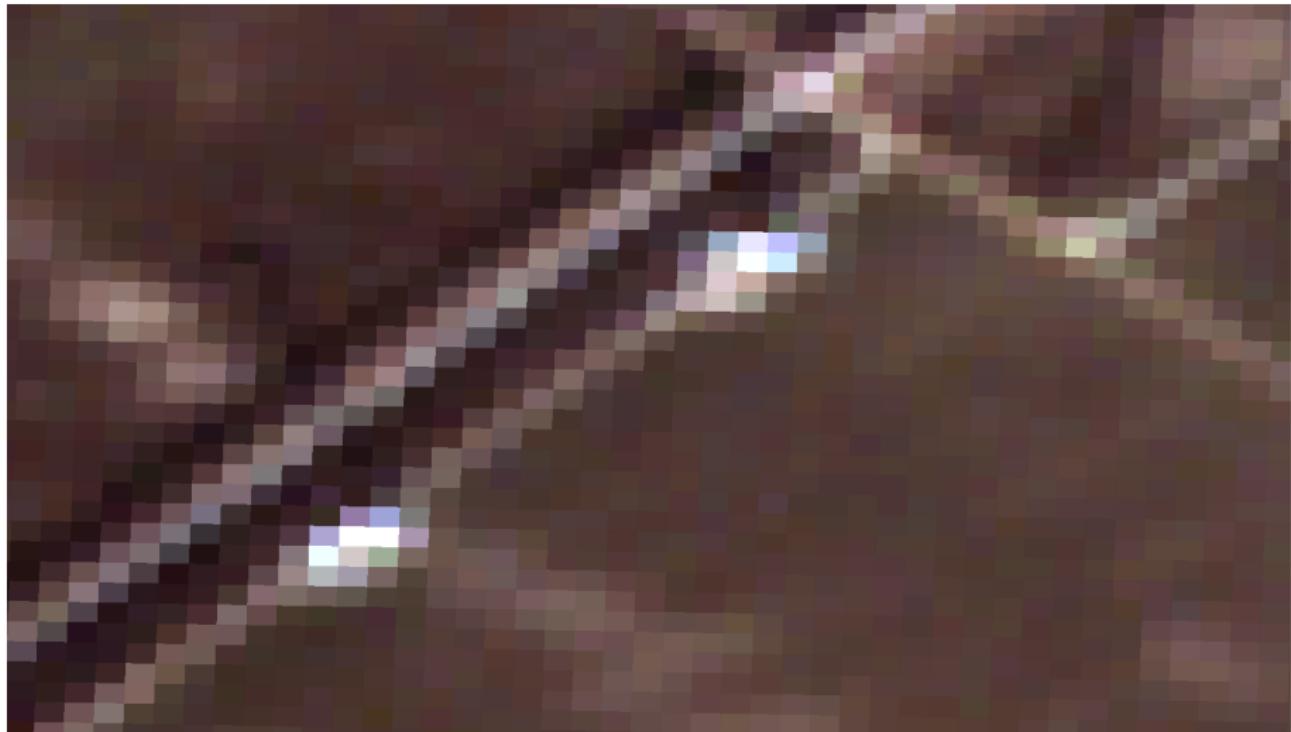
1. Assess global bio-physical potentials of renewable energies (wind and PV) and associated synthetic fuel production
2. Assess options for spatial arbitrage by trade
3. Understand associated *land-use issues*
4. Look into *past* land-use change caused by renewable energy development

# Global Power Plant Database: Need for Validation



# Satellite data for validation

Sentinel-2 (10 meter resolution), free

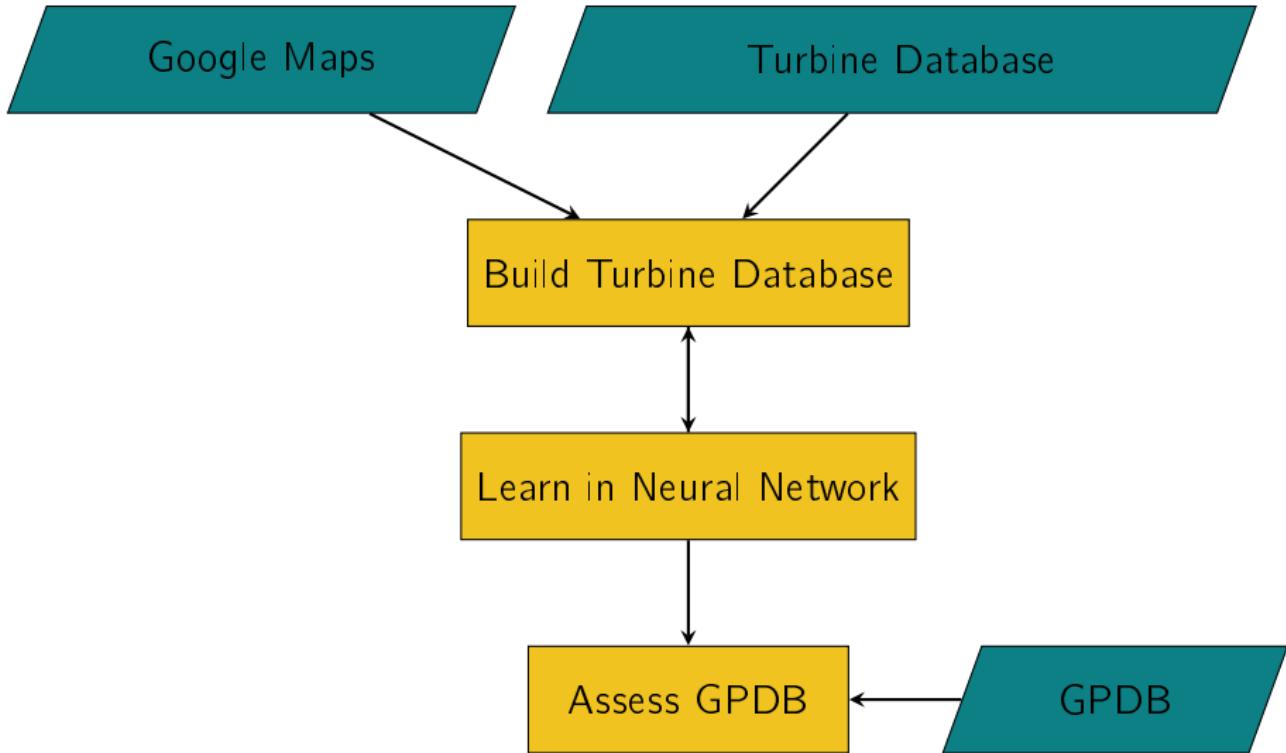


# Satellite data for validation

WorldView-4, Quickbird (0.31-0.65 meter resolution) used by e.g. Google Maps



# Approach



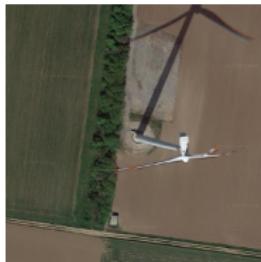
# Software

- ▶ Downloading of data from google maps: R (RGdal, tidyverse, raster)
- ▶ Machine Learning: Python (keras, scikit-image, gdal)
- ▶ Mixing R and Python: not a brilliant idea. Just lazy.

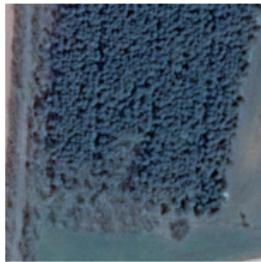
# Create Sample Database

## Manual quality control

Positives

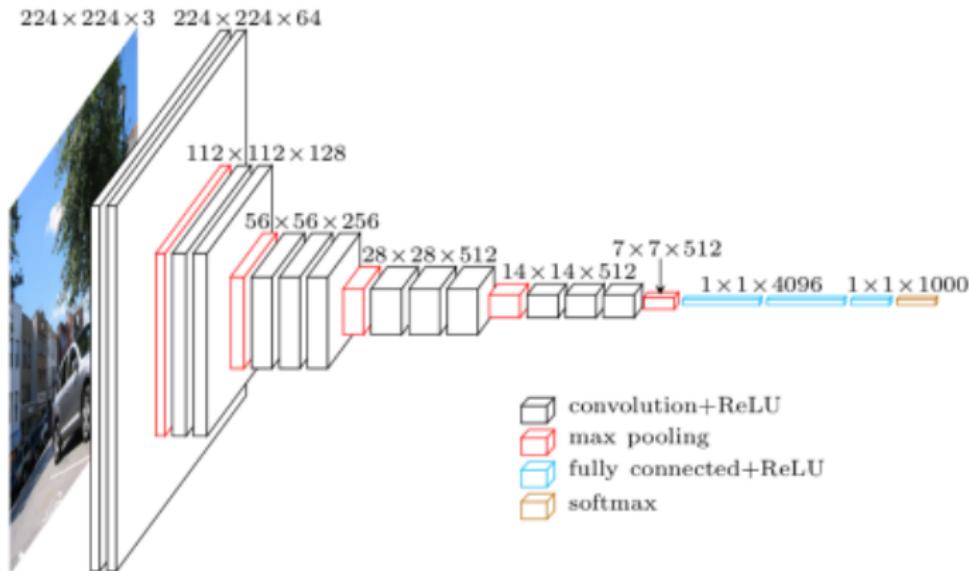


Negatives



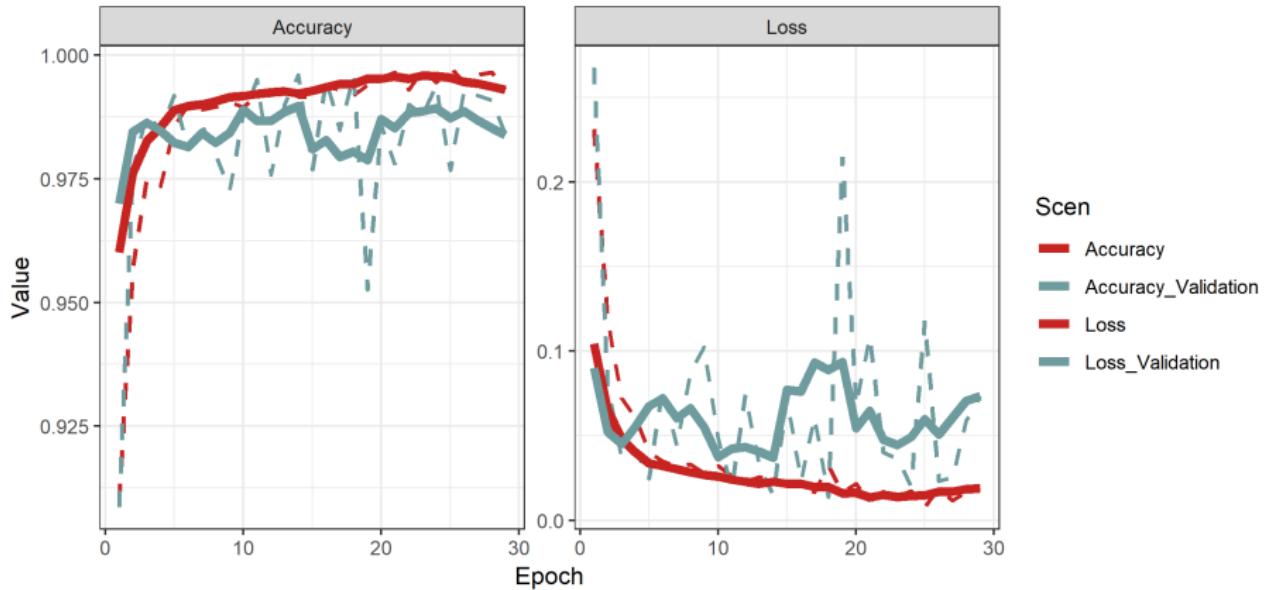
# Learn

## Unfreezing prelearned Neural Network with Convolutional Layers



# Results Learning

## Quality Measures

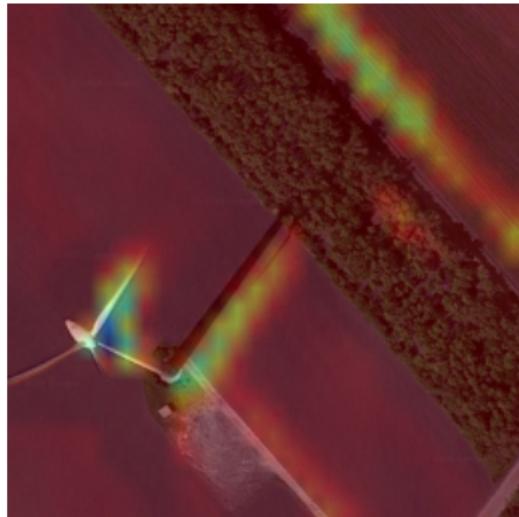


$$\text{Accuracy} = \frac{\text{Number of Correct Predictions}}{\text{Number of Total Predictions}}$$

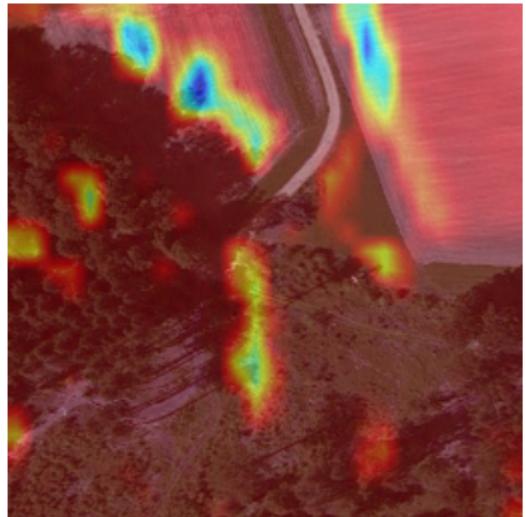
$$\text{Loss} = -1 \frac{1}{N} \sum_{i=1}^N y_i \log(p(y_i)) + (1 - y_i) \log(1 - p(y_i))$$

# Results Learning

## Class Activation Map



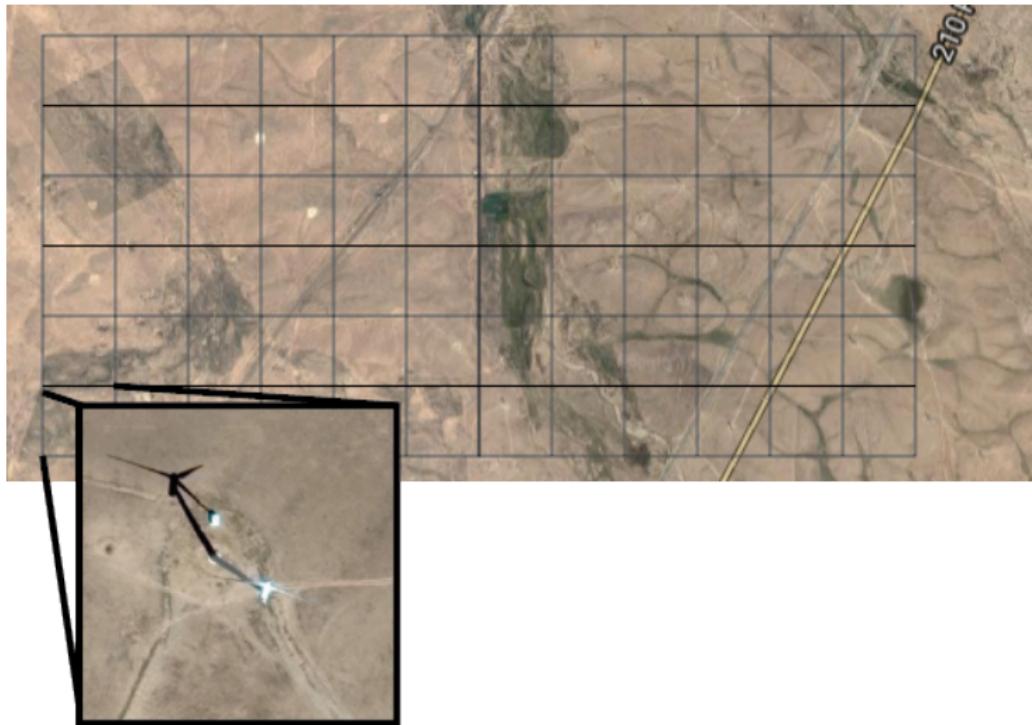
Probability of being wind turbine 0.999



Probability of being wind turbine 0.001

# Searching wind turbines

Create grid and check image by image



# Application

## Chinese and French Wind Parks

**China:** 28 parks assessed. Found 266 turbines. 83 wrongly classified (manual control). At 7 out of 28 park locations no turbine was found.

**France:** 10 parks assessed. Found 27 turbines. 15 wrongly classified (manual control). At 5 out of 10 park locations no turbine was found.

# Discussion

- ▶ Dates of satellite photos and dates of turbine installations unknown
- ▶ Shaky legal conditions
- ▶ Runtime prohibitive on desktop computers. For a full global check would need cluster computing with high bandwidth (but see shaky legal conditions...)
- ▶ Classification has to be extended to be applicable to different world regions. Applying Austrian/Brazilian conditions to China/France does not work very well (i.e. classification errors).

# Conclusions

- ▶ Full validation of GPDB due to missing temporal information not fully possible. However, first assessments of accuracy possible.
- ▶ Current best commercial satellite data allows identification of single turbines with good accuracy by using machine learning approaches
- ▶ Research limited by Data availability. Public domain data is low resolution (i.e. Sentinel-2) or small spatial domain (free ortho-photos, like basemmap.at).

# Thank you!

For updates on the project, check **refuel.world**

For source-code, check

**github.com/joph/MachineLearningCourse**

mail: johannes.schmidt@boku.ac.at

