PORTFOLIO ALIGNEMENT MEASUREMENT

RESIDENTIAL REAL ESTATE MORTGAGES PROPOSED METHODOLOGY

July 2021

Lead Authors:
Imène Ben Rejeb-Mzah, BNP Paribas
Paul Grimal, Société Générale

Acknowledgments:
Patrick Bader, FBF
Delphine Bartre, BPCE
Marie-Eve Chardon, Crédit Agricole
Julien Delorme, Sebastien Depailler, HSBC
Denis Majou, Crédit Mutuel
Zineb Tazi, La Banque Postale

Table of content

1. INTRODUCTION

This note offers a first attempt to define the KPIs and a methodology to measure the alignment of French retail real estate mortgages. The objective of this memo will therefore specifically be to:

- define the KPIs (scope, indicators, scenario) used for the retail real estate portfolio;
- offer methodological guidance to measure the retail real estate portfolio's alignment and the difficulties encountered, especially in the handling of the ADEME's[1] Energy Performance Certificates (EPC) database;
- propose guidance for calculating absolute greenhouse gas (GHG) emissions ('financed emissions');
- discuss future research and next steps to improve this measurement.

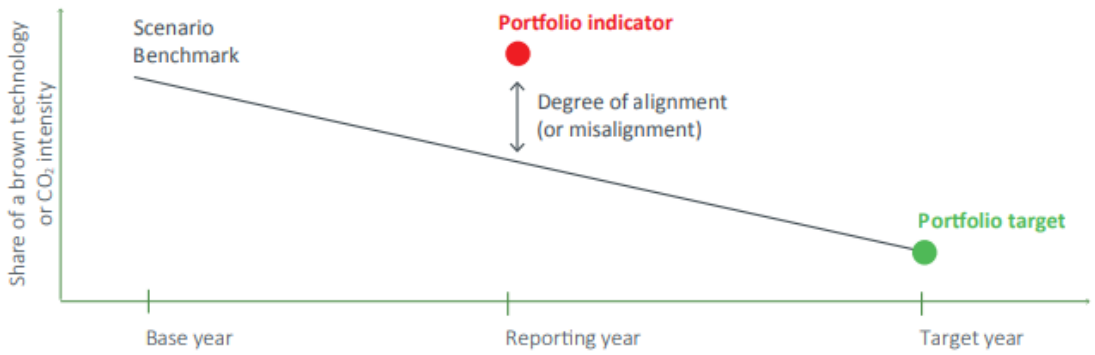
1.1. Portfolio alignment measurement

Alignment can be defined as a process that requires measuring the performance of the portfolio compared with the scenario metrics (like CO2 intensity of the sector portfolio at a point in time compared with sustainable development scenario required CO2 intensity at that point of time).

It also requires setting targets derived at least from a 'below 2° scenario and it requires reorienting or steering financial flows so that they stay on track with the trajectory and the final goal. The last step of the process is to track progress to ensure that the portfolio remains on track to meet the final target. [1]

It is important to note that the purpose of this note is to propose a methodology to measure the alignment of residential real estate mortgage portfolio with a defined scenario compatible with the Paris Agreement. It is by no means a guide on how to achieve such alignment, as each bank will have to determine how it intends to achieve such alignment based on the characteristics of its portfolio, its commercial policiesand public policies and regulatory requirements. It is important to note that the residential real estate portfolio is highly illiquid in the sense that banks commit to provide mortgages for a very long duration (up to 25 years). As such the implementation of such alignment is expected to be achieved on a medium to long term basis and the banks are only at the beginning of such journey.

Figure 1: Illustration of alignement at portfolio level

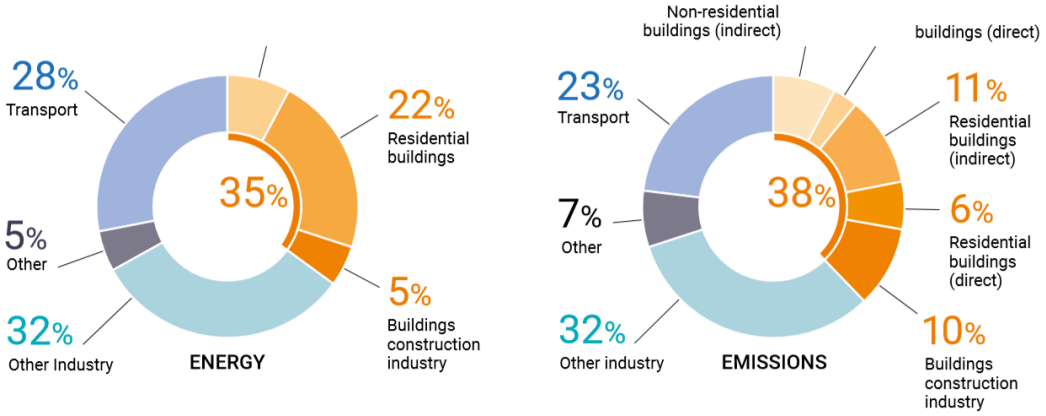


[1] *ADEME (*Agence de l'environnement et de la maîtrise de l'énergie*) is the French Environment protection agency.*

1.2. Real estate sector climate relevance

35% of global energy demand and 38% of global GHG emissions are due to the real estate sector, according to the IEA. [2]

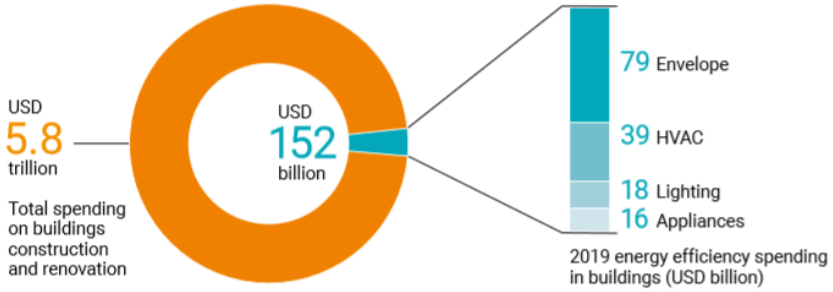
Figure 2: Global share of buildings and construction final energy and emissions, 2019



IEA 2020d; IEA 2020b

At the same time, the share of energy efficiency among real estate investments is very small. It should double on average in the next 10 years, according to the SDS scenario. [2]

Figure 3: Building construction and energy efficiency investment and breakdown



Source: IEA (2019b). All rights reserved. Adapted from "Energy efficiency policies: Buildings"

2. KPIS FOR THE RETAIL REAL ESTATE PORTFOLIO

Energy intensity by floor area (in kWh/sqm/yr) and GHG emissions intensity by floor area (kgCO2e/sqm/yr) are clearly the most relevant metrics to measure alignment of real estate mortgages.

The rationale for selecting these KPIs is based on the fact that around 86% of real estate energy consumption and 74% of real estate GHG emissions are attributable to building utilisation. In addition, most scenarios and counterparty data are expressed in the same units. This is the case for example of the Stratégie Nationale Bas Carbone (SNBC) and ADEME's EPC database. It is also the indicator used in the EU taxonomy.

In addition, banks may find it also useful to report the portfolio's absolute emissions (or financed emissions). A section also proposes how to measure this indicator

Regarding Energy intensity, two metrics could be considered: final energy intensity and primary energy intensity. Since the latter is used to compute energy label for the EPC, it is generallty more frequently filled out and its data are of better quality.

3. ALIGNMENT METHODOLOGY

3.1. Data inputs

- The bank's retail real estate loans database (with addresses or GPS coordinates)
- ADEME's EPC database[2] contains addresses, coordinates, the associated KPIs above and several other columns that give additional information about the accommodation. The new release of this database contains coordinates for around 10 million lines corresponding to around 4 million different addresses (to be compared with around 20 million French building addresses according to the French national land registry. This is to be compared with a 36 million housing identified by the French statistics agency [INSEE[3]])
- Benchmark scenario: scenarios are used to measure alignment. Most banks prefer using the SNBC pathway because it is France-specific and hence it is the most suitable to assess a French mortgage portfolio and easier to communicate in France. It provides a residential/tertiary distinction (and even a segmentation within tertiary, e.g. offices, hotels, commercial centres). Other acceptable options include the IEA's SDS, etc…

3.2. Modelling structure

The figure that follows gives a summary of the modelling flow.

First, Real Estate Mortgage (REM) deals are matched with ADEME's EPC database. Two techniques are possible for the matching step:

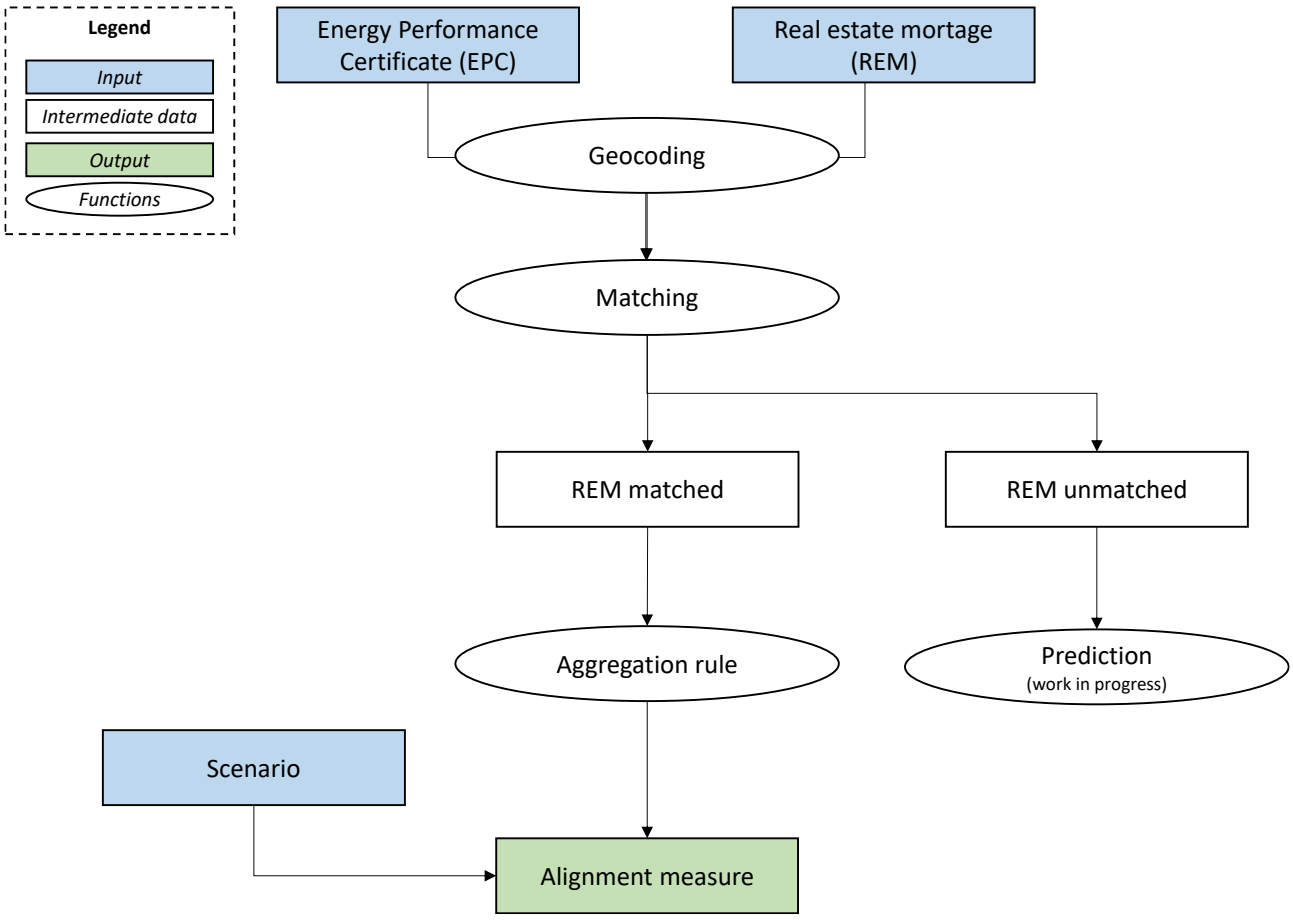
- The addresses in the REM and EPC are geocoded (i.e. the precise GPS coordinates are found) and the two databases are matched on this basis (see 3.3 Geocoding REM); or
- The addresses are matched based on character strings (e.g. « 10 bld Saint Louis » should be matched with « 10 boulevard Saint Louis »). Approximate string matching (often colloquially referred to as 'fuzzy string searching') is a technique for finding strings that match a pattern approximately (rather than exactly). It should be noted that the algorithm output may need some manual validation when the matching score is not 100 (100 being the score associated to a perfect match). This is a highly time-consuming process. Second, for REM that are unmatched, several options are possible to attempt to predict the energy/emission intensity of the dwellings, typically relying on proxies.

Second, the average energy/carbon intensity of the portfolio is calculated using an aggregation formula.

Third, this average portfolio's energy/carbon intensity is then compared to the scenario to evaluate alignement.

[2] *The DPE (diagnostics de performance énergétique) database is available online: https://data.ademe.fr/datasets/dpe-france*
[3] *https://www.insee.fr/fr/statistiques/*

Figure 4: Modelling flow & structure



3.3. Geocoding REM

ADEME's EPC database comes with addresses and geocodes (GPS coordinates). To match based on geocodes, REM are geocoded using the same geocoding API.

However, even if a geocode is returned for most addresses, the geocoding score[4] (indicator of quality for the geocode) may not be high enough to ensure that the coordinates correspond correctly to the geocode. It is then important to evaluate what constitutes a good geocoding score. Banks still need to identify what a reliable geoscore means.

Alternatively, to improve the matching rate, the remaining addresses can be matched using exact or approximate string matching (fuzzy string searching).

[4] For the address '1 rue de Rivoli 75004 Paris', the geocoding API (*https://geo.api.gouv.fr/adresse*) returns the following JSON object:

{"type": "FeatureCollection", "version": "draft", "features": [{"type": "Feature", "geometry": {"type": "Point", "coordinates": [2.359276, 48.855534]}, "properties": {"label": "1 Rue de Rivoli 75004 Paris", "score": 0.705720826446281, "housenumber": "1", "id": "75104_8249_00001", "name": "1 Rue de Rivoli", "postcode": "75004", "citycode": "75104", "x": 652987.24, "y": 6861912.49, "city": "Paris", "district": "Paris 4e Arrondissement", "context": "75, Paris, \u00cele-de-France", "type": "housenumber", "importance": 0.67202, "street": "Rue de Rivoli"}}], "attribution": "BAN", "licence": "ETALAB-2.0", "query": "1 rue de Rivoli 75004 Parislimit=1", "limit": 5}

The score, a variable between 0 and 1, refers to the relevance of the geocoding.
The geocoding API may return several results for the same address, we choose the one with the highest score.

3.4. Processing duplicates

The ADEME database comes with duplicate EPCs by addresses; there are two reasons for the presence of such duplicates:

- For individual residences (houses), this could be due to EPCs being renewed as part of a new transaction. Indeed, EPCs become out of date after 10 years and homeowners have an obligation to renew them if they wish to sell their homes. Here the only available approach to deal with duplicates is to retain the most recent EPC value.
- For collective residences (apartments), the same reason may apply. Alternatively, it could be due to EPCs belonging to different dwellings in the same building. Here most banks prefer to take a simple average of the EPCs, as it is simpler to communicate. Some banks may prefer the surface-weighted average though the accuracy of this approach needs to be further tested.

It should be noted that if the REM's surface data is available in the Bank's information system, it could be used to find the specific apartment in the ADEME's database. If more than one apartment has the same surface, the average of their climate performance indicator could be considered.

NB: Banks only using EPC labels in their portolio climate analysis could use the modal value of the residence.

3.5. Predicting unmatched REM

Matched addresses are those that are "perfectly" linked to the ADEME database: i.e identified using addresses or geocodes (with 6 decimals) – though there may be false positives given that geocoding may have been erroneous. These matched addresses may be complemented with EPCs available in the bank's internal IT.

However, a portion of REM may remain unmatched. This may be explained by:

- REM adresses that are not in ADEME's database.
- Some REMs are without (properly written) addresses and therefore cannot be geocoded or matched.

To deal with these issues and improve the EPC coverage of the portfolio, some banks prefer to exclude unmatched addresses from the computation until they develop a robust proxy based on third-part data (using a statistical or machine-learning estimate). Other banks use a simple proxy (neighbours' average indicator).

3.6. Calculating the portfolio average emission and energy intensity using an aggregation rule

Once we have the corresponding EPC for each REM, we calculate the portfolio average emission intensity and energy intensity.

There are two options for this aggregation rule:

- **Option 1:** A PACTA-like or portfolio-weighted approach. This is a simple average of the energy/emission intensity weighted by loan size. The advantage of this approach is that it is simple to implement as it does not require information on a dwelling's surface or the Loan-to-Value (LTV) ratio. However, this approach gives more weight to expensive dwellings (which are generally associated with bigger loans), thereby introducing a significant bias.

$$Portfolio\ average\ emission\ intensity = \ gCo2e/m2_{pf,y} = \frac{\sum_a (gCo2e/m2_{a,y} * LoanSize_{a,y})}{\sum_a LoanSize_{a,y}}$$

where pf=portfolio, a=asset, y=year, $gCo2e/m2_{a,y}$ = emission intensity of the asset expressed in gCO2e per square metre and per year

$$Portfolio\ average\ energy\ intensity = \ KWHep/m2_{pf,y} = \frac{\sum_a (KWH/m2_{a,y} * LoanSize_{a,y})}{\sum_a LoanSize_{a,y}}$$

where pf=portfolio, a=asset, y=year, $KWH/m2_{a,y}$ = energy intensity of the asset expressed in KWHper square metre and per year

- **Option 2:** A PCAF-like or financed emissions approach. The average energy/emission intensity is calculated on the basis of the dwelling surface attributed to the portfolio. The attribution factor is the Loan-to-Value (LTV) ratio. This approach requires more information (surface and value of the asset financed), which may not always be available or of optimal quality. Nonetheless, it does not introduce the asset price bias that comes with option 1.

$$Portfolio\ average\ emission\ intensity = \frac{gCo2e}{m2}_{pf,y} = \frac{\sum_a \frac{gCo2e}{m2}_{a,y} * m2_{a,y} * LTV_{a,y})}{\sum_a (m2_{a,y} * LTV_{a,y})}$$

where pf=portfolio, a=asset, y=year, LTV = Loan-to-Value ratio at the time of engagement: outstanding amount divided by the property value at the origination, $m2_{a,y}$ = asset surface area, $gCo2e/m2_{a,y}$ = emission intensity of the asset a expressed in gCO2e per square metre and per year

$$Portfolio\ average\ energy\ intensity = KWH/m2_{pf,y} = \frac{\sum_a KWH/m2_{a,y} * m2_{a,y} * LTV_{a,y})}{\sum_a (m2_{a,y} * LTV_{a,y})}$$

where pf=portfolio, a=asset, y=year, LTV = Loan-to-Value ratio at the time of engagement, $m2_{a,y}$ = asset surface area, $KWH/m2_{a,y}$ = energy intensity of the asset a expressed in KWH per square metre and per year

NB: if banks have LTV data available, the original home value would avoid volatily of the LTV due to changes in asset prices. However, this may depend on the data that banks have available on their systems. Alternatively, the last LTV value based on the most updated property value could be used.

Though the portfolio-weighted approach is simpler to implement, most banks prefer to use the PCAF-like approach (emissions attributed to the portfolio divided by the square metres attributed to the portfolio) because there is no bias towards higher asset price. It also enables computation of total attributed absolute emissions and is easier to communicate. However, it requires additional data (asset surface, Loan-to-Value ratio).

3.7. Alignment

The degree of alignment between the portfolio average EPC can be assessed against the level of the benchmark from the chosen climate scenario. The portfolio average EPC is considered 'aligned' if the level of the indicator is below that of the benchmark for decreasing benchmarks ('brown' activities like REM). Figure 2 provides an illustration of a portfolio being misaligned with a benchmark. Alignment can also be measured at portfolio, client or asset level. For assets and clients, it provides information that can be used to inform and eventually accompany clients where misalignment is greatest.

3.8. Absolute emissions methodology

The PCAF approach to measuring absolute emissions could be used to compute the portfolio's attributed emissions

$$Portfolio\ carbon\ footprint\ = \sum_{a} \frac{gCo2e}{m2}_{a,y} * m2_{a,y} * LTV_{a,y})$$

where pf=portfolio, a=asset, y=year, LTV = Loan-to-Value ratio at the time of engagement outstanding amount divided by the property value at the origination, $m2_{a,y}$ *= asset surface area,* $gCo2e/m2_{a,y}$ *= emission intensity of the asset a expressed in gCO2e per square metre and per year*

4. NEXT STEPS AND FUTURE RESEARCH

- Liaise with ADEME to share the methodology, the difficulties encountered and ways of improving the EPC database
- Identify solutions to improve the data quality of internal database
- Fine-tune the SNBC benchmark pathway to derive a target pathway expressed in the same units and computed on the same scope as the EPC indicators.
- How to include work loans (prets travaux) in the portfolio
- How to project the portfolio performance indicator (Should we take into account refinancing (Rachats)…)
- Adapt to corporate portfolio

5. REFERENCES

[1] CREDIT PORTFOLIO ALIGNMENT – AN APPLICATION OF THE PACTA METHODOLOGY BY KATOWICE IN PARTNERSHIP WITH THE 2 DEGREES INVESTING INITIATIVE

[2] 2020 GLOBAL STATUS REPORT FOR BUILDINGS AND CONSTRUCTION

[3] STRATEGIE NATIONALE BAS-CARBONE – MARS 2020