ENGAGE
for ESG

The ENGAGE for ESG Data Extraction Process from a Dutch Perspective: The WOONNU Experience

Wednesday, 18 September 2024

Disclaimer

This presentation (the "Presentation") is being made available for information purposes only. No representation, warranty or undertaking, express or implied, is made as to the accuracy, completeness or appropriateness of the information and opinions contained in this Presentation. Under no circumstances shall the authors or their organisations have any liability for any loss or damage that may arise from the use of this Presentation or the information or opinions contained herein. Certain information contained in this file may include assumptions, opinions, analysis and views of the authors as of September 2024. The assumptions, opinions, analysis and views contained herein are solely opinions which are uncertain and subject to risks. The information contained herein is not intended to constitute investment, legal, regulatory, nor any other sort of advice.

Today’s Speakers

Marco Angheben
Head of Business Development
& Regulatory Affairs
European DataWarehouse
ENGAGE Project Coordinator

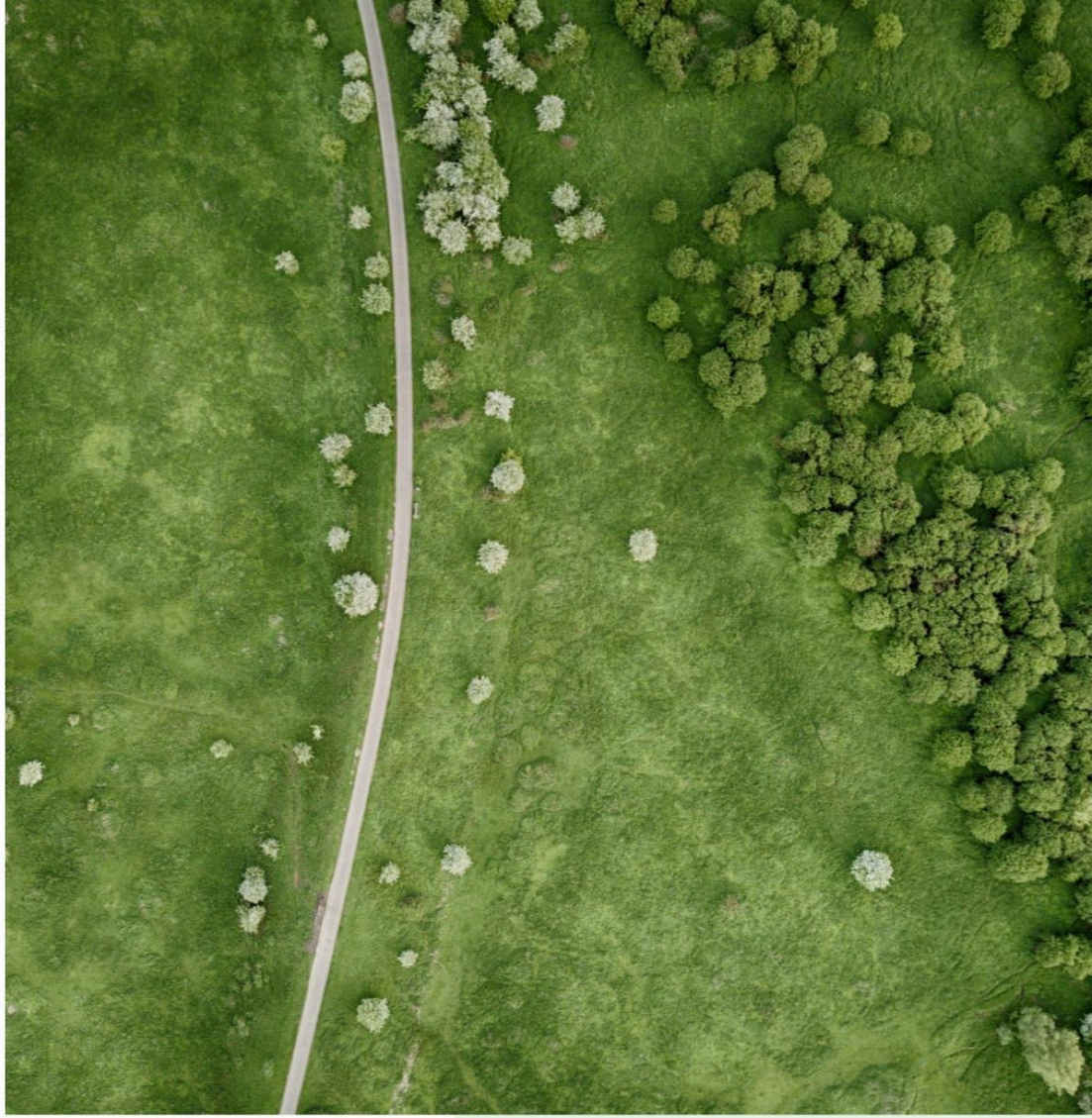
Martijn Breed
MT Lead
Woonnu
NN Bank

Koen Cop
Project & Change
Management
NN Bank

Michele Costola
Researcher
Ca’ Foscari University

Agenda

10:00	Welcome and introduction
10:05	The ENGAGE Templates version 1.1
10:15	Data extraction process to use the ENGAGE Templates: the Woonnu experience
10:30	2024 Research Report on energy efficiency in residential buildings
10:50	How to become a Test User of the ENGAGE Templates via the Portal
10:55	Next webinars on 25 September, 24 October and 21 November 2024



Introduction of the ENGAGE for ESG initiative

Marco Angheben

European DataWarehouse & ENGAGE Project Coordinator

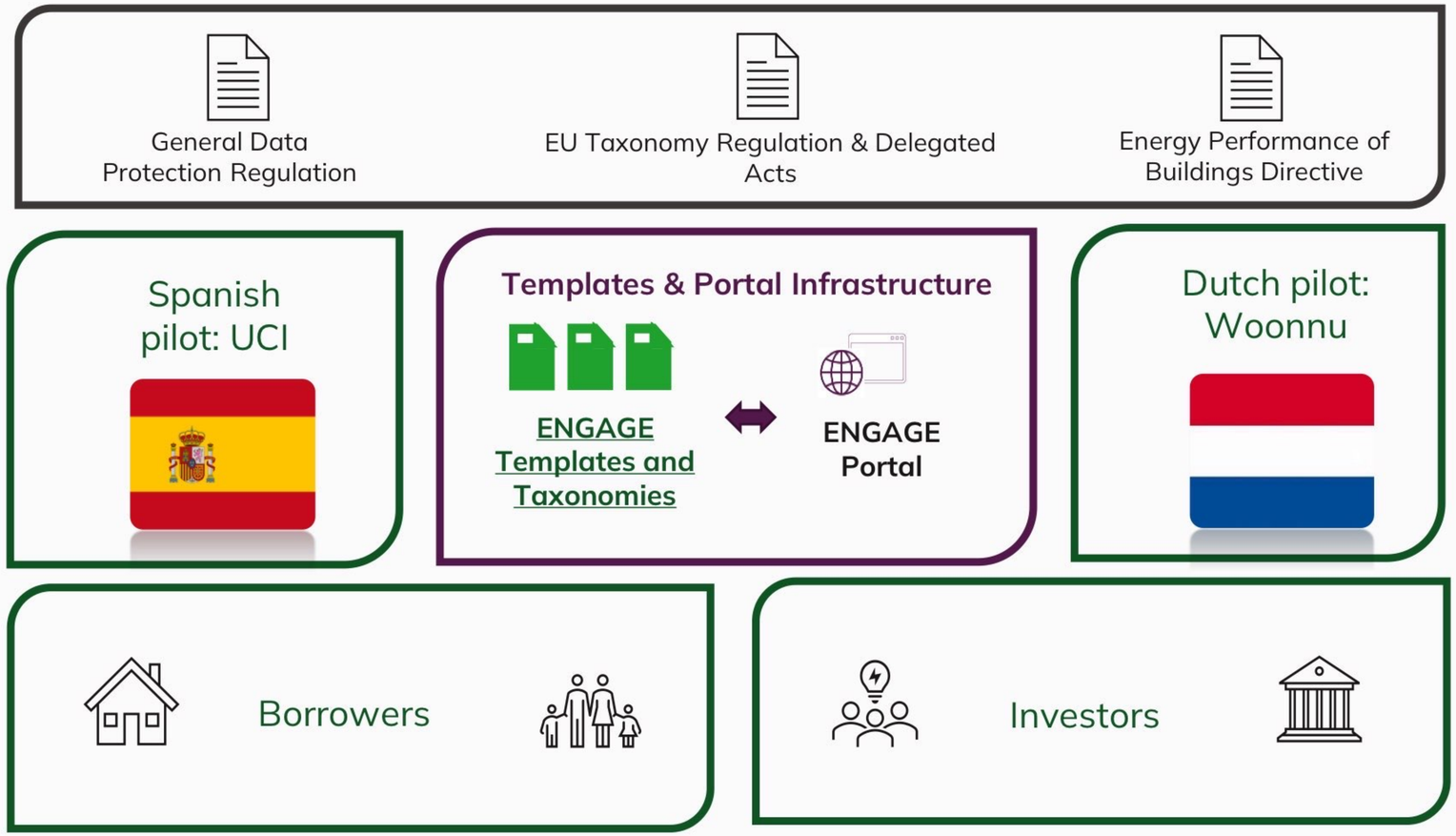
The ENGAGE for ESG initiative

- The ENGAGE for ESG initiative aims to contribute to the goals of the EU Green Deal, namely, to the activation of sustainable investments in the building sector, by promoting ESG transparency for residential mortgages and home renovation loans.
- The ENGAGE Templates 1.1, released in August 2024, include data elements that allow financial institutions to disclose the alignment of their mortgages with the EU Taxonomy requirements in line with the Substantial Contribution Criteria and Do Not Significant Harm of the Climate Delegated Act for the economic activities of acquisition and ownership of real estate, as well as the minimum safeguards.
- The Templates will also enable the assessment of the degree of sustainability for mortgages and the classification of investments for certain mortgages as “sustainable” according to the EU Taxonomy.
- The Templates will be updated and expanded over the coming years to incorporate the most relevant European sustainability regulations. They will be operationalised through the ENGAGE Portal, an IT infrastructure available from November 2024.
- All institutions are invited to test the Templates and the Portal upon request to engage4esg@eurodw.eu

EUROPEAN DATAWAREHOUSE HYPOPORT UCI woonnu DEXAI ARTIFICIAL ETHICS Università Ca’Foscari Venezia

ENGAGE: The High-Level Concept

- A future proof format for real estate data encompassing Europe's most relevant regulatory and sustainable finance requirements
- Converting regulatory sustainable finance regulation into requirements incorporating both the consumer and the financial institution perspective



ENGAGE Templates Structure

Building block 1 created

Building blocks 2 & 3 under construction

Building block 0

- Current ESMA underlying exposures for residential mortgages (Annex 2) and consumer loans (Annex 6);

Building block 1

- Compliance with the EU Taxonomy Regulation for:

  a) Mortgages for the acquisition of new buildings (already built or under construction) and existing buildings;
  b) Loans for the renovation of existing buildings.

Building block 2

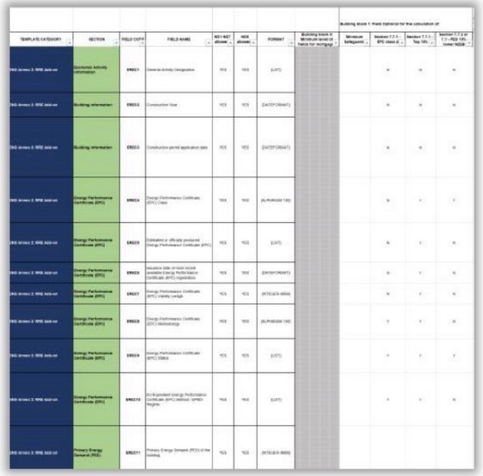
- European Investment Bank information requirements to capture funds from green financing programmes (e.g., ELENA) for the financing of the renovation wave.

Building block 3

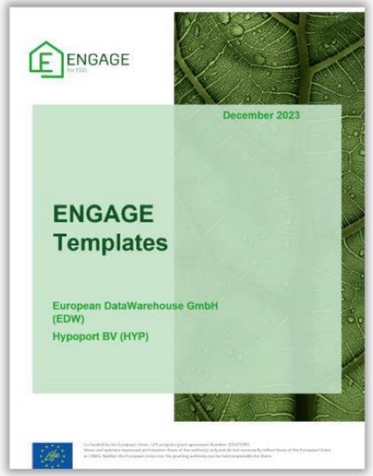
- European Central Bank stress test fields related to climate change (such as flooding and heating) for financing activities related to the building stock.

Accessing the ENGAGE Templates

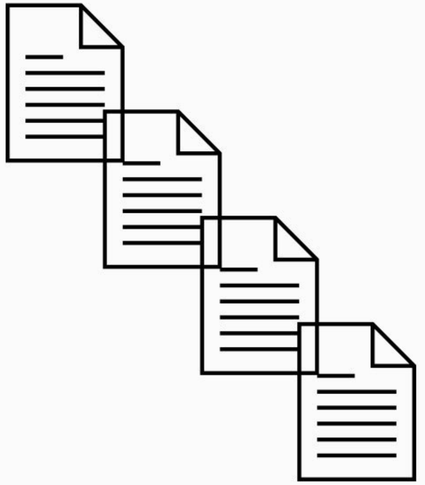
Template



Guiding Documentation



Sample files



Available via the Access Request Form: https://forms.office.com/e/td14aYsMQS

The ENGAGE Templates Version 1.1

Marco Angheben

European DataWarehouse & ENGAGE Project Coordinator

The ENGAGE Templates Structure

Information Type	EU Securitisation Regulation disclosure Annex	Field Code Designator	Section	Total # Fields	Data Level	ENGAGE Add-On
Assets	Annex 2 :RRE	RREL	Underlying exposures information section	82	Loan-level	
		RREC	Collateral information section	23	Building-unit level	
	ENGAGE specific (new files)	EREC	ENGAGE Extended Collateral File		Quantitative	Additional information for the checking of SCC and DNSH criteria to identify the relevant energy performance metrics on a building (unit) level.
Documentation & Transaction Structure	ENGAGE specific (new files)	EGFF	ENGAGE Governance File (aggregated information)		Qualitative	Qualitative Information with reference towards Minimum Safeguards, top-15% and other relevant documentation

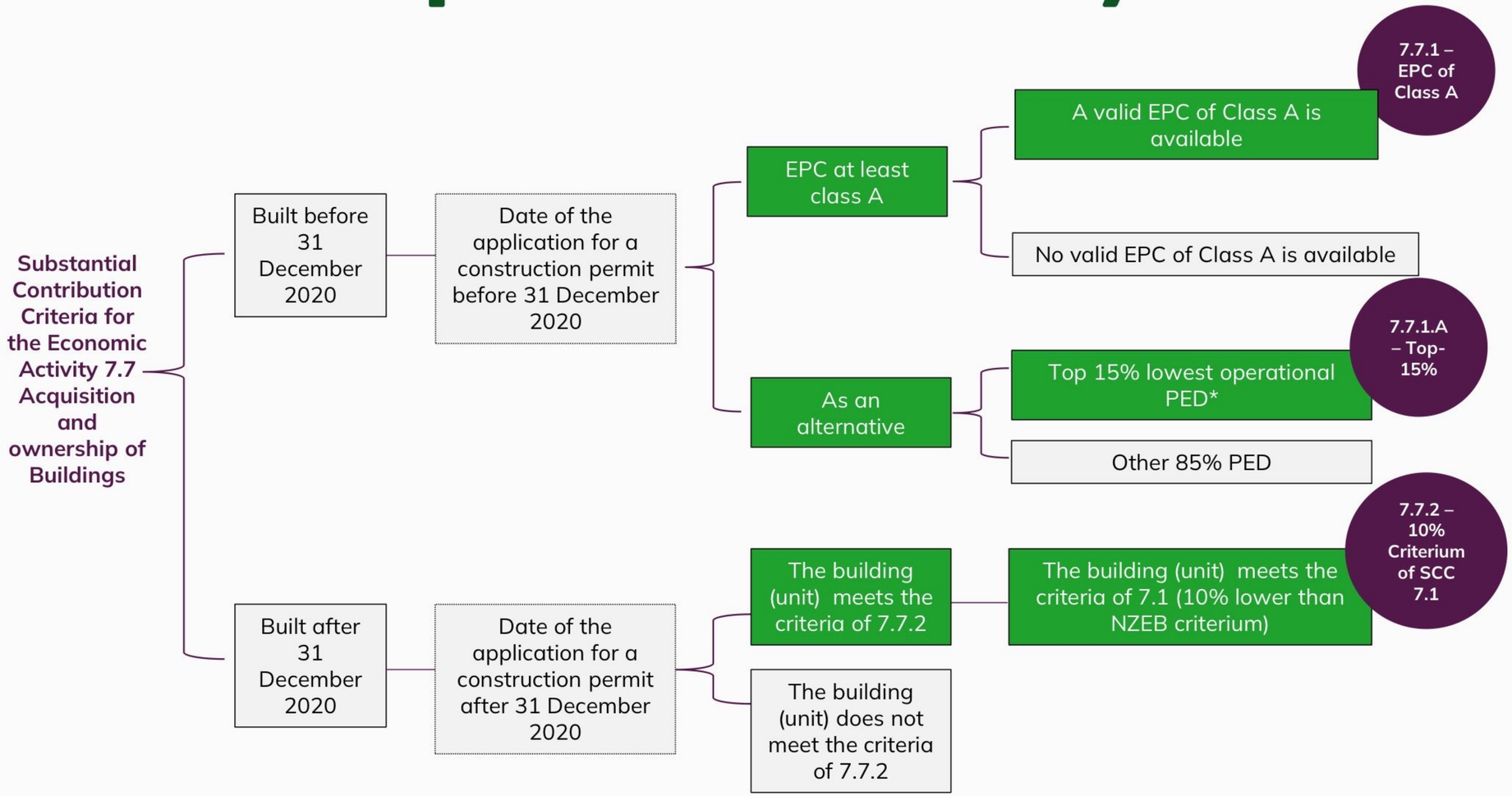
For a mortgage portfolio only, these sections are relevant, because there is no liability or transaction structure.

The ENGAGE Templates are transaction-agnostic. They can be used for many transaction structures such as ABS, RMBS, covered bonds and mortgage loan portfolios. The ENGAGE Templates are flexible and modular.

The ENGAGE Templates Summary

When developing **version 1.1** of the ENGAGE Templates, the ENGAGE consortium has considered:

- ✓ The ENGAGE Templates are **an add-on to existing market best practices**
- ✓ ENGAGE Templates are **scalable**. Meaning that if in the first phase the focus is on the EU Taxonomy, other elements can be incorporated at a later timeframe.
- ✓ **Continuous improvement and innovation**: The framework is designed to facilitate continuous improvement and innovation. This includes mechanisms for gathering and incorporating user feedback, as well as staying up-to-date with technological advancements and industry best practices.



The EU Taxonomy is the common denominator of many sustainable finance regulations

ENGAGE: A solution based on the EC Q&A of December 2022

Section	NACE	Substantial contribution to Climate Change Mitigation of Annex I	Footnote
7.7 Acquisition and ownership of buildings	L68	1. For buildings built before 31 December 2020, the building has at least an Energy Performance Certificate (EPC) class A. As an alternative, the building is within the top 15% of the national or regional building stock expressed as operational Primary Energy Demand (PED) and demonstrated by adequate evidence, which at least compares the performance of the relevant asset to the performance of the national or regional stock built before 31 December 2020 and at least distinguishes between residential and non-residential buildings. 2. For buildings built after 31 December 2020, the building meets the criteria specified in Section 7.1 of this Annex that are relevant at the time of the acquisition.	Not Applicable

7.7.1 – EPC of Class A

7.7.1.A – Top-15%

7.7.2 – 10% Criterium of SCC 7.1

In this version of the template, we have focused on the Substantial Contribution Criteria that are relevant for energy efficient buildings – both existing and new buildings. As we follow a phased approach, we deemed it most appropriate to begin here, with economic activity 7.7 as:

- (existing) real estate tends to be the largest part of the balance sheet of European lending institutions.
- The whole (current) balance of the loan can be attributed towards Taxonomy alignment or the Green Asset Ratio (GAR) if the TSC are met.
- These criteria are relatively straightforward to apply contrary to more challenging criteria such as the TSC for renovation loans and DNSH criteria.
- In addition, it is important that the criteria for new properties can be directly applied to identify and fund energy efficient new constructions.

We have divided activity 7.7 into 3 sub-sections

ENGAGE Templates: SCC 7.7.1

Section	NACE	Substantial contribution to Climate Change Mitigation of Annex I	Footnote
7.7 Acquisition and ownership of buildings	L68	1. For buildings built before 31 December 2020, the building has at least an Energy Performance Certificate (EPC) class A.	Not Applicable

Interpretation:

- At the reporting or assessment date the EPC of the building unit should be of Class A.
- The application date of the construction permit is needed to assess if the building is built before 31 December 2020.
- An EPC of Class A is needed (A, A+, A++, A+++, A++++ also satisfies this condition).
- A certificate should be present with a valid validity date, as of the assessment date, irrespective of the methodology.

Field Code	Field Name
RREL1	Unique Identifier
RREL2	Original Underlying Exposure Identifier
RREL3	New Underlying Exposure Identifier
RREL5	New Obligor Identifier
RREL6	Data Cut-Off Date
RREL30	Current Principal Balance
RREC2	Underlying Exposure Identifier
RREC4	New Collateral Identifier
EREC1	Unique Identifier
EREC2	Underlying Exposure Identifier Building block 0
EREC3	Collateral Identifier Building block 0
EREC4	General Activity Designation
EREC5	Construction Year
EREC6	Construction permit application date
EREC7	Energy Performance Certificate (EPC) Class
EREC8	Estimated or officially produced Energy Performance Certificate (EPC)
EREC9	Issuance date of most recent available Energy Performance Certificate (EPC) registration
EREC10	Energy Performance Certificate (EPC) original validity
EGFF1	Unique Identifier

Key considerations:

- The application date of the construction permit is needed to assess if the building is built before 31 December 2020. For some it is clear that the building was built before 31 December 2020.
- Note Answer 104 of the Q&A: EPC methodologies differ per country or sometimes within a country. Some jurisdictions use energy demand instead of energy consumption. As long as it is an official EPC, this does not matter.

ENGAGE Templates: SCC 7.7.1 Alternative (top-15%)

Section	NACE	Substantial contribution to Climate Change Mitigation of Annex I	Footnote
7.7 Acquisition and ownership of buildings	L68	As an alternative, the building is within the top 15% of the national or regional building stock expressed as operational Primary Energy Demand (PED) and demonstrated by adequate evidence, which at least compares the performance of the relevant asset to the performance of the national or regional stock built before 31 December 2020 and at least distinguishes between residential and non-residential buildings.	Not Applicable

Key considerations:

- The application date of the construction permit is needed to assess if the building is built before 31 December 2020.
- There are a range of varying estimation techniques available to perform a top-15% study.
- We have facilitated flexibility of the method in the ENGAGE Templates and we have emphasised the message of the European Commission in the Q&A that the methodology should be public and transparent.
- Therefore, we have incorporated the numerator and denominator that is used in the top-15% assessment and some background information in the ENGAGE Templates, so that stakeholders can understand the methodology and study that is applied.

Field Code	Field Name
RREL1	Unique Identifier
RREL2	Original Underlying Exposure Identifier
RREL3	New Underlying Exposure Identifier
RREL5	New Obligor Identifier
RREL6	Data Cut-Off Date
RREL30	Current Principal Balance
RREC2	Underlying Exposure Identifier
RREC4	New Collateral Identifier
EREC1	Unique Identifier
EREC2	Underlying Exposure Identifier Building block 0
EREC3	Collateral Identifier Building block 0
EREC4	General Activity Designation
EREC5	Construction Year
EREC6	Construction permit application date
EREC21	Building unit in top-15% indicator
EREC22	Top15_Explanatory Variable
EREC23	Top15_Object Reference Value
EREC24	Top15_Object Threshold Value
EGFF1	Unique Identifier
EGFF2	Top15_Document Name
EGFF3	Top15_Document Issuance Date
EGFF4	Top15_Document URL
EGFF5	Top15 Document Geographic Scope
EGFF6	Top15 Numerator
EGFF7	Top15 Denominator
EGFF8	Top15 Methodology Description

ENGAGE Templates: SCC 7.7.2 (1/2)

Section	NACE	Substantial contribution to Climate Change Mitigation of Annex I	Footnote
7.7 **Acquisition and ownership of buildings**	L68	2. For buildings built after 31 December 2020, the building meets the criteria specified in Section 7.1 of this Annex that are relevant at the time of the acquisition.	Not Applicable

Section	NACE	Substantial contribution to Climate Change Mitigation of Annex I	Footnote
7.1 **Construction of New Buildings**	F41.1, F41.2, F43	Constructions of new buildings for which: The Primary Energy Demand (PED)[282], defining the energy performance of the building resulting from the construction, is at least 10 % lower than the threshold set for the nearly zero-energy building (NZEB) requirements in national measures implementing Directive 2010/31/EU of the European Parliament and of the Council [283]. The energy performance is certified using an as built Energy Performance Certificate (EPC).	[282]: The calculated amount of energy needed to meet the energy demand associated with the typical uses of a building expressed by a numeric indicator of total primary energy use in kWh/m^2 per year and based on the relevant national calculation methodology and as displayed on the Energy Performance Certificate (EPC). [283]: Directive 2010/31/EU of the European Parliament and of the Council of 19 May 2010 on the energy performance of buildings (OJ L 153, 18.6.2010, p. 13).

Interpretation:

- The EU Taxonomy requires building units with a construction permit application date after 31 December 2020 to be built according to the NZEB criteria and the PED should be 10% less than the locally applicable threshold value.
- As of 31 December 2020, NZEB should be implemented in the EU according to the Energy Performance of Buildings Directive (EPBD III). As part of this Directive, the PED should be recorded and displayed on the EPC.

ENGAGE Templates: SCC 7.7.2 (2/2)

Section	NACE	Substantial contribution to Climate Change Mitigation of Annex I	Footnote
7.1 **Construction of New Buildings**	F41.1, F41.2, F43	Constructions of new buildings for which: The Primary Energy Demand (PED)[282], defining the energy performance of the building resulting from the construction, is at least 10 % lower than the threshold set for the nearly zero-energy building (NZEB) requirements in national measures implementing Directive 2010/31/EU of the European Parliament and of the Council [283]. The energy performance is certified using an as built Energy Performance Certificate (EPC).	[282]: The calculated amount of energy needed to meet the energy demand associated with the typical uses of a building expressed by a numeric indicator of total primary energy use in kWh/m^2 per year and based on the relevant national calculation methodology and as displayed on the Energy Performance Certificate (EPC). [283]: Directive 2010/31/EU of the European Parliament and of the Council of 19 May 2010 on the energy performance of buildings (OJ L 153, 18.6.2010, p. 13).

Key considerations:

- The application date of the construction permit is needed to assess if the building is built after 31 December 2020, see answer 106 of the Q&A.
- Assess if NZEB is incorporated in the jurisdiction. Also see answer 105 of the Q&A.
- Answer 109 of the Q&A states: "*For the energy threshold, this depends on national regulations, i.e. if the EPC applies to the whole building, or to each apartment. Whichever is the requirement at national level, it should apply for both residential and non-residential buildings. The correct EPC will be provided in any case, in line with the national regulations. For identical apartments, having normally identical EPCs, a limited sub-set can be used. However, if there are different types of apartments, with different EPCs, all types need to be checked.*"
- Answer 114 of the Q&A states that the TSC applicable at the time of the building permit should be used (i.e. the date of the complete application for receiving the building permit).
- Answer 115 of the Q&A states that for new buildings, either an EPC (valid for 10 years) or an EPC as-built are valid.

Field Code	Field Name
RREL1	Unique Identifier
RREL2	Original Underlying Exposure Identifier
RREL3	New Underlying Exposure Identifier
RREL5	New Obligor Identifier
RREL6	Data Cut-Off Date
RREL30	Current Principal Balance
RREC2	Underlying Exposure Identifier
RREC4	New Collateral Identifier
EREC1	Unique Identifier
EREC2	Underlying Exposure Identifier Building block 0
EREC3	Collateral Identifier Building block 0
EREC4	General Activity Designation
EREC5	Construction Year
EREC6	Construction permit application date
EREC8	Estimated or officially produced Energy Performance Certificate (EPC)
EREC9	Issuance date of most recent available Energy Performance Certificate (EPC) registration
EREC10	Energy Performance Certificate (EPC) original validity
EREC11	Energy Performance Certificate (EPC) Methodology
EREC13	EU-Equivalent Energy Performance Certificate (EPC) Method / EPBD-Regime
EREC14	Primary Energy Demand (PED) of the building
EREC15	Estimated or officially produced Primary Energy Demand (PED)
EREC18	Nearly zero-energy building (NZEB) threshold
EGFF1	Unique Identifier

ENGAGE Templates: Do Not Significant Harm (DNSH)

According to section 7.7 of Annex I o f the Climate Delegated Act, the acquisition and ownership of buildings (either built before or after 31 December 2020) shall be deemed as not significantly harming climate change mitigation if the activity complies with the criteria set out in Appendix A to Annex I of the Climate Delegated Act for the environmental objective of **climate change adaptation**.

Field Code	Field Name
EREC25	DNSH EO2 Indicator
EGFF15	DNSH Compliance Reference

ENGAGE Templates: Minimum Safeguards

Article	Text
18	Minimum safeguards 1.The minimum safeguards referred to in point (c) of Article 3 shall be procedures implemented by an undertaking that is carrying out an economic activity to ensure the alignment with the OECD Guidelines for Multinational Enterprises and the UN Guiding Principles on Business and Human Rights, including the principles and rights set out in the eight fundamental conventions identified in the Declaration of the International Labour Organisation on Fundamental Principles and Rights at Work and the International Bill of Human Rights. 2.When implementing the procedures referred to in paragraph 1 of this Article, undertakings shall adhere to the principle of 'do no significant harm' referred to in point (17) of Article 2 of Regulation (EU) 2019/2088.

Interpretation:

- When addressing the question of who undertakes this economic activity, we arrive at the conclusion that in our specific context, it is the (prospective) homeowner. The term "undertaking" is not defined in the context of the Taxonomy Regulation. The term "undertaking" is commonly understood to refer to a corporation, business entity, or an organized enterprise, rather than an individual.
- The (prospective) building owner is exercising ownership – and thus carrying out the economic activity. A financial institution is facilitating this via a mortgage loan – financing the economic activity - of buying real estate.

Field Code	Field Name
EGFF9	Description on how Minimum Safeguards are complied with
EGFF10	URL towards MSS Issuer Statement
EGFF11	Environmental Objective
EGFF12	OECD Guidelines for MNE Reference(s)
EGFF13	UN GP Reference(s)
EGFF14	Bill of Human Rights Reference(s)

We deem, with the current guidance, the Minimum Safeguards of the Taxonomy Regulation in the context of (mortgage) lending for residential properties to homeowners - more concretely for the economic activities of 7.2 – 7.7 of the Climate Delegated Act – Annex I, not to be applicable to households as we do not consider these to be undertakings.

The ENGAGE Templates: Change log v1.0 vs v1.1

Building block 0

- Revision of the **optional fields** to check the alignment of loans with the Climate Delegated Act.
- Field **RREL1 (Unique Identifier):** revision of the field description to cover non-securitised loans.

Building block 1

- ENGAGE **Additional Collateral-level Information**
  - ✓ Introduction of **three new fields** (**EREC1, EREC2, EREC3**) to track the transactionID, the loanID and the collateralID between building blocks 0 and 1.
  - ✓ Adjustments to the **wording** of the fields' description.
- ENGAGE **Governance File**
  - ✓ Introduction of **one new field** (**EGFF1-Unique Identifier**) to track the transactionID between building blocks 0 and 1.
  - ✓Adjustments to the **wording** of the fields' description.
- DNSH principle

Data Extraction Process to use ENGAGE Templates: The Woonnu Experience

Martijn Breed and Koen Cop

NN Bank (Woonnu)

- Woonnu was founded in 2019 as 100% subsidiary of NN Bank
- Recently merged into NN Bank as the green mortgage label of NN
- Customer journey with tooling and sustainability partners
- Introduction of a 3D pricing scheme in the Dutch market with discounts for better EPC labels
- All EPC labels of our clients are known
- NN Bank current GAR: 64%

Dutch Perspective

NUMBERS

- 8.1 million homes
- 60% valid EPC (as of 1 January 2023)
- 34.5% energy label A or better (A+, A++, A+++)

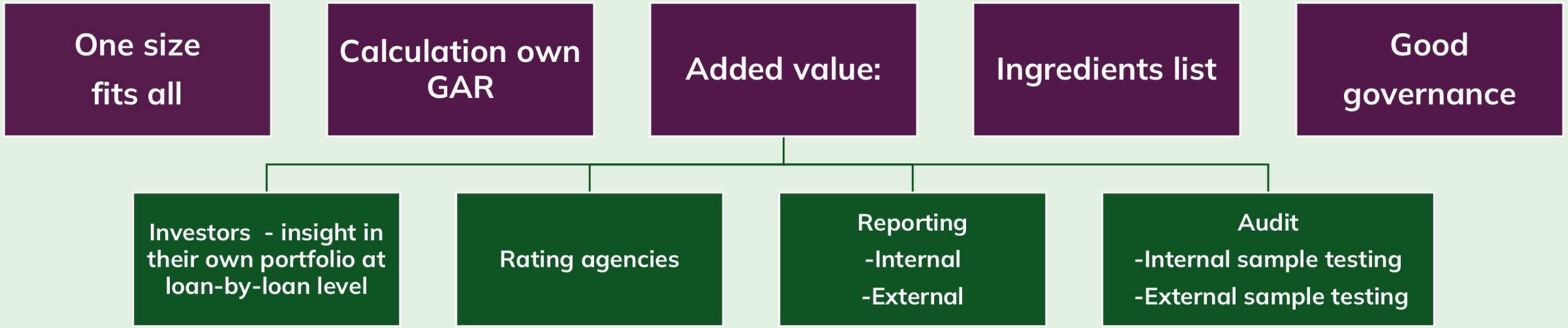
DATA SOURCES

- *EP online*
- *Klimate Effect Atlas*

MEMBERS OF EEM NL Hub

- Dutch Energy Efficient Mortgage Framework (DEEMF)

ENGAGE: success factors



Data extraction process data sources

NN Bank & NN group reported on EUT alignment in the 2023 annual statement: required template data was available.

Also applicable for the data governance: most of the required policies are in place.

THE MAJORITY OF THE DATA IN THE ENGAGE TEMPLATES COMES FROM TWO PUBLIC DATA SOURCES:

Energy certificate data: EP-online

Every official and valid energy performance certificate should be registered in the official Dutch database.

Cadastral data: National building administration (Basis Administratie Gebouwen (BAG))

Every building in the Netherlands should have a registration in BAG.

The public sources are considered as the "Truth".

Public climate maps from *stichting CAS* are used to assess climate risks exposed to a building & compliance with climate adaption DNSH criteria

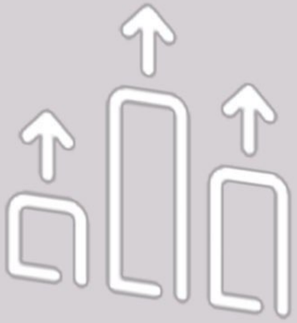
NN Bank data challenges & improvements

CHALLENGES

- Biggest challenge is to link the public sources to our mortgage administration (based on postal code, house number and house number addition):
  - New built homes lack postal code/address.
  - Difference in the records of house numbers addition (approx. 7% of our portfolio).
  - Not in all cases a 1-to-1 relation: 1 collateral in our administration exist out of 2 or more building units.
  - Only 60% of the buildings have a valid energy certificate (Woonnu > 90%).

IMPROVEMENTS

- New built homes: the postal code / address is recorded at an early stage (when deed is passed).
- Improved house numbers records and available in near future in the unique BAG_id in mortgage administration.
- Pilot to engage customers to apply for an official energy certificate.



EPC data availability

EP-ONLINE (ISSUANCE YEAR > 2020)	13%	Energy Performance Certificate (EPC) Class	Estimated or officially produced EPC	Issuance date of most recent EPC	EPC original Validity	EPC Methodology	Energy Performance Certificate (EPC) Status	EU-Equivalent EPBD-Regime
		Primary Energy Demand (PED) of the building	Estimated or officially produced Primary Energy Demand (PED)	Primary Energy Demand (PED) Based on Reference Building	Primary Energy Demand (PED) Based on Building or Building unit	Nearly zero-energy building (NZEB) theshold		
EP-ONLINE (ISSUANCE YEAR <= 2020)	44%	Energy Performance Certificate (EPC) Class	Estimated or officially produced EPC	Issuance date of most recent EPC	EPC original validity	EPC methodology	Energy Performance Certificate (EPC) Status	EU-Equivalent EPBD-Regime
PROVISIONAL LABEL	40%	Energy Performance Certificate (EPC) Class	Estimated or officially produced EPC					
BUILDING YEAR LABEL	3%	Energy Performance Certificate (EPC) Class	Estimated or officially produced EPC					

Cadastral data availability

• BAG*	98%	General Activity Designation	Construction Year	Building unit in top 15% indicator	Top 15% Explanatory Variable	Top 15% Object Reference Value	Top 15% Object Threshold Value	Construction permit application date
• NO BAG MATCH	2%	General Activity Designation	Construction Year	Building unit in top 15% indicator	Top 15% Explanatory Variable	Top 15% Object Reference Value	Top 15% Object Threshold Value	
• CLIMATE RISK DATA	99%*	DNSH EO2 Indicator						

*BAG: Basis Administratie Gebouwen

ENGAGE impact on NN Bank: Benefits, challenges, and conclusion

BENEFITS

- Enables investors to report on the GAR and creates insight of how this evolves
- Detailed insights for investors (and auditors) on loan level
- Helps to check consistency of the data and detect DQ issues
- When a few extra data fields are added to the ENGAGE Templates it would also enable investors to report on their emissions (CSRD) and monitor progress in reaching ESG targets.

CHALLENGES

- Further improvement of matching our portfolio with the public data sources
- Percentage of building with a valid Energy performance certificate must increase
- Extension of the ENGAGE Templates for emission data is desired to support ESG reporting (CSRD).

"ENGAGE helps us to better understand the current limitations in our data and where to improve and enables our investors in ESG reporting and to create necessary and detailed insights."

2024 Research Report on Energy Efficiency in Residential Buildings

Michele Costola

Università Ca'Foscari

Learning from Experts: Energy Efficiency in Residential Buildings

Monica Billio[a] Roberto Casarin[a,c,d]

Michele Costola[a] Veronica Veggente[b,a]

[a] *Ca' Foscari University of Venice, Venice, Italy*
[b] *Imperial College Business School, London, UK*
[c] *Venice Center for Risk Analytics (VERA)*
[d] *European Center for Living Technologies (ECLT)*

ENGAGE for ESG

18th September, 2024

Funding Acknowledgement

This research was carried out within the project funded by Next Generation EU - Project PE0000018 *GRINS - Growing Resilient, INclusive and Sustainable*, National Recovery and Resilience Plan (NRRP) - EP9 - Mission 4, C2, Intervention 1.3.

The views and opinions expressed are only those of the authors and do not necessarily reflect those of the European Union or the European Commission. Neither the European Union nor the European Commission can be held responsible for them.

This research used the HPC multiprocessor cluster system of the *Venice Center for Risk Analytics* (VERA) at Ca' Foscari University of Venice.

Motivation - Green Transition

- The **building sector** has a relevant impact on global warming;
  → more than 40% of energy consumption and 33% of the CO2 emissions globally

- Policymakers aim at reducing GHG in the residential building sector;
  → *Energy Performance of Buildings Directive* (2002), *Green Homes Directive* (2024)

- GHG emissions are one of the main **drivers of transition risk**;
  → *Paris Agreement* (2015)

- Reduced energy consumption relevant for financial risk management;
  → Low energy efficiency and Lower solvency risk in the mortgage market are correlated (Billio et al. (2021), Guin et al. (2022))

Contribution: Energy Performance Data

We leverage on the expertise of the assessors contained in **Energy Performance Certificates** (EPC data) to train machine learning models predicting **potential energy efficiency increase**.

Our contribution is twofold:

- Studies **forecasting energy efficiency and energy demand in buildings** (Gómez-Omella et al., 2021), (Mehmood et al., 2019).
- Literature on **machine learning techniques for modelling big data and non-linear relationships** in different fields of science (social sciences (Gupta et al., 2016), genetics (Bleich et al., 2014), engineering (He et al., 2017)).

Contribution: EP determinants

Leverage on expert's opinions in EPC to:

1. **Predict** potential increase in energy efficiency levels of buildings;
2. Identify the **determinants** of building energy efficiency;
3. Analyse the **non-linear** effects of such determinants;
4. Design **green policies** informed by machine learning;
5. Evaluate the **performance** of these policy actions.

Energy Performance Certificates

- EPCs are the main tool to measure energy performance in buildings;

  $\rightarrow$ introduced in the EU with the **Energy Performance of Buildings Directive - EPBD** (2002/91/EC), then revised with *EPBD* (2010/31/EU) and the *Clean Energy Package* within the Dir. (2018/844/EU) amending 2010/31/EU and the *Energy Efficiency Directive* (2012/27/EU). EPB is then regulated by the *Energy Efficiency Directive* (EU/2023/1791) and the *Energy Performance of Buildings Directive* (EU/2024/1275)

- These documents contain specific information on:

  a. **Structural characteristics** of buildings and services installed;

  b. **Current** energy efficiency level (EE);

  c. **Recommendation** of the experts;

  d. **Potential** energy efficiency level (EE_{pot}).

- We use these information to train our set of machine learning models.

A measure of Energy Efficiency (I)

Target variable Y included in our study is the **logistic transformation** of the **ratio** EE_{pot}/EE, where low EE and EE_{pot} indicates high performance.

$$Y = log\left(\frac{EE_{pot}}{EE - EE_{pot}}\right)$$

- defined in $(-\infty, +\infty)$;
- inversely related to the relative energy performance increase.

A measure of Energy Efficiency (II)

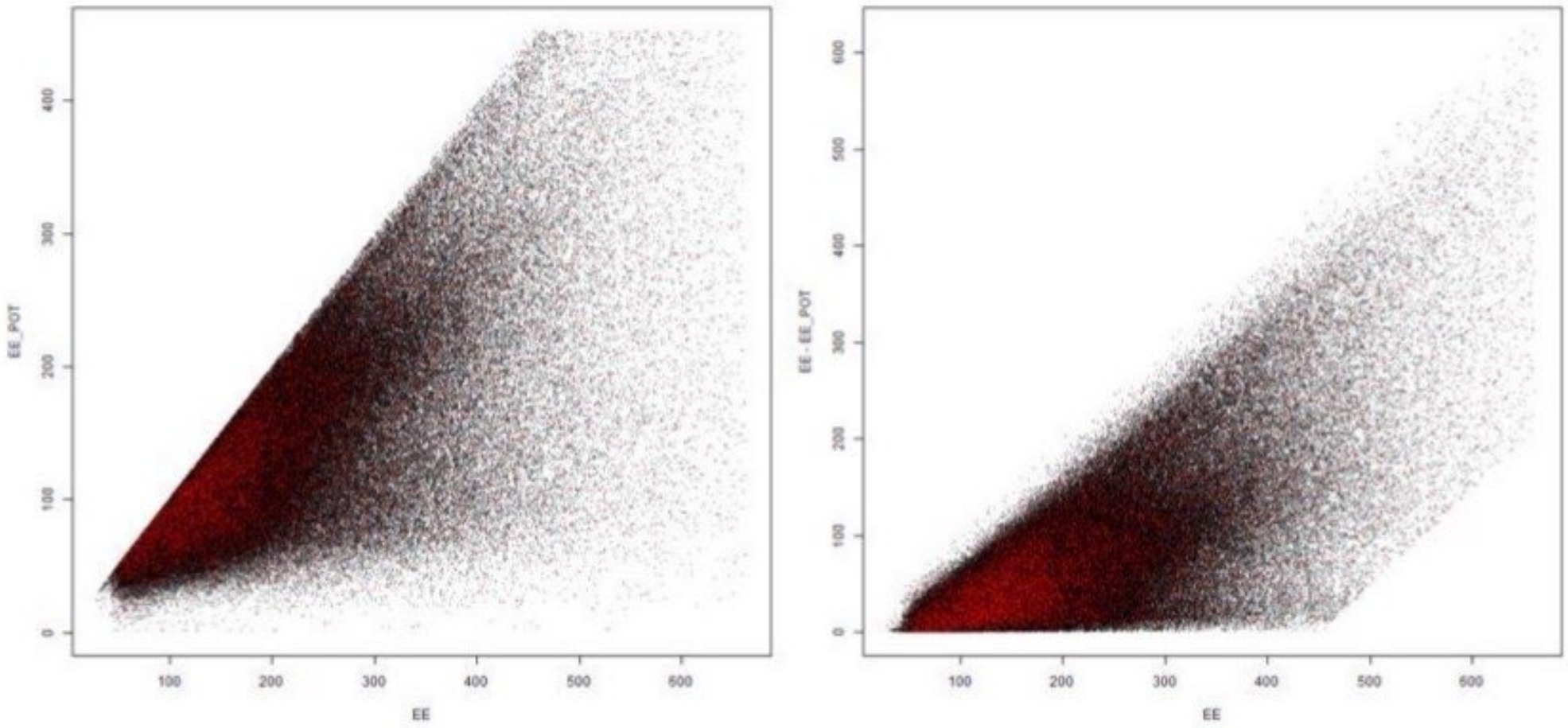


Figure: The case of Lombardy (Italy). The initial non-renewable energy performance index EE (horizontal axis) versus the final index EE_{pot} (vertical axis, left panel), the expected performance difference $EE - EE_{pot}$ (vertical axis, right panel). The entire sample involves 205,049 buildings (gray dots) and a sub-sample of 10,000 buildings (red dots).

A measure of Energy Efficiency (II)

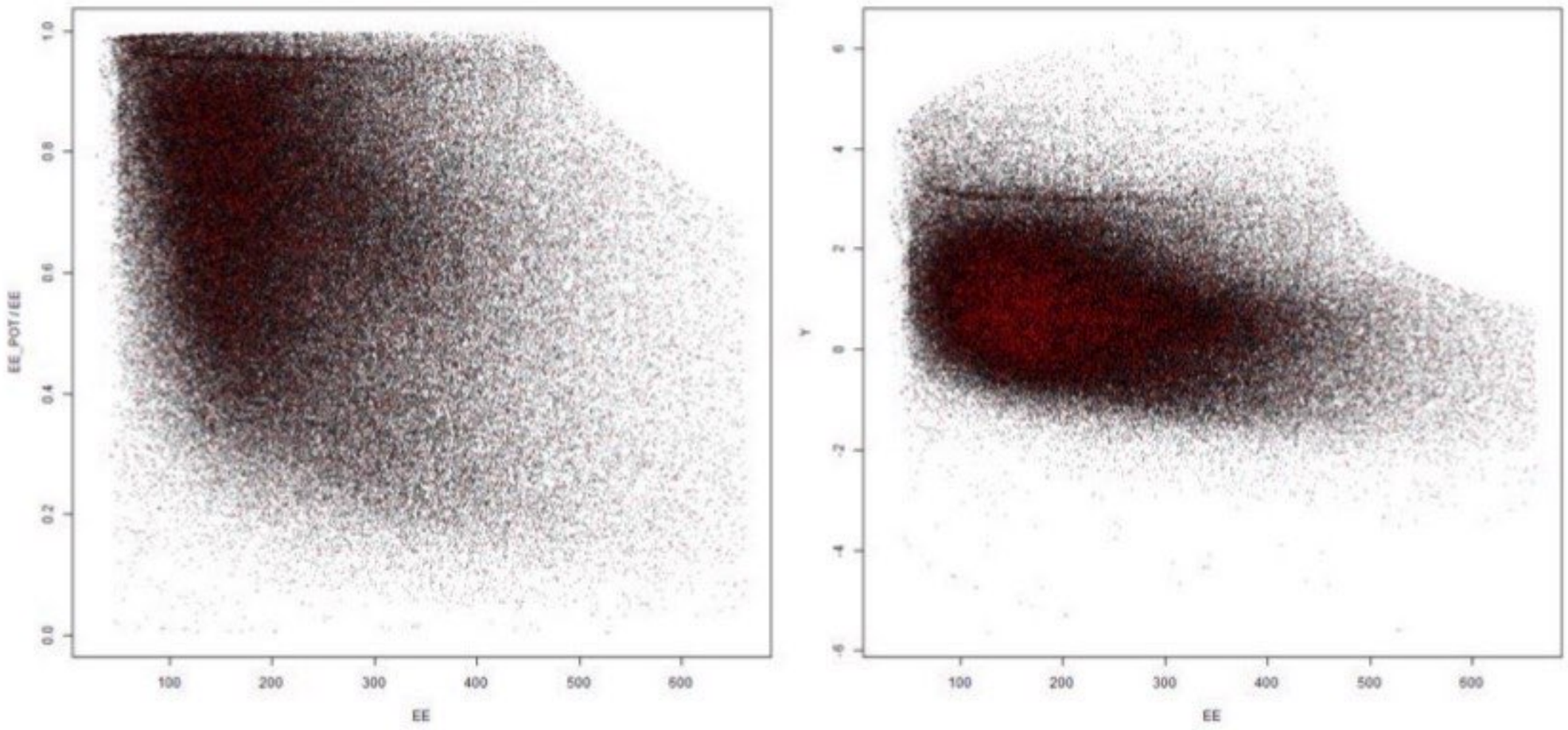
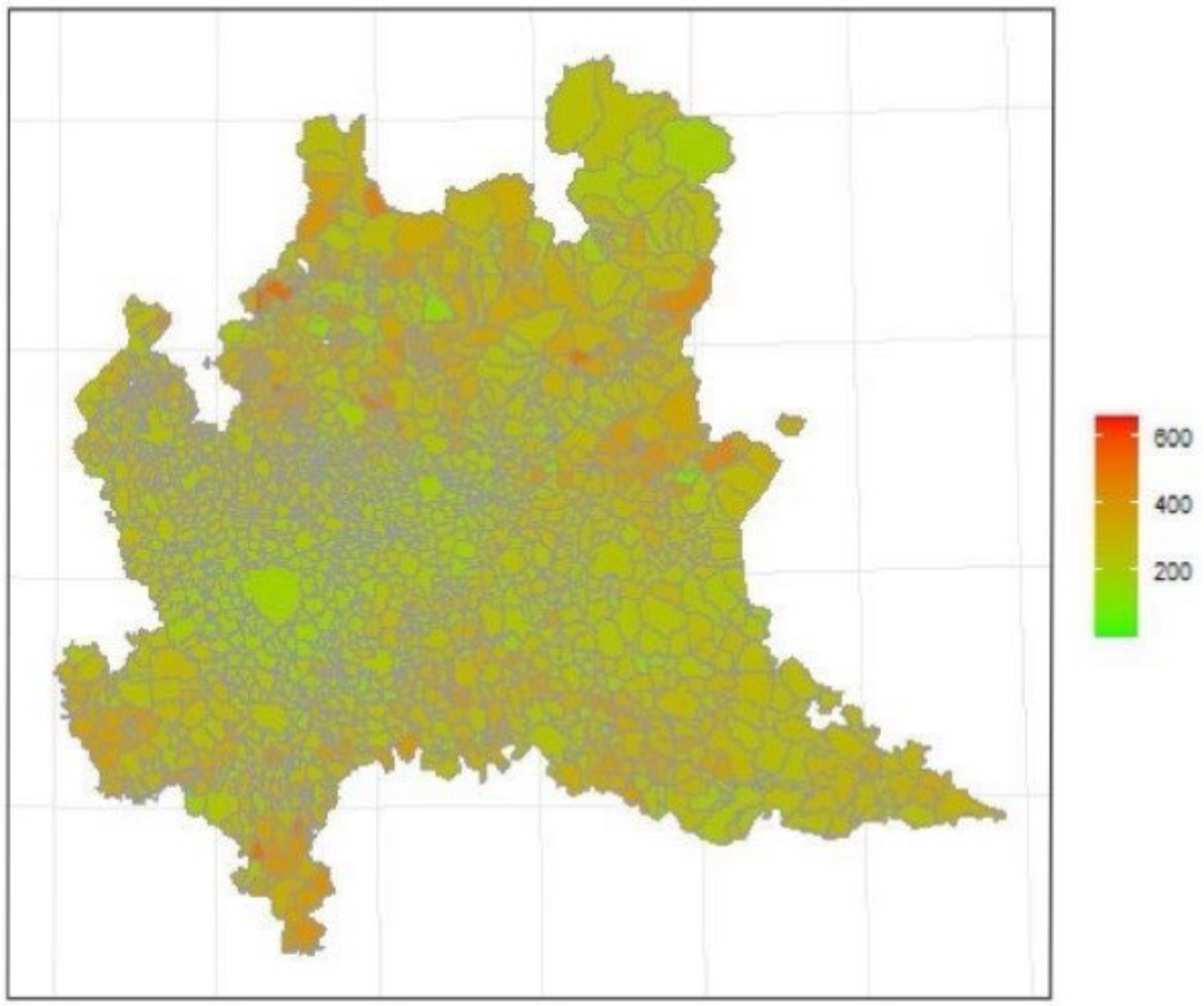


Figure: The case of Lombardy (Italy). The initial non-renewable energy performance index EE (horizontal axis) versus the expected performance difference ratio EE/EE_{pot} (left axis, right panel) and the final indicator Y (right) panel. The entire sample involves 205,049 buildings (gray dots) and a sub-sample of 10,000 buildings (red dots).

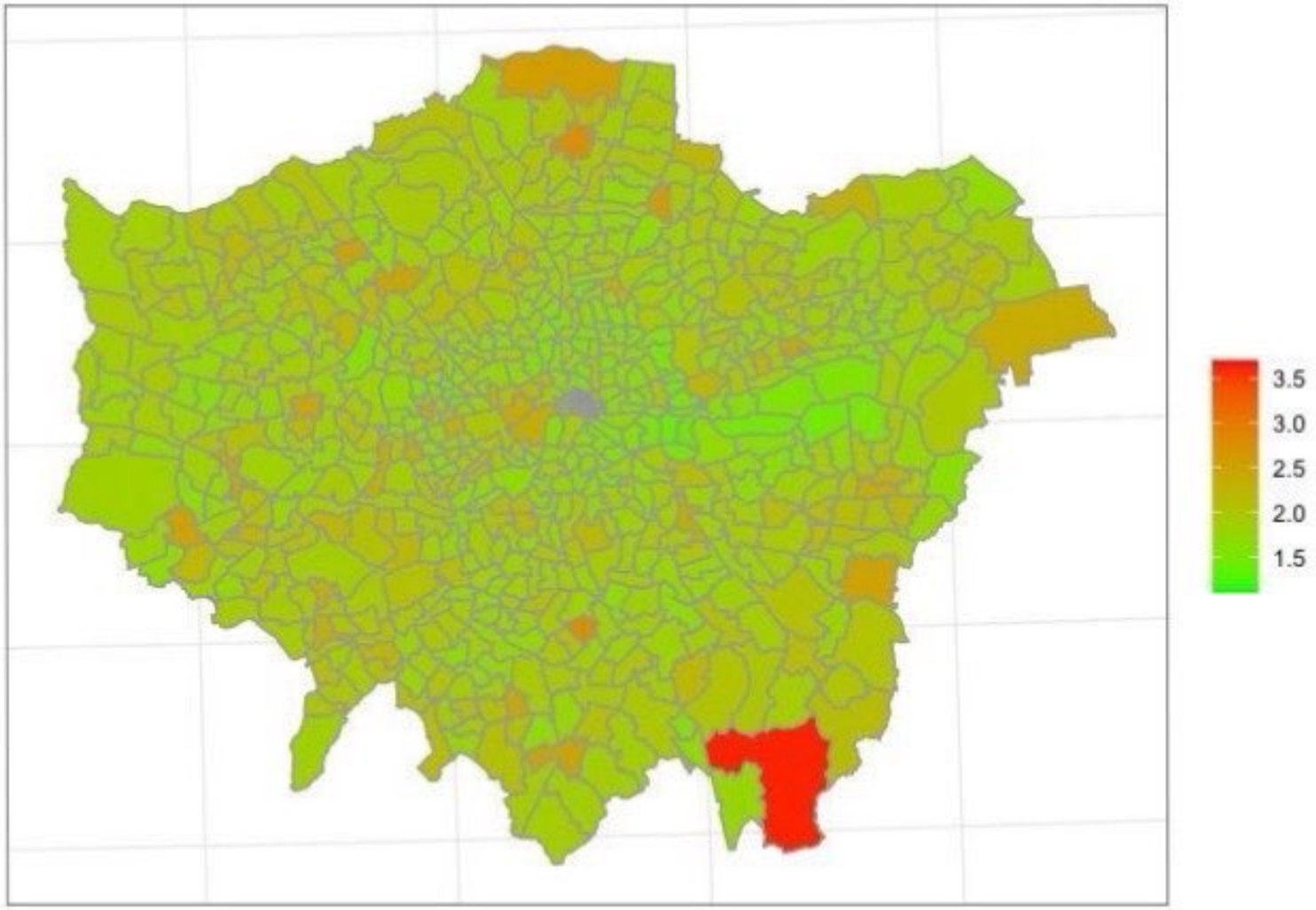
Data: Lombardy (IT)

Open Access micro data on residential buildings: **CENED+2**, Lombardy (Italy), 2015-2021 (downloaded on 2022): 205,049 observations, 42 features.
Source: *Catasto Energetico Edifici Regionale, Regione Lombardia, ARIA Spa.*



Data: Greater London (UK)

Open Access micro data on residential buildings: **EPBD UK**, Greater London (UK), 2015-2020 (downloaded on 2022): 445,661 observations, 82 features. Source: *Department for Levelling Up, Housing and Communities and Ministry of Housing, Communities & Local Government*.



Borough of Bromley is one of the districts, it has a total of 47 conservation areas. Strict requirements to make alterations to your property.

Models

Multiple machine learning models to predict energy efficiency increase:

- **Linear models**:
  - **a.** Lasso
  - **b.** Ridge
  - **c.** Elastic Net
- **Non-linear models**:
  - **a.** Bayesian Additive Regression Tree (BART)
  - **b.** Random Forest (RF)
  - **c.** Extreme Gradient Boosting (XGBOOST)

Application on full and reduced set of observations; 70% training and 30% test set.

Linear Models

Linear parametric models with penalization to shrink the coefficients estimate to reduce model complexity.

1. **Lasso**: shrinks some coefficients to zero using an ℓ_1 penalization;
2. **Ridge**: reduces the impact of some predictors using an ℓ_2 penalization;
3. **Elastic Net**: combines Lasso and Ridge regression.

Linear Model Settings: Lasso, Ridge, Elastic Net

- General minimization problem:

$$||y - \beta_0 - X\beta||_2^2 + \lambda \left[\alpha ||\beta||_1 + (1 - \alpha) \frac{1}{2} ||\beta||_2^2 \right]$$

- where $X \in \mathbb{R}^k$ is a covariate vector.
- Three different specifications using `glmnet` R-package (Friedman et al., 2010) with Lasso ($\alpha = 1$), Ridge ($\alpha = 0$) and Elastic Net ($0 < \alpha < 1$), e.g. $\alpha = 0.5$.
- λ is the regularization parameter chosen using cross-validation.

Non-linear Models: Tree Based ensemble models

1. **Bayesian Additive Regression Trees**:
   - Flexible model which approximates the unknown function as a sum of decision trees.
   - Regularization prior on the tree structure, the characteristic values at each node and the error variance, to prevent a single tree to dominate the whole fit.
   - The posterior is non-tractable and approximated numerically by a MCMC algorithm.
2. **Random Forests**:
   - Combination of a collection of decision trees, trained in parallel on subset of the data;
   - Random subset of m variables out of p is selected as candidates at each node, to handle correlation among factors.
3. **Extreme Gradient Boosting**:
   - Trees are grown sequentially based on the residuals of the previous tree (i.e a modified version of the original dataset).
   - Avoid to overfitting the data by using a model that learns slowly.

Tree Based Ensemble Models Settings: BART, RF, EGB

- Nonlinear regression model (e.g. BART):

$$y = \sum_{j=1}^{J} g_j(X, \mathcal{T}_j) + \varepsilon, \quad \varepsilon \sim (0, \sigma^2)$$

$$g_j(X, \mathcal{T}_j) = \sum_{n=1}^{N_j} \beta_{jn} \mathbb{I}(X \in \mathcal{X}_n(\mathcal{T}_j))$$

- where $\mathcal{T}_j$ is a sequence of trees, $\beta_{jn} \in \mathbb{R}^k$ a sequence of parameter vectors, $X \in \mathbb{R}^k$ is a covariate vector, $(\mathcal{X}_1(\mathcal{T}_j), \ldots, \mathcal{X}_{N_j}(\mathcal{T}_j))$ is a partition of the covariate space generated by a tree $\mathcal{T}_j$.
- Different R-packages: `randomForest` (Breiman, 2001, Liaw and Wiener, 2002), `BayesTree` (Chipman and McCulloch, 2016), `BART` Sparapani et al. (2021), `xgboost` (Chen and Guestrin, 2016)
- the number of trees J is chosen using cross-validation.

Empirical Results

Empirical results - Italy and UK

Prediction MSE: Linear vs Non-Linear - Italy

	Full sample		Subsample	
	In sample	Out of sample	In sample	Out of sample
	Mean Square Error			
LASSO	0.7732	0.7795	0.7998	0.8350
RIDGE	0.8471	0.8552	1.2682	1.3210
ELASTIC NET	0.7729	0.7792	0.8095	0.8429
BART	0.6125	0.6415	0.6085	0.7197
RANDOM FOREST	0.6583	0.6613	0.6716	**0.7087**
XGBOOST	**0.5275**	**0.6126**	**0.4778**	0.7347

Table: **Mean Square Error** between actual and predicted values estimated by Lasso, Ridge, Elastic Net, BART, Random Forest, and XGBoost. In-sample and out-of-sample results for the whole sample (first and second column) and a random subsample (third and fourth column) for the Italian case.

Non-linear models have higher predictive performance

Prediction MAE: Linear vs Non-Linear - Italy

	Full sample		Subsample	
	In sample	Out of sample	In sample	Out of sample
	Mean Absolute Error			
LASSO	0.6731	0.6755	0.6867	0.7057
RIDGE	0.7096	0.7139	0.8794	0.9049
ELASTIC NET	0.6730	0.6755	0.6912	0.7102
BART	0.5956	0.6071	0.5997	0.6524
RANDOM FOREST	0.6184	0.6197	0.6251	**0.6413**
XGBOOST	**0.5513**	**0.5914**	**0.5315**	0.6580

Table: **Mean Absolute Error** between actual and predicted values estimated by Lasso, Ridge, Elastic Net, BART, Random Forest, and XGBoost. In-sample and out-of-sample results for the whole sample (first and second column) and a random subsample (third and fourth column) for the Italian case.

Non-linear models have higher predictive performance

Partial Dependence: Non-Linearities - Italy

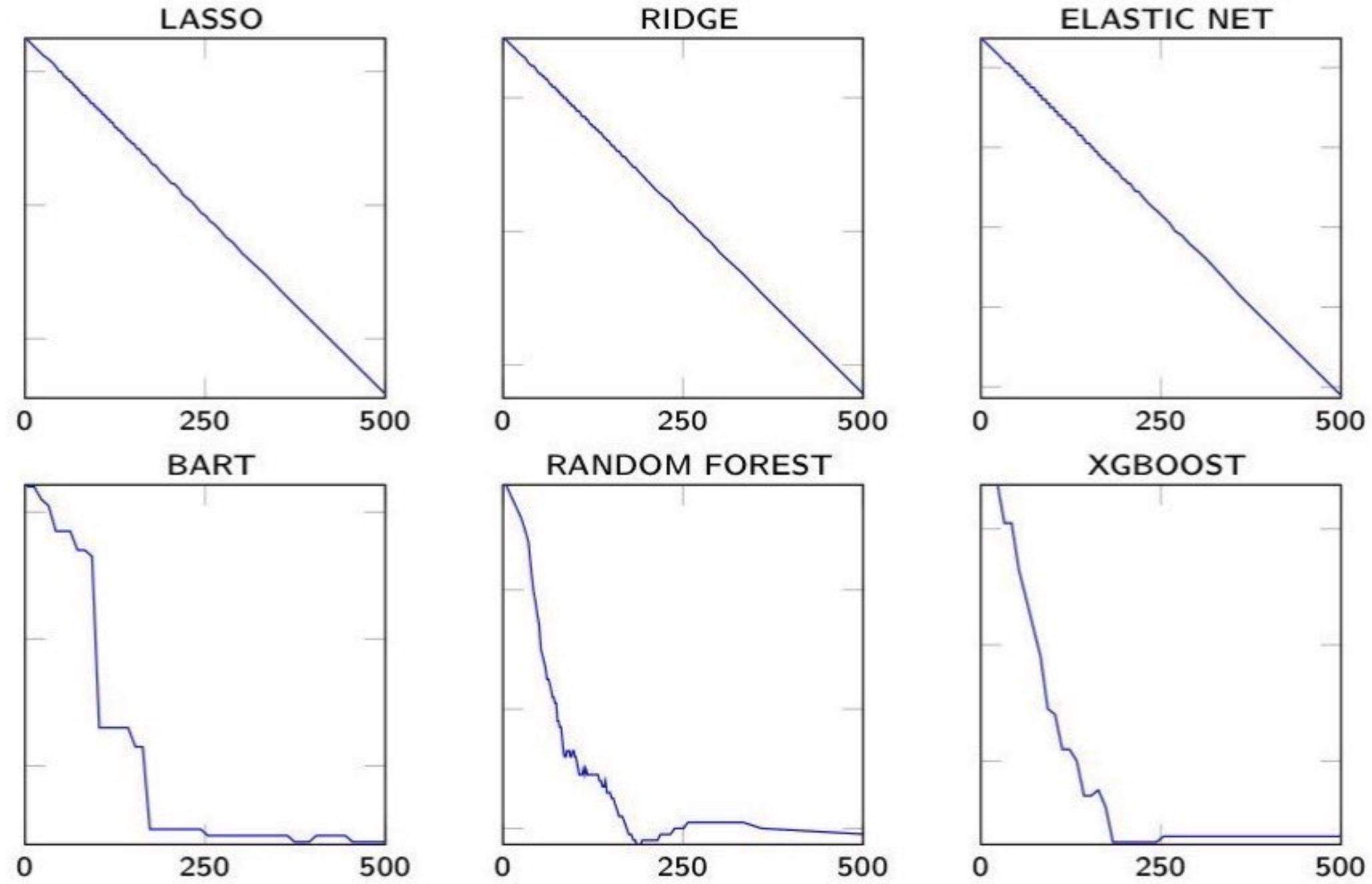


Figure: Partial Dependence Plots (PDPs) in the Italian case for the "THERMAL_EFFICIENCY" characteristic across LASSO, RIDGE, ELASTIC NET, BART, RANDOM FOREST, and XGBOOST. Plots illustrate the relationship between thermal efficiency (measured in kWh/m^2) on the x-axis and the predicted potential variation of energy performance on the y-axis.

Feature importance: Energy Efficiency Determinants - Italy

BART	RANDOM FOREST	XGBOOST
Current Environmental Impact	Current Environmental Impact	Current Environmental Impact
R30: Flat roof insulation	Number of Recommendations	Number of Recommendations
R7: 50 mm internal or external wall insulation	Current Heating Cost	Main Heating System Energy Efficiency
Main Heating System Energy Efficiency	Current Energy Consumption	Current Heating Cost
Current Heating Cost	Local Authority Label	R23: Low energy lighting for all fixed outlets
Number of Recommendations	Main Heating System Energy Efficiency	R7: 50 mm internal or external wall insulation
R40: High heat retention storage heaters	R7: 50 mm internal or external wall insulation	Current Energy Consumption
R39: HHRS heaters - dual immersion cylinder	Current CO2 Emissions	Main Heating System Environmental Efficiency
Walls Energy Efficiency	Walls Energy Efficiency	R30: Flat roof insulation
R5: Increase loft insulation to 270 mm	Hot Water Energy Efficiency	Walls Energy Efficiency
R15: Replace boiler with new condensing boiler	Walls Environmental Efficiency	Hot Water Energy Efficiency
R6: Cavity wall insulation	Current Cost Hot Water System	Current Cost Hot Water System
Current Cost Hot Water System	Construction Period	R39: HHRS heaters - dual immersion cylinder
Main Heating System Environmental Efficiency	R23: Low energy lighting for all fixed outlets	R34: Heat recovery system for mixer showers
Current CO2 Emissions	Low Energy Lighting	Low Energy Lighting

Table: **Variable Importance Rankings, displaying bar chart weights for the top 15 variables across BART, RANDOM FOREST, and XGBOOST models in the UK case.**

Extract the most relevant feature, including recommendations

Selected features consistent across models and countries

Prediction: MSE - UK

	Full sample		Subsample	
	In sample	Out of sample	In sample	Out of sample
	Mean Square Error			
LASSO	0.1635	0.1650	0.1630	0.1771
RIDGE	0.1706	0.1721	0.1705	0.1860
ELASTIC NET	0.1633	0.1647	0.1655	0.1789
BART	0.0690	0.0735	0.0689	0.1026
RF	0.1166	0.1185	0.1393	0.1478
XGBOOST	**0.0402**	**0.0556**	**0.0268**	**0.1000**

Table: **Mean Square Error** between actual and predicted values estimated by LASSO, RIDGE, ELASTIC NET, BART, RANDOM FOREST, and XGBOOST. In-sample and out-of-sample results for the whole sample (first and second column) and a random subsample (third and fourth column).

Non-linear models have higher predictive performance

Prediction: MAE - UK

	Full sample		Subsample	
	In sample	Out of sample	In sample	Out of sample
	Mean Absolute Error			
LASSO	0.3085	0.3092	0.3112	0.3192
RIDGE	0.3151	0.3160	0.3172	0.3269
ELASTIC NET	0.3083	0.3090	0.3139	0.3213
BART	0.1895	0.1945	0.1961	0.2302
RF	0.2497	0.2514	0.2697	0.2768
XGBOOST	**0.1383**	**0.1613**	**0.1177**	**0.2231**

Table: **Mean Absolute Error** between actual and predicted values estimated by LASSO, RIDGE, ELASTIC NET, BART, RANDOM FOREST, and XGBOOST. In-sample and out-of-sample results for the whole sample (first and second column) and a random subsample (third and fourth column) for the UK case.

Non-linear models have higher predictive performance

Partial Dependence: Non-Linearities - UK

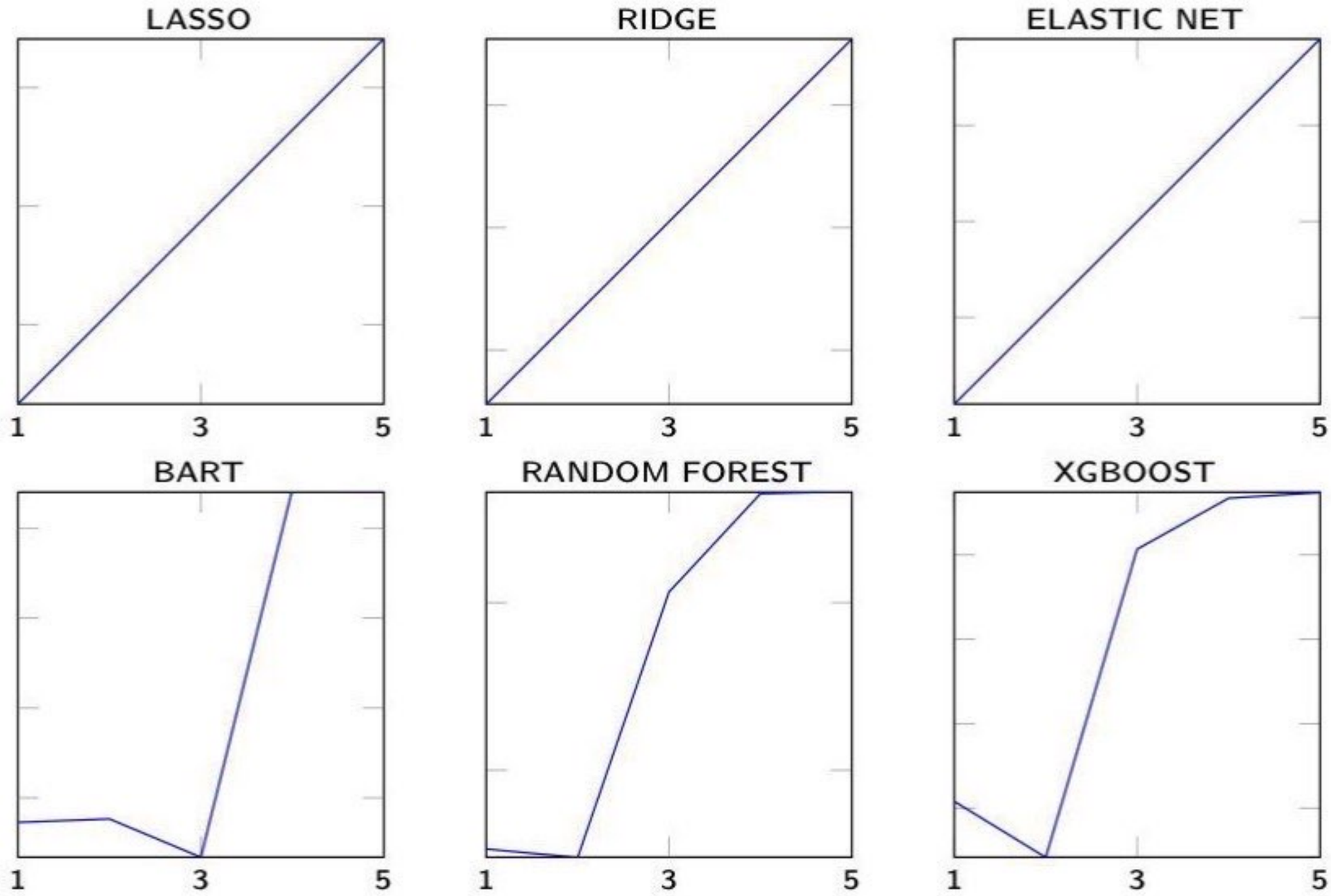


Figure: Partial Dependence Plots (PDPs) in the UK case for the "WALLS_ENERGY_EFF" characteristic across LASSO, RIDGE, ELASTIC NET, BART, RANDOM FOREST, and XGBOOST. Plots illustrate the relationship between the rating of a building's wall efficiency on the x-axis and the predicted potential variation of energy performance on the y-axis.

Policy Design and Evaluation

Policy Design and Evaluation - Italy and UK

Policy Design - Italy

We set up two comparable policies:

a. **Model Driven**: the government subsidises interventions selected by XGBOOST

b. **Benchmark**: the government subsidises interventions for the least efficient buildings

Given number of subsidized buildings

- Policies target a similar number of buildings (around 165k);
- Model Driven: **one selected interventions**;
- Benchmark: subsidies to **class F and G buildings**.

Policy Evaluation - Italy

Table: Results for the Model Driven and Benchmark policies based on Italian data include the number of subsidized buildings, the average payback period of funded interventions in years, and the average energy performance increase per intervention in kWh/m^2 per year.

Policy	N.	Average Payback Period	Average Δ EE
Model Driven	164,210	12.04 years	83.31 kWh/m^2
Benchmark	167,333	14.94 years	59.44 kWh/m^2
Percentage Deviation	-1.87%	-19.41%	40.16%

The policy based on our analysis subsidises interventions that are more convenient in terms of payback period and energy performance increase.

Policy Design - United Kingdom

We set up two comparable policies:

a. **Model Driven**: the government subsidises interventions selected by XGBOOST

b. **Benchmark**: the government subsidises interventions for the least efficient buildings

Given subsidy budget

- Policies have the same allocated **budget** (£10 million);
  $\rightarrow$ leverage on data for intervention specific cost
- Model Driven: **five selected interventions**;
  $\rightarrow$ weights based on variable importance measure for XGBOOST
- Benchmark: subsidies to **class G buildings**.

Policy Evaluation - United Kingdom

Table: Results for the Model Driven and Benchmark policies based on the average cost per intervention in £, the average energy performance increase at aggregate levels in kWh/m^2 per year, and aggregate reduction in CO_2 emissions in tonnes per year.

Policy (Budget £10M.)	**Average Cost**	**Average Δ EE**	**Average Δ CO2**
Model Driven	£86.40	1,043,702.06 kWh/m^2	10,512.95 tonnes
Benchmark	£3,486.34	196,651.2 kWh/m^2	2,982.71 tonnes
Percentage Deviation	-97.52 %	+469.26%	+252.46%

The policy based on our analysis achieves higher aggregate increase in energy efficiency and reduction in CO2 emissions.

Conclusions

1. We leverage on the expertise of assessors stored in EPCs to propose a coherent framework for potential energy performance prediction;
2. Our analysis uncovers non-linearities between features and target variable, which are well approximated by tree based models;
3. We are able to select the most relevant features and recommendation among a large number of candidates;
4. Building on this, we design ML-based policies that achieve good results in terms of cost and effectiveness.

Thank you!

Monica Billio, Roberto Casarin, Michele Costola, Veronica Veggente (2024). *Learning from experts: Energy efficiency in residential buildings*, **Energy Economics**, forthcoming. Qcode.p

References I

Monica Billio, Michele Costola, Loriana Pelizzon, and Max Riedel. Buildings' energy efficiency and the probability of mortgage default: The dutch case. *The Journal of Real Estate Finance and Economics*, May 2021. ISSN 1573-045X. doi: 10.1007/s11146-021-09838-0.

Justin Bleich, Adam Kapelner, Edward I George, and Shane T Jensen. Variable selection for BART: an application to gene regulation. *The Annals of Applied Statistics*, pages 1750–1781, 2014.

Leo Breiman. Random forests. *Machine Learning*, 45:5–32, 2001.

Tianqi Chen and Carlos Guestrin. Xgboost: A scalable tree boosting system. In *Proceedings of the 22nd acm sigkdd international conference on knowledge discovery and data mining*, pages 785–794, 2016.

Hugh Chipman and Robert McCulloch. *BayesTree: Bayesian Additive Regression Trees*, 2016. URL `https://CRAN.R-project.org/package=BayesTree`. R package version 0.3-1.4.

Jerome Friedman, Robert Tibshirani, and Trevor Hastie. Regularization Paths for Generalized Linear Models via Coordinate Descent. *Journal of Statistical Software*, 33(1):1–22, 2010. doi: 10.18637/jss.v033.i01.

Meritxell Gómez-Omella, Iker Esnaola-Gonzalez, Susana Ferreiro, and Basilio Sierra. k-nearest patterns for electrical demand forecasting in residential and small commercial buildings. *Energy and Buildings*, 253:111396, 2021.

References II

Benjamin Guin, Perttu Korhonen, and Sidharth Moktan. Risk differentials between green and brown assets? *Economics Letters*, 213:110320, 2022.

Rangan Gupta, Christian Pierdzioch, and Marian Risse. On international uncertainty links: BART-based empirical evidence for canada. *Economics Letters*, 143:24–27, 2016.

Jichao He, David W Wanik, Brian M Hartman, Emmanouil N Anagnostou, Marina Astitha, and Maria EB Frediani. Nonparametric tree-based predictive modeling of storm outages on an electric distribution network. *Risk Analysis*, 37(3):441–458, 2017.

Andy Liaw and Matthew Wiener. Classification and Regression by randomForest. *R News*, 2(3): 18–22, 2002. URL `https://CRAN.R-project.org/doc/Rnews/`.

Muhammad Uzair Mehmood, Daye Chun, Hyunjoo Han, Gyuyeob Jeon, Kuan Chen, et al. A review of the applications of artificial intelligence and big data to buildings for energy-efficiency and a comfortable indoor living environment. *Energy and Buildings*, 202: 109383, 2019.

Rodney Sparapani, Charles Spanbauer, and Robert McCulloch. Nonparametric machine learning and efficient computation with Bayesian additive regression trees: The BART R package. *Journal of Statistical Software*, 97(1):1–66, 2021. doi: 10.18637/jss.v097.i01.

Steps to become an ENGAGE Portal Test User

Marco Angheben

European DataWarehouse & ENGAGE Project Coordinator

Why Become an ENGAGE Test User?

How to ENGAGE as a Test User

1. Request the ENGAGE Templates:
https://forms.office.com/e/td14aYsMQS

2. Set up a 30 minute one-to-one demo with the ENGAGE Team (upon request)

Optional step

3. Sign the ENGAGE agreement to submit the sample data

4. Submit your ENGAGE Templates (mortgage sample file) into the ENGAGE Portal

5. Receive the ENGAGE report on EU Taxonomy compliance

All interested institutions are invited to test the ENGAGE Templates and Portal upon request to engage4esg@eurodw.eu
A specific legal framework has been prepared for the safe and lawful processing of the data.
The testing of the ENGAGE Templates and Portal is free of charge until 31 October 2025.

How Does it Work in Practice?

Request ENGAGE Templates

- Version 1.1 available now through Request for Access to the ENGAGE Templates
- More than 50 institutions have requested access to the ENGAGE Templates

ENGAGE Demonstration

- The ENGAGE Team is available for clarifications on the ENGAGE Templates
- Several sessions have already been set up with interested parties

ENGAGE Template Submission to Portal

- Available from **November 2024**, subject to a standard legal arrangement
- Free of charge until October 2025
- Possibility to submit the ENGAGE sample files various times enriching the information

ENGAGE Report on EU Taxonomy Compliance

- For each sample file submission a compliance report is generated

ENGAGE Webinar for the Dutch market

Registration through email to events@hypoport.com

ENGAGE Autumn Webinar Series

CLICK BELOW TO REGISTER:

- Thursday, **24 October** 15:00 (CET)
- Thursday, **21 November** 10:00 (CET)

The ENGAGE Consortium Partners

EUROPEAN DATAWAREHOUSE

HYPOPORT

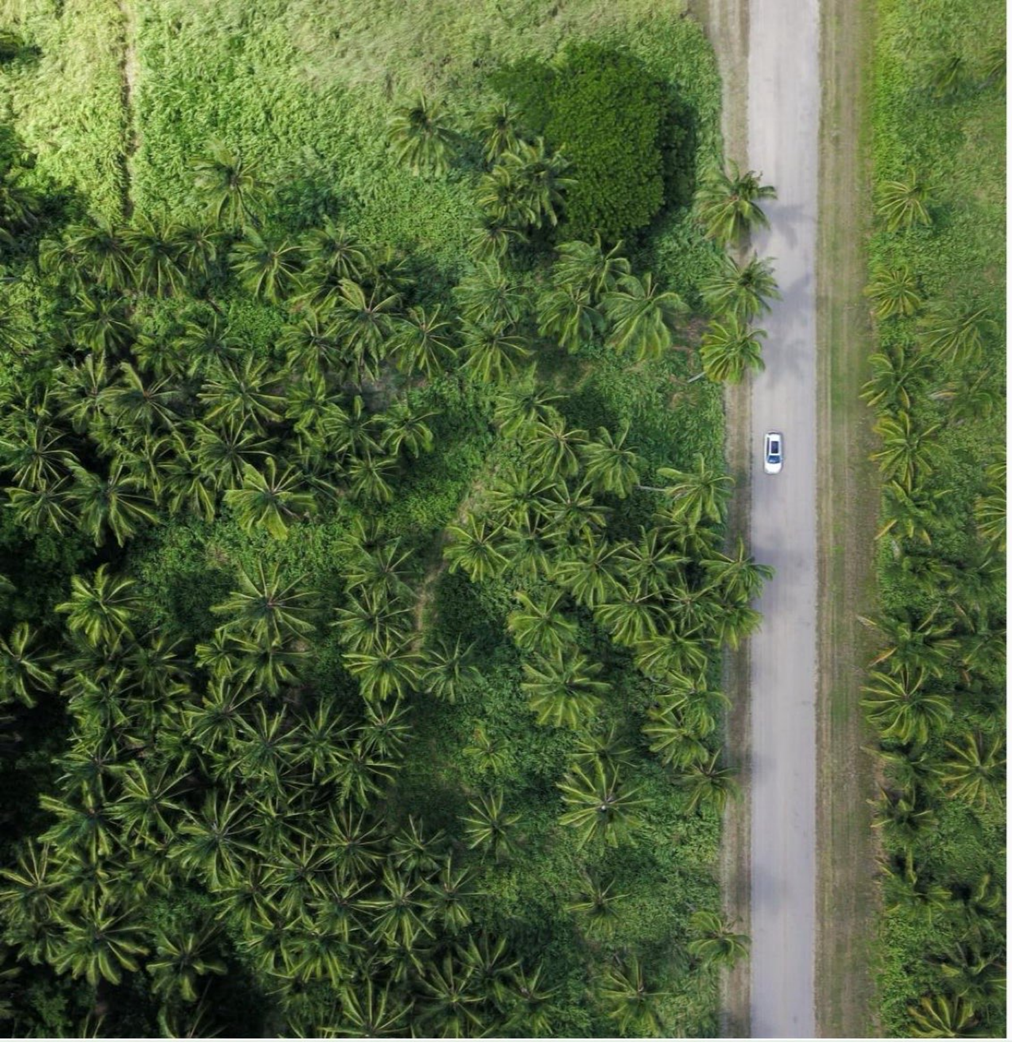
UCI

woonnu

VENETIARVM VNIVERSITAS IN DOMO FOSCARI

Ca' Foscari University of Venice

DEXAI ARTIFICIAL ETHICS



ENGAGE
for ESG

Website:
engage4esg.eurodw.eu/

Social Media:
https://www.linkedin.com/company/engage-for-esg-activation-investments/

Co-funded by the European Union. Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or CINEA. Neither the European Union nor the granting authority can be held responsible for them.

Project Coordinator:
Marco Angheben
marco.angheben@eurodw.eu

ENGAGE General Contact:
engage@eurodw.eu

Communications
Carla Scarsella
carla.scarsella@eurodw.eu