

## REduceData

Leveraging model accuracy of PV forecasts with an reduced database by a satellite-heavy multi-modular AI approach

<b>Programm / Ausschreibung</b>	Weltraum, Weltraum, ASAP Ausschreibung 2022	<b>Status</b>	abgeschlossen
<b>Projektstart</b>	01.09.2023	<b>Projektende</b>	28.02.2025
<b>Zeitraum</b>	2023 - 2025	<b>Projektlaufzeit</b>	18 Monate
<b>Keywords</b>	Energy meteorology, satellite data, forecasting models, photovoltaik, machine learning		

### Projektbeschreibung

#### Motivation

Die Energieversorgung unserer zukünftigen Gesellschaften ist abhängig von präzisen Intraday- und Day-Ahead-Wettervorhersagen unter Verwendung einer Vielzahl von Informationsquellen (Satellitendaten, bodengestützte Messungen, PV-Anlagenmonitoring). Neben den allgemeinen Wetterprognosen werden zielgruppenspezifische Prognosen von Prognoseanbietern erstellt wie auch Wetterinformationen und Vorhersagen für die Erzeugung erneuerbarer Energien. Bessere Erzeugungsprognosen erneuerbarer Energie (hier Photovoltaik) ermöglichen eine schnellere Transformation in eine CO<sub>2</sub>-neutrale Wirtschaft, da erneuerbare Energien besser geplant, aufgrund verringerter Risiken günstiger gehandelt und in größeren Anteilen am Stromnetz beteiligt werden können.

#### Hauptziel von REduceData

Ziel dieses explorativen Projekts ist es, satellitendatenbasierte modulare Solarstromerzeugungsprognosen mit expertenunterstützten KI-Modellen mit klassischen monolithischen KI-Modellen zu vergleichen. Zusätzlich zu einem Hybrid-Satelliten-NWP-datengesteuerten meteorologischen Prognosemodellen werden verschiedene andere Datenquellen und eine nahtlose Kopplung an die Prognose der Solarstromerzeugung die Auswirkungen der Datensparsamkeit quantifizieren können, Daten verfügbarer Zeitbereich und Datenqualität über den Output des KI-Modells und den weiteren Modellentwicklungsbedarf.

#### Innovation von REduceData

Unser Ziel ist es, modulare Lösungen zu implementieren, um sowohl die meteorologischen Prognosen als auch die direkten PV-Prognosen in einer nahtlosen Modellierungslösung zu verbessern. Die erste Neuheit besteht darin, satellitendatenbasierte numerische Wetter Prädiktoren (NWP-Modelle) mit einer Vielzahl von Daten, meteorologischen Beobachtungen und Fernerkundung zu kombinieren. Die Methoden werden verwendet, da sie erwarten lassen, mit spärlichen Datensätzen, fehlenden Daten und einer physikalischen Komponente gut zu funktionieren. Hier liegt die zweite Neuheit, da die Prognoselösungen vom Nowcasting-Bereich auf einen nahtlosen Prognosebereich für den nächsten Day-Ahead erweitert werden. Als Zieldaten für die Strahlung werden die APOLIS-Daten sowie Punkt- und Gitterbeobachtungen verwendet. Die dritte Neuheit liegt in der Reduktion des Lerndatenbedarfs und der Akzeptanz der Datenmangel durch systematische

Modularisierung in autonom anlernbare Teilmodelle sowie für die satellitengestützte meteorologische und PV-Produktionsprognose. Die vierte Neuerung liegt in der Nutzung der Modularisierung und Data Spasticity Acceptance zur Bereitstellung von Transfer-Metrik und Funktionen für die Modelle an neuen Standorten, an denen kein Training stattfand.

## **Abstract**

### Motivation

Modern societies energy provision relies on accurate intra-day and days-ahead weather forecasts using a variety of sources of information (e.g. satellite, ground based, PV-plant monitorings). Besides the more general weather forecasts, specialized forecasts tailored for specific goals are generated by forecast providers as predictions for renewable energy production. Better renewable power production predictions (here: photovoltaic) enable a faster transformation into a carbon-neutral economy, as renewable energy can be better planned for, traded cheaper due to decreased risk, and partake in the power grid in larger fractions.

### Main objective of REduceData

Goal of this exploratory project is to compare satellite-data based modular solar electric energy production forecasting expert assisted AI models to classical monolithic AI models. Additional to a hybrid satellite-NWP-data driven meteorological forecast model, various other data sources and a seamless coupling to solar electric energy production forecasting will allow to quantify the impact of data sparsity, available data time range and data quality on the AI model's output and further model development needs.

### Innovation by REduceData

We aim at implementing modularised solutions to improve both the meteorological forecasts and direct PV forecasts in a seamless modeling solution. The first novelty lies in combining satellite based numerical weather predictors (NWP) with a wider variety of data, meteorological observations and remote sensing. The methods are used as they will allow for approaches performing well with sparse data sets, missing data and include a physics component. Here lies the second novelty as these two methods will be extended from the nowcasting range to a seamless prediction range for the next day-ahead. As target data for radiation the APOLIS data as well as point and grid-based observations are used. The third novelty lies in the reduction of learning data requirements and the acceptance of data sparsity by systematic modularization into individually trainable sub-models as well for the satellite based meteorological and PV production forecast. The fourth novelty lies in the use of the modularization and data spasticity acceptance for deploying transferability metrics and functions to the models for new sites where no training took place.

## **Endberichtkurzfassung**

The primary goal of REduceData was to investigate and validate options for a modular, data-sparse capable forecasting framework for photovoltaic (PV) production. The framework integrates meteorological and PV modeling components and is capable of delivering robust forecasts even with incomplete or limited observational data.

Key Achievements of the project run time were:

Modular modeling architecture was implemented, combining satellite-based nowcasting (IrradPhyDNet) with day-ahead LSTM-based forecasts, adaptable to various input data configurations.

Synthetic PV production data generation was successfully integrated to support model training in data-sparse environments, increasing forecast quality even with short historical records.

High-quality datasets for both meteorology and PV production were compiled, processed, and harmonized to support reproducible machine learning workflows.

Sparse PV production data sets : for testing PV modular machine learning modelling was generated as well as applied from monitoring data.

Comprehensive validation was carried out using a rich KPI framework, demonstrating strong performance of the modular models in comparison to monolithic benchmarks.

Transferability and interoperability were tested across different sites and configurations, showing that model components can be reused with minimal performance loss.

Operational insights were gained through analysis of forecasting benefits for grid flexibility, balancing costs, and ramp detection, emphasizing the economic and ecological relevance of the forecasting approach.

## **Projektkoordinator**

- AIT Austrian Institute of Technology GmbH

## **Projektpartner**

- GeoSphere Austria - Bundesanstalt für Geologie, Geophysik, Klimatologie und Meteorologie