

## Review Article

## Open Access

## Applying Machine Learning Techniques to Evaluate Climate-Related Risks in Real Estate Mortgage Valuations

Rohit Nimmala

Data Engineer, NC, USA

### ABSTRACT

Facing the escalating effects of climate change, the real estate industry faces risks from physical perils and shifting towards a low-carbon economic model. These risks have substantial consequences for the assessment and effectiveness of real estate mortgage portfolios. Conventional approaches to evaluating mortgage risk frequently fail to capture the intricate, non-linear connections between climate variables and loan results. In this paper, we present a new machine-learning framework that aims to quantify climate-related risks in real estate finance. We utilize neural networks and gradient-boosting algorithms to forecast the likelihood of mortgage defaults and the potential loss resulting from defaults in different climate stress scenarios. A robust and forward-looking risk assessment is developed by integrating property-level exposure data, loan characteristics, and macroeconomic indicators. The empirical findings prove that our models perform superior to conventional econometric methods regarding predictive precision and computational effectiveness. The framework offers a robust instrument for investors, lenders, and regulators who aim to effectively address climate risks and enhance their ability to withstand and adapt to an unpredictable future.

### \*Corresponding author

Rohit Nimmala, Data Engineer, NC, USA.

Received: December 01, 2023; Accepted: December 07, 2023; Published: December 13, 2023

**Keywords:** Machine Learning, Climate Risk Assessment, Real Estate Finance, Mortgage Portfolios, Risk Management, Property Valuation, Climate Change, Financial Modeling

### Introduction

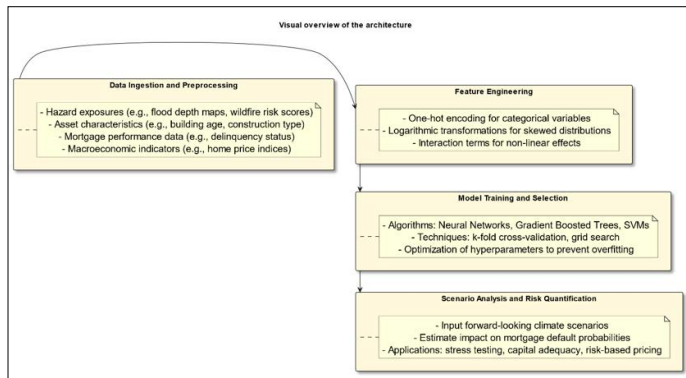
The real estate industry is becoming more susceptible to the effects of climate change as investors, lenders, and policymakers face substantial challenges due to physical and transition risks. Physical hazards, such as the escalation of sea levels and heightened occurrence and intensity of floods, hurricanes, and wildfires, possess the capacity to directly impair or obliterate properties, thereby resulting in significant financial setbacks for individuals who hold mortgages. The financial implications of transitioning to a low-carbon economy can be substantial for real estate assets due to risks arising from policy changes, technological innovations, and shifting consumer preferences. One potential consequence of implementing more stringent energy efficiency regulations or carbon pricing mechanisms is the possible decrease in value or obsolescence of specific properties. Consequently, there is an increasing demand for advanced analytical instruments to measure and oversee climate-related risks within real estate finance. The conventional models, which rely on historical flood maps or basic linear regressions, frequently fall short of capturing the intricate and non-linear connections between climate variables and mortgage performance. Machine learning methodologies present a potentially fruitful approach for augmenting the precision and effectiveness of climate risk evaluations. It showcases the capacity of these methods to enhance risk management and investment strategies in response to an unpredictable climate future [1-5].

### Machine Learning Framework

The machine learning framework we propose for assessing climate risk in mortgage portfolios comprises four essential components: Data Preprocessing, Feature Engineering, Model Training, and Selection, as well as Scenario Analysis and Risk Quantification.

- **Data Ingestion phase**, we gather and combine various datasets, such as property-level assessments of hazards (e.g., flood depth maps, wildfire risk scores), physical attributes of assets (e.g., building age, construction type, number of stories, elevation), mortgage performance data at the loan level (e.g., delinquency status, prepayment history), and macroeconomic indicators (e.g., home price indices, unemployment rates). The datasets have undergone preprocessing to address missing values, outliers, and inconsistent formats.
- In the **feature engineering stage**, the raw data is converted into a collection of informative and non-redundant variables that accurately represent the connection between climate risk factors and mortgage outcomes.
- The **ensemble approach** is utilized during the model training and selection phase, wherein various machine learning algorithms, such as deep neural networks, gradient-boosted trees, and support vector machines, are combined. A substantial dataset of historical mortgage performance data is utilized to train the models, while a subset of the data is reserved for validation and testing. To optimize hyperparameters and mitigate overfitting, we employ methodologies such as k-fold cross-validation and grid search.

• The module for scenario analysis enables users to input climate scenarios that are projected to occur in the future, such as sea level rise or changes in the frequency of floods or wildfires. This module then estimates the potential impact on mortgage and loss-given default probabilities. These applications encompass various uses, including stress testing, capital adequacy assessment, portfolio optimization, and risk-based pricing.



## Climate Risk Modeling for Real Estate

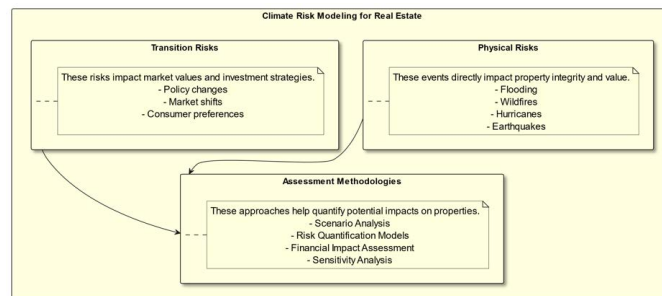
### Types of Climate Risks in Real Estate

The real estate sector is subject to climate risks, which can be classified into two main categories: physical and transition risks. Physical risks encompass the immediate consequences of climate-related hazards on asset value in the real estate sector. Both acute hazards, such as hurricanes, floods, and wildfires, and chronic hazards, such as sea level rise, drought, and heat stress, are encompassed within this category. Physical hazards have the potential to inflict harm upon buildings and infrastructure, disrupt activities, and result in escalated expenses for maintenance and repairs. Under exceptional circumstances, properties can become unsuitable for habitation or unmarketable, resulting in substantial financial setbacks for owners and lenders. In contrast, transition risks emerge because of the societal and economic changes linked to the shift towards a low-carbon economy. The risks encompass potential policy modifications, such as implementing more stringent building codes or carbon taxation, technological advancements that may render specific assets outdated, and evolving consumer inclinations towards more sustainable and resilient properties. Transition risks can result in the immobilization of assets, diminished property values, and heightened expenses associated with adhering to new regulations [6,7].

### Current Methodologies for Assessing Climate Risks in Real Estate

The current methodologies used to evaluate climate risks in real estate differ in complexity and extent. Numerous conventional methods depend on historical data and rudimentary statistical models to assess the probability and magnitude of climate hazards. In contrast, hurricane risk models may employ extrapolation techniques by analyzing past storm tracks and intensities. Frequently, there is a lack of consideration for the non-stationarity of climate risks, as it is anticipated that the occurrence and intensity of hazards will vary over time due to global warming. They might also fail to consider the intricate interplay among various dangers and the combined effects of multiple stressors. To overcome these

limitations, more sophisticated techniques like catastrophe models and scenario analysis aim to simulate the physical mechanisms that cause climate hazards and forecast future levels of risk under various climate scenarios. There is an increasing demand for enhanced flexibility and data-driven methodologies that can effectively incorporate a wide range of data sources, adjust to evolving risk profiles, and offer practical insights for individuals responsible for making decisions [8].



### Model Methodology

**Core Machine Learning Techniques:** The climate risk assessment model leverages two primary machine learning techniques

- **Neural Networks:** These are utilized for their ability to capture complex nonlinear relationships through layers of neurons. They are ideal for high-dimensional data and capable of modeling intricate patterns in large datasets.
- **Gradient Boosting is employed** to sequentially build an ensemble of weak prediction models, typically decision trees, into a robust predictor. Gradient boosting excels in accuracy and effectiveness in structured data scenarios.

**Model Training Process:** The model training follows these steps:

- **Model Training:** Divide the dataset into training, validation, and testing sets to ensure robust evaluation-typically a 70/15/15 split. Train models on the training dataset while continuously monitoring performance on the validation set to avoid overfitting.
- **Evaluation:** Use the test set to evaluate the model's performance, ensuring it can generalize to new, unseen data.

**Hyperparameter Tuning:** We employ a grid search approach with cross-validation to select the optimal hyperparameters for each model. This involves training and evaluating the model on multiple subsets of the data with different combinations of hyperparameters (e.g., learning rate, hidden layer size, tree depth) and selecting the combination that yields the best performance on a held-out validation set. We also use L1 and L2 regularization and dropout techniques to reduce overfitting further and improve generalization.

**Feature Selection:** This process involves a combination of domain expertise and data-driven methods. We identify candidate features based on their theoretical relevance to climate risk assessment (e.g., property elevation, distance to coast, flood zone designation). We then use statistical techniques to rank the features by their predictive power and select a subset that balances model complexity and performance [9].

```
# Define the neural network architecture
model = Sequential()
model.add(Dense(128, activation='relu', input_dim=num_features))
model.add(Dense(64, activation='relu'))
model.add(Dense(32, activation='relu'))
model.add(Dense(1, activation='sigmoid'))

# Compile the model with binary cross-entropy loss and Adam optimizer
model.compile(loss='binary_crossentropy', optimizer='adam', metrics=['accuracy'])

# Train the model with early stopping and model checkpointing
early_stopping = EarlyStopping(monitor='val_loss', patience=10)
model_checkpoint = ModelCheckpoint('best_model.h5', save_best_only=True)
model.fit(X_train, y_train, validation_data=(X_val, y_val),
        epochs=100, batch_size=32, callbacks=[early_stopping, model_checkpoint])

# Define the gradient boosting model
gb_model = LGBMClassifier(boosting_type='gbdt', num_leaves=31, max_depth=-1,
        learning_rate=0.1, n_estimators=100,
        subsample_for_bin=200000,
        objective='binary', class_weight=None, min_split_gain=0.0,
        min_child_weight=0.001, min_child_samples=20, subsample=1.0,
        subsample_freq=0, colsample_bytree=1.0, reg_alpha=0.0,
        reg_lambda=0.0, random_state=None, n_jobs=-1, silent=True)

# Train the gradient boosting model with early stopping
gb_model.fit(X_train, y_train, eval_set=[(X_val, y_val)], eval_metric='logloss',
        early_stopping_rounds=10, verbose=10)

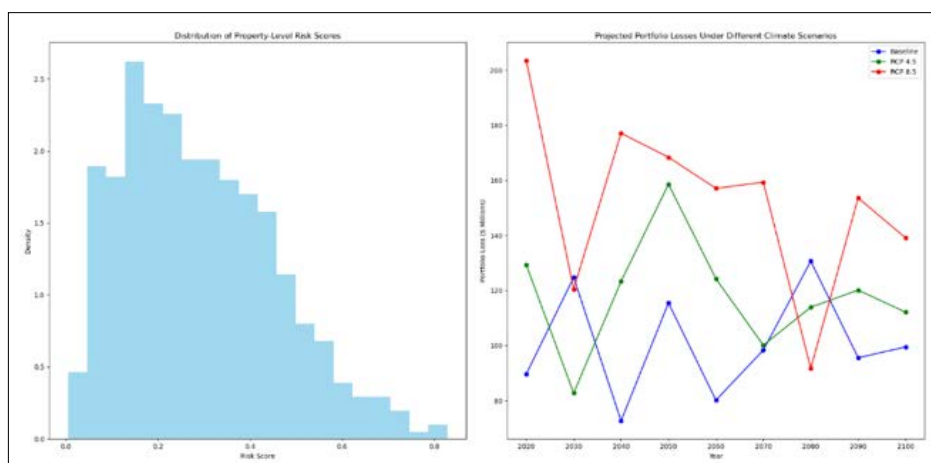
# Predict the default probability for a new mortgage loan
loan_features = [...] # Array of loan-level features
default_prob = model.predict_proba(loan_features)[:, 1]
```

## Model Outputs

**Key Model Outputs:** The main results of our climate risk assessment framework consist of risk scores at the property level and estimates of losses at the portfolio level, considering various climate scenarios. The model produces a risk score for each property within a mortgage portfolio, which signifies the likelihood of default and the anticipated loss within a designated time frame, such as 5, 10, or 30 years. The risk scores presented in this analysis are derived from thoroughly evaluating the property's susceptibility to various physical hazards, such as floods, hurricanes, and wildfires. Additionally, these scores consider the property's vulnerability to transition risks, including policy developments and market fluctuations. The model combines the risk scores of individual properties at the portfolio level to calculate the anticipated losses and the distribution of losses across various climate scenarios. The scenarios above can be tailored to incorporate particular risk factors or policy assumptions.

**Visualizations of Model Results:** A collection of interactive visualizations has been devised to enhance the comprehension and dissemination of model outcomes. These visualizations enable users to investigate the spatial and temporal trends of climate risk within their mortgage portfolios. The visual representations depict the spatial arrangement of risk scores, scatter plots that juxtapose the risk profiles of various properties or loan categories, and line charts that monitor the progression of portfolio losses over time across diverse climate scenarios.

## Here are Some Example Visualizations of Model Results



## Limitations and Areas for Enhancement

**Critical Limitations of the Modeling Approach:** Although our climate risk assessment framework is a notable improvement compared to conventional methods, it is crucial to recognize its constraints. A significant obstacle lies in the accessibility and reliability of data about climate hazards and their ramifications on the real estate sector. The model relies on datasets derived from historical observations or simplified physical models, such as flood maps or wildfire risk scores. However, these datasets may not comprehensively capture the intricate dynamics of climate change. Furthermore, the datasets may lack adequate spatial and temporal resolution to facilitate property-level risk assessments, especially in regions with limited monitoring or modeling capabilities. One additional constraint pertains to the inherent uncertainty associated with long-term climate projections and their subsequent conversion into financial risks. Despite including various climate scenarios and uncertainty estimates in our model, there remains a considerable disparity in the anticipated consequences of climate change across diverse models and assumptions. Uncertainty can pose challenges in effectively communicating and implementing the model's outcomes, especially when dealing with extended timeframes or more extreme circumstances.

**Areas for Future Model Improvement:** Notwithstanding these constraints, numerous prospects exist for augmenting and expanding our climate risk assessment framework in forthcoming investigations. A potentially fruitful avenue involves incorporating additional detailed and varied data sources into the model, such as high-resolution satellite imagery, flood observations obtained through crowdsourcing, or energy efficiency data at the building level. Using these datasets can enhance the precision of risk assessments by improving the spatial and temporal resolution, thereby enabling the capture of more intricate dimensions of property vulnerability and resilience. An additional aspect that could be enhanced involves the investigation of more sophisticated machine learning methodologies, such as deep learning or transfer learning. These techniques can autonomously extract characteristics from unstructured data, such as images or text, and subsequently adjust to evolving risk profiles as time progresses. Implementing these techniques has the potential to mitigate the need for manual feature engineering and enhance the model's capacity to acquire knowledge from intricate, high-dimensional datasets. In conclusion, future research endeavors may integrate our climate risk assessment framework with various mortgage lending and servicing facets, including loan origination, property valuation, and loss mitigation. Lenders and investors can adopt a comprehensive and proactive approach by incorporating climate risk considerations throughout the mortgage lifecycle to enhance the management of financial risks and opportunities related to climate change.

## Conclusion

In summary, our study showcases the considerable capacity of machine learning models in evaluating and controlling climate-related risks within the mortgage sector. Our framework utilizes sophisticated algorithms and extensive datasets to offer a precise, detailed, and proactive evaluation of the physical and transition risks that individual properties and entire portfolios may encounter. This data can assist lenders, investors, and regulators make well-informed choices regarding underwriting, pricing, capital allocation, and risk management in light of an unpredictable future climate. Nevertheless, to fully harness the capabilities of machine learning models in the context of climate risk assessment, it is imperative to maintain continuous investment and foster collaboration within

the mortgage industry. This entails securing sustained funding for research and development, establishing partnerships among academic institutions, government agencies, and private sector firms, and enhancing standardization and transparency in climate risk transparency and reporting frameworks. To advance the field of climate risk analytics, it is imperative to emphasize the significance of industry collaboration and data sharing. This entails consolidating data and expertise to generate comprehensive and dependable datasets for training and validating models. Furthermore, it involves collaborating on developing standardized scenarios and risk metrics and providing support for creating open-source datasets and modeling frameworks. Through collaborative efforts aimed at enhancing the caliber and availability of climate risk data and analytics, the mortgage industry has the potential to contribute to the development of a housing finance system that is both resilient and sustainable in the context of the 21<sup>st</sup> century [10,11].

## References

1. Climate Change 2021 – The Physical Science Basis. Intergovernmental Panel on Climate Change (IPCC) <https://www.cambridge.org/core/books/climate-change-2021-the-physical-science-basis/415F29233B8BD19FB55F65E3DC67272B>.
2. Bakkensen L, Barrage L (2017) Flood Risk Belief Heterogeneity and Coastal Home Price Dynamics: Going Under Water? National Bureau of Economic Research [https://www.nber.org/system/files/working\\_papers/w23854/w23854.pdf](https://www.nber.org/system/files/working_papers/w23854/w23854.pdf).
3. Hino M, Burke M (2021) The effect of information about climate risk on property values. Proceedings of the National Academy of Sciences 118.
4. Goldsmith-Pinkham PS, Gustafson M, Lewis R, Schwert M (2019) Sea Level Rise and Municipal Bond Yields. SSRN Electronic Journal. Elsevier BV [https://papers.ssrn.com/sol3/papers.cfm?abstract\\_id=3478364](https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3478364).
5. Baldauf M, Garlappi L, Yannelis C (2020) Does Climate Change Affect Real Estate Prices? Only If You Believe In It. The Review of Financial Studies 33: 1256-1295.
6. Keys B, Mulder P (2020) Neglected No More: Housing Markets, Mortgage Lending, and Sea Level Rise. National Bureau of Economic Research [https://www.nber.org/system/files/working\\_papers/w27930/w27930.pdf](https://www.nber.org/system/files/working_papers/w27930/w27930.pdf).
7. Bernstein A, Gustafson MT, Lewis R (2019) Disaster on the horizon: The price effect of sea level rise. Journal of Financial Economics 134: 253-272.
8. Hallegatte S, Green C, Nicholls RJ, Corfee-Morlot J (2013) Future flood losses in major coastal cities. Nature Climate Change 3: 802-806.
9. Friedman JH (2001) Greedy function approximation: A gradient boosting machine. The Annals of Statistics 29.
10. Chiu CC, Wu CM, Chien TN, Kao LJ, Li C (2023) Integrating Structured and Unstructured EHR Data for Predicting Mortality by Machine Learning and Latent Dirichlet Allocation Method. International Journal of Environmental Research and Public Health 20: 4340.
11. (2019) United States: Argo and The Climate Service Join the Wells Fargo Startup Accelerator. MENA Report <https://www.businesswire.com/news/home/20190910005379/en/Argo-and-The-Climate-Service-Join-the-Wells-Fargo-Startup-Accelerator>.

**Copyright:** ©2023 Rohit Nimmala. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.