

# Regression Tests for Reproducibility

**Modernize a complex data pipeline without breaking everything**

Mika Pflüger  
Climate Resource

# Motivation - Climate Change

- Paris Agreement: well below 2 °C global warming, preferably limit to 1.5 °C
- Mechanism: parties (states) make promises, called *nationally determined contributions* (NDCs)
- Our work: quantify NDCs
  - Who promises what?
  - Is it enough?

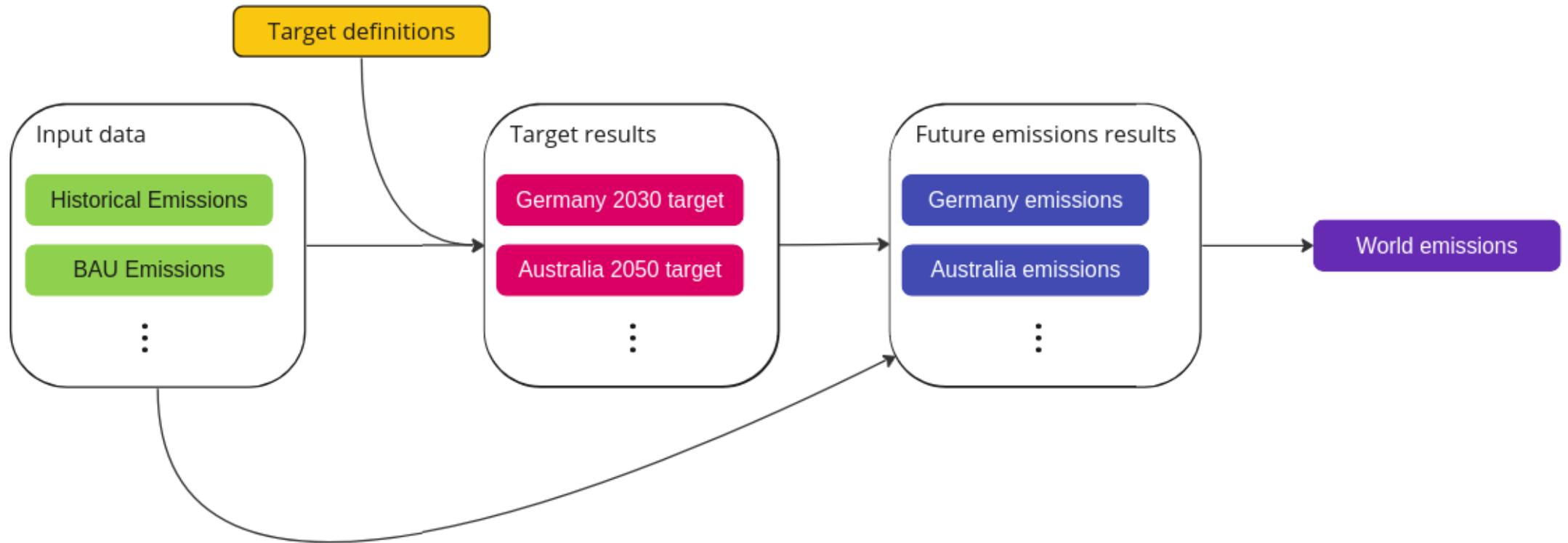
# Status quo

- multiple publications<sup>1</sup>
- published dataset: [NDC factsheets](#)
- several new/updated promises (NDCs) per year
- new input data every year

<sup>1</sup> Meinshausen, M., Lewis, J., McGlade, C. et al. Realization of Paris Agreement pledges may limit warming just below 2 °C. Nature 604, 304–309 (2022). [doi:10.1038/s41586-022-04553-z](https://doi.org/10.1038/s41586-022-04553-z)



# Status quo: Data Pipeline (simplified)



# Status quo: Implementation

- One big Excel workbook
- Hard to understand
- Not user friendly
- Difficult to add new features

Representative formula:

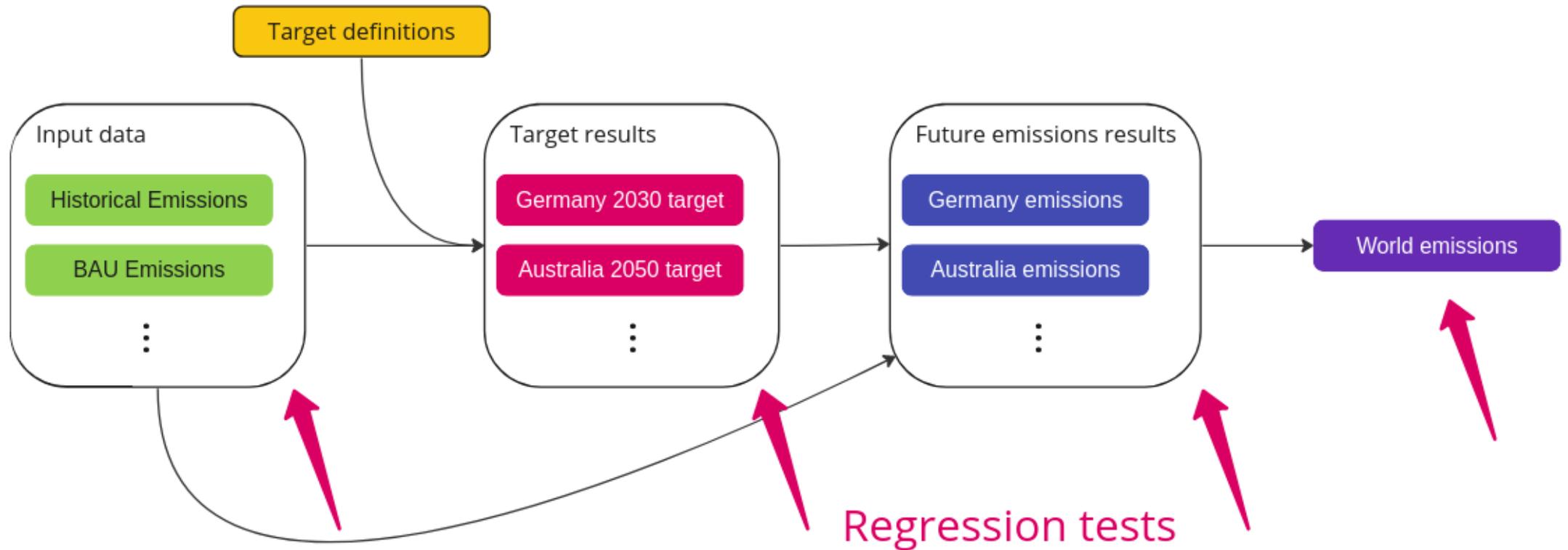
```
=IF(LY21=1,LZ21,IF(MA21=1,MB21,IF(MC21=1,MD21,IF(MF21=1,MK21,NA()))))+IF(AND($ML$9,NOT(LY21=1),NOT(MA21=1)),ML21,0)
```

# Project: New Architecture

- Re-build as a web application, with Python backend
- Requirements:
  - Reproducible: new tool is consistent with old tool
  - Transparent: all updates / changes are explainable

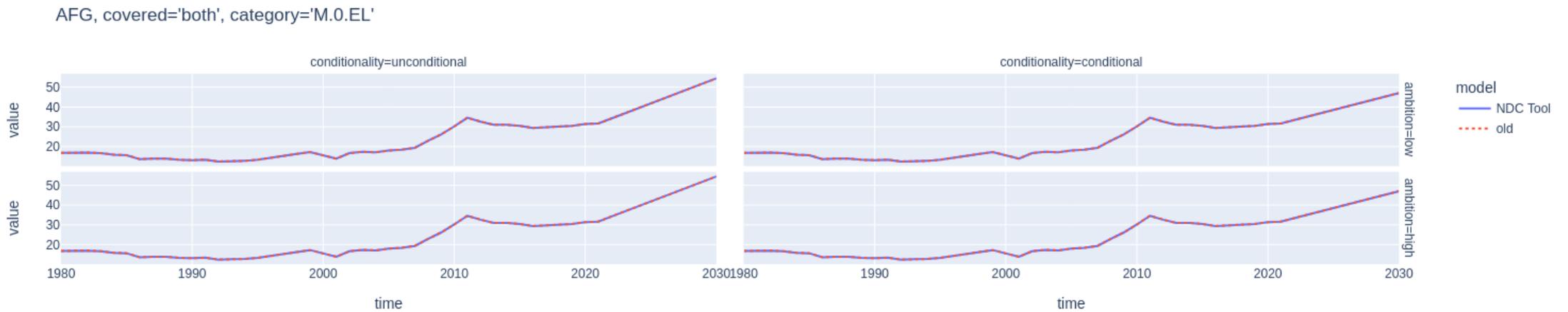
# Approach: Automated Regression Testing

- Extract final and intermediary results from old tool
- Compare with new results



# Results

## First fully reproduced country emissions: Afghanistan



# Results

Not everything is fully reproduced, yet

- ✓ (GEO-Ambition\_LOW-Conditionality\_UNCONDITIONAL-2050)
- ✓ (GEO-Ambition\_HIGH-Conditionality\_UNCONDITIONAL-2050)
- ✗ (GHA-Ambition\_MEDIAN-Conditionality\_UNCONDITIONAL-2030)
- ✗ (GHA-Ambition\_MEDIAN-Conditionality\_CONDITIONAL-2030)
- ✓ (GIN-Ambition\_MEDIAN-Conditionality\_UNCONDITIONAL-2030)
- ✓ (GIN-Ambition\_MEDIAN-Conditionality\_CONDITIONAL-2030)
- ✗ (GMB-Ambition\_MEDIAN-Conditionality\_UNCONDITIONAL-2025)
- ✗ (GMB-Ambition\_MEDIAN-Conditionality\_CONDITIONAL-2025)
- ✗ (GMB-Ambition\_MEDIAN-Conditionality\_UNCONDITIONAL-2030)
- ✗ (GMB-Ambition\_MEDIAN-Conditionality\_CONDITIONAL-2030)
- ✓ (GMB-Ambition\_MEDIAN-Conditionality\_UNCONDITIONAL-2050)

# Conclusion

- Automated testing necessary - manual comparisons hopeless
- Especially to catch newly introduced errors
- Comparison at multiple points in the pipeline super helpful
- Higher effort than expected

Find the slides at [https://uncertainty.es/2024-osd/ndcs\\_regression\\_tests.html](https://uncertainty.es/2024-osd/ndcs_regression_tests.html)