Explainable artificial intelligence modeling for bankruptcy prediction under climate change risks

Nicolae Stef

CEREN EA 7477, Burgundy School of Business, Université Bourgogne Franche-Comté, Department of Accounting, Finance & Law, Dijon, France. E-mail: nicolae.stef@bsbeducation.com

Sami Ben Jabeur

Institute of Sustainable Business and Organizations, Sciences and Humanities Confluence Research Center -UCLY, ESDES, France. E-mail: sbenjabeur@univ-catholyon.fr

Pedro Carmona

University of Valencia, Department of Accounting, Av. Tarongers s/n. 46022 Valencia.

Spain. E-mail: pedro.carmona@uv.es

Robert F. Scherer

Trinity UniversityChapman 320, 210.999-7281, One Trinity Place, San Antonio, TX 78212.

E-mail: rscherer@trinity.edu

Abstract: This study examines the performance of climate change risk (CCR) variables in predicting corporate failure modeling. The experimental findings, which are supported by real-world datasets from France, demonstrate that CCR variables combined with accounting- based data increase prediction modeling accuracy. When we use the XGBoost model-based climate change risk variables (XGBoost-CCR), the impact becomes more prominent. Moreover, explainable artificial intelligence framework (XAI) and feature importance plots are used to enhance model interpretability. Findings show that our proposed model correctly predicted 94.76% of the liquidation cases, whereas the accuracy decreased to 90.35% when predicting bankruptcy only with accounting ratios. Additionally, the results indicate that incorporating CCR variables enhances the overall accuracy of prediction models by an average of 4.41%, and increases the area under the ROC curve by 6.2%.

Keywords: Climate change; Bankruptcy; Firm; Temperature; Weather

JEL codes: Q54; G33

1. Introduction

Temperature is raising across the world affecting not only the local biodiversity and population, but also the business environment. According to the 2019 report of the *International Labor Organization*, an estimated increase of 1.5ºC in the global temperature by the end of 2100 could lead to a labor productivity loss in 2030 representing about 80 million full-time jobs. By assuming that the activities of agriculture and construction would be carried out in the sun, that loss could reach 136 million full-time jobs. Moreover, the heat stress would also generate economic losses of $2.400 billion in 2030 that are almost 8 times higher than the similar losses incurred in 1995 (ILO, 2019). However, supporting the climate change resilience of individuals, firms and systems can be a solution to reduce the severity of those losses. According to Neufeldt et al. (2021), the costs of adaptation to actual and expected climate effects will almost double for the developing countries during 2030 and 2050. Even in the case of a strong financial capacity to implement adaptation measures, some degree of climate change cannot be avoided through mitigation policies (European Investment Bank, 2012).

Hence, extreme changes in the local temperature can lead to the destruction of full-time jobs and productivity losses. Such consequences were empirically confirmed by the environmental literature. For instance, Cai et al. (2018) reported an inverted U-shaped relationship between the manufacturing worker productivity and the daily maximum temperature in China. The cold and the heat stress can negatively affect the efficiency of manual-labor intensive tasks generating significant economic losses for the manufacturing firms. Liu et al. (2023) also reached a similar result for Chinese listed firms over the period 2011-2019 for which the labor productivity should be expected to decrease by 1.64% following an increase of one standard deviation in the high-temperature weather. More interestingly, Zhao et al. (2021) estimated that the global economic losses caused by the heat-related employees' productivity losses should range between 0.31% and 2.6% of the 2100 global gross domestic product (GDP) with South Asia, Sub-Saharan Africa and Central America incurring the largest economic losses. In this frame of temperature-firm relationship, the business failure can be accelerated by the productive inefficiency measured by the distance from the firm's optimal operational approach (Becchetti and Sierra, 2003). As the losses in labor productivity can be associated with the operational inefficiencies, an important question for the corporate risk literature can be raised that is how the weather conditions should relate to the bankruptcy risk of firms.

This paper addresses that question by examining the bankruptcy prediction power of climate change risk assessed at local and national level. In this regard, our analysis is performed on an original sample composed of non-financial and non-listed French firms over the period 2010-2022 among which 13975 healthy firms and 2360 financially distressed firms for which a liquidation procedure was triggered in court. The existence of a nexus between weather conditions and corporate bankruptcy can be explained mainly through the physical damage and operational disruptions that the climate change events can have on a firm’s assets (Campiglio et al., 2022; Fuss, 2016). In this framework, the environmental determinism (Bourgeois III, 1984; Whittington, 1988) argues that the environment limits the number of alternatives or strategies a firm can adopt to overcome the financial distress or to increase its performance. Consequently, the risk of failure should be higher for firms with a limited capacity to overcome or to avoid the physical damage that weather events can have on its assets.

In terms of novelty and contributions, our research expands three major strands of academic literature. First, the ecological literature investigated mainly the consequences of the environmental degradation on a firm’s financial performance or productivity (Cevik and Miryugin, 2023) Huang et al., 2018; Pankratz et al., 2023; Zhang et al., 2018). This study offers a new research perspective by focusing for the first time on the relationship between local and global weather conditions and the failure risk of firms. Second, the corporate bankruptcy literature acknowledged that non-financial variables can be used to better explained the outcome of firms confronted with financial issues (Buehler et al., 2012; Cooper and Uzun, 2019; ;Stef, 2021; Stef and Bissieux, 2022; Stef and Zenou, 2021; Zorn et al., 2017). In this research, we propose original and new determinants of corporate bankruptcy constructed using local climatological data relying on the daily average temperature (1), the humidity ratio (2), the daily precipitation (3) and the average wind speed (4). Additionally, we also explore the prediction relevance of the climate change risk measured at national level through the *Climate Altering Land Cover Index*, the *Climate Risk Index*, the *Vulnerability Index* and the annual change in the surface temperature. Third, our focus on bankruptcy prediction and climate change concerns, combined with the incorporation of explainable artificial intelligence (XAI) techniques, enables us to provide novel insights into the role of climate change on firm liquidation while maintaining model transparency and interpretability.

This paper is organized as followed. Section 2 discusses the potential impact of weather conditions on a firm's bankruptcy risk. We briefly present the artificial intelligence methodology in section 3 while the data set and the variables are described in section 4. The empirical results are shown in section 5. We conclude and highlight the major caveats of this study in the last section.

2. Climate change risk and corporate bankruptcy risk

The causes of corporate bankruptcy are various and multiple as it was suggested by Blazy et al. (2013) and Blazy and Stef (2020) that classified such causes in different categories covering aspects dealing with strategic behavior, production operations, financial policy, management performance, accidents, outlets and macroeconomic shocks. Ooghe and De Prijcker (2008) proposed a different classification of bankruptcy causes that included the firm's general environment based on external factors, the immediate environment that is determined by the interactions between firm and its stakeholders, the characteristics of management, the corporate policy and the firm's characteristics. Among those typologies, one may ask if the global warming as captured by changes in temperature should be considered as a default cause and how it should be associated with the likelihood of a firm's bankruptcy. The answers to those questions may rely on the impact that severe temperature changes can have on the value and the performance of a firm's assets.

In this regard, Fuss (2016) argued that climate change events can harm the value of financial assets through destruction (1), accelerated depreciation (2) and/or operational disruptions (3). The financial pressure exercised by the climate change on firm's financial strength has also been addressed by Campiglio et al. (2022) that identified four main transmission channels. First, capital assets can be destroyed or becoming less profitable following a climate-related event. Second, high mean temperature can induce a physical risk by forcing firms to upgrade their logistic process. Third, the demand patterns and the supply chains can be disrupted by extreme weather events leading to revenue losses and high operating expenditures. Fourth, the cost of debt can increase for firms that are exposed to physical risks related to climate events.

Similarly, Huang et al. (2018) suggested that the fixed assets can be physically damaged by extreme weather. Consequently, the firm's earnings and thus the economic performance can decrease. Using a large sample of firms operating in 55 countries over the period 1993-2012, their panel regressions with year and industry fixed effects confirmed that the firm's performance is negatively affected by the climate risk. Additionally, Zhang et al. (2018) pointed out that high temperatures not only impact the labor productivity but can also harm the machine performance and capital productivity. By examining half million manufacturing plants, their study revealed that extremely high temperatures can significantly decrease the output and the total factor productivity of Chinese firms. Additionally, (Pankratz et al., 2023) performed an empirical analysis on a larger sample composed of more than 17000 firms from 93 countries confirming that the financial performance tends to decrease because of an increased heat exposure. They suggested that investors may not fully anticipate the economic impact of the physical climate risk.

In this framework, the physical effects of climate change concern the operations, the distribution activities and the firm's access to resources (Linnenluecke et al., 2013). According to Tzouvanas et al. (2019), hot temperature shocks can produce systemic risks by aggravating the losses incurred by financially distressed firms. Their empirical analysis based on 600 listed European firms over the period 1990-2017 revealed that the systemic risk can go up by 0.24 basis points following a 1ºC raise in the temperature. Tzouvanas et al. (2019) explained those findings through the direct exposure of assets to environmental shocks and the degradation of the interconnection between firms with climate sensitive assets and other businesses (Battiston et al., 2017). Hence, unexpected temperature changes can generate losses to firms operating with climate sensitive assets that can also generate negative externalities for the local economy.

From a different theoretical perspective, the relationship between the temperature and a firm's default risk can be explained through the environmental determinism. According to this Bourgeois (1984), determinism view implies that the environment imposes some constraints that dramatically reduce the number of organizational actions that can produce the first best outcome for the firm. In other words, the features of the environment define a limited number of alternative decisions (strategies) a firm can adopt to prosper and survive (Whittington, 1988). However, the environmental determinism perspective sustains that organizations are subject to inertial structures that reduce the capacity of firms to benefit from limited environmental resources (Gopalakrishnan and Dugal, 1998). Therefore, firms that can promptly react to the environmental changes have more chances to survive in the long run. Compared to other environmental forces such as legal policies, market competition, consumer behavior, political stability or access to credit, the variation in temperature represents an irreversible phenomenon with a long-term persistence even after the stop of carbon emissions at a global scale (Solomon et al., 2009). In the light of the environmental determinism theory, a firm that is not able to adapt to the unanticipated changes in temperature may be obliged to incur additional costs and losses of operational performances that can affect its financial health.

However, the changes in climatic conditions may not always lead to harmful financial impacts. Some studies such as Smit et al. (2000) and Linnenluecke and Griffiths (2012) suggested that organizations are subject to a coping range that is composed of circumstances related to climate events that will not have adverse consequences on the organizational activities. Resilient firms with a wide coping range can tolerate the effects of climate change without incurring major damages or additional costs (Linnenluecke & Griffiths, 2012). In this regard, Addoum et al. (2020) examined the consequences of the average daily temperature and the temperature extremes on U.S. economic establishments over the period 1990-2015. Their estimates reported a non-significant relationship between the abnormal temperature exposures and firm-level profitability measured through the operating income, net income and earnings announcement returns. As highlighted by Linnenluecke and Griffiths (2010), organizations more exposed to extreme weather must face new challenges that will require the development of new capabilities and slack resources. Hence, firms with a high degree of resilience to environmental impacts and extreme changes in local weather should better preserve their financial health compared with firms that have a narrow coping range.

To strengthen the resilience to global warming, Winston (2014) suggested that firms must reconsider the valuation of unpriced costs and benefits, but also embrace radical innovations. Nevertheless, firms may be reluctant to replace their assets because of the uncertainty on how the climate will change and how the future benefits will be affected by those environmental changes (Mendelsohn, 2012). Consequently, the firm's adaptation to global warming may be slow. If the costs of adaptation are significant, a firm less resilient to climate change may be subject to a weaker financial health than a firm with less climate sensitive assets. Overall, a positive impact of the extreme changes in the local weather conditions on the likelihood of a firm's financial failure can be explained by the physical degradation of assets (1) and the weak capacity of firms to adapt to climate change (2).

3. Artificial intelligence modeling

3.1. Logistic model

As opposed to discriminant analysis, logistic regression was proposed by Ohlson (1980). The logistic regression model estimates the coefficients of these independent variables, which indicate the strength and direction of the relationship between each independent variable and the probability of bankruptcy. The score function of the probability of failure is calculated as follows:

$$Z = \frac{1}{1+e^{-(\sum_{i=1}^{N}\gamma_i x_i+\gamma_0)}} \quad (1)$$

where x_i is the explanatory factors, and γ_i are the coefficients of the estimated function. The Ohlson model has been shown to have better predictive power than the Altman Z-score model. Nevertheless, logit model may be less accurate in predicting bankruptcy during periods of economic instability or financial crisis, when market conditions can change rapidly and unpredictably (Du Jardin, 2015).

3.2. Random forest

Random forest (Breiman, 2001) is a popular machine learning algorithm that is widely used for both classification and regression problems. Random forest (RF) can identify the key drivers of bankruptcy risk in a dataset. For a given firm, the bankruptcy likelihood stated as a z score, can be expressed as follows:

$$Z = argmax\frac{1}{T}\sum_{t=1}^{T} q_t(y/x) \quad (2)$$

where $q_t(y|x)$ is the probability distribution of each tree (t), and x is a set of test observations. To increase performance accuracy, RF combines a decision tree with an ensemble architecture. Each tree in the procedure casts a unit vote, allocating each variable to the output class with the highest likelihood of success (Jabeur et al., 2021).

3.3. Gradient boosting machine

The Gradient Boosting Machine (GBM) was introduced by Friedman (2001). The core concept of GBM is to repeatedly fit ever more simplistic prediction models to the residual errors of the prior models. Each weak model is often a decision tree with a limited number of terminal nodes, with the purpose of minimizing a loss function that evaluates the difference between predicted and true values. The z score function is calculated as follows:

$$Z = \sum_{j=1}^{M} \gamma_j k(x; c_j) \qquad (3)$$

where $k(x; c_j)$ is the base learner, x is the explanatory variables, γ_j is the expansion coefficients, and c_j is the parameters of the model. Several variants of GBM have also been cretaed over the years, including XGBoost, LightGBM, and CatBoost, which improved performance and scalability over the original algorithm.

3.4. Extremely randomized trees

Extremely Randomized Trees (ERT) is an ensemble learning method that combines multiple decision trees to improve the accuracy (Geurts et al., 2006). It mainly consists of heavily randomizing both attribute and cut-point selection when splitting a tree node. It creates completely randomized trees with topologies that are independent of the learning sample's output values. The z score output is calculated using the average of the probabilities over all trees is defined as follows:

$$Z = argmax \frac{1}{T} \sum_{t=1}^{T} q_t(y/v) \qquad (4)$$

where T denotes the total number of trees, $q_t(y|x)$ is the conditional probability of class y given a vector v of a set sample. ERT employes a higher degree of randomization by splitting nodes based on a randomly selected subset of features and threshold values. This can further reduce the variance of the model and improve performance. To avoid overfitting, the method also includes a number of regularization procedures, such as restricting the maximum depth of the trees.

3.5. Extreme Gradient Boosting

Extreme Gradient Boosting (XGBoost) models have recently brought a lot of interest to in bankruptcy prediction and credit scoring field (Jones, 2017; Fontecha et al., 2021). XGBoost is more suitable for high dimensional, nonlinear analysis, which arguably better reflects the real-world context of corporate bankruptcy. The gradient boosting approach, in contrast to more traditional methods such as logit, is not sensitive to the presence of outliers and multicollinearity. The z score output is expressed as follow:

$$Z = \sum_{j=1}^{K} f_j\,(x_i) \qquad (5)$$

where x_i are the explanatory factors, and fj(xi) is the function that determines the output of each tree. As compared to other machine learning algorithms, XGBoost excels in a number of ways. It's lightning fast, and it can process massive datasets with a plethora of features. Additionally, XGBoost provides several hyperparameters that can be tuned to optimize the model's performance.

3.6. Deep neural networks

Deep neural networks (DNNs) are a type of artificial neural network (ANN) that are characterized by having multiple hidden layers (Kraus et al., 2020). In such networks, there are more free parameters, and the network is more able to represent highly non-linear functions. The input layer receives input data, and the output layer produces the final output of the network. Kraus et al. (2020), compute a deep neural network with k layer functions as follows:

$$\emptyset(\mathrm{x})_{DNN} = \underbrace{\emptyset_{1NN}(\emptyset_{1NN}(\ldots \emptyset_{1NN}(X)))}_{k} \qquad (6)$$

where $\emptyset_{1NN}$ is a single-layer perceptron computed via a a linear combination embedded activation function. The input layer is referred to the first layer, the output layer is the final layer, and the remaining levels are called hidden layers.

In classification tasks, the network learns to map input data to output labels via an iterative process of forward and backpropagation in which the network's weights are modified to minimize a loss function that assesses the difference between the predicted and true labels. By using a softmax activation in the output layer, classification tasks typically provide a discrete probability distribution across all possible classes (Russell, 2010 ; Gao and Wang, 2022).

4. Data and variables

Using the *Diane* database of the *Bureau Van Dijk*, we have constructed a sample of non-listed French firms composed of failed and healthy firms. We had to address the issue of unbalanced observations, where the majority class of healthy firms (n=354,152) significantly outweighs the minority class of failed firms (n=2,360), resulting in a low proportion of the latter class (0.00666). Additionally, the dataset contains many missing values, requiring data pre-processing steps. To tackle this problem, we reduced the unbalanced dataset by eliminating missing values of healthy firms and selecting only 10% of the remaining healthy firm observations. Subsequently, we imputed missing values of failed firms. Following these data pre-processing steps, the data distribution was as follows: 2,360 firms were classified as failed, and 13,975 as healthy (yielding a proportion of the minority class of 0.1688). The sample remains imbalanced, although the level of imbalance has reduced significantly compared to the initial dataset. We have included only firms with available financial data for the last 3 years prior to the liquidation triggering date. Our sample covering the period 2010-2022 is based on two sets of variables. The first set is based on annual financial variables that are commonly used in the bankruptcy prediction literature dealing with the firm's size (*TA*), indebtedness degree (*Leverage*), liquidity (*CR*, *Cash* and *C/TA*), operational performance (*EBITDA*, *EBITDA/TA*, *WC* and *WC/TA*), equity level (*Equity* and *GR*), sales (*Sales*, *I/S* and *TS/TA*). Additionally, we have also added *Age* as a proxy of firm's experience.

The second set relates to variables capturing the climate change effects at global and local scales. First, we address the global environmental degradation using the Climate Altering Land Cover Index (*CALCI*), the Climate Risk Index (*CRI*), Vulnerability Index (*VUL*) and Annual Surface Temperature Change (*ASTC*). Second, we have relied on the database of *Meteo France* to build four variables measuring the annual weather conditions at French county level, namely temperature (*TEMP*), humidity (*HUM*), daily precipitation (*DP*) and wind speed (*WS*). The table 1 provides detailed definitions of those variables.

Table 2 shows a summary of the most important descriptive statistics of failed and healthy firms. Some of the key insights from the table include that failed firms have a significantly higher leverage than healthy firms, with a mean of 2799.93. The equity of failed firms also has a smaller mean value of 242.93 compared to the healthy firms with a mean value of 8672.16. The mean value of sales for failed firms is 3654.70, which is much lower than the healthy firms with a mean value of 25426.48. Healthy firms have significantly higher total assets (*TA*) than failed firms, with a mean value of 24577.24. In terms of the environmental variables, failed firms tend to have a slightly higher Climate Risk Index (*CRI*) with a mean value of 45.61

compared to healthy firms with a mean value of 45.72. There are no major differences between the healthy and failed firms for the remaining climate change variables.

As a precautionary measure, we ensure there was not a severe correlational problem among the different predictors and the response variable. Given the size of the dots and the absence of an intense color, it indicates the nonexistence of this potential problem (see correlogram plot in figure 1). A correlogram is a graph that illustrates the correlation matrix, representing each correlation coefficient with a dot color and size according to its value. Nevertheless, it suggests that *Liquidation* (failed firms) has a weak negative correlation with *TA*, *CR*, *Cash*, and *EBITDA*, meaning that an increase in these variables is associated with a decrease in the likelihood of liquidation. The correlogram also shows very weak correlations between *Liquidation* and environmental variables.

(Insert Table 1 here)

(Insert Table 2 here)

(Insert Figure 1)

5. Empirical results

5.1. Prediction performance using climate changes risk variables

This research analyzes whether climate change factors give substantial information that may be used to forecast the failure of a company, as well as the impact that CCR variables play in determining the risk that a company would collapse. So, we evaluate our proposed model, which incorporates CCR factors, against the outcomes of the conventional business failure model. We created two models: one that depends simply on financial ratios (Model A) and another that uses CCR variables in conjunction with financial data (Model B). These models were applied to the six different classifiers overviewed in section 4, namely the LR, GBM, DNN, RF, ETR, and XGBoost. We used cross-validated trials to carry out parameter optimization for the six techniques. Additionally, all models were fitted in *R* (R Core Team, 2022) version 4.1.3, and for model interpretability, we used DALEX package version 2.4.0 (Biecek et al., 2022).

To evaluate the effectiveness of the models, we used three evaluation metrics commonly applied in the literature (Du Jardin, 2015): Accuracy (ACC), area under the ROC curve (AUC), and F-score (F). Table 3 presents the correct classification rates. The results indicate that CCR variables can predict corporate failure and provide valuable information about a firm's financial condition.

(Insert Table 3 here)

Table 3 compares the accuracy of several models in forecasting corporate failure, using Model A (financial ratios) and Model B (financial ratios and CCR variables). The table shows the accuracy percentages for each model, as well as differences in percentage points between Model B and Model A. In each of the models, except for DNN, Model B demonstrated superior performance to Model A. XGBoost achieved the highest accuracy percentages for both Model A (90.35%) and Model B (94.76%), representing an increase in accuracy of 4.41 percentage points from Model A to Model B. GBM exhibited a smaller increase in accuracy of 0.63

percentage points, while RF and ETR experienced only marginal improvements in accuracy. LR demonstrated a very slight increase in accuracy of 0.58 percentage points. These results highlight the superior predictive capability of Model B, particularly when employing the XGBoost algorithm, and demonstrate the incremental gains in accuracy that can be obtained using additional variables. The model has only misclassified 5.24% of observations. This result corroborates the findings from previous bankruptcy studies that the gradient boosting model reports higher accurate performance (e.g, Jones, 2017).

(Insert Table 4 here)

To get an insight about the added value of CCR variables, we calculated the difference between the accurate classification rates achieved with XGboost-based CCR variables (XGboost-CCR) and the rates achieved with all other models. We first evaluated the differences by modeling method, then by type of model (A and B), in addition to these computations, we supplemented our investigation with a two-sided test for differences between proportions to determine which of the changes were statistically significant. Results are shown in Table 4. In this table, Panel A indicates the magnitude of the differences rates while Panel B the *p*-values of a two-sided test for differences. These panels demonstrate that XGBoost-CCR are considerably higher than those estimated with any other model (*p*-value 0.0001), with differences ranging from 3.94 percentage points all the way up to 14.78 percentage points.

The findings suggest that variables related to CCR can predict corporate failure, and the decisions made by firms regarding CCR offer crucial insights into their financial health. It is not entirely surprising to discover this connection, as existing research has already established a link between firm performance and climate risk (Ozkan et al., 2022; Pankratz et al., 2023). This could clarify why such specific information is useful in differentiating between failed and non-failed firms.

4.2. Additional analysis

We further enhanced the prior findings by incorporating an additional performance metric, the area under the receiver operating characteristic curve (AUC). This measurement offers a comprehensive evaluation of model accuracy that is entirely unaffected by the distribution of observations across different classes (du Jardin, 2021) . The area under the ROC curve can be used to judge a model overall performance without assuming a relative cost structure (Mai et al., 2019).

(Insert Table 5 here)

Table 5 presents the AUC values for two different models (Model A and Model B) across various types of algorithms, as well as the percentage point difference between Model B and Model A. XGBoost-based CCR variables outperforms Model A with an AUC of 0.969 compared to 0.907, representing an increase in accuracy of 4.41 percentage points from Model A to Model B. For GBM, Model B also surpasses Model A with an AUC of 0.923 versus 0.908. The difference between the two models is a 1.5 percentage point increase in favor of Model B. Conversely, the DNN algorithm shows a higher AUC for Model A (0.816) than Model B (0.768). The RF algorithm exhibits a marginal difference in performance between Model A (0.899) and Model B (0.903). Similar to RF, the LR algorithm shows a slight increase in AUC for Model B (0.821) compared to Model A (0.818). Finally, the ERT algorithm presents a higher AUC for Model B (0.911) compared to Model A (0.904). Overall, the results of Model B perform better than Model A that is based on the accounting ratios only.

(Insert Table 6 here)

To deepen the analysis, we analyzed the differences between the outcomes determined with XGBoost-based CCR variables and those predicted with the other algorithms. These are shown in Panel A of Table 6. Panel A of this table is supplemented by Panel B, which contains the *p*-values of a test for differences between proportions. This test was performed to evaluate the statistical significance of the differences that were described in Panel A. These panels demonstrate that XGBoost-CCR consistently outperforms single all model results; the differences vary between 4.6 and 15.3 percentage points, and each difference is statistically significant at a very low level (less than 0.1%).

4.3. Model explainability and evaluation of explainability

Explainable Artificial Intelligence (XAI) is a subfield of artificial intelligence (AI) that focuses on creating machine learning models that can provide human-understandable explanations for their predictions and decisions (Chakraborty et al., 2021). The primary goal of XAI is to enhance the transparency, trustworthiness, and interpretability of AI systems. The main objective of XAI is to make artificial intelligence systems more open, reliable, and easy to understand.

In this research, we employed the SHAP (SHapley Additive exPlanations) model by (Lundberg et al., 2020) to analyze the predictions of the XGBoost models. The SHAP "global interpretability" analysis shown in Figure 2 reveals the relative order of importance of accounting and CCR variables (Model B), but it also provides the well-informed conscious predictions made by the models. The variables are ranked in the order of their normalized SHAP value. It can be seen from figure 2 that 20 of the 23 predictor variables have nonzero SHAP value. This means that all 20 variables contributed to out-of-sample predictive power. As can be seen by figure 2, a several range of bankruptcy predictors dominate the analysis, including the accounting and CCR dimensions of corporate failure. The predictor factors that are featured most prominently in the analysis include the following: (1) a number of accounting variables and ratios including equity, sales, total assets, EBITDA to total assets, leverage, EBITDA, age, interest to sales, working capital, total sales to total assets and cash. This findings are in line with previous studies on corporate failure prediction (e.g, (Kumar and Ravi, 2007) ; Jones, 2017); (2) a range of climate change risks variables including climate altering land cover, vulnerability, wind speed, climate risk index, humidity and temperature.

(Please Insert Figure 2 here)

Figure 2 indicates that the strongest predictor overall is equity ratio. The second strongest variable is the *Climate Altering Land Cover Index*. Higher and lower feature values are represented by red and blue dots, respectively. Higher values of this feature result in higher SHAP values, which correspond to a higher probability that a failure has occurred. Severe changes in the share of climate altering French land cover include shifts in temperature and precipitation, as well as a rise in the frequency of severe weather events. Such changes might impact the stability of the systems or the likelihood of failures in infrastructure and resources. Those results provide some support to the argument of Fuss (2016) and Campiglio et al. (2022) arguing that climate change events can damage the value of financial assets leading to operational disruptions.

The next strongest variable is the *Vulnerability Index* that is a metric determining how susceptible a nation is to the unfavorable consequences of climate change. This index determines a country's total vulnerability by considering food, water, health, ecosystem service, human habitat, and infrastructure. Higher values of vulnerability correspond to a higher probability of corporate failure. Firms operating in an environment more sensitive to the global warming effects may face interruptions to their supply chains, manufacturing facilities, or distribution networks because of severe weather events such as floods, storms, or droughts. These operational interruptions might result in higher expenses, lower production, or even temporary or permanent firm closures. Overall, the findings on *VUL* are in line with Huang et al. (2018) that pointed out that the fixed assets can be physically damaged by extreme weather.

Other high impacting CCR variables include climate risk index (*CRI*). Higher values of *CRI* correspond to a higher chance of failure occurrence. The *CRI* serves as an indicator of the vulnerability of a region or system to climate change risks, such as extreme weather events, rising sea levels, and shifts in precipitation patterns. Firms located in areas with high *CRI* values are exposed to a greater risk of disruptions to their operations, supply chains and infrastructure. This result is not completely surprising; available literature has documented that climate risk reduces firm performance (e.g, Huang et al., 2018; Ozkan et al., 2022). Other variables also appear to with nonzero importance, such us *HUM* and *TEM*. For example, lower values of *TEM* correspond to a higher chance of corporate failure occurrence, suggesting that cold can negatively affect the efficiency of the productivity and generating significant economic losses for the manufacturing firms (Cai et al., 2018).

(Insert Figure 3 here)

In Fig. 3, we assess the outcome of our analysis for a particular firm predicted as failed. Using the SHAP force plot, we can easily illustrate why each company in the dataset is likely to go bankrupt. We randomly select a failed company to determine which financial and CCR variables (Model B) play an important role in the prediction of the XGBoost model. As shown in Fig. 3, *VUL* (*VUL*=0.313), *EBIDTA* divided by *TA* (*EBIDTA_TA*=-0.216), and *CRI* (*CRI*=27.83) are the three variables that have the greatest impact on the firm's outcome. This confirms once more that the CCR variables can have a bankruptcy prediction power.

4.4. Further evaluation

In this section, we will enhance our initial analysis along the following lines. Table 7 presents the F-scores of Model A and B. The F-score is a performance indicator that combines accuracy and recall, making it valuable for assessing classification models, especially those with imbalanced datasets (Goutte and Gaussier, 2005). The results indicate that Model B performs significantly better with XGBoost, slightly better with GBM and XRT, almost the same with logistic models, and worse with deep learning and Random Forest models. Considering the F-score, XGBoost based CCR seems to be the most effective option among the all models.

(Insert Table 7 here)

(Insert Table 8 here)

To conduct a more thorough investigation, we compared results predicted by various algorithms to those derived using XGBoost-based CCR variables. Table 8 demonstrates that the XGBoost-CCR model, which incorporates climate change risk variables, is more effective at classification tasks than the other models tested, whether based solely on financial ratios (Model A) or a combination of financial ratios and climate change risk variables (Model B).

(Insert Figure 4 here)

To deepen our analysis, we used a variable importance measure by Friedman, (2001). In Fig. 4, we present variable importance measures for the 10 most important variables. We can see that vulnerability (*VUL*), equity, leverage and climate altering land cover (*CALCI*) have the most impact on corporate failure than other features. The explanation of variables importance is broadly in accordance with previous findings confirming the effectiveness of using climate change variables in predicting corporate failure and the impact of climate change in determining the risk of a company collapsing.

5. Conclusion and implications

In this paper, we have examined how climate change risk affects corporate failure. Our analysis confirms that CCR is associated with firm survival and improves model prediction performance. Incorporating relevant nonfinancial variables into company failure prediction models tends to generate a more accurate forecast. In this regard, our study created two models, one using financial ratios and another using climate change variables along with financial data, and applied them to six different classifiers. The results indicated that Model B, which included climate change variables, demonstrated superior performance to Model A. We also found that variables related to climate change can predict corporate failure, and decisions made by firms regarding climate change offer crucial insights into their financial health. It appears that the best model was fitted using the XGBoost algorithm, and it incorporated climate change variables in addition to financial ratios to improve its predictive power. This suggests that the model was able to leverage the relationships between financial performance and climate change factors to better predict firm's outcome that is either survival or liquidation.

Our focus on bankruptcy prediction and climate change risk allows us to provide novel insights into the role of climate change on firm's liquidation. These outcomes contribute to the ongoing discussion in the academic and policy communities about the economic effects of climate change risks that could have substantial implications for the stability of the financial system. First, although previous research has investigated the impact of climate hazards on firm performance (e.g, Ozkan et al., 2022 ; Pankratz et al., 2023), to the best of our knowledge, this is the first attempt that directly considers the relationship between climate risk and corporate failure. Moreover, our study delivers novel research perspectives about the capacity of climate change risks–based variables to improve bankruptcy prediction models. These results provide evidence of improvements to predicting models that integrate information on climate change as well as financial factors. This finding is interesting in light of the channels through which extreme weather conditions relate to insolvency risk. Second, our findings add to the continuing debate in the academic and policy communities concerning the economic consequences of climate change concerns and the advantages of utilizing XAI approaches in modeling high-dimensional and nonlinear manufacturing data (Senoner et al., 2022). As pointed by Bauer et al. (2023) from a cognitive perspective, explanations generally serve to enhance people's understanding, improve reasoning, and facilitate learning. As a result, explanations can facilitate the learning of machine knowledge, enabling users to access new insights autonomously derived from Big Data by AI systems and previously overlooked by domain experts (Teodorescu et al., 2021).

From a practical perspective, the findings of our analysis are useful to financial institutions, banks, and investors, for evaluating the expected costs of climate change risks. First, the use of XAI based CCR variables in bankruptcy prediction models can help identify specific climate change risk factors that contribute to a higher likelihood of insolvency. This information can be valuable for firms, as it enables them to focus on targeted adaptation and mitigation strategies to address the most critical climate-related risks. In turn, this can help improve firms' resilience to climate change and reduce the probability of liquidation. Second, banks and lending institutions should address climate change issues in order to price more effectively the risks they are taking on. Understanding the specific climate-related causes driving insolvency risk allows regulators to target their actions to address the most pressing challenges, ultimately improving the financial system's overall stability. Overall, by incorporating XAI based CCR approaches into bankruptcy prediction models, we may get a better understanding of the intricate linkages between climate change threats, financial considerations, and corporate insolvency. By making these models more interpretable and transparent, stakeholders such as regulators, financial institutions, and investors would be able to better understand and trust the models' forecasts, resulting in more informed decision-making and risk management.

Our study has some limitations, but it's no different than any other research out there. We think these constraints provide opportunities for further study, especially considering the growing literature of climate change concerns. First, our study primarily focuses on the impact of accounting ratios and climate change risk on corporate failure, leaving other non-financial factors and their potential effects on bankruptcy prediction unexplored. Future research could incorporate a wider range of non-financial variables to better understand their influence not only on corporate survival, but also on the likelihood of firm’s restructuring. Second, our findings may be limited to the French context. Future studies could investigate different countries, regions and industries. Third, Capasso et al. (2020) found that rising temperatures could disrupt financial markets and the banking system, as evidenced by firms' creditworthiness being affected by climate risks. In a similar vein, Chava (2014) demonstrated that businesses that have various environmental problems must pay more for bank loans. Therefore, the bankruptcy relevance of CCR should also be addressed to determine the risk failure or the financial degradation of banks.

References

Addoum, J.M., Ng, D.T., Ortiz-Bobea, A., 2020. Temperature Shocks and Establishment Sales. The Review of Financial Studies 33, 1331–1366. https://doi.org/10.1093/rfs/hhz126

Battiston, S., Mandel, A., Monasterolo, I., Schütze, F., Visentin, G., 2017. A climate stress-test of the financial system. Nature Climate Change 7, 283–288. https://doi.org/10.1038/nclimate3255

Bauer, K., von Zahn, M., Hinz, O., 2023. Expl (AI) ned: The Impact of Explainable Artificial Intelligence on Users' Information Processing. Information Systems Research.

Becchetti, L., Sierra, J., 2003. Bankruptcy risk and productive efficiency in manufacturing firms. Journal of banking & finance 27, 2099–2120.

Blazy, R., Chopard, B., Nigam, N., 2013. Building legal indexes to explain recovery rates: An analysis of the French and English bankruptcy codes. Journal of Banking & Finance 37, 1936–1959.

Blazy, R., Stef, N., 2020. Bankruptcy procedures in the post-transition economies. European Journal of Law and Economics 50, 7–64. https://doi.org/10.1007/s10657-019-09634-5

Bourgeois III, L.J., 1984. Strategic management and determinism. Academy of Management review 9, 586–596.

Bourgeois, L.J., 1984. Strategic Management and Determinism. AMR 9, 586–596. https://doi.org/10.5465/amr.1984.4277347

Breiman, L., 2001. Random forests. Machine learning 45, 5–32.

Buehler, S., Kaiser, C., Jaeger, F., 2012. The geographic determinants of bankruptcy: evidence from Switzerland. Small Business Economics 39, 231–251.

Cai, X., Lu, Y., Wang, J., 2018. The impact of temperature on manufacturing worker productivity: evidence from personnel data. Journal of Comparative Economics 46, 889–905.

Campiglio, E., Daumas, L., Monnin, P., von Jagow, A., 2022. Climate-related risks in financial assets. Journal of Economic Surveys.

Capasso, G., Gianfrate, G., Spinelli, M., 2020. Climate change and credit risk. Journal of Cleaner Production 266, 121634.

Cevik, S., Miryugin, F., 2023. Rogue Waves: Climate change and firm performance. Comparative Economic Studies 65, 29–59.

Chakraborty, D., Başağaoğlu, H., Winterle, J., 2021. Interpretable vs. noninterpretable machine learning models for data-driven hydro-climatological process modeling. Expert Systems with Applications 170, 114498. https://doi.org/10.1016/j.eswa.2020.114498

Chava, S., 2014. Environmental Externalities and Cost of Capital. Management Science 60, 2223–2247. https://doi.org/10.1287/mnsc.2013.1863

Cooper, E., Uzun, H., 2019. Corporate social responsibility and bankruptcy. Studies in Economics and Finance.

du Jardin, P., 2021. Forecasting corporate failure using ensemble of self-organizing neural networks. European Journal of Operational Research 288, 869–885.

Du Jardin, P., 2015. Bankruptcy prediction using terminal failure processes. European Journal of Operational Research 242, 286–303.

European Investment Bank. (2012). EIB Working Papers 2012/05 - The costs of climate-change adaptation in Europe: a review. In Journal of Addiction Disorder and Rehabilitation. European Investment Bank. https://doi.org/https://doi.org/

Fontecha, J.E., Agarwal, P., Torres, M.N., Mukherjee, S., Walteros, J.L., Rodríguez, J.P., 2021. A Two-Stage Data-Driven Spatiotemporal Analysis to Predict Failure Risk of Urban Sewer Systems Leveraging Machine Learning Algorithms. Risk Analysis 41, 2356–2391. https://doi.org/10.1111/risa.13742

Friedman, J.H., 2001. Greedy function approximation: a gradient boosting machine. Annals of statistics 1189–1232.

Fuss, S., 2016. Substantial risk for financial assets. Nature Climate Change 6, 659–660.

Gao, S., Wang, Y., 2022. Explainable deep learning powered building risk assessment model for proactive hurricane response. Risk analysis.

Geurts, P., Ernst, D., Wehenkel, L., 2006. Extremely randomized trees. Machine learning 63, 3–42.

Gopalakrishnan, S., Dugal, M., 1998. STRATEGIC CHOICE VERSUS ENVIRONMENTAL DETERMINISM: A DEBATE REVISITED. The International Journal of Organizational Analysis 6, 146–164. https://doi.org/10.1108/eb028882

Goutte, C., Gaussier, E., 2005. A Probabilistic Interpretation of Precision, Recall and F-Score, with Implication for Evaluation, in: Losada, D.E., Fernández-Luna, J.M. (Eds.), Advances in Information Retrieval. Springer Berlin Heidelberg, Berlin, Heidelberg, pp. 345–359.

Huang, H.H., Kerstein, J., Wang, C., 2018. The impact of climate risk on firm performance and financing choices: An international comparison. Journal of International Business Studies 49, 633–656. https://doi.org/10.1057/s41267-017-0125-5

Jabeur, S.B., Gharib, C., Mefteh-Wali, S., Arfi, W.B., 2021. CatBoost model and artificial intelligence techniques for corporate failure prediction. Technological Forecasting and Social Change 166, 120658.

Jones, S., 2017. Corporate bankruptcy prediction: a high dimensional analysis. Review of Accounting Studies 22, 1366–1422.

Kraus, M., Feuerriegel, S., Oztekin, A., 2020. Deep learning in business analytics and operations research: Models, applications and managerial implications. European Journal of Operational Research 281, 628–641. https://doi.org/10.1016/j.ejor.2019.09.018

Kumar, P.R., Ravi, V., 2007. Bankruptcy prediction in banks and firms via statistical and intelligent techniques – A review. European Journal of Operational Research 180, 1–28. https://doi.org/10.1016/j.ejor.2006.08.043

Linnenluecke, M., Griffiths, A., 2010. Beyond Adaptation: Resilience for Business in Light of Climate Change and Weather Extremes. Business & Society 49, 477–511. https://doi.org/10.1177/0007650310368814

Linnenluecke, M.K., Griffiths, A., 2012. Assessing organizational resilience to climate and weather extremes: complexities and methodological pathways. Climatic Change 113, 933–947. https://doi.org/10.1007/s10584-011-0380-6

Linnenluecke, M.K., Griffiths, A., Winn, M.I., 2013. Firm and industry adaptation to climate change: a review of climate adaptation studies in the business and management field. WIREs Climate Change 4, 397–416. https://doi.org/10.1002/wcc.214

Liu, X., Zhang, K., Ren, Y., 2023. Does climate warming affect labour productivity in emerging economies?—Evidence from Chinese-listed firms. Applied Economics 55, 2801–2814.

Lundberg, S.M., Erion, G., Chen, H., DeGrave, A., Prutkin, J.M., Nair, B., Katz, R., Himmelfarb, J., Bansal, N., Lee, S.-I., 2020. From local explanations to global understanding with explainable AI for trees. Nature Machine Intelligence 2, 56–67. https://doi.org/10.1038/s42256-019-0138-9

Mai, F., Tian, S., Lee, C., Ma, L., 2019. Deep learning models for bankruptcy prediction using textual disclosures. European Journal of Operational Research 274, 743–758. https://doi.org/10.1016/j.ejor.2018.10.024

MENDELSOHN, R., 2012. THE ECONOMICS OF ADAPTATION TO CLIMATE CHANGE IN DEVELOPING COUNTRIES. Climate Change Economics 03, 1250006. https://doi.org/10.1142/S2010007812500066

Neufeldt, H., Christiansen, L., Dale, T.W., 2021. Adaptation Gap Report 2021-The Gathering Storm: Adapting to climate change in a post-pandemic world. UNEP DTU Partnership.

Ohlson, J.A., 1980. Financial ratios and the probabilistic prediction of bankruptcy. Journal of accounting research 109–131.

Ooghe, H., De Prijcker, S., 2008. Failure processes and causes of company bankruptcy: a typology. Management Decision 46, 223–242. https://doi.org/10.1108/00251740810854131

Ozkan, A., Temiz, H., Yildiz, Y., 2022. Climate Risk, Corporate Social Responsibility, and Firm Performance. British Journal of Management.

Pankratz, N., Bauer, R., Derwall, J., 2023. Climate change, firm performance, and investor surprises. Management Science.

Pankratz, N., Bauer, R., Derwall, J., n.d. Climate Change, Firm Performance, and Investor Surprises. Management Science 0, null. https://doi.org/10.1287/mnsc.2023.4685

Russell, S.J., 2010. Artificial intelligence a modern approach. Pearson Education, Inc.

Senoner, J., Netland, T., Feuerriegel, S., 2022. Using Explainable Artificial Intelligence to Improve Process Quality: Evidence from Semiconductor Manufacturing. Management Science 68, 5704–5723. https://doi.org/10.1287/mnsc.2021.4190

Smit, B., Burton, I., Klein, R.J.T., Wandel, J., 2000. An Anatomy of Adaptation to Climate Change and Variability. Climatic Change 45, 223–251. https://doi.org/10.1023/A:1005661622966

Stef, N., 2021. Institutions and corporate financial distress in Central and Eastern Europe. European Journal of Law and Economics 52, 57–87.

Stef, N., Bissieux, J.-J., 2022. Resolution of corporate insolvency during COVID-19 pandemic. Evidence from France. International Review of law and Economics 70, 106063.

Stef, N., Zenou, E., 2021. Management-to-staff ratio and a firm’s exit. Journal of Business Research 125, 252–260.

Teodorescu, M.H., Morse, L., Awwad, Y., Kane, G.C., 2021. Failures of Fairness in Automation Require a Deeper Understanding of Human-ML Augmentation. MIS Quarterly 45.

Tzouvanas, P., Kizys, R., Chatziantoniou, I., Sagitova, R., 2019. Can Variations in Temperature Explain the Systemic Risk of European Firms? Environmental and Resource Economics 74, 1723–1759. https://doi.org/10.1007/s10640-019-00385-0

Whittington, Richard, 1988. Environmental structure and theories of strategic choice. Journal of Management studies 25, 521–536.

Winston, A. (2014). Resilience in a hotter world. Harvard Business Review, 92(4), 56–64, 132. https://europepmc.org/article/med/24830282

Zhang, P., Deschenes, O., Meng, K., Zhang, J., 2018. Temperature effects on productivity and factor reallocation: Evidence from a half million Chinese manufacturing plants. Journal of Environmental Economics and Management 88, 1–17.

Zhao, M., Lee, J.K.W., Kjellstrom, T., Cai, W., 2021. Assessment of the economic impact of heat-related labor productivity loss: a systematic review. Climatic Change 167, 1–16.

Zorn, M.L., Norman, P.M., Butler, F.C., Bhussar, M.S., 2017. Cure or curse: Does downsizing increase the likelihood of bankruptcy? Journal of Business Research 76, 24–33.

Table 1. Variables definition	
Financial variables	
TA	Total assets
Leverage	Ratio between firm's total debt and TA
Current Ratio	Ratio between current assets and current liabilities
Cash	Amount of firm's cash
C/TA	Ratio between Cash and TA
EBITDA	Earnings before interest, taxes, depreciation, and amortization
EBITDA/TA	Ratio between EBITDA and TA
WC	Working capital
WC/TA	Ratio between WC and TA
Equity	Total equity value
GR	Ratio between total debt and Equity
Sales	Sales
I/S	Ratio between interests and sales
TS/TA	Ratio between Sales and TA
Age	Firm's age in years
Climate change variables	
CALCI	Climate Altering Land Cover Index (CALCI) assesses the changes in the share of climate altering land cover as compared to the base year, 2015. Source: International Monetary Fund: https://climatedata.imf.org/
CRI	Global Climate Risk Index captures to what extent countries have been affected by impacts of weather-related loss events (storms, floods, heat waves etc.). Source: Eckstein et al. (2021), Greenwatch database: https://www.germanwatch.org/en/cri.
VUL	Vulnerability index measures a country's sensitivity to the adverse effects of climate change by relying on six dimensions, namely food, water, health, ecosystem service, human habitat, and infrastructure. Source: University of Notre Dame. https://gain.nd.edu/our-work/country-index/
ASTC	Annual mean surface temperature change estimated with respect to a baseline climatology. Source: International Monetary Fund: https://climatedata.imf.org/
TEM	Average temperature of the county in Celsius
HUM	Humidity ratio measured as the weight of water vapor per unit weight of dry air
DP	Amount of daily precipitation in mm that has accumulated within the last 24 hours
WS	Average wind speed for the last 10 minutes in m/s
Notes: The observations of the financial variables were gathered from the *Diane* database of *Bureau Van Dijk* while the local weather variables (*TEM*, *HUM*, *DP* and *WS*) from Meteo France.	

Table 2. Descriptive statistics					
Failed firms					
Name	**mean**	**sd**	**median**	**min**	**max**
Leverage	2799.93	122084.14	16.09	-41021.20	5911290.00
TA	2914.48	11797.50	927.01	0.10	285320.40
CR	1.99	4.88	1.21	0.00	94.86
Equity	242.93	4220.68	75.20	-48350.40	79743.89
Cash	153.08	1165.25	30.55	-3403.98	28925.19
EBITDA	-75.48	743.55	0.00	-13726.31	13406.08
WC	334.25	4484.31	30.31	-18847.01	167592.80
GR	103.85	177.57	17.88	0.00	986.88
Sales	3654.70	11724.27	1393.12	-15.56	308628.52
I/S	762.70	4907.38	0.28	-3.52	77732.96
Age	17.00	14.13	12.00	-1.00	70.00
CRI	45.61	7.56	42.67	27.83	61.17
CALCI	99.71	0.24	99.78	99.37	100.00
VUL	0.31	0.00	0.31	0.31	0.31
ASTC	1.66	0.45	1.69	0.33	2.41
C/TA	0.08	0.16	0.04	-0.55	1.00
EBITDA/TA	-0.06	0.41	0.00	-13.22	1.27
WC/TA	-0.27	7.99	0.04	-375.99	0.96
TS/TA	2.02	2.88	1.78	-0.16	92.96
TEM	13.86	3.00	13.09	8.40	27.35
HUM	73.74	5.67	75.02	57.80	85.75
DP	2.01	0.81	1.89	0.47	9.71
WS	3.82	1.00	3.65	1.56	7.31
Healthy firms					
Name	**mean**	**sd**	**median**	**min**	**max**
Leverage	113.84	1435.93	29.47	-22037.29	122892.70
TA	24577.24	208343.31	913.00	3.08	9575055.00
CR	2.85	4.94	1.79	0.00	90.68
Equity	8672.16	95617.48	353.06	-1006943.00	7577739.00
Cash	1613.96	14099.85	114.13	-141107.30	702511.00
EBITDA	1458.29	15697.83	72.00	-494220.70	677380.00
WC	2496.94	39649.64	67.25	-922908.50	2611862.00
GR	60.24	113.73	17.71	0.00	988.31
Sales	25426.48	258235.37	1160.30	0.00	14795377.00
I/S	6.22	358.22	0.18	-1.64	38200.00
Age	20.02	15.45	16.00	-1.00	120.00
CRI	45.72	7.33	42.67	37.33	61.17
CALCI	99.72	0.24	99.78	99.37	100.00
VUL	0.31	0.00	0.31	0.31	0.31
ASTC	1.65	0.46	1.69	0.33	2.41
C/TA	0.19	0.21	0.13	-1.66	1.00
EBITDA/TA	0.11	0.21	0.09	-4.47	15.03

WC/TA	0.10	0.34	0.10	-13.18	1.70
TS/TA	2.03	1.90	1.73	0.00	60.57
TEM	13.87	3.17	13.05	7.77	27.83
HUM	73.85	5.58	75.02	57.80	86.97
DP	2.06	0.87	1.93	0.47	11.30
WS	3.83	0.93	3.68	1.52	7.31

Table 3. Model accuracy (%) and significant differences (%).			
	Model A	Model B	Differences (% point): Model B-A
XGBoost	90.35	94.76	4.41
GBM	90.19	90.82	0.63
DNN	83.29	79.98	-3.31
RF	89.80	89.06	-0.74
LR	84.94	85.52	0.58
ETR	89.78	89.89	0.11

Note: Results reported on cross-validation data (5-fold)

Table 4. Differences between correct classification rates achieved with XGBoost-based CCR variables, and those achieved with other models.

		Panel A : Differences (% point)		Panel B: P-values of a test for differences	
	Methods	Model A	Model B	Model A	Model B
XGBoost-CCR	XGBoost	4.41***	-	0.000	-
	GBM	4.57***	3.94***	0.000	0.000
	DNN	11.47***	14.78***	0.000	0.000
	RF	4.96***	5.7***	0.000	0.000
	LR	9.82***	9.24***	0.000	0.000
	ETR	4.98***	4.87***	0.000	0.000

Notes: Model A is constructed solely on financial ratios, Model B is constructed on financial ratios and CCR. XGBoost: Extreme Gradient Boosting, DNN : deep neural network, RF: Random Forest, LR : Logistic regression. Results reported on cross-validation data (5-fold)

Table 5. AUC (Area under the ROC curve) by type of model			
	Model A	Model B	Differences (% point): Model B-A
XGBoost	0.907	0.969	6
GBM	0.908	0.923	1.5
DNN	0.816	0.768	-4.8
RF	0.899	0.903	0.4
LR	0.818	0.821	0.3
ERT	0.904	0.911	0.7
Note: Note: Results reported on cross-validation data (5-fold)			

Table 6. Differences between Area under the ROC curve (AUC) estimated by the XGBoost-CCR, and those achieved with other models.					
		Panel A : Differences (% point)		Panel B: P-values of a test for differences	
	Methods	Model A	Model B	Model A	Model B
XGBoost-CCR	XGBoost	6***	-	0.000	-
	GBM	6.1***	4.6***	0.000	0.000
	DNN	15.3***	20.1***	0.000	0.000
	RF	7***	6.6***	0.000	0.000
	LR	15.1***	14.8***	0.000	0.000
	ETR	6.5***	5.8***	0.000	0.000

Notes: Model A is constructed solely on financial ratios, Model B is constructed on financial ratios and climate change risk variables. XGBoost: Extreme Gradient Boosting, GBM:Gradient Boosting Machine, DNN : deep neural network, RF: Random Forest, LR : Logistic regression

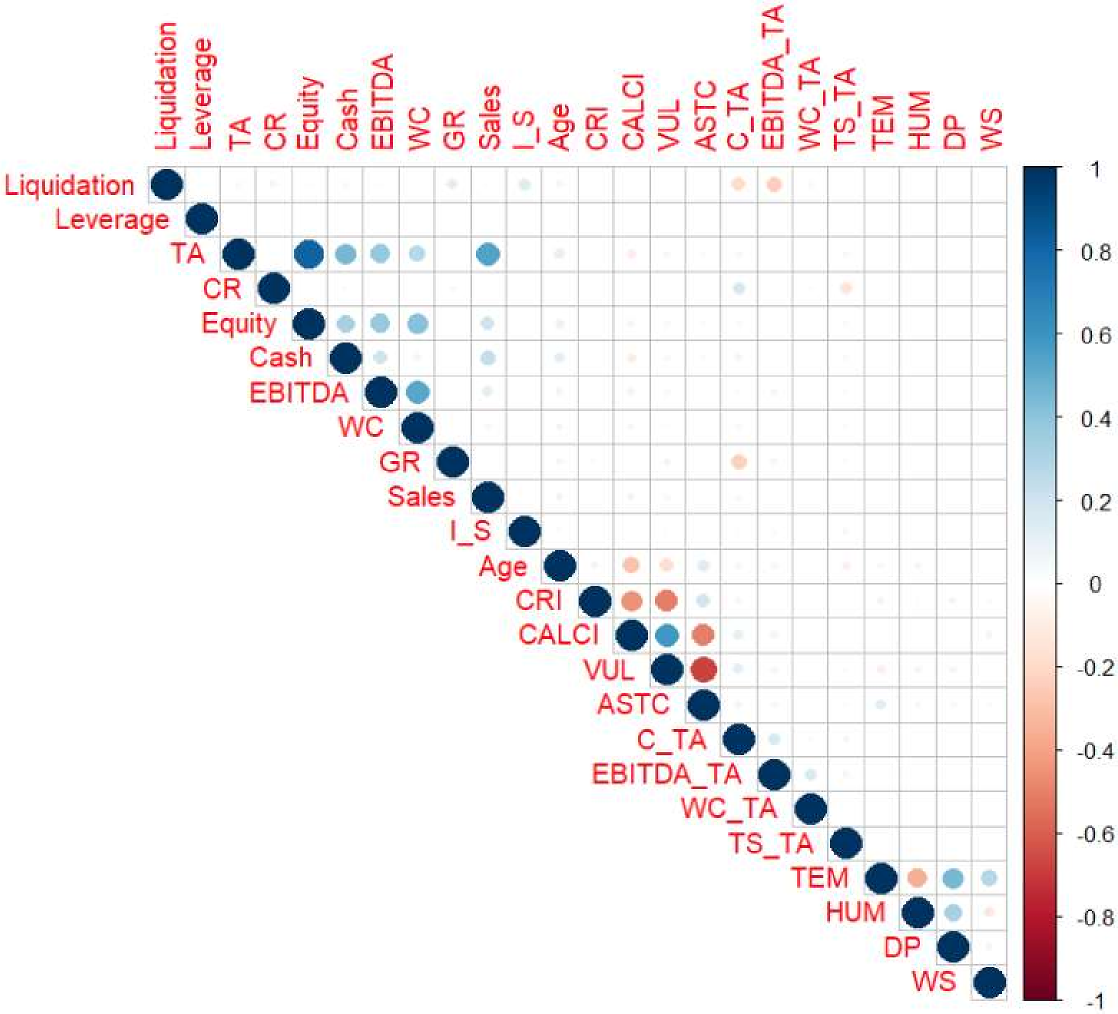

Fig. 1. Correlogram plot of all predictors

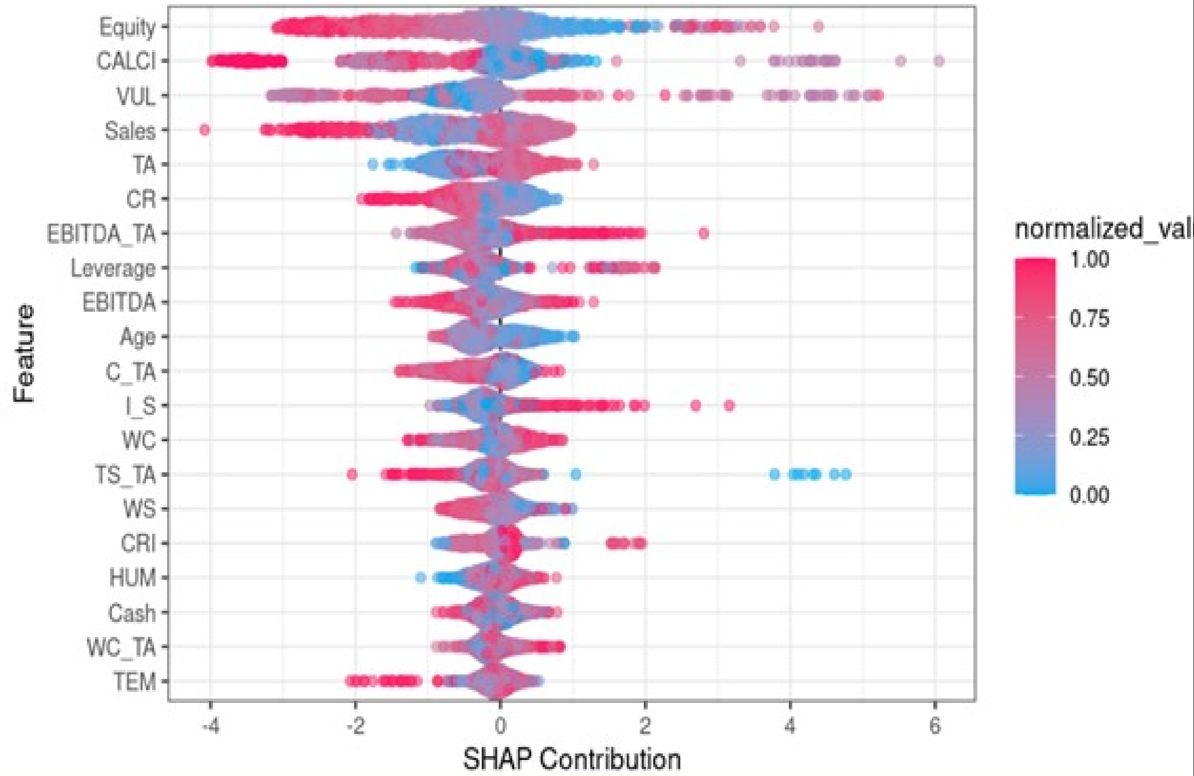

Fig .2. Global interpretation plots of the XGBoost.

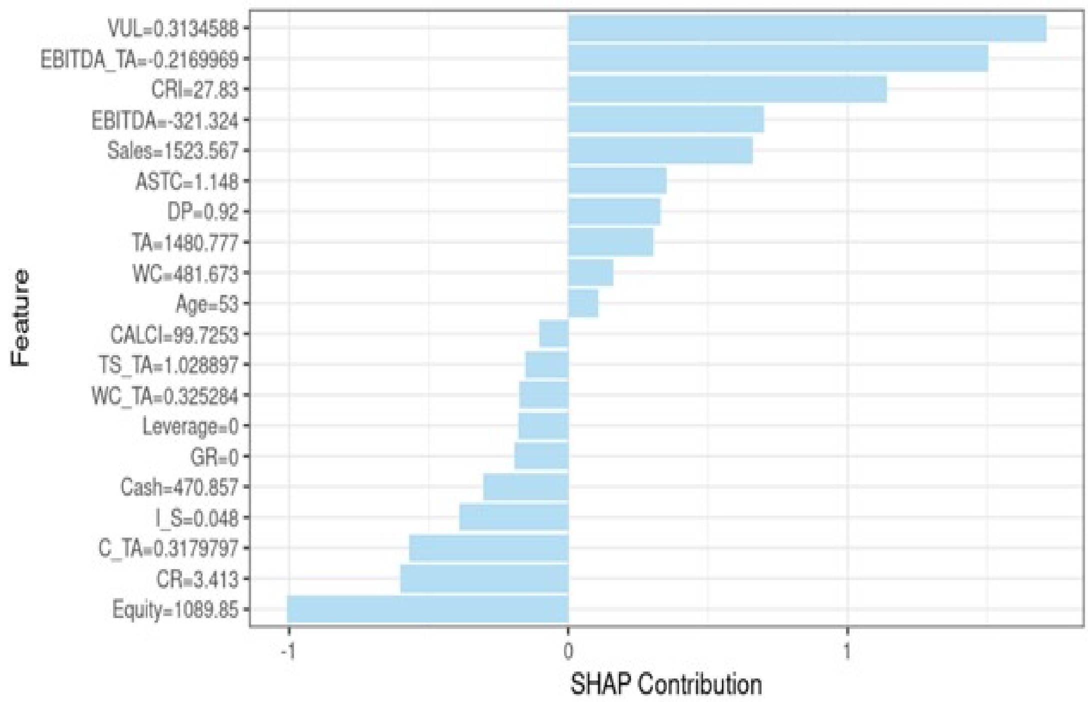

Fig. 3. SHAP contribution for an observation with a high predicted probability of failure

Table 7. F-score by type of model			
	Model A	**Model B**	Differences (% point): Model B-A
XGBoost	0.656	0.809	15.3
GBM	0.657	0.684	2.7
Deep learning	0.492	0.432	-6
Random Forest (DRF)	0.644	0.640	-0.4
Logistic model	0.497	0.496	-0.1
Extremely Randomized Trees (XRT)	0.651	0.662	1.1

Table 8. Differences between F-score estimated by the XGBoost-CCR, and those achieved with other models.

		Panel A : Differences (% point)		Panel B: P-values of a test for differences	
	Methods	Model A	Model B	Model A	Model B
XGBoost-CCR	XGBoost	15.3***	-	0.000	-
	GBM	15.2***	12.5***	0.000	0.000
	DNN	31.7***	37.7***	0.000	0.000
	RF	16.5***	16.9***	0.000	0.000
	LR	31.2***	31.3***	0.000	0.000
	ETR	15.8***	14.7***	0.000	0.000

Notes: Model A is constructed solely on financial ratios, Model B is constructed on financial ratios and climate change risk variables. XGBoost-CCR: Extreme Gradient Boosting based climate change risk XGBoost: Extreme Gradient Boosting, GBM:Gradient Boosting Machine, DNN : deep neural network, RF: Random Forest, LR : Logistic regression

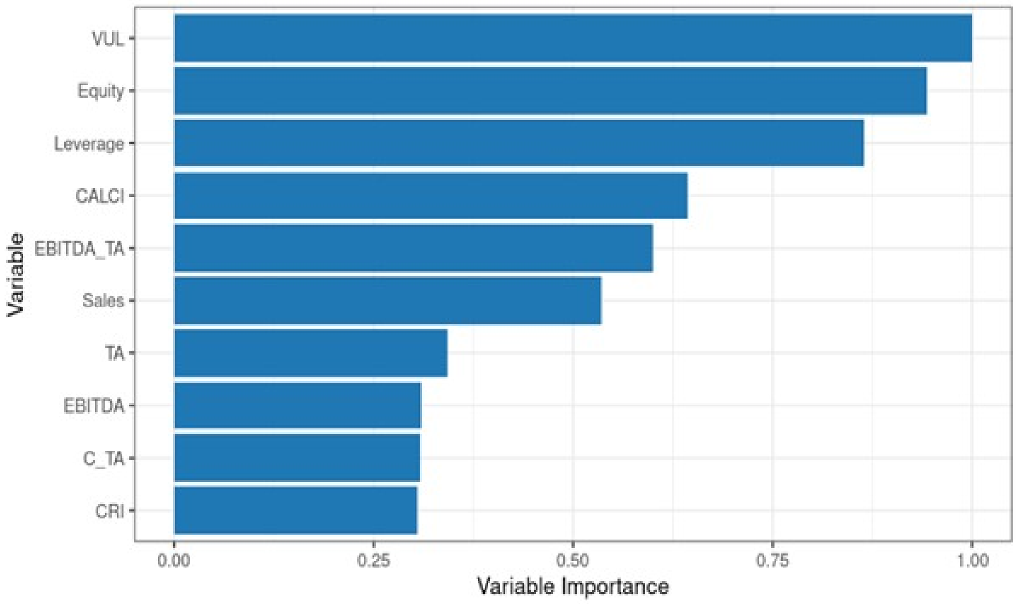

Fig. 4. Variable importance