POLITECNICO
MILANO 1863

SCUOLA DI INGEGNERIA INDUSTRIALE
E DELL'INFORMAZIONE

Sovereign Rating : Assessment with parsimonious models and link with Climate Change

Tesi di Laurea Magistrale in
Mathematical Engineering - Ingegneria Matematica

Author: **Adrien HALLARD**

Student ID: 10784448
Advisor: Prof. Massimo TAVONI
Co-advisor: Matteo CALCATERRA
Academic Year: 2021-22

Beyond a critical point within a finite space, freedom diminishes as numbers increase. This is as true of human in the finite space of a planetarium ecosystem as it is of gas molecules in a sealed flask. The human question is not how many can possibly survive within the system, but what kind of existence is possible for those who do survive

Pardot Kynes, First Planetologist of Arrakis

Appendix I, The Ecology of Dune

Dune, Frank Herbert

Abstract

Climate Change is the biggest challenge of the century, and finance has also to become sustainable. However, a bridge should be built between climate science and finance. Moreover, decision makers need a clear and easy interpretable information to know how climate change could be translated in risk. Based on the work from Klusak et al., parsimonious models are built to reconstruct the sovereign rating. The idea is to use few variables in input and a machine learning tool to obtain the sovereign rating in output. Sovereign rating is used since it is one of the clearest and most used parameter by decision makers, portfolio managers etc in finance. Some machine learning tools are explored in this work, with a special focus on interpretable tools. First the model is built using only macroeconomic parameters as Klusak et al., then using only data for which SSP projections are available. This work underlines the weakness from Klusak et al., in particular in their way to make projections for macroeconomic variables and shows that their model is not robust. Besides, this work emphasises an other point of view on sovereign rating, without direct link with debts but with development parameters. This new model emphasises that fighting climate change is not bad for sovereign ratings, leading also to some upgrades compared to today. Clearly, harsher the climate damage function, the better the effects of fighting climate change on the sovereign ratings. Eventually, this work underlines also the weakness and the difficulties of this kind of approach to reconstruct the rating in particular to make projections.

Keywords: Climate Finance, Machine Learning, Sovereign Ratings, Climate Change

Abstract in lingua italiana

Il cambiamento climatico è la più grande sfida del secolo e anche la finanza deve diventare sostenibile. Tuttavia, è necessario creare un ponte tra la scienza del clima e la finanza. Inoltre, i decisori pubblici e privati hanno bisogno di informazioni chiare e facilmente interpretabili per sapere come il cambiamento climatico possa tradursi in rischio. Sulla base del lavoro dell'autore di Klusak et al., sono stati costruiti modelli parsimoniosi per ricostruire il rating sovrano. L'idea è quella di utilizzare poche variabili in ingresso e uno strumento di apprendimento automatico per ottenere il rating sovrano in uscita. Il rating sovrano viene utilizzato in quanto è uno dei parametri più chiari e più utilizzati dai decisori, dai gestori di portafoglio, ecc. in finanza. In questo lavoro vengono esplorati alcuni strumenti di apprendimento automatico, con particolare attenzione a quelli interpretabili. In primo luogo il modello viene costruito utilizzando solo i parametri macroeconomici come Klusak et al., quindi utilizzando solo i dati per i quali sono disponibili le proiezioni SSP. Questo lavoro sottolinea la debolezza di Klusak et al., in particolare nel loro modo di fare proiezioni per le variabili macroeconomiche e dimostra che il loro modello non è robusto. Inoltre, questo lavoro sottolinea un altro punto di vista sul rating sovrano, senza un legame diretto con il debito ma con i parametri di sviluppo. Questo nuovo modello sottolinea che la lotta al cambiamento climatico non è negativa per i rating sovrani, portando anche ad alcuni upgrade rispetto ad oggi. Chiaramente, più severa è la funzione di danno climatico, migliori sono gli effetti della lotta al cambiamento climatico sui rating sovrani. Infine, questo lavoro sottolinea anche la debolezza e le difficoltà di questo tipo di approccio per ricostruire il rating, in particolare per fare proiezioni.

Parole chiave: Finanza climatica, machine learning, rating sovrani, cambiamento climatico

Abstract en langue française

Le changement climatique est le plus grand défi du siècle et la finance doit elle aussi devenir durable. Cependant, un pont doit être construit entre les sciences du climat et la finance. De plus, les décideurs ont besoin d'informations claires et faciles à interpréter pour savoir comment le changement climatique peut se traduire en risques. Sur la base des travaux du Klusak et al., des modèles parcimonieux sont construits pour reconstruire la notation souveraine. L'idée est d'utiliser peu de variables en entrée et un outil d'apprentissage automatique pour obtenir la notation souveraine en sortie. La notation souveraine est utilisée car il s'agit de l'un des paramètres les plus clairs et les plus utilisés par les décideurs, les gestionnaires de portefeuille, etc. en finance. Certains outils d'apprentissage automatique sont explorés dans ce travail, avec un accent particulier donné sur les outils interprétables. Tout d'abord, le modèle est construit en utilisant uniquement les paramètres macroéconomiques, comme Klusak et al., puis en utilisant uniquement les données pour lesquelles des projections SSP sont disponibles. Ce travail souligne les faiblesses de Klusak et al., en particulier dans leur façon de faire des projections pour les variables macroéconomiques et montre que leur modèle n'est pas robuste. Par ailleurs, ce travail met en avant un autre point de vue sur la notation souveraine, sans lien direct avec les dettes, mais avec des paramètres de développement. Ce nouveau modèle montre que la lutte contre le changement climatique n'est pas mauvaise pour les notations souveraines, conduisant également à certaines améliorations par rapport à aujourd'hui. Il est clair que plus la fonction de dommage climatique est sévère, meilleurs sont les effets de la lutte contre le changement climatique sur les notations souveraines. Enfin, ce travail souligne également la faiblesse et les difficultés de ce type d'approche pour reconstruire la notation, en particulier pour faire des projections.

Mots Clefs: Finance Climatique, Apprentissage Automatique, Notations Souveraines, Changement Climatique

Contents

Introduction

Climate change is the biggest challenge of our century. It has already some impacts on all the society, increasing environmental, political, social, health, economic, and financial risks. This risks include physical damage from extreme events, litigation risk and consumers movement. Moreover, according to climate economic models, the gross world product will lose between 2% and 22% by 2100 [4, 10, 17]. However, we can keep some hope. According to the Intergovernmental Panel on Climate Change (IPCC), *"Climate resilient development is enabled when governments, civil society and the private sector make inclusive development choices that prioritise risk reduction, equity and justice, and when decision-making processes, finance and actions are integrated across governance levels, sectors and timeframes"* [1]. Even if some works exist to estimate some risk, like 'Climate Value at Risk' [13] public and private investors need more information on how to green finance, that is how to make finance efficient to face climate change, what are the risks of the climate change, what are the investment 'good' for the planet and the society.

Then, greening the finance system is a very actual challenge, and the European Central Bank (ECB) in particular wants to become a leader of climate finance creating a framework to assess, analyse and act to face climate risk [8] (2021). However, even if greening finance is a more and more popular theme, and that more and more people want to act, some problem remains. Although Environmental, Social, and Governance (ESG) criteria appears to help investor to act [3], this criteria may raise some problems given how they are built. For example, the controversial company Totalenergies (with a new project in Uganda[1] and exploration for gas in Artic and other places[2]) has obtained an A grade according to the MSCI index[3], that can raise some questions on the construction of this kind of criteria. Indeed, it is really hard to translate climate science into some acceptable metric for investors and decisions makers.

[1] https://www.la-croix.com/Economie/profits-embarrassants-TotalEnergies-2022-02-10-1201199604

[2] https://totalenergies.com/fr/expertise-energies/exploration-production/petrole-gaz/notre-ambition-petrole-et-gaz

[3] https://sustainable-performance.totalenergies.com/fr/reporting/indices-esg-de-reference#nasdaq

The goal is thus to build a way to rely on climate science and real-world financial indicators, based on the article *Rising Temperatures, Falling Ratings: The Effect of Climate Change on Sovereign Creditworthiness* [19] (2021). The general goal is to simulate the effect of climate change on sovereign ratings for more than hundred countries. The idea is first to reconstruct a method for the sovereign rating which should be parsimonious and reliable with climate science. Then climate model and S&P's own natural disaster risk assessments Kraemer et al. are used to build climate adjusted parameters which feed the model built in first part. Finally, it is possible to project the additional cost of the debt due to climate-induced sovereign downgrades. The idea is also to make some constructive remarks on the work done by Klusak et al. (2021) and build a parsimonious model to build again the sovereign rating in an other way.

This work focuses on the sovereign ratings since it is a well known indicator, easily readable, used by portfolio managers, investors, authorities etc. Moreover, the part outstanding sovereign debt was about 44 trillion USD in 2020 according to the Bank for International Settlement (BIS) data, representing a large part of total asset under management. Therefore, the ratings have a big impact on the market. In fact, a downgrade means an increase of the debt cost (either public and private) and has an impact on the whole economy of the country. Besides, it can have indirect impact on other asset classes, for example Fitch, a Credit Rating Agency, regularly downgrades bank of a country after having downgraded the sovereign debt, the last example being Russia[4].

Finally, it must be underlined that climate change impact the whole economy even the whole system, that is why we focus on a global indicator to evaluate the impact of the climate change for finance. Indeed, it is crucial to first have a global view and idea and what could happen according to different scenarios to be able to act.

This thesis focuses in particular on the parsimonious way to build again the sovereign rating, to provide a very accurate model and also try to obtain an interpretable model, using additional methods compared to the basis work from Klusak et al. (2021). Moreover, this thesis is also a critic of Klusak et al. (2021). Besides, this thesis uses other indicators to build an other parsimonious model to build again the sovereign rating and make projections.

[4] Here the example with Russia :`https://www.fitchratings.com/research/banks/fitch-downgrades-russian-banks-following-sovereign-downgrade-15-03-2022` notice an other example with Bolivia in 2020 : `https://www.fitchratings.com/research/banks/fitch-downgrades-two-bolivian-banks-following-sovereign-rating-downgrade-08-10-2020`

In the following, the thesis is divided in 3 main chapters and a Chapter for the conclusions and discussions :

- Chapter 1 which is going to present the climate economic models, the sovereign rating and the machine learning basis we are going to use
- Chapter 2 where the idea of Klusak et al. is used to to build the Sovereign Rating Model, making some works around and see the limits
- Chapter 3 where new ways are presented to build the Sovereign Ratings with new projections
- Chapter 4 where we conclude and discuss the results and present some perspectives

1 | State of the art

In this chapter, a general presentation of the state of the art about climate models, sovereign ratings and machine learning used in next chapters are presented.

1.1. Climate Models

In macroeconomic models, two categories could be made: in one hand the Integrated Assessment Models (IAM) such DICE for which Bill Nordhaus received a Nobel Prize in 2018, in the other hand, new macroeconomic models estimating the long run impact of change in temperature and precipitation at country level.

1.1.1. Integrated Assessments Models

The IAM are used to:

1. Generate emissions scenarios and their climate and economic consequences
2. Assess the implication and costs of climate change policy making a Cost-effectiveness analysis
3. Define the optimal mitigation level using a Cost Benefit Analysis

However, many IAMs exist and are not necessary alike, using or a dynamic time treatment, or a static time treatment. They can be used or for policy evaluation or for policy optimisation. The technology representation can be very different between two IAMs, etc.

Moreover, IAMs depend on many input parameters, such as parameters specifying preferences, discount rates, elasticity of substitution between factors of production, climate change economic impacts and adaptation cost. Thus, the IAMs are very sensitive to the input parameters. Besides, IAMs take some inputs from scenarios like Shared Socioeconomic Pathways (SSP) scenarios. The SSP are based on 5 narratives describing broad socioeconomic trends that could shape future society.

The five narratives are:

- SSP1: Sustainability pathway, with relatively low challenges for climate mitigation and adaptation
- SSP2: the "middle of the road" scenario
- SSP3: the Regional Rivalry scenario (A Rocky Road), with High socioeconomic challenges both for mitigation and adaptation
- SSP4: the Inequality (A Road divided) scenario, with high socioeconomics challenges for adaptation but few for mitigation
- SSP5: the : Fossil-fueled Development (Taking the Highway) scenario, with high socioeconomics challenges for mitigation but few for adaptation

Therefore, IAMs are not projection tools, but guidance for policy, based on cost optimization.

The other limitation of the IAMs for the application for next chapters is that they aggregate many countries together to build some regions[1], while here the goal is to obtain the effect of climate on sovereign creditworthiness at a country level. It could be complex to estimate from a region how the risks are allocated, for instance Dietz et al. (2016) estimated a global 'climate value at risk' using DICE without commenting the distribution among countries.

Informations about IAM can be found on `https://www.iamconsortium.org/`.

1.1.2. More Recent Models

New models are trying to combine climate science with long run macroeconomic analysis, combining relationship between the Gross Domestic Product (GDP) and temperatures, [10–12, 17].

The idea behind is to be able to make some GDP projections according to a Representative Concentration Pathway (RCP) scenarios, which represent a greenhouse gas concentration trajectory.

First, Dell et al. (2012) considers the effect on the GDP/capita of the temperature in this way:

$$g_{i,t} = g_i + (\beta + \gamma)T_{i,t} - \beta T_{i,t-1} \tag{1.1}$$

[1]Even with RICE model even if we can have now until 57 regions Gazzotti (2022)

where $g_{i,t}$ corresponds to the growth rate of per-capita output (for country i at time t), β to the "level effect" of weather shock on output, and γ to the "growth effect" on weather shock. With the equation 1.1 the identification of level effect and growth effect could be separated through the examination of transitory weather shocks. The level effect can reverse itself when the temperature returns to his prior state, while the growth effect corresponds to the summation of the temperature effect over time. Dell et al. extend the reasoning this reasoning since the temperature effects can play out more slowly, adding some lag structure.

$$g_{i,t} = \theta_i + \theta_{rt} + \sum_{j=0}^{L} \rho_j T_{i,t-j} + \varepsilon_{i,t} \tag{1.2}$$

where θ_i corresponds to country fixed effect, θ_{rt} to time fixed effects (interacted separately with region dummies and a poor country dummy in our main specifications), $\epsilon_{i,t}$ to an error term, $T_{i,t}$ to a vector of annual average temperature and precipitation with up to L lags included.

Burke et al. (2015) relax the linear hypothesis made by Dell et al. and include fixed effect in a different way.

$$g_{i,t} = h(T_{i,t}) + \lambda_1 P_{i,t} + \lambda_2 P_{i,t}^2 + \mu_i + \nu_t + \theta_i t + \theta_{i2} t^2 + \varepsilon_{i,t} \tag{1.3}$$

where $h()$ is a quadratic function (or more complex), μ_i corresponds to specific countries fixed effect, ν_t to time effect (like shocks or global recession), P to precipitation and $\theta_i t + \theta_{i2} t^2$ to flexible time trend country specific (to catch for example governmental change, demographic change etc).

Kahn et al. (2021) make different consideration to include the temperature. They use this AutoRegressive Distributed Lag (ARDL) model :

$$\Delta g_{i,t} = a_i + \sum_{l=1}^{p} \varphi_l \Delta g_{i,t-l} + \sum_{l=1}^{p} \beta_l^{'} \Delta x_{i,t-l} + \varepsilon_{i,t} \tag{1.4}$$

where a_i corresponds to the fixed effect, $x_{i,t} = [(\mathcal{C}_{i,t} - \mathcal{C}_{i,t-1}^{*})^{+}, (\mathcal{C}_{i,t} - \mathcal{C}_{i,t-1}^{*})^{-}]$, $\mathcal{C}_{i,t} = (\mathcal{T}_{i,t}, \mathcal{P}_{i,t})^{'}$, $\mathcal{C}_{i,t-1}^{*} = (\mathcal{T}_{i,t-1}^{*}, \mathcal{P}_{i,t-1}^{*})^{'}$. $\mathcal{T}_{i,t}$ and $\mathcal{P}_{i,t}$ corresponds to the population weighted average temperature and precipitation of country i in year t, respectively, and $\mathcal{T}_{i,t-1}^{*}$ and $\mathcal{P}_{i,t-1}^{*}$ are the historical norms of climate variables. For the historical norms, they consider the moving averages of temperature and precipitation of country i based on the past m years (typically $m = 30$).

Kalkuhl and Wenz (2020) make different considerations (but follow the idea of Dell et al. and Burke et al.) in this way:

$$g_{i,t} = \sum_{s=0}^{S} \alpha_{s+1} \Delta\Omega_{i,t-s} + \sum_{s=0}^{S} \beta_{s+1} \Delta\Omega_{i,t-s} \tilde{\Omega}_i + F(\tilde{\Omega}_i) + p_i(t) + \mu_t + \epsilon_{i,t} \qquad (1.5)$$

where $g_{i,t}$ is the per-capita growth rate (logarithmic change) of Gross Regional Product (GRP) in region i, compared to GRP in the previous year, $\Omega_{i,t}$ corresponds to a vector of annual mean temperature levels and annual total precipitation values within year t at region i, and then $\Delta\Omega_{i,t} = \Omega_{i,t} - \Omega_{i,t-1}$ describes the respective change of weather. The term $\tilde{\Omega}_i$ refers to average temperature and precipitation levels in region i, either for year t, (contemporaneous weather), or for previous years (lagged weather), depending on the model specification. The term $p_i(t)$ refers to region-specific polynomial time trends which control for various possible slow-moving regional changes that affect growth. μ_t corresponds to the year fixed effect and $\epsilon_{i,t}$ is the error term.

In any cases, with this kind of model, it is possible to make economic projections based on climate science and climate science scenarios.

1.2. Sovereign Ratings

In this section is presented what is the Sovereign Rating and the link with Climate Change.

1.2.1. General Notion

A sovereign credit rating is an independent assessment which measure the creditworthiness of a country or a sovereign entity. This measure is made thanks to some fundamental analysis and not on a market basis. Thus the sovereign credit rating is an indicator which measure the capacity of the obligor to pay back the debt.

The sovereign rating is provided by Credit Rating Agencies (CRA), the most famous and important for the sovereign credit rating are *Standard & Poors (S&P)*, *Moody's KMV* and *Fitch*.

In the following the data from S&P will be used, since S&P rates the most countries, and the data is easily available. Note that S&P adds a little more details on their way of rating, in particular, they take into consideration:

- The likelihood of payment, ie the capacity and willingness of the obligor to meet its financial commitments on an obligation in accordance with the terms of the obligation

- The nature of the financial obligation, and the provisions
- The protection given, and relative position of, the financial obligation in the event of a bankruptcy, reorganization, or other arrangement under the laws of bankruptcy and other laws affecting creditors' rights

S&P precises that the issue rating may take into account an assessment of relative seniority or ultimate recovery in the event of default. At the end, the highest rating given is 'AAA', that means that the obligor will meet his obligation with a very high probability, to D, which means a default of the obligor. This ratings may also be converted on a numerical scale, and separated in investment-grade and non-investment grade (for example). The detailed table of the grades and the 20-notches scale used are in Appendix A in figure A.1.

Notice that ratings are non linear and even if the ratings should appear as a class (like 'AAA') it is more 'continuous'. Moreover the fact that some input are subjective leads us to find some way to 'reconstruct' the data from objective data (even if the methodology is public see figure 1.1), in particular thanks to some Machine Learning techniques, presented in section 1.3.

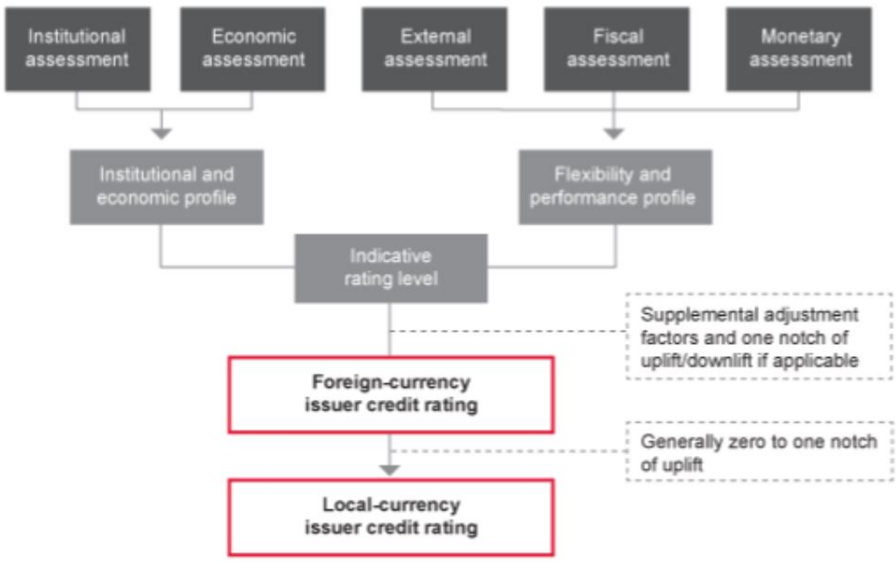


Figure 1.1: S&P rating methodology

Notice that the effects of downgrade on the sovereign debt cost have been already studied, especially by Afonso et al. (2012) and Gande and Parsley (2005). In Klusak et al. they use them to the conversion between sovereign downgrades into yields, using for lower bond Afonso et al. and for upper bond Gande and Parsley whereby 1 notch sovereign downgrade increases sovereign bond spread by 0.08% and 0.12% respectively.

1.2.2. Linked to Climate Change

According to Klusak et al. (2021) not a lot of studies exist making a direct link between climate change and sovereign ratings. They are the first to try to study the impact of climate change on the future ratings. However, as they cite, S&P has studied the impact of natural disasters on ratings and quantified also this impact on key variables (Kraemer et al. [20]).

Kraemer et al. have evaluated the impact of different natural disasters such as earthquakes, tropical storms, floods, and winter storms on the ratings. According to their study, tropical storms and earthquakes have the biggest impact on the rating. They proceed in three steps : quantifying the direct damages, simulating the macroeconomic impacts and then simulating sovereign ratings outcome. However, this study considers only the natural disaster having a big impact and that may occur every 250 years, and is not focused on climate change. it is still interesting since this study quantifies the influence of natural disaster on some key macroeconomics parameters which are important to rate a country.

Klusak et al. (2021) are thus the first to evaluate the climate change impact on sovereign ratings. First, they construct a parsimonious model to build the rating thanks to a machine learning technique. This model is built using some key parameters that can be evaluate thanks to a projection according to climate change. Then, they use the GDP projection from Kahn et al. (2021) and evaluate the key parameters projections according to Kraemer et al. (S&P study) to evaluate the impact of climate change on the sovereign ratings. Since here their ideas are used for this work, more details will be provided in particular in Chapter 2.

1.3. Machine Learning

In this section is presented some basis about Machine Learning and some techniques used in the following. This section is mainly based on the books from Bishop [9] and Mitchell [21] and the Machine Learning Course from prof. Daniele Loiacono.

1.3.1. General Notion

The machine learning is a sub-field of the Artificial Intelligence where knowledge comes from induction and experience. A definition could be : *"A computer program is said to learn from experience E with respect to some class of tasks T and performance measure P, improves with experience E"* Mitchell (1997)

In Machine Learning there are:

- **Supervised** learning, where with a labeled data the computer learns to produce the correct output given a new set of input
- **Unsupervised** learning, where the computer exploits the regularities in Data to be used for reasoning or prediction
- **Reinforcement** learning, where producing actions which affect the environment and receiving rewards according to the actions taken learn to act in order to maximize the rewards in the long term

In the following, we will focus on Supervised Learning, since it is what is used in Chapter 2 and 3.

In supervised machine learning, we have the data $\mathcal{D}$ from an unknown function f that map an input x to an output t. The input variables are called the features and the output variables the labels. The goal is then to learn f. There are three possible tasks in supervised machine learning :

- **Classification** if t is discrete
- **Regression** if t is continuous
- **Probability Estimation** if t is a probability

To approximate f from $\mathcal{D}$,we define a loss function $\mathcal{L}$, choose the hypothesis space $\mathcal{H}$ and find in $\mathcal{H}$ an approximation h of f that minimizes $\mathcal{L}$. It could be possible to approximate f without error. However, it could often occur the overfitting problem. it is when h corresponds too closely or exactly to f on a particular set of data $\mathcal{D}$, and may therefore fail to fit to additional data or predict future observations reliably. To avoid this problem it is common to separate $\mathcal{D}$ in training and test data. To choose the appropriate parameters for a machine learning model the training data is separated in training and validation set. More details about how it will be proceed for this work are in Chapter 2.

1.3.2. Machine Learning Methods

In this subsection is presented some machine learning methods which will be used in the next chapters. All are supervised machine learning methods. These methods will be used to 'reconstruct' sovereign ratings, since Credit Ratings Agencies make their methodology to grade countries public but some inputs are subjective. Besides, ratings are non linear. Thus, to reconstruct sovereign rating from objective data machine learning is useful. it is important to underline that implicitly we make an induction assumption, what was done in the past is what will be done in the future.

First, Decision Tree is a method creating a model which predicts the value of a label variable by learning simple decisions rules inferred from the data features. This method can be used either for classification or regression task. This method is a white box, easy to interpret, easy to visualize, as you can see an example on figure 1.2 (simple example with the python library sklearn[2]).

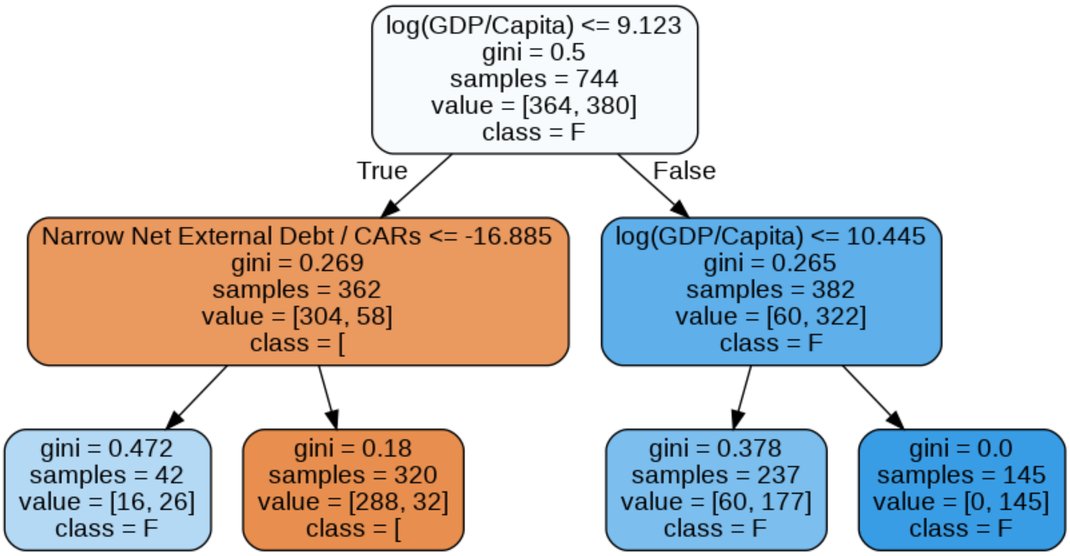


Figure 1.2: Decision Tree Example

However, Decision Tree method is prompt to overfitting, unstable (that means that a small change in the data could result a big change in the tree), so have a high variance. Besides, in practice, decision tree learning algorithms are based on heuristic algorithms where optimally decisions are made at each nodes. But this does not guarantee to return the globally optimal decision tree.

To reduce these problems, and try having low variance and low bias, there are some

[2] https://scikit-learn.org/stable/index.html

ensemble methods. The goal is to combine the predictions of several based estimators such decision trees. There are two main families of ensemble method : bagging methods and boosting methods.

First, bagging stands for bootstrap integration. With bagging N datasets are generated applying random sampling with replacement. Then a model is trained thanks to the based estimator from each dataset generated. At the end, to compute the prediction for new samples, all the based models are applied and their outputs combined, with majority voting for classification or averaging for regression. Although the sample datasets are not independent, bagging is useful to reduce variance, with respect to unstable learner like decision trees and reduce overfitting. Some algorithms use the idea of bagging but are specifically design to have decision trees in based estimator, like random forest algorithm. The random forest procedure is shown in figure 1.3 from wikipedia[3] for a classification task (otherwise it is very similar just average instead of majority voting).

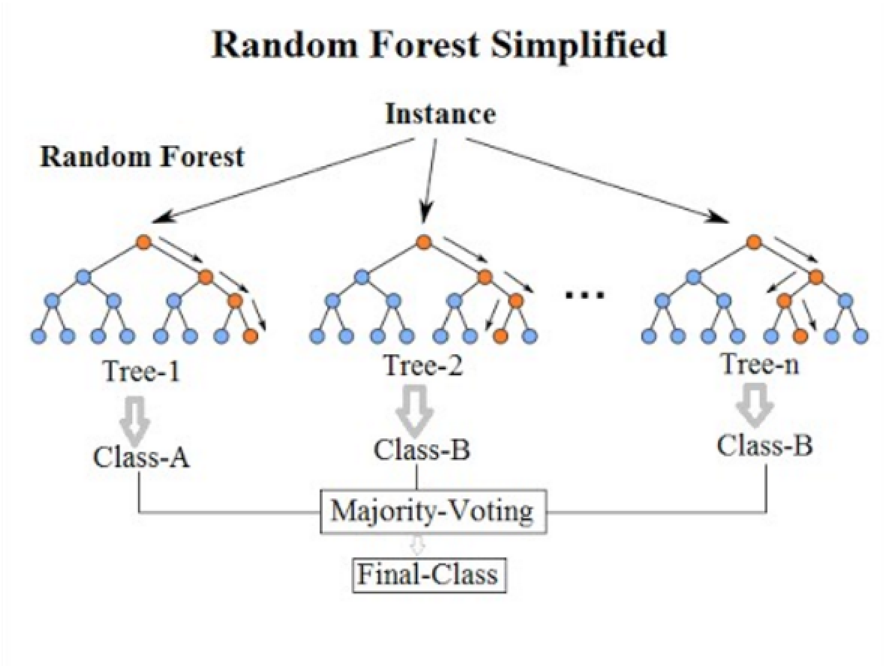


Figure 1.3: Random Forest Procedure

With a naive bagging, the decision trees could be very correlated and thus the variance and the overfitting problem are not reduced a lot. With random forest, some randomness is included in the construction of the decision tree. This leads to reduce the correlation

[3]By Venkata Jagannath - https://community.tibco.com/wiki/random-forest-template-tibco-spotfirer-wiki-page, CC BY-SA 4.0,https://commons.wikimedia.org/w/index.php?curid=68995764

between all the decision trees used to make the final prediction and thus reduce more effectively the variance and the overfitting problem.

In the other hand, the other ensemble method family is the boosting one. Its goal is to achieve a small bias by using simple (weak) learners like decision trees. At the same time, it aims also to keep a small variance. This is achieved by training weak learners sequentially:

1. Give an equal weight to all the samples in the training set
2. Train a weak learner on the weighted training set
3. Compute the error of the trained model on the weighted training set
4. Depending on how well a weak learner did on its predictions, it gets assigned an importance/weight or amount of say. A weak learner that outputs very good predictions will have a high amount of say in the final decision.
5. Repeat from 2. until a criteria is met

The ensemble of models learned can be applied on new samples by computing the weighted prediction of each model (more accurate models weight more). Some algorithms of boosting exist (AdaBoost, Gradient Boosting, XGBoost etc) but they almost all follow the same pattern presented above. We show the AdaBoost algorithm in the following to show an example on how could be computed the different weights (for classification with x for the features and y the targets, with M features):

1. Initialize the observation weights $w_i = \frac{1}{N}$ for $i \in [1, N]$
2. For m = 1 to M
   (a) Fit a classifier $G_m(x)$ to the training data using the weights w_i
   (b) Compute $err_m = \frac{\sum_{i=1}^{N} w_i \mathbf{I}(y_i \neq G_m(x_i))}{\sum_{i=1}^{N} w_i}$
   (c) Compute $\alpha_m = \log(\frac{1-err_m}{err_m})$
   (d) Set $w_i \leftarrow w_i \cdot \exp\left[\alpha_m \cdot \mathbf{I}(y_i \neq G_m(x_i))\right]$, $i \in [1, N]$
3. Output $G(x) = sign(\sum_{m=1}^{M} \alpha_m G_m(x))$

Thanks to the ensemble methods, better results are obtained and that reduce variance and/or bias of simple methods like decision tree. However, this leads to obtain at the end some black box models where the easy way to interpret the models like decision trees is lost. That is the reason why in next subsection some recent research are presented. This research try to use the knowledge of the ensemble model to build very accurate white box model like decision trees.

1.3.3. Recent Research

The need of interpretability of ensemble method has well increased and thus some recent studies try to make the ensemble tree method interpretable. The general idea is to decompose each decision tree in set of rules, then find a way to merge them and prune the result. If the result is a decision tree it is called a 'born again tree'. However, the challenge is to find an efficient way to proceed, since it needs some memories and computing efforts. Merging decision classifier tree was explored by Andrzejak et al. (2013).

One way used by Vidal et al. (2020) is to develop a dynamic-programming based algorithm that exploits sophisticated pruning and bounding rules to reduce the number of recursive calls.

An other way was proposed by Sagi and Rokach (2020) is to create a set of rule conjunctions that represent the original decision forest; the conjunctions are then hierarchically organized to form a new decision tree. Sagi and Rokach (2021) have also worked on the interpretability of boosting algorithm like XGBoost.

To give interpretability to the random forest, Fernández et al. (2020) proposed to extract counterfactual sets from a random forest based on a partial fusion of tree predictors from a random forest into a single Decision Tree and it obtains a counterfactual set that contains the optimal counterfactual.

2 | Sovereign Ratings Model with the idea of **Klusak et al.**

In this chapter the work is introduced done following the idea of Klusak et al.. First, the Data used is presented, making some remarks about. Then the model used for assess ratings is built and some improvements are tested. At the end, some interpretable models are built, if possible with the same performance. Besides, the methodology used for the projection is presented and we make some remarks about it.

2.1. Data

The data used come from the S&P rating Database[1] with data downloaded the 24/02/2022 (having a last update the 13/12/2021). The rating history comes also from S&P website[2]. The alphabetical rating is converted into notches using the conversion usually used and presented in Appendix A. To have an idea of the countries used, see figure 2.1 :

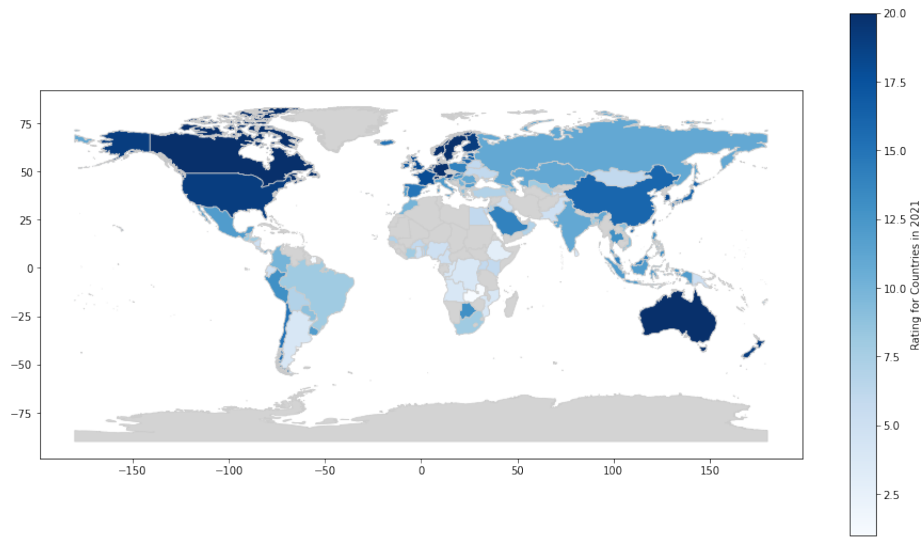


Figure 2.1: Map of the Rating of the Countries in 2021

[1]https://disclosure.spglobal.com/sri/

[2]https://www.spglobal.com/ratings/en/research/articles/210726-sovereign-ratings-history-1202943

Since the willing is to build a parsimonious model, and to avoid overfitting, only 6 macroeconomic variables have been selected, basis on the capacity to rely them with climate change (it will be explained how in the next section). Some statistics from our data are shown[3] in table 2.1. Note that the data contains 872 samples from 128 countries for years between 2015 and 2021. It is important to notice that some countries 'appear' or 'disappear' according to from when they have get rated for the first time or are not rated anymore. it is important that the 2021 data is a forecast one but we consider it as very accurate and it is used just to test our models.

Variable	Mean	Std	Min	25%	50%	75%	Max
Rating (Notch)	11.06	5.13	1.00	7.00	11.00	16.00	20.00
Log(GDP/Capita) (Current USD)	9.17	1.29	6.06	8.26	9.18	10.20	11.77
Real GDP Growth	3.11	3.06	-10.19	1.74	2.96	4.63	25.18
Current Account Balance / GDP	-0.98	8.41	-47.85	-4.05	-1.32	2.61	71.34
Net General Government Debt / GDP	35.24	67.41	-558.87	20.85	39.94	65.26	267.47
Narrow Net External Debt / CARs	58.56	130.90	-682.05	-0.62	50.03	119.58	533.36
General Government Balance / GDP	-3.31	4.62	-28.5	-5.73	-3.10	-0.79	21.80

Table 2.1: Summary Statistic

A global visualisation of the data is available in Appendix B. Something interesting to observe is the correlations between all our data, which is on figure 2.2

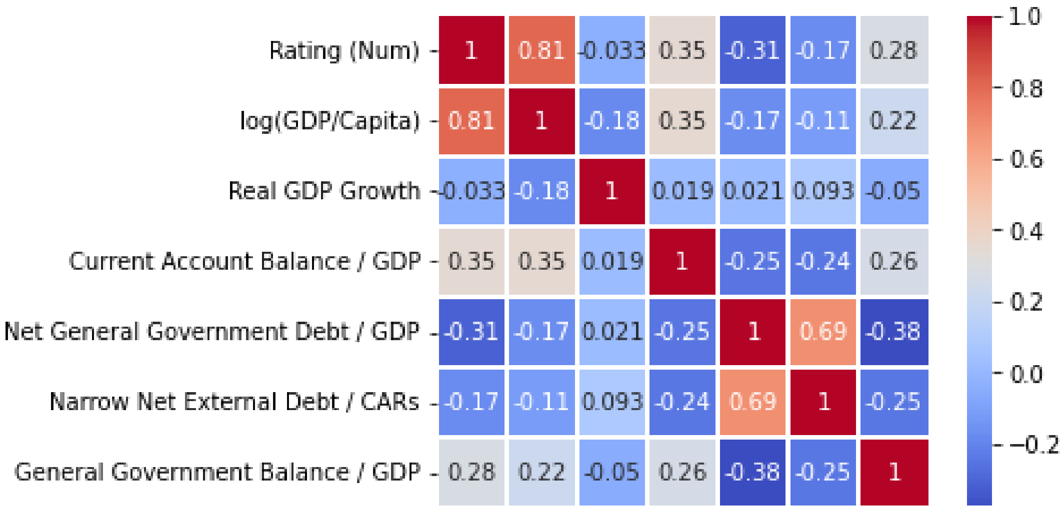
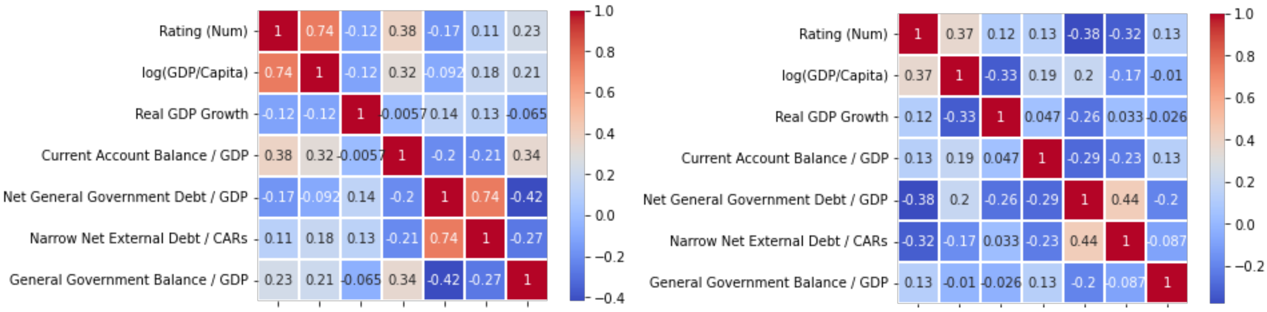


Figure 2.2: Correlation Matrix of the Data

[3]Note that CARs is for Current Account Receipts (CARs)

it is interesting to notice the correlation between log(GDP/Capita) and the Ratings (see figure B.2 available in Appendix B). Otherwise, the variables are not really correlated (except the debt macroeconomic variables). it is also interesting to plot the correlations matrices for investment and for non investment grades on figure 2.3.



(a) (b)

Figure 2.3: Correlations Matrices for Investment (a) and Non Investment Grades (b)

It must be underlined that the correlation between log(GDP/Capita) and Ratings is clearly more important investment grades. Some remarks of this type could be also done if OECD and non OECD countries are separated, the correlations matrices are in Appendix B with also the distribution of the ratings in our data.

2.2. Methodology and Models used

In this section the methodology and the results using the Klusak et al. to build the parsimonious model for the rating (and near models) and the model used to make the predictions are introduced.

2.2.1. Methodology to build the Model

A parsimonious model is built with only the 6 variables presented above. They were selected since they are meaningful and because we will be able according to Klusak et al. to use them to make predictions. Note that we use regression model, since it allows to keep the hierarchical idea of the rating, something that would be lost if a classification model were used. Even if at the end the CRA give a class, since it is easier to understand and use, they use also some regressions to grade a country.

The idea is to build a random forest regression thanks to this 6 variables to find the rating. Some parameters in a random forest are important, like the number of trees used and the maximum depth of the trees used. To select this parameters, we proceed as the following :

1. Creation of the list with containing the value we want to test for one parameter
2. Set N (precision)
3. Err = Empty list
4. for value in list:
   (a) Sum_error = 0
   (b) for i in range(N)
      i. split data in train and test set (with a split size of 20%[4])
      ii. Train the model on the training set
      iii. Test the model using the test set
      iv. Sum_error += Error done on the test set (1-R^2) (R^2: Determination coefficient)
   (c) Err.append(Sum error/N)

With that, a graph like in figure 2.4 is obtained

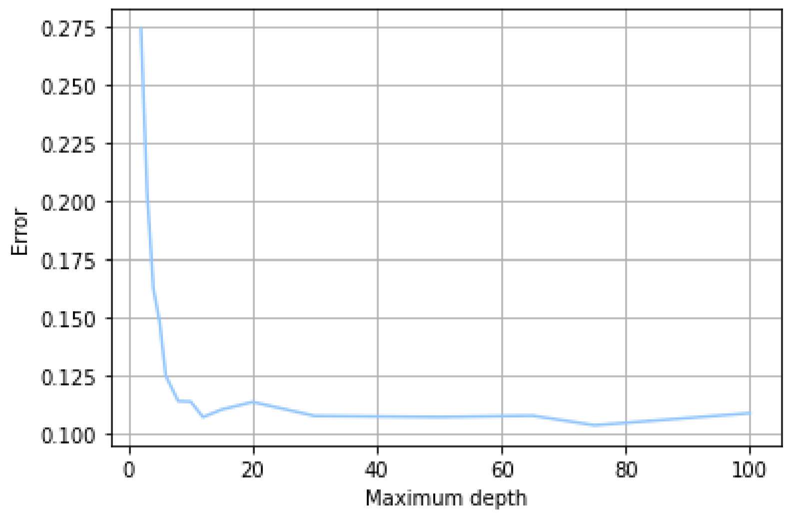


Figure 2.4: Error according to the value of the parameter tested (here maximum depth with random forest)

[4]Usual practice in Machine Learning

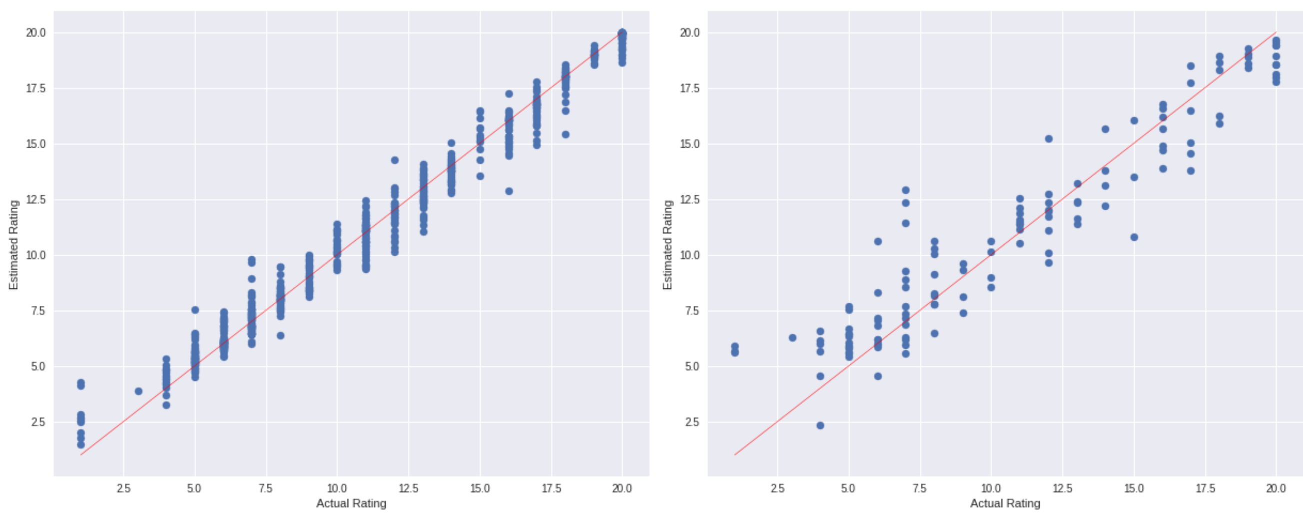
This method allows to avoid overfitting and also to find the better trade off between precision and complexity of the model. For example, in figure 2.4, one can observe that the best trade off is to select a Maximum Depth equal to 10, while here for this case the best parameter is 75.

After having used the random forest, other models have been tested to try to have better result (using AdaBoost for example). The same technique to select the key parameters is kept. Notice that always the mean squared error for the loss function is used, which is a usual practice. The results are presented in the next subsection.

2.2.2. Results

In this section first is detailed the result given thanks to the random forest training on the whole sample, then with test with training on particular test set (2021 set). A table comparing all the different results is presented at the end of the subsection. Also some machine learning algorithms are tested to try having better results and the results are presented at the end of the subsection.

Training on the whole sample (corresponding to the data between 2015 and 2020) we obtain a mean error of just 0.47 notches, and a mean error of 1.31 notches on the test data (2021). The graphs are presented in figure 2.5



(a) (b)

Figure 2.5: Random Forest fitting on the training set (a) and on test set (b)

These results can also be visualised on map, on figure 2.6 for the training.

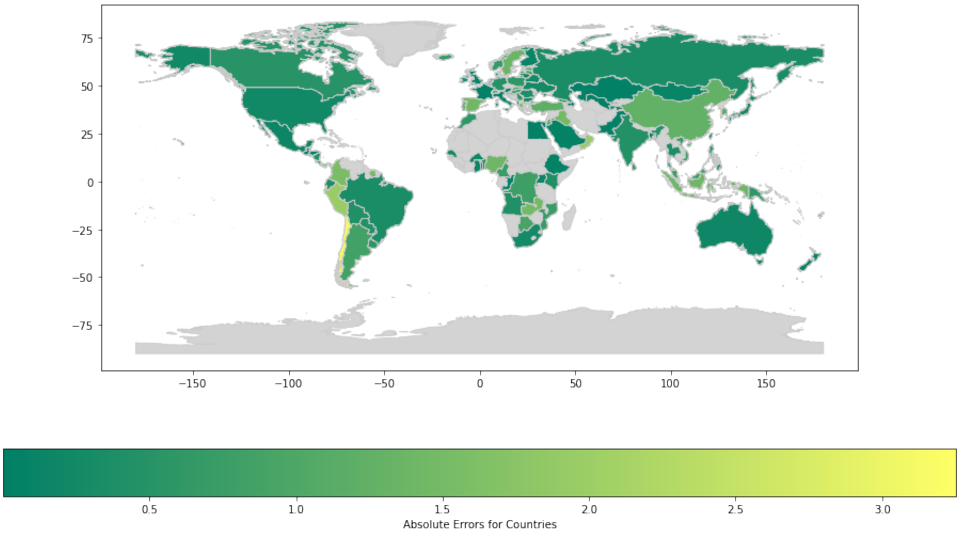


Figure 2.6: Map Error done on the training set

On the figure 2.7 one can visualize the error done for the test set (notice that the scale of color change):

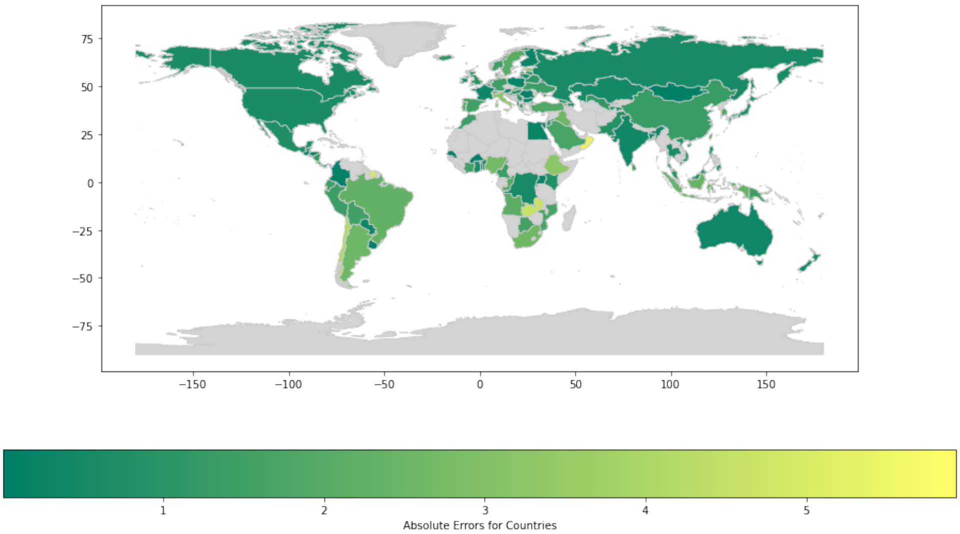
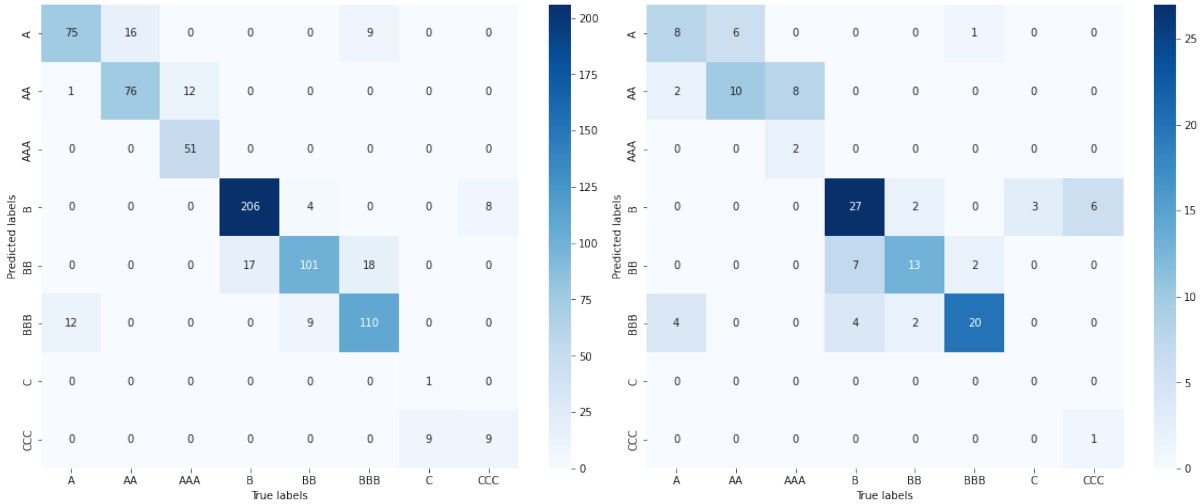


Figure 2.7: Map Error done on the test set

Something interesting is that it is possible after using regression to come back to class (since the correspondence between notch and rating given by CRA). Then, one can visualize the confusion matrices in figure 2.8. (Notice that here we plot the 'global' confusion matrices (just with AAA, AA, A etc), the 'basic' confusion matrices are in Appendix B.)



(a) (b)

Figure 2.8: Random Forest confusion matrix on the training set (a) and on test set (b)

As it could be expected, the confusions are mainly made between near class (like BB and B).

Then the random forest regression have been trained and tested using only particular set (always with data from 2015 to 2020). The results are detailed in the table 2.2. Notice that the results presented correspond to the average result resulting where the model was trained with 80% of the set and test it with the 20% left. The results in the tables correspond to the % predicted (except the mean error). The prediction is considered exact if the error is less than 0.5 notches (since the good rating is found if one convert it)

Set	**Mean Error**	**Exact**	**Within 1**	**Within 2**	**Within 3**
Whole	*1.16*	*36.26*	*58.71*	*82.18*	*92.10*
Investment Grade	0.93	42.29	64.46	87.87	95.97
Non Investment Grade	0.82	44.18	72.49	92.91	97.16
G20 countries	1.07	44.52	61.30	82.0	92.65
No G20 countries	1.13	36.33	59.90	83.06	92.74
OECD countries	0.86	54.24	72.33	88.26	93.48
No OECD countries	1.07	37.19	61.56	85.52	93.78

Table 2.2: Results of Random Forest using different set

From table 2.2 one may observe that training just on particular set could be interesting, in particular for Investment Grade countries and Non investment grade countries. However, it is not possible to use that to make projection since it would mean to make the strong assumption that the past investment grade countries (and past non investment countries) will be the future investment countries (and future non investment countries). A good trade off could be to separate OECD countries from non OECD countries. However, there is no clear improvement for non OECD countries. And the OECD countries are among the richest countries and it is easier for the model to grade them, since for the richest country in practice the economic variables are very important for the rating, while for the others some subjective parameters in real practice could have many importance but are not in the model due to the methodology adopted.

Notice that except for non OECD countries, the log(GDP/Capita) is the most important feature (it is the Narrow Net External Debt/CARs for Non OECD countries). The feature importance are the most 'balanced' for non investment grade countries. Since not a lot of data is available, in the following the whole set is kept and make no separation to avoid loss of information and because when we split the training set, the model is focus only on a partial range of ratings.

In the following the results are compared given by some Machine learning algorithms using also some bagging and/or some boosting technique. The same way as before is used to show the results. The whole sample is used. The results are in table 2.3. In any cases regression problems and regression algorithms are considered.

Algorithm	**Mean Error**	**Exact**	**Within 1**	**Within 2**	**Within 3**
Random Forest	*1.16*	***36.26***	*58.71*	*82.18*	*92.10*
Extra-trees	1.19	32.10	56.89	80.65	93.38
AdaBoost with Decision Tree	1.29	23.76	46.35	83.22	93.67
AdaBoost with Random Forest	**1.10**	32.33	**60.54**	**85.19**	**94.36**
Bagging with Decision Tree	1.15	36.13	58.39	82.15	92.10
Bagging with Random Forest	1.28	30.87	53.94	78.74	90.81
Gradient Boosting	1.26	31.67	53.63	79.66	91.32
Hist Gradient Boosting	1.19	32.46	55.91	81.45	92.68
Light Gradient Boosting	1.18	32.01	56.31	80.78	92.51
XGBoost	1.18	35.68	58.34	81.36	91.97

Table 2.3: Machine Learning Algorithm Comparison

Note that in table 2.3, when AdaBoost or Bagging with random forest is used, it is not with the same parameters than for the first random forest but it is a small random forest (with only 30 decision trees in the forest compared to the 1000 for the first algorithm). From the table 2.3, one can deduce that random forest produce results hard to beat and is a good choice for this particular problem.

2.2.3. Interpretable Models

In the last subsection very accurate results were obtained thanks to some machine learning algorithms, like random forest. However, all the algorithms presented are 'black-box' models, hard to interpret. However, it would be interesting to understand how a country is graded, that is why this part present what was done to obtain interpretable models.

First, basis the subsection 1.3.3 which present some recent researches to make interpretable some tested algorithms before, the codes sharing by Sagi and Rokach, Vidal et al. were tried. However, here we have already 6 features (while in general to test their techniques they use 2 features) with already quite big weak learner inside random forest or XGBoost. Thus, no results were obtained basis on that although some tests. Thus, a come back to simple algorithms and simple ideas is done.

The first idea is simply to use decision tree (regressor). To obtain easily readable and interpretable result is imposed a small depth. With a maximum depth equal to 5, on training an error of 1.35 is obtained (while it was near 0.4 for random forest). However it is possible to visualize the tree and understand how it works, see the example for a maximum depth of 5 and maximum leaf nodes of 20 in Appendix C. Different maximum depths were tested and results are in table 2.4 presented as in the previous tables.

Depth	Mean Error	Exact	Within 1	Within 2	Within 3
Random Forest	*1.16*	*36.26*	*58.71*	*82.18*	*92.10*
5	1.63	24.75	45.23	69.84	83.15
7	1.46	35.50	56.74	76.26	85.99
10	1.40	42.53	64.96	77.87	85.94
20	1.35	44.60	70.36	80.21	87.16

Table 2.4: Decision Trees results

After the maximum depth of 20 there are no more clear improvements. Besides, even if decision tree is able to beat random forest on exact prediction and for prediction within 1 notch, decision tree is always beaten otherwise. Here is met the problem of overadaptation of decision tree. Moreover, more the depth increases, less it is easy to interpret the decision tree and to visualize it. As for random forest, some tests were done just on particular set. Some results are quite interesting (in particular with OECD countries and investment and non investment grade countries), but for the same reasons as before, that will be not used. The results for this particular set are in Appendix C in table C.1.

An idea could have been to select the best tree among all the trees of the random forest according to a criteria (like least error on the training or on the test set). However, the probability is very high that the best tree in the random forest overfits the data.

An other idea is to use a part of the knowledge of the random forest. First, extracting all the threshold and then 'observing' that like in figure 2.9 define the thresholds of the tree, or using some statistical criteria for each feature define the thresholds in the tree.

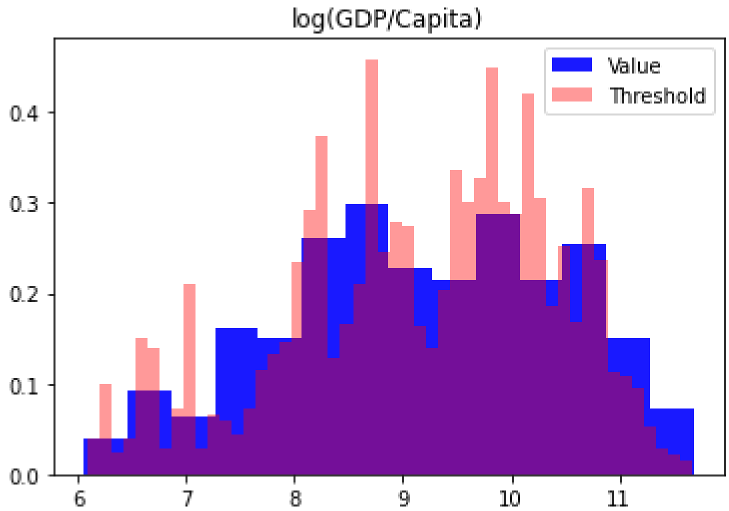


Figure 2.9: Threshold distribution in the random forest and feature distribution

The statistical criteria can be chosen according to the feature importance in the tree, if the feature is important, one may take the value of each decile for the threshold, if less important, each quartile. However, the results show a very little improvement compared to the classical decision tree algorithm used in sklearn.

Then, a simple idea is to make a brute force algorithm using some statistical thresholds we define thanks to the data.

The procedure is:

1. Definition of list of threshold for each feature
2. for each permutation of threshold
    (a) val = ?, list = []
    (b) Browse our data
    (c) If samples in our data falling in the threshold, save the target value in list
    (d) if list non empty, make the mean of the list
    (e) Save result in table

To give a clear idea, in this case we have 6 features and one target (the rating). For each feature is selected some threshold like for example for log(GDP/Capita) can be select [$-\infty$, 8.26, 9.18, 10.2, $+\infty$] (quartile), for Real GDP growth may be selected [$-\infty$, 2.96, $+\infty$] etc. Then we look in the data all the samples such that the log(GDP/Capita) is between $-\infty$ and 8.26, the Real GDP Growth is between $-\infty$ and 2.96, etc and if in the data some samples correspond we save the target and average it. That is why it is a brute force algorithm, all possibilities are tested.

As result a table is obtained, containing lower bonds and upper bonds for each feature and a column with the corresponding rating. With this method, even if it could be thought that they could raise an overfiting issue, better results even on test set compared to the random forest can be obtained. However, the results in table are not very easily interpretable (the tables with 'good' results have from 320 to 520 rows). Besides, to make prediction, if the sample is not in thresholds found before exactly some averages are made (with the features where it corresponds to the thresholds observed before).

A last remark, the idea of last paragraph was used using the thresholds from a simple tree and then apply the brute force algorithm. However, the tables obtained remain very high (and results are similar than before). The advantage of this method is that give a more general way on how select the thresholds. However, this kind of method is possible since the data is not huge.

2.3. Methodology for the prediction

Now the model is trained, it is interesting to know what it could predict across different scenarios. Klusak et al. use the data from Kahn et al. (and then test robustness of their result with Burke et al.).

To obtain the log(GDP/Capita) projection and the Real GDP projection, Klusak et al. use directly Kahn et al.. However, there are not direct projections for the macroeconomic variables. To obtain them, Klusak et al. make 2 steps. Using Kraemer et al., they make a direct link between GDP/Capita loss and the macroeconomic variable, making linear and polynomial fit, that enable to derivate from the GDP/Capita Loss the macroeconomic variable. They use ANOVA test to know whether increasing polynomial order is significant or not. The process is illustrated on figure 2.10 (and other examples in Appendix D).

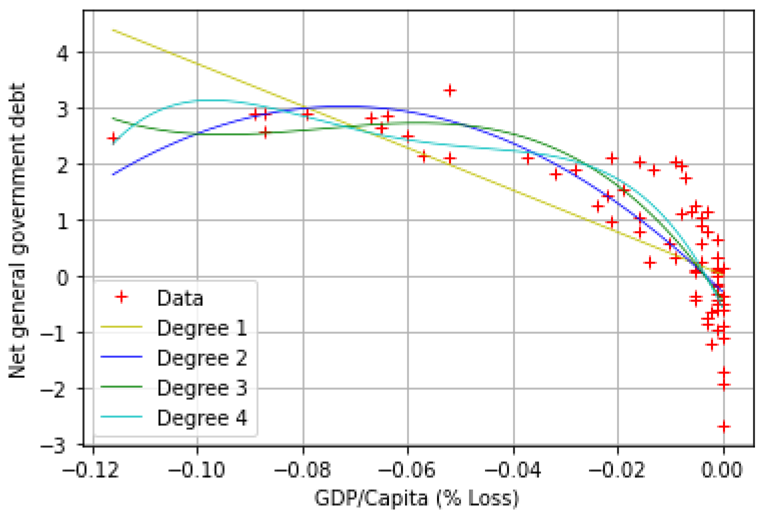


Figure 2.10: Fitting models of the effect of GDP loss on Net General Government Balance

The benefit of this method is to make a direct link between GDP/capita Loss and the macroeconomic variable. With that, just with a GDP Loss it is possible to derive the macroeconomic variables. For the fit, a polynomial 3 is the best trade off between fit and complexity.

However, this approach is weak and tricky and raise some problems. Indeed the GDP loss is capped between 0% and 12%, that could raise problems, and the fitting is not so obvious according to the different figures. Moreover, some problems may be raised using the global approach from Klusak et al.. Therefore a new approach will be tested. All of that is presented in next Chapter, the Chapter 3.

3 | New Models to assess Sovereign Ratings

In this Chapter the approach and model made in Chapter 2 is discussed, then what has been made to face the problems raised is presented.

3.1. Small Variations

In this section first the approach made in Chapter 2 is discussed and some variations and tests are made.

First it is important to underline that we are not able to reproduce the results of Klusak et al. since the Kahn et al. data neither the Burke et al. data are available.

Thus the first idea is to use the SSP to obtain the log(GDP/Capita) and real GDP Growth, combining that with the GDP loss from Kahn et al. to obtain the projection of the macroeconomic variables. To test that the data from [2] is used, where can be obtained the SSP projections with the GDP in current USD and with the population. However, in SSPs some development are expected, and even if damages are added for this particular case, in any cases a global increase of the GDP/Capita is obtained that lead to produce many upgrades. However, that shows how the log(GDP/Capita) is a very important feature. Moreover, this approach is weak but underline some weakness of Klusak et al..

Indeed, Klusak et al. use the log(GDP/Capita) using the GDP in current USD, something which is not commonly used. For example, the SSP use the GDP in constant 2005 USD PPP (Purchasing Power Parity (PPP)). Indeed, this unit makes more sense if the goal is to compare the countries between them and during the time.

That is why the log(GDP/Capita) in current USD is substituted by log(GDP/Capita) in constant 2005 USD PPP, using the data from the World Bank Database. Naturally, the fit of the random forest using this new unit is very similar than using the other unit.

However it gives the opportunity to try to use the projections given thanks to the SSP trough RICE.

Remember that RICE is an IAM (described by Gazzotti [16]). Using RICE, we are able to test the model using only the SSP. Indeed, RICE has inside some damage functions implemented, thus the GDP Loss according to different damage functions can be obtained (with Kahn et al. and Burke et al. damage functions) but others too. However, RICE can have at maximum 57 regions, 44 of them are countries.

However, even using RICE with Kahn damage function, we are not able to find the results of Klusak et al., and not also using different damage functions. The two biggest limit of Klusak et al. paper are met here:

- The model is mainly driven by the log(GDP/Capita)
- The approach to estimate the macroeconomic variable is too weak

The approach of Klusak et al. is then clearly not robust, although they claim the opposite. First because the way to find the projection of macroeconomic data is very limited. They have decided to cap the GDP loss between 0 and 0.12 since the fit is done with a GDP loss of this kind (precision put in Annex). This assumption appears to be very strong, since using RICE with other damage functions it is common to have bigger GDP loss than 12% and more of the double. The initial idea is interesting (make a direct link between the GDP loss and the macroeconomic variable) but not well applicable in this way at least.

To face the first problem, many tests were done like instead of using directly log(GDP/Capita), use the deviation of this parameter to the mean, or rescale this parameter (between 0 and 1 for example). Even if on the training we have good results, this kind of approach is tricky and does not provide results, since or we have many upgrade and many downgrade (with mainly just two grades given by the projection), or an averaging (in the sense that with the projection, all countries obtain the same grade).

Therefore, an other approach is tested to assess the sovereign rating.

Notice that some tests were also done to cancel the parameters with less importance. For example, we can run random forest without the Real GDP Growth or without the General Governance Balance / GDP and keep similar results as before.

3.2. New Approach

In this section a new idea and a new way to provide the sovereign rating thanks to a parsimonious model is presented.

3.2.1. Idea and Methodology

The idea is to build a model which could be already used with the SSP framework. Thus, no macroeconomic variables could be taken as before. However, log(GDP/Capita) and Real GDP growth may be conserved, since are in the SSP framework. log(GDP/Capita) will be conserved for one model but not Real GDP Growth since the little importance of the Real GDP Growth in previous chapter.

Then, the idea is that to grade countries, CRA take into account also some government variables. Klusak et al. do not try to take some of them since they were not able to make predictions about. However, with SSP, some links could be made.

Indeed, thanks to the SSP framework, it could be possible to take into account inequalities as Rao et al. (2019) did or take into account the governance (Andrijevic and Crespo Cuaresma).

Thus, using knowledges from Rao et al. and Andrijevic and Crespo Cuaresma, parameters are selected. This parameters are for sure correlated with the sovereign ratings, since they reflect governance performance, development and inequalities.

The goal then is to find data for which projections are available, from the Wittgenstein Centre, and past Data, from Human Development Report (from United Nations Program for Development). Features select are:

- Life expectancy at birth (one feature for men, one for women)[1]
- Urbanization
- Mean years of schooling (Average number of completed years of education of a country's population aged 25 years and older) (one feature for men, one for women)

This parameters are usually used when we want to measure the development of a country.

The idea then is like Klusak et al. did first, build a random forest and put this data in input, fit the data with the ratings and see what could occur in the future. To find the

[1]Taking separately the data for men and women allow the model to implicitly take into account the inequalities

parameters of the random forest, the same methodology is used (see subsection 2.2.1).

3.2.2. Data

In this subsection the data used to train the new model is presented. In table 3.1, some statistics about the variables are shown. Notice that this time the data training covers the period between 2015 to 2019 for 115 countries (it is very near what we had in Chapter 2).

Variable	Mean	Std	Min	25%	50%	75%	Max
Rating (Notch)	10.99	5.09	1.00	7.00	10.0	15.0	20.0
Log(GDP/Capita) (2005 USD PPP)	9.53	0.98	6.91	8.95	9.63	10.32	11.44
Mean Year Schooling 25+ (Men)	9.89	2.51	2.0	8.20	10.30	12.10	14.60
Mean Year Schooling 25+ (Fem)	9.39	3.03	1.0	7.50	10.10	11.80	13.90
Urbanization (% of pop)	65.64	20.21	13.01	54.62	66.95	81.38	100
Life Expectancy at Birth (Men)	72.35	6.41	52.30	68.70	73.20	77.75	81.90
Life Expectancy at Birth (Fem)	77.51	6.57	54.0	75.35	79.10	82.45	87.70

Table 3.1: New Summary Statistic

Here again it is interesting to observe on figure 3.1 the correlation matrix.

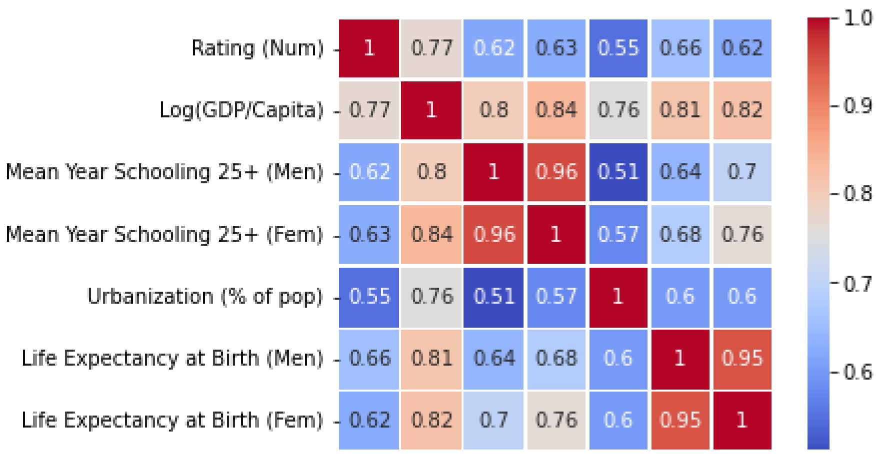


Figure 3.1: Correlation Matrix of the new data

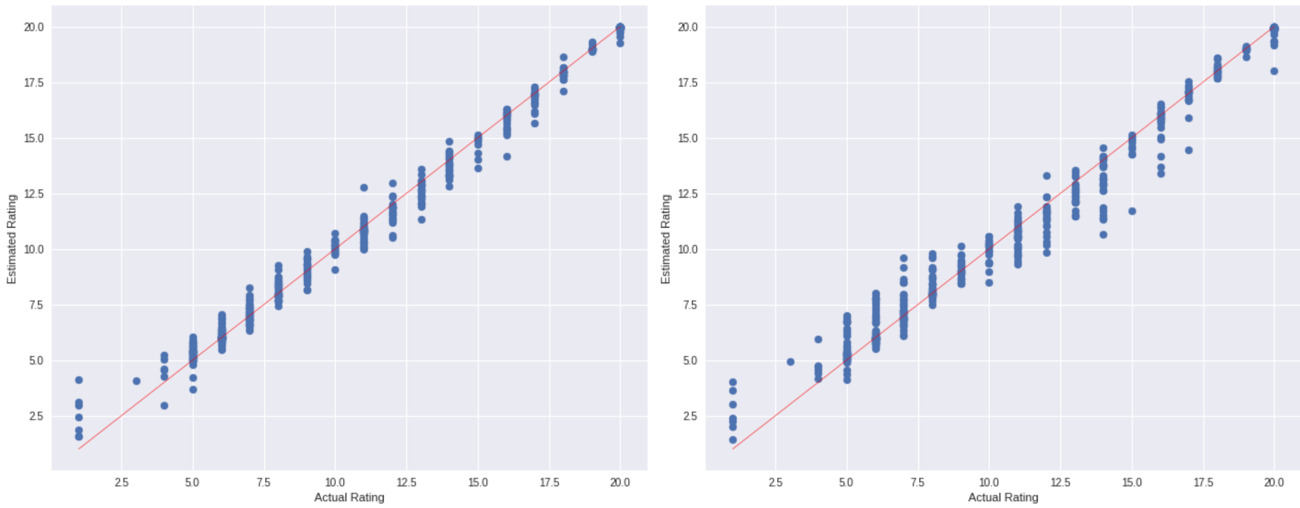
This data is more correlated than before and naturally they are high correlation between the same index feature when is taken one for the women one for the men. However, women and men are taken separately since the gap will be implicitly taken into account in the model and it is interesting to have both indicators.

The global Data visualization is available in Appendix E.

3.2.3. Results

Since the discussion of the section 3.1 about the log(GDP/Capita) for the projection, one model (named Model 1) is trained with the inputs presented in previous subsection and the other model (named Model 2) without the log(GDP/Capita).

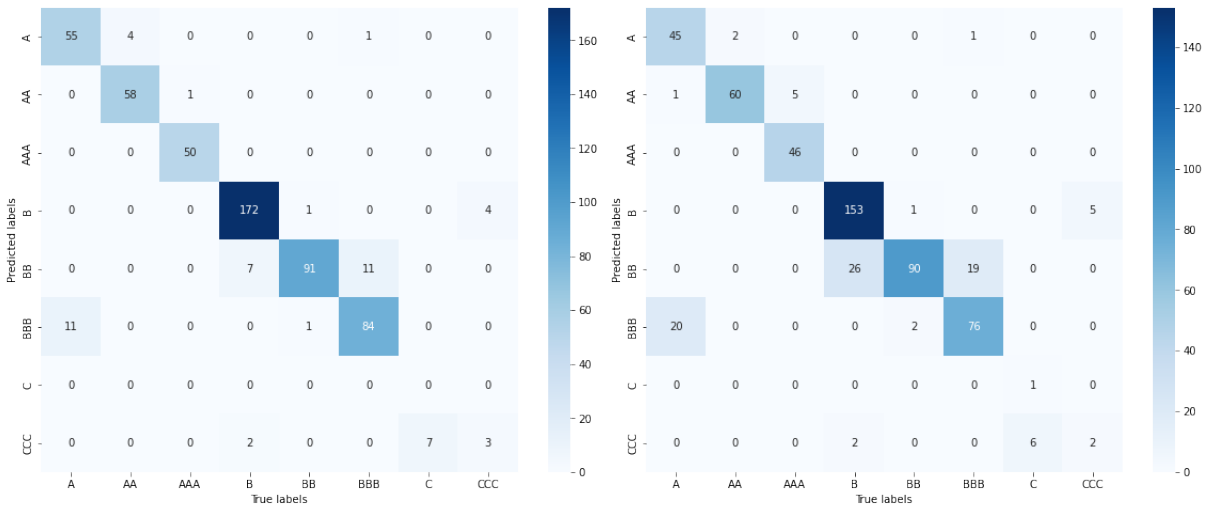
First, in any cases, one can fit well the data as it is shown in figure 3.2.



(a) (b)

Figure 3.2: Random Forest fitting with Model 1 (a) and Model 2 (b)

As done before in subsection 2.2.2, one can observe the 'global' confusion matrix, for our case in figure 3.3 (the 'basic' one is available in Appendix E). Here again we have few confusion and most of them are between B and BB or BB and BBB. (Notice about the figure 3.3 that the data is not sorted as we could expect, on the first column, we have also some confusion with line 6 but it is normal since it is between A and BBB).



(a) (b)

Figure 3.3: Random Forest confusion matrix with Model 1 (a) and Model 2 (b)

Then, the most interesting is to compare the models using the same way than in Chapter 2 (explained in subsection 2.2.2). The results are in table 3.2.

Model	Mean Error	Exact	Within 1	Within 2	Within 3
Previous Model	*1.16*	*36.26*	*58.71*	*82.18*	*92.10*
Model 1	0.87	46.04	70.27	89.35	96.09
Model 2	1.00	42.95	65.30	86.08	94.12

Table 3.2: Models Comparisons

Thus, the new approach is very accurate even more than Klusak et al., even without any direct economic parameters in model 2 (but it is obvious that education and health parameters are correlated with the GDP).

Now these models are projected within SSP1, SSP2 and SSP3 to see what could occur in the future. Notice that here is just tested what happen if the SSP is followed (there are no damages). The results are in table 3.3. No damages are included since only on GDP/Capita it could be possible to consider the damages (whereas the damages from climate change would touch also health and education and influence inequalities).

SSP1	2030	2050	2070	2100
Model 1	6.65	7.92	8.06	6.28
Model 2	2.57	5.36	6.61	6.27
SSP2	**2030**	**2050**	**2070**	**2100**
Model 1	6.51	7.10	7.55	7.65
Model 2	1.71	3.83	5.58	6.30
SSP3	**2030**	**2050**	**2070**	**2100**
Model 1	6.38	6.50	6.51	6.81
Model 2	0.87	1.78	2.97	3.88

Table 3.3: Mean Upgrade compared to end 2021 Rating according to the Scenarios and the models

The results of the predictions for the Model 1 are shown to underline how tricky it is to include log(GDP/Capita) in absolute value. Indeed, with SSP, in any cases and in particular for the less developed countries, the GDP/Capita is expected to increase. However, if some relative notion is introduced, it is difficult to justify how we normalise this parameter. Besides, since it is a key driver to grade a country, relative can condemn some countries to still have low grades. Moreover, if relative and absolute are both included in the model, problematic aspect about both would be met, since the model will privilege one of the two feature during the training.

However, some interesting remarks have to be done regarding the results of table 3.3. First, both model results reflect the storyline of the SSP. Indeed, the upgrades are higher and come quicker for SSP1 than for SSP2 and the same between SSP2 and SSP3. That make sense and respect the meaning of the SSP presented in subsection 1.1.1. However the Model 1 is not so clear about that in particular at the end of the period (for 2070 and 2100).

The model 2 is a clear reflection of the development and the development rate expected for each SSP. it is something that was expected due to the way this model was built.

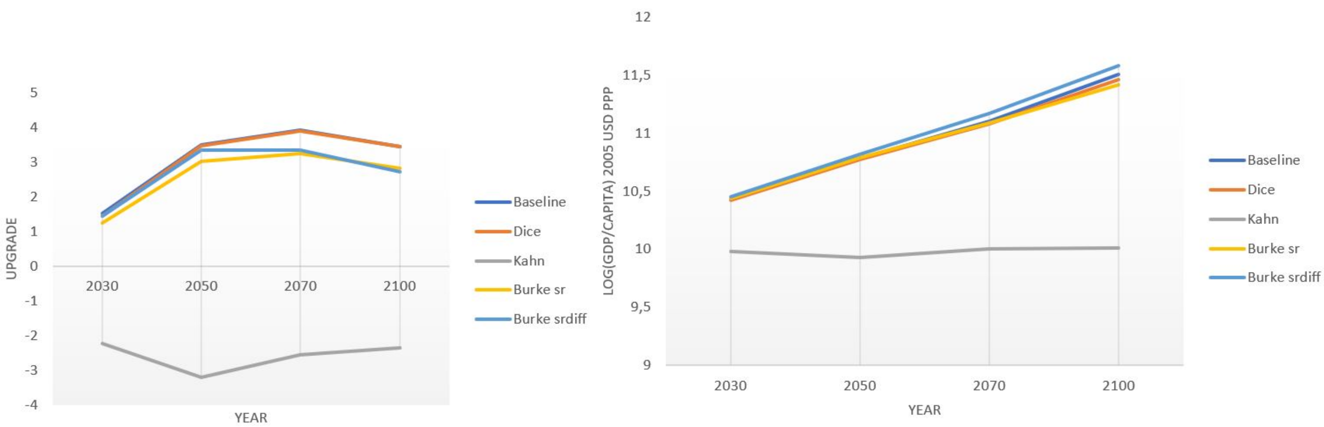
Besides, it is very important to remark that in this way we obtain the opposite result compared to Klusak et al., since we have upgrade in any cases while they obtain downgrade in any cases but as presented before they use different way of doing. Moreover, it should be underlined that no climate damage are considered in this case.

To end we try to add climate damage in Model 1 using RICE (thus we just have 44 countries, notice that almost all are rich). Unfortunately, we can have climate damages only for log(GDP/Capita). Using RICE we are able to use different damages functions. The results of the projections using RICE in non cooperative business as usual impact mode are in table 3.4.

SSP2	2030	2050	2070	2100
NO Damage	1.51	3.49	3.91	3.45
Dice	1.43	3.46	3.88	3.45
Kahn	-2.23	-3.20	-2.57	-2.37
Burke sr	1.25	3.02	3.25	2.81
Burke srdiff	1.45	3.33	3.33	2.72

Table 3.4: Mean Upgrade compared to end 2021 Rating according to the damage function used

It must be emphasized that only with Kahn we obtain downgrades. Thus this result underlines the importance of the damage function used, since that would impact a lot the log(GDP/Capita) and then the rating. Besides, this result is quite coherent with the idea of the results in Klusak et al., however, it underlines that the Kahn framework is maybe too pessimist. Notice that the result is presented only for SSP2 and for a specific run of RICE. It could be noticed that comparing Burke with no damage, we obtain around 0.7 less upgrade.

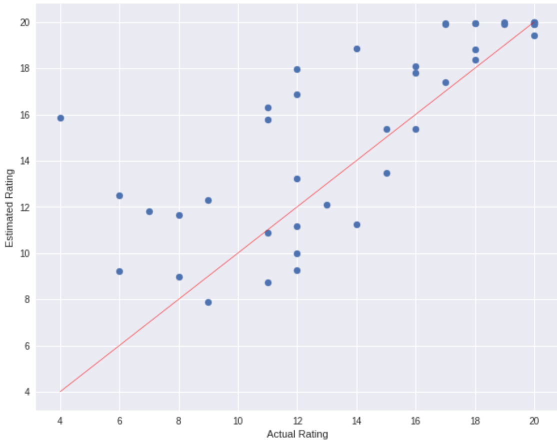


(a) (b)

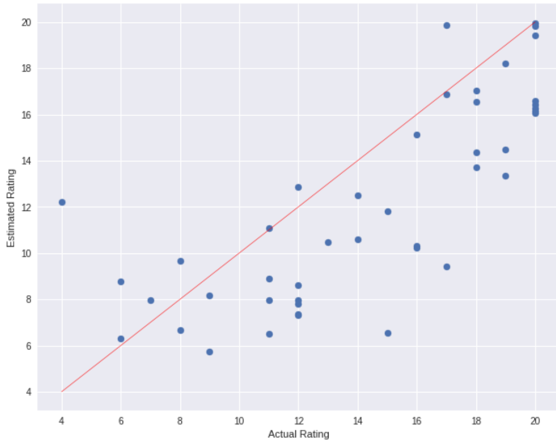
Figure 3.4: Upgrade compared to 2021 (a) and log(GDP/Capita) (b) with different damages functions

On figure 3.4 we may observe the link between the log(GDP/Capita) and the upgrades (or downgrades). Low GDP/Capita means downgrade while higher GDP/Capita means upgrade even if at one point we attain a maximum of upgrade (due to the fact a country can obtain 20 at maximum and that there are other parameters in the model that play a little role).

It is interesting to look at the results more deeply, looking at what happen in 2030 and 2100 with the Kahn damage function and an other one (Burke srdiff for example). The results for 2030 are in figure 3.5, with the corresponding map in figure 3.6 and 3.7 and for 2100 in figure 3.8, with the corresponding map in figure 3.9 and 3.10.



(a)



(b)

Figure 3.5: Projection in 2030 in non coop mode for BAU-impact scenario with Burke (srdiff) damage (a) and Kahn damage (b)

Notice that on figure 3.5 the abscissa corresponds to the rating in 2021 and the ordered to the estimation for 2030. Thus, if we are above the red line, we project an upgrade, otherwise, we project a downgrade.

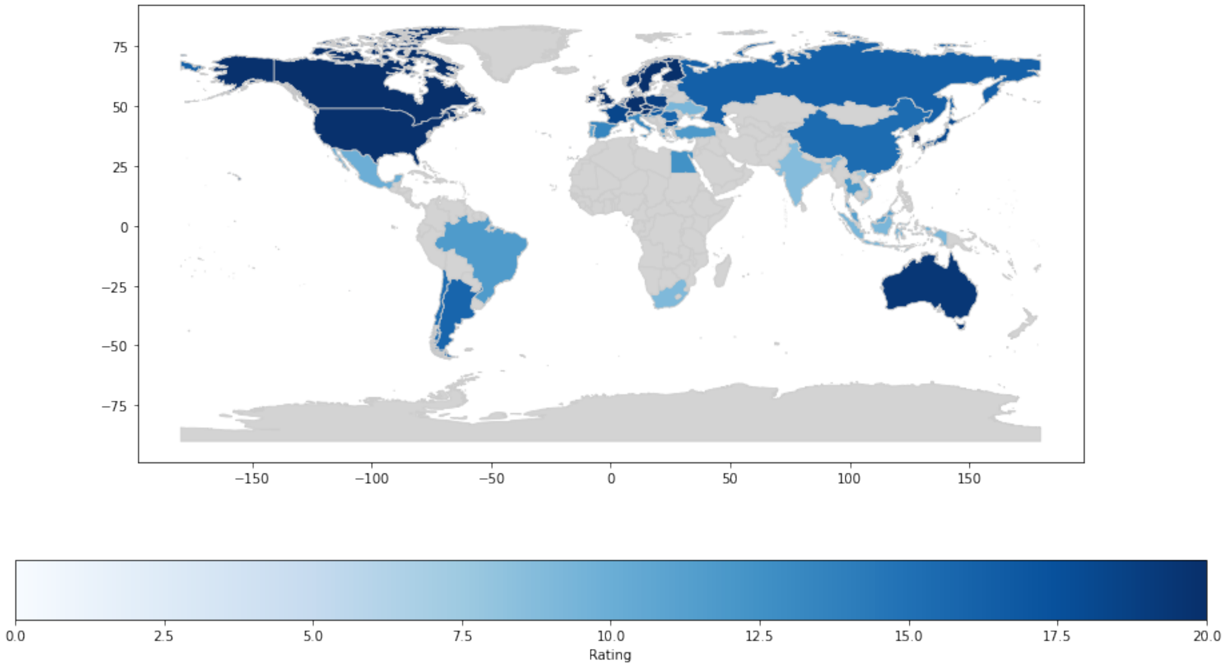


Figure 3.6: Map of Ratings in 2030 in non coop mode for BAU-impact scenario with Burke (srdiff) damage

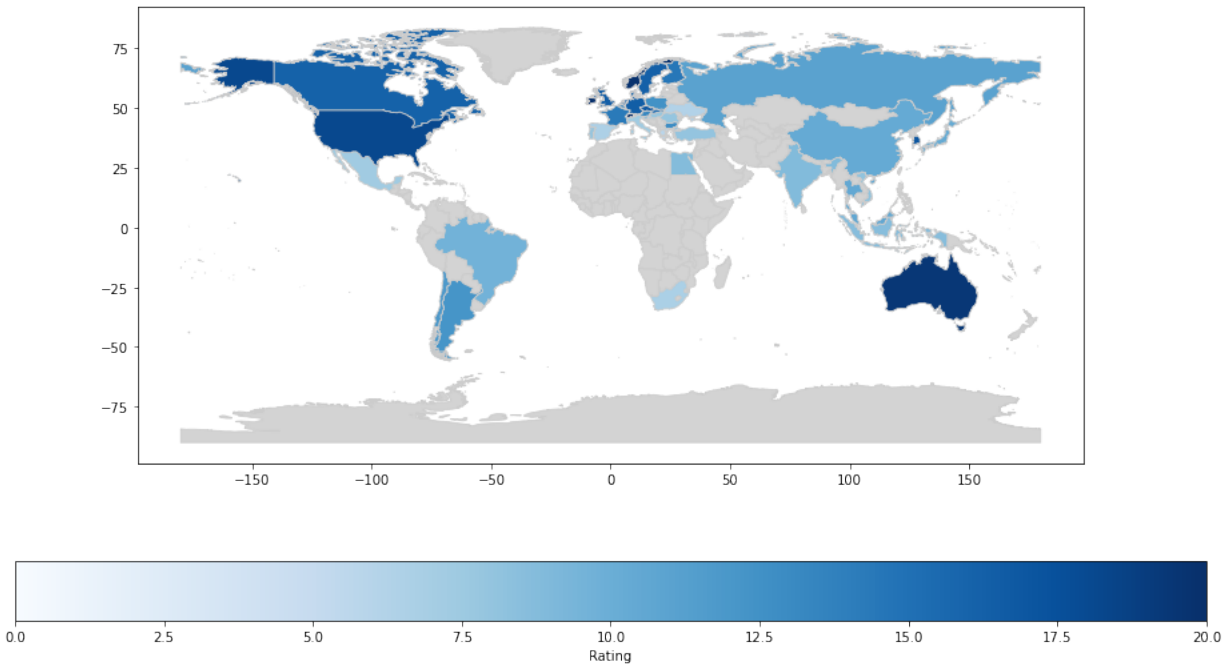
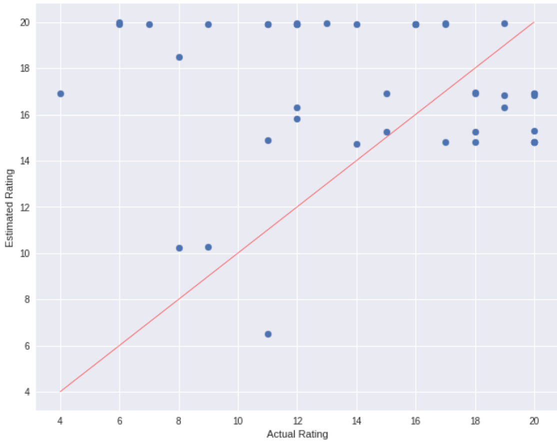
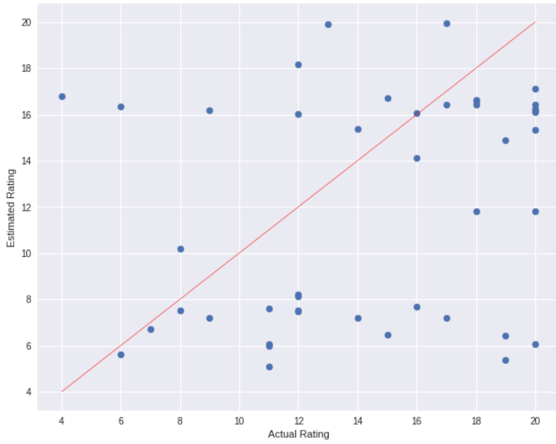


Figure 3.7: Map of Ratings in 2030 in non coop mode for BAU-impact scenario with Kahn damage



(a)



(b)

Figure 3.8: Projection in 2100 in non coop mode for BAU-impact scenario with Burke (srdiff) damage (a) and Kahn damage (b)

Notice that on figure 3.8 the abscissa corresponds to the rating in 2021 and the ordered to the estimation for 2100. Thus, if we are above the red line, we project an upgrade, otherwise, we project a downgrade. It is important to notice that it is quite normal that the actual 20 grade countries could have little downgrade. Indeed, since the model has learnt to grade until 20, and it is mainly driven by the log(GDP/Capita), it is easy for the richest countries to obtain a lower grade. (However, notice that a downgrade from AAA to AA+ has less impact than a downgrade from BB- to B+ in general).

These results show how the choice of the damage function is crucial, since that could change the analysis. Burke and Kahn are compared for two reasons, first because it is what Klusak et al. did to assess that their results are robust and also because Burke is near the other damages tested.

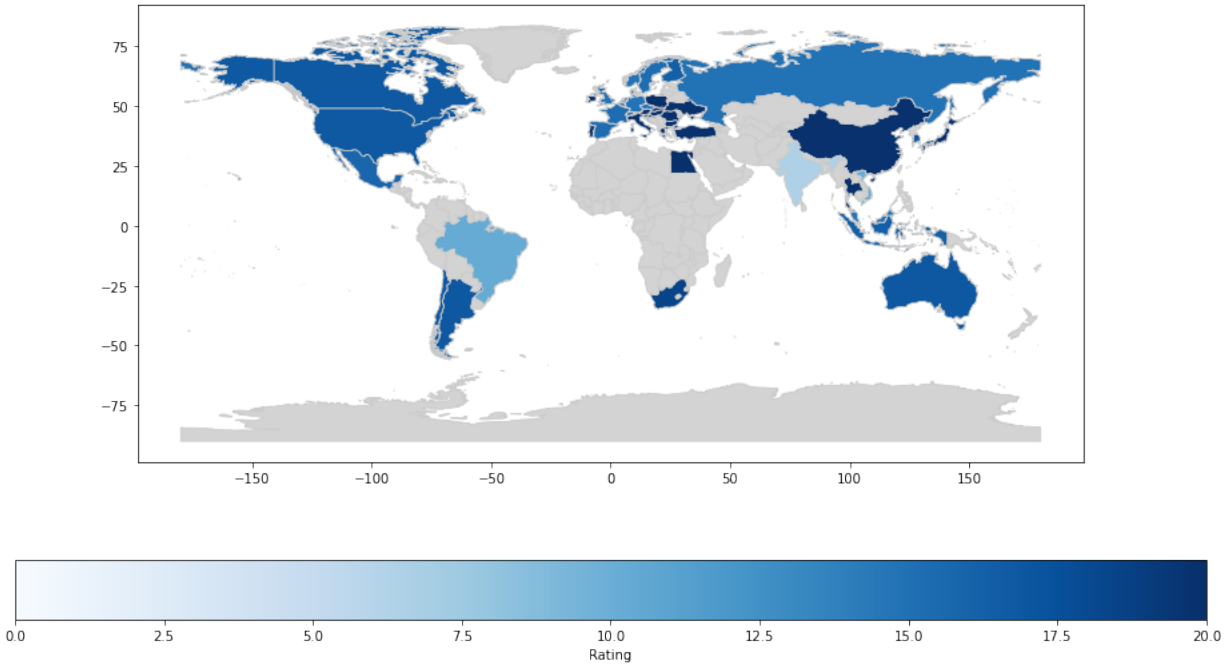


Figure 3.9: Map of Ratings in 2100 in non coop mode for BAU-impact scenario with Burke (srdiff) damage

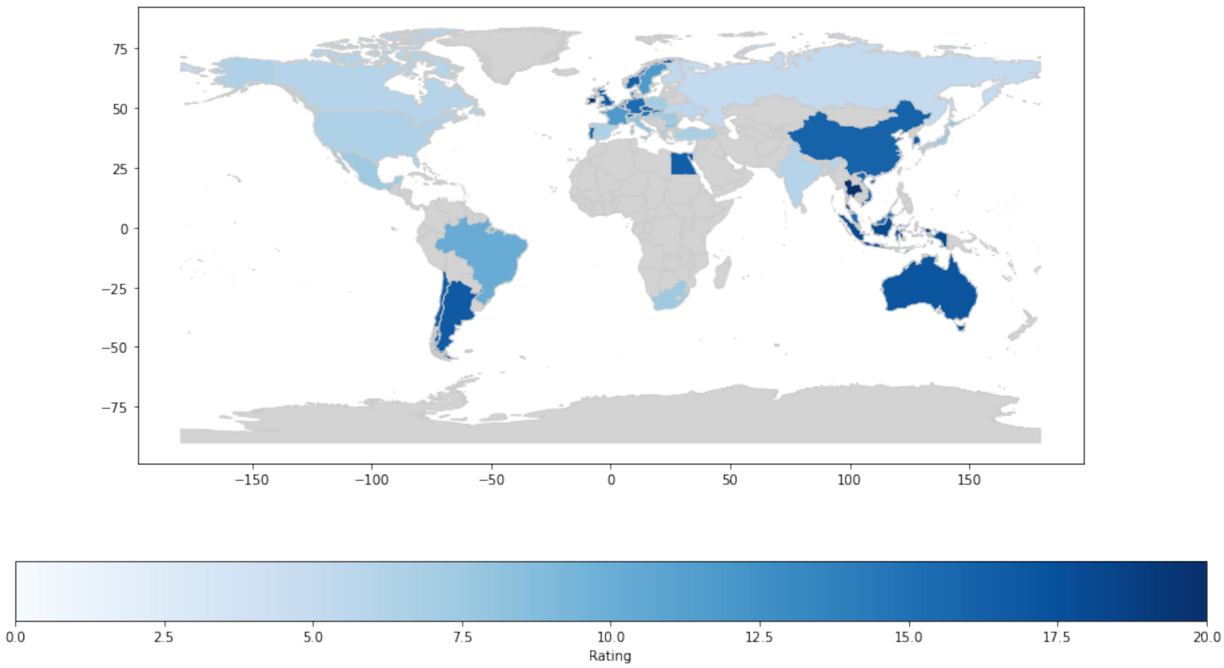


Figure 3.10: Map of Ratings in 2100 in non coop mode for BAU-impact scenario with Kahn damage

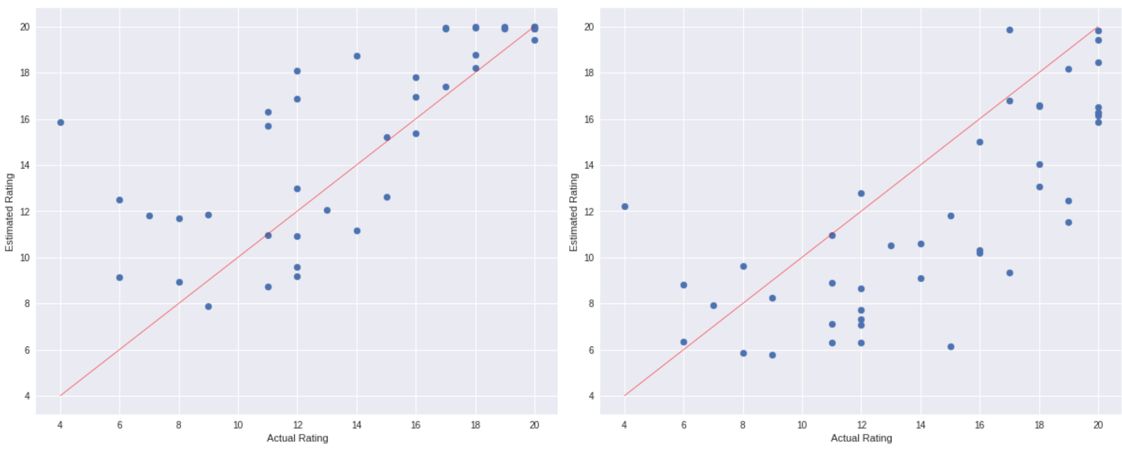
If we use RICE with the same parameters than before except that now we are in cooperative mode, we obtain results that are in table 3.5.

SSP2	2030	2050	2070	2100
NO Damage	1.51	3.49	3.91	3.45
Kahn	-2.57	-3.44	-0.40	2.07
Burke srdiff	1.36	3.18	3.35	2.99

Table 3.5: Mean Upgrade compared to end 2021 Rating according to the damage function used - Cooperative mode

The results from table 3.5 are very interesting in the sense that we can observe that in cooperative mode in any cases (for what was tested) we obtain better results at long term and just upgrades in 2100. Notice that in any cases, we have on average less upgrade compared to the baseline (No damages).

As before, it is interesting to look at the results more deeply, looking at what happen in 2030 and 2100 with the Kahn damage function and an other one (Burke srdiff for example). The results for 2030 are in figure 3.11, with the corresponding map in figure 3.12 and 3.13 and for 2100 in figure 3.14, with the corresponding map in figure 3.15 and 3.16.



(a) (b)

Figure 3.11: Projection in 2030 in coop mode for BAU-impact scenario with Burke (srdiff) damage (a) and Kahn damage (b)

Notice that on figure 3.11 the abscissa corresponds to the rating in 2021 and the ordered to the estimation for 2030. Thus, if we are above the red line, we project an upgrade, otherwise, we project a downgrade.

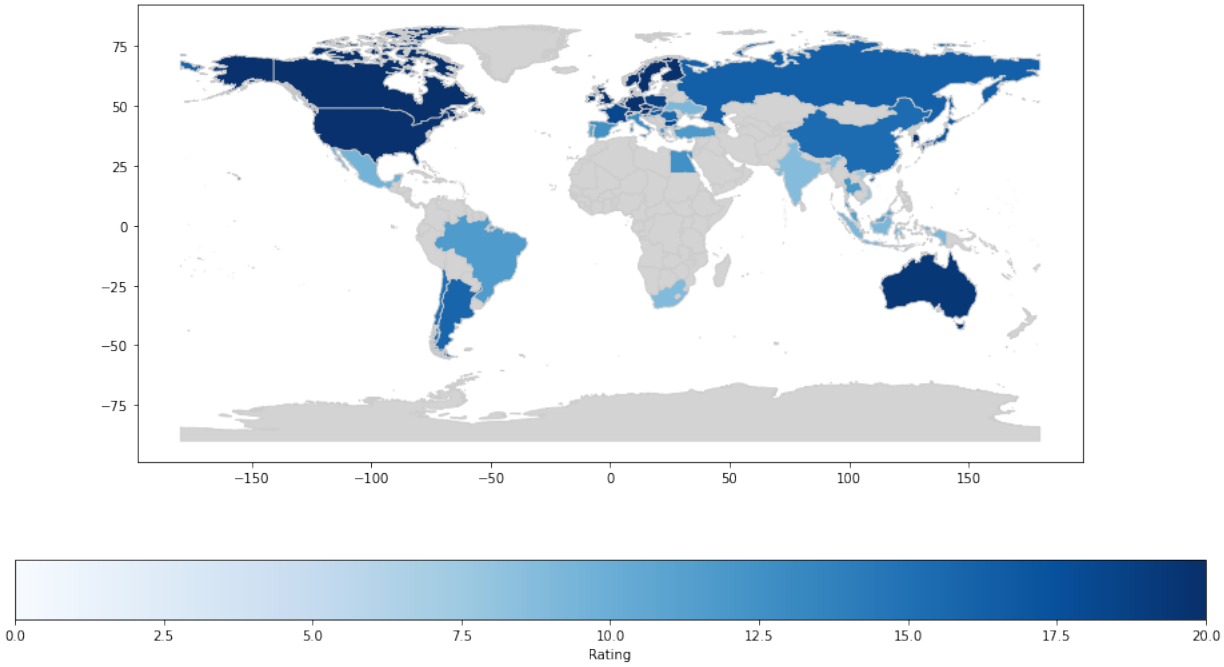


Figure 3.12: Map of Ratings in 2030 in coop mode for BAU-impact scenario with Burke (srdiff) damage

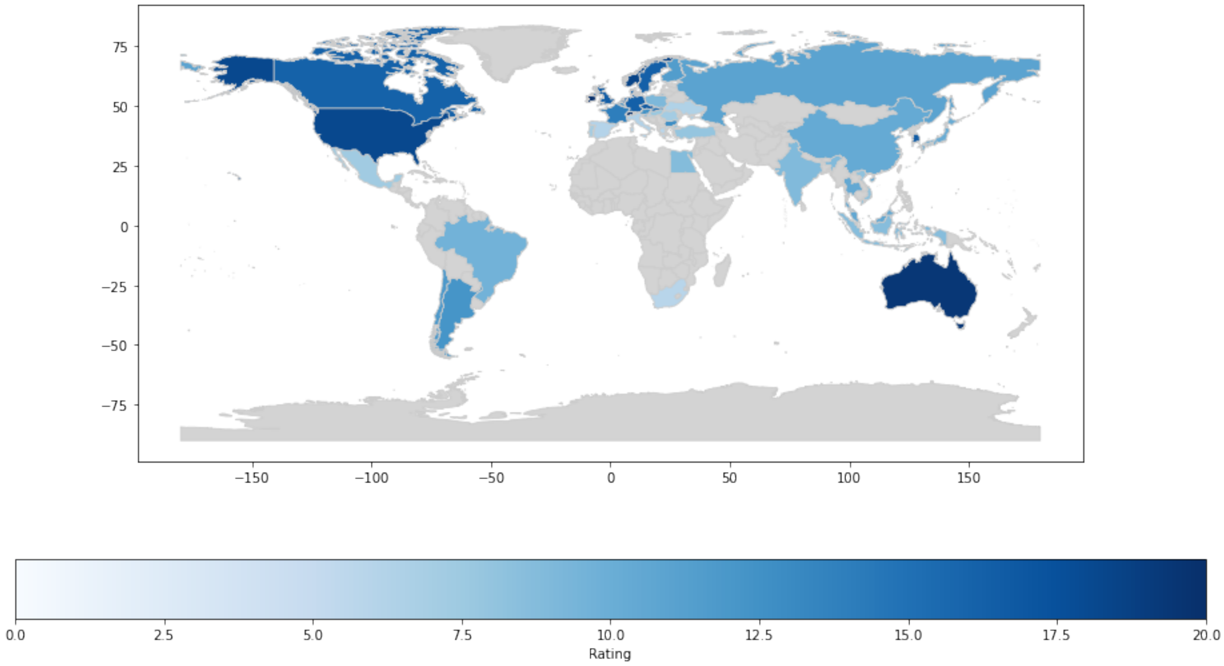
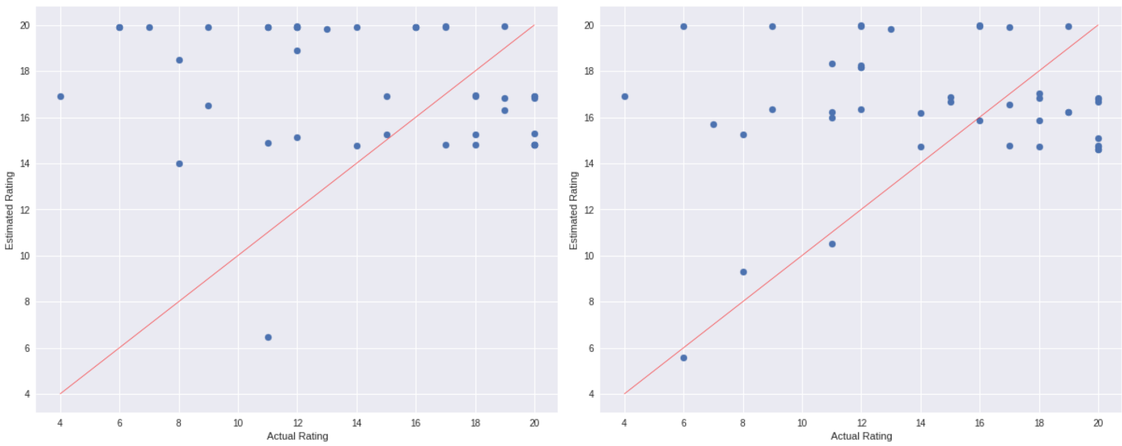


Figure 3.13: Map of Ratings in 2030 in coop mode for BAU-impact scenario with Kahn damage



(a) (b)

Figure 3.14: Projection in 2100 in coop mode for BAU-impact scenario with Burke (srdiff) damage (a) and Kahn damage (b)

Notice that on figure 3.14 the abscissa corresponds to the rating in 2021 and the ordered to the estimation for 2100. Thus, if we are above the red line, we project an upgrade, otherwise, we project a downgrade.

This results show also the importance of cooperation if we want to have more upgrade meaning a more developed world. They also indicate that to fight climate change, it is clearly better to cooperate. Indeed, without cooperation, the Temprature of the atmosphere increase of around 2°C in 2050 and 3°C in 2100 (with Kahn or Burke) when it is about 1.78°C in 2050 and 1.70°C in 2100 with cooperation.

To end this chapter it is important to notice that it was shown that some parsimonious model to grade a county can be made, even without macroeconomic parameters, since at the end the grade of a country reflects also his development. This could raise some questions about the construction of parsimonious models to grade countries, that will be discussed in next Chapter. An other important result is that fighting climate change is not bad for the economy. More precisely, fighting climate change leads also to upgrade and good results (for sovereign ratings) but it is not particularly better at the end (it is particularly better just for Kahn).

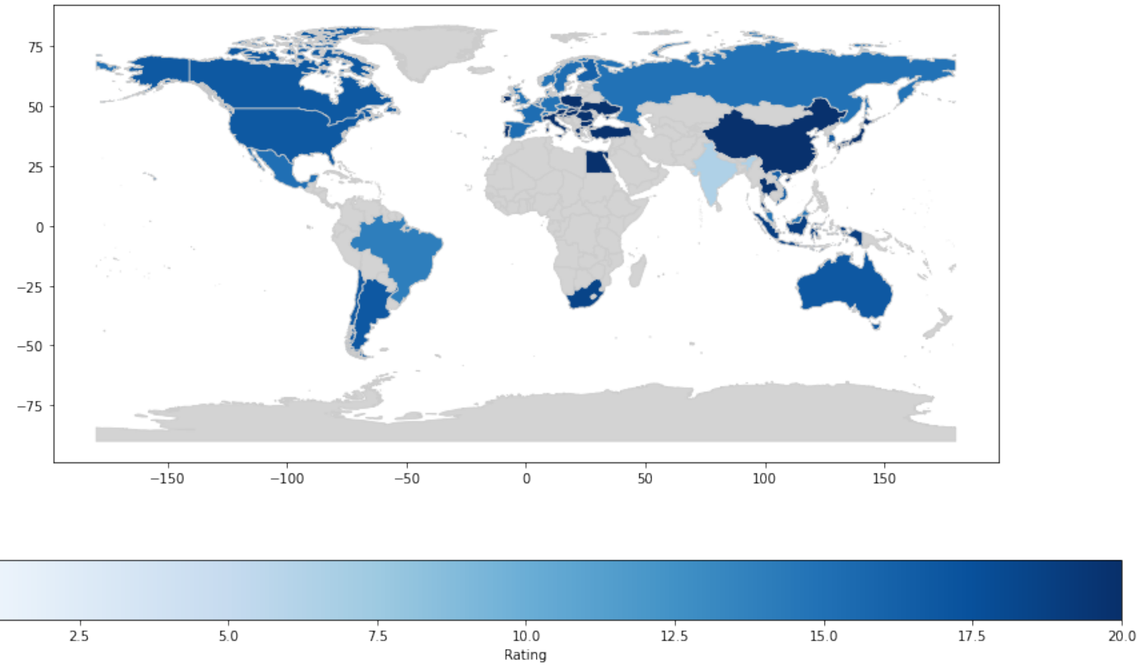


Figure 3.15: Map of Ratings in 2100 in coop mode for BAU-impact scenario with Burke (srdiff) damage

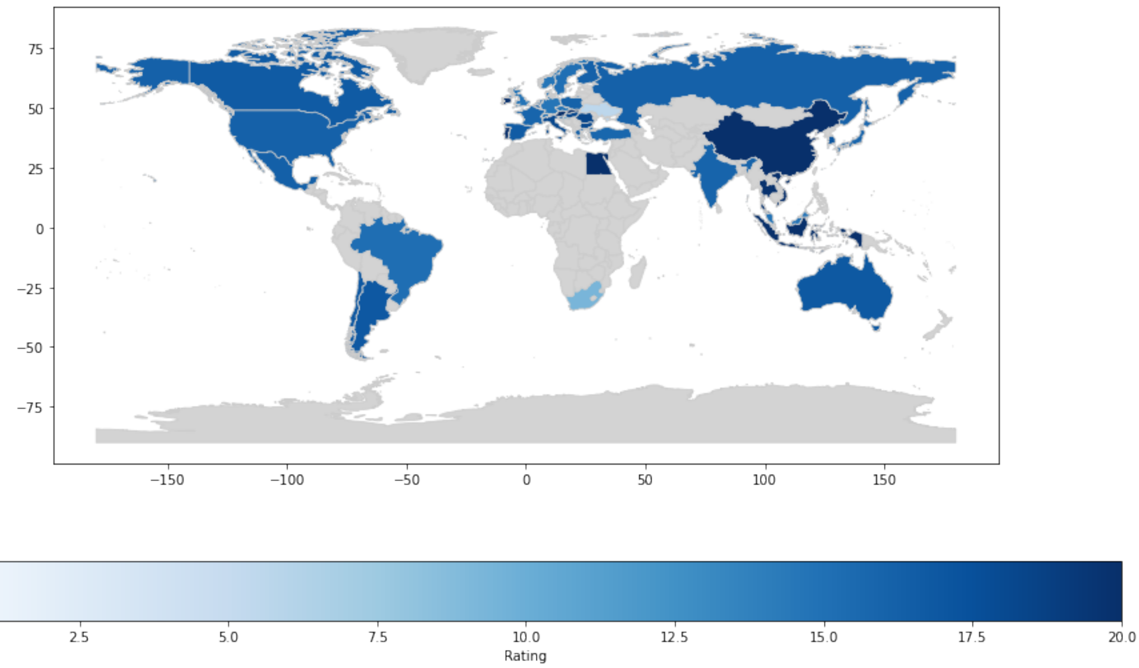


Figure 3.16: Map of Ratings in 2100 in coop mode for BAU-impact scenario with Kahn damage

4 | Conclusions and future developments

This Chapter draws a conclusion to the thesis. First, a General Conclusion is drawn, second a discussion about the work is presented and third ideas for future developments are made.

4.1. General Conclusion

This research contributes to the understanding of Sovereign Ratings and the link that could be made between climate science and climate finance. Taking the Klusak et al. work, a parsimonious model using macroeconomic variables was created to build again the way of grading a country. The procedure to build the rating is simple, there are 6 variables, a machine learning algorithm is used, and the rating is obtained. Some way were explored to have a better machine learning tool (using some Boosting algorithms for examples instead of the random forest) and others to obtain an interpretable machine learning tool (which still accurate).

One of the main result is that Klusak et al. paper is clearly not robust, although they claim the opposite. Two mains weakness were underlined, first, the way to project the macroeconomic variables according to a GDP/Capita loss and second the use of the log(GDP/Capita) in current USD. To face the weaknesses and to add some understanding to the meaning of sovereign rating a new approach was tested.

The innovative part is thus building a new model to assess sovereign ratings with just governance performance indicator. Keeping the idea of Klusak et al. of building a parsimonious model with a machine learning tool, the inputs were changed to take more into account governance action (taking health and education parameters). This leads to a new point of view about the sovereign ratings, which could even be seen like a development parameter non directly correlated with the debt. The inputs were also chosen to be able to use SSP projections.

Besides, the last results show clearly that fight the climate change make sens also from an economic point of view. Also, it shows that fighting climate change leads also to upgrade and good results (for sovereign ratings) but it is not particularly better at the end than if nothing is done (it is particularly better just for Kahn).

Last this study shows eventually the difficulty to chose parameters to build a model to assess the sovereign rating and even at the end shows the limitations and weakness of this approach, that are more detailed in next section.

4.2. Discussion

It must be underlined that implicitly, the approach uses an induction assumption since supervised machine learning is used. That means that the way a country is graded today (and in the past) will be the same in the future. It could be seen as a reasonable assumption. However, that could raise problem with absolute parameters (like log(GDP/Capita), and also education and health parameters in the new approach).

Indeed, in particular with the log(GDP/Capita), first we must be insure that make sense to compare its value across time. Whatever could be the unit, this question is crucial to make the projection. Moreover, the countries are also graded compared to the others, and since it is a key parameter, it could be also interesting to rescale it. However, because it is a key parameter, that could raise problems. For example, is the country with the lowest log(GDP/Capita) condemned to obtain one of the lowest grades ? For the opposite, the assumption that always AAA grade on the market is available could be reasonable, (since it could be seen as the safest investment compared to the others), but what about the mean grading ? Should be of BB, BBB, A ? Note that taking two parameters for log(GDP/Capita), one with absolute value, the other with a relative value, could not be a solution, since the problems would remain.

The approach based on socioeconomic SSP projections gives the opportunity to change the point of view about the sense of the grading, but it takes some absolute parameters and making them relative could be tricky (as for log(GDP/Capita)). Then, one of the weaknesses of this kind of approach is that if one wants to try making projection, one has to be sure to compare exactly the same thing across the time.

Therefore, two remarks have to be made. First, as did Klusak et al., only data since 2015 is used. Second, and more important, with this kind of models, it is not possible to learn how to change internal value across time. That means that it is not possible to catch evolution if there are evolutions across time in the parameters to grade a country. For example, if for each decision tree in the random forest the first split correspond to whether log(GDP/Capita) is more or less than around 9, the model can not learn (even adding data) that some years ago this value could have been more similar to 8. This argument could be also true for relative parameters (like Narrow Net External Debt / GDP). This is linked naturally to the induction assumption and implicitly a stationary assumption was made (more precisely non stationary do not play a role in our case).

Besides, the new approach is interesting to underline an other way to interpret the grade of a country. It is also its weakness, since that does not rely on any macroeconomic parameters, whereas it should be an indicator for the debt. One can wonder how his pertinent to assess the sovereign rating without parameters linked to the debt, and lead to an other weakness, the idea to have a very parsimonious model to make prediction to grade countries. Indeed, according to the choices of the parameters, it is possible to obtain opposite conclusions if we make projections.

A last remark, the new approach underlines with the last results that Kahn seems to be a very pessimist framework, that could explain the resuts from Klusak et al. and underlines also how importance is the macroeconomic framework considered in this kind of approach. Besides, the last results shows that Kahn is not consistent with other approach like Burke or Dice.

4.3. Future Developments

Even if the model of Klusak et al. has some weaknesses, it is still interesting. Some tests were done with RICE and if there was the possibility to include, alongside SSP projections, some macroeconomic variables linked to the debt, first their results may be found and second it would be immediately known whether their model is robust or not. Indeed, many damage functions are included in RICE thus some studies could compare the results of the prediction according to the damage function.

Besides, some work has to be made to know whether it is pertinent the way they include the log(GDP/Capita) or how we include it. More generally, for each absolute parameters (even for relative), some works could be made whether it is consistent or not to keep the same criteria across time, and if not, should we add a kind of discount factor to the future data or should we scale the data and how.

Moreover, it could be interesting to include more data, since it should improve our model (at least the robustness). It could also be interesting to include other new inputs or to combine more the economic and non economic inputs.

Eventually, remaining update with machine learning research it could be possible in the future to build more accurate and easier to understand machine learning model.

About the new approach, it should be interesting to observe also the damage effect due to climate change on all the criteria. Thus, one can use the SSP projection with some damage on the criteria chosen. To know whether it is credible or not, it could be compared to the work that could be done with SSP with the Klusak et al. model.

To conclude, some future developments can be made : first, select well the criteria in input and/or add a factor to the input for the prediction, second, include in RICE the two approaches and look whether they lead to the same conclusion or not and whether they could be robust, third, stay connected to the research in machine learning to obtain more accurate interpretable results. Moreover, the training data could be extended.

Bibliography

[1] AR6 Climate Change 2022: Impacts, Adaptation and Vulnerability — IPCC, . URL https://www.ipcc.ch/report/sixth-assessment-report-working-group-ii/.

[2] CEPII - Documents de travail - Immigration, unemployment and GDP in the host country: Bootstrap panel Granger causality analysis on OECD countries, . URL http://www.cepii.fr/CEPII/fr/publications/wp/abstract.asp?NoDoc=3938.

[3] Environmental, Social, & Governance (ESG) Criteria Definition, . URL https://www.investopedia.com/terms/e/environmental-social-and-governance-esg-criteria.asp.

[4] S. Acevedo, M. Mrkaic, N. Novta, E. Pugacheva, and P. Topalova. The Effects of Weather Shocks on Economic Activity: What are the Channels of Impact? *Journal of Macroeconomics*, 65:103207, Sept. 2020. ISSN 0164-0704. doi: 10.1016/j.jmacro.2020.103207. URL https://www.sciencedirect.com/science/article/pii/S0164070420301336.

[5] A. Afonso, D. Furceri, and P. Gomes. Sovereign credit ratings and financial markets linkages: Application to European data. *Journal of International Money and Finance*, 31(3):606–638, Apr. 2012. ISSN 0261-5606. doi: 10.1016/j.jimonfin.2012.01.016. URL https://www.sciencedirect.com/science/article/pii/S0261560612000320.

[6] M. Andrijevic and J. Crespo Cuaresma. Governance in socioeconomic pathways and its role for future adaptive capacity. URL https://core.ac.uk/display/286187477?utm_source=pdf&utm_medium=banner&utm_campaign=pdf-decoration-v1.

[7] A. Andrzejak, F. Langner, and S. Zabala. Interpretable models from distributed data via merging of decision trees. In *2013 IEEE Symposium on Computational Intelligence and Data Mining (CIDM)*, pages 1–9, Apr. 2013. doi: 10.1109/CIDM.2013.6597210.

[8] E. C. Bank. Climate Change and Central Banks: Analysing, Advising and Acting.

July 2021. URL https://www.ecb.europa.eu/press/key/date/2021/html/ecb.sp210711~ffe35034d0.en.html.

[9] C. Bishop. Pattern Recognition and Machine Learning. Jan. 2006. URL https://www.microsoft.com/en-us/research/publication/pattern-recognition-machine-learning/.

[10] M. Burke, S. M. Hsiang, and E. Miguel. Global non-linear effect of temperature on economic production. *Nature*, 527(7577):235–239, Nov. 2015. ISSN 1476-4687. doi: 10.1038/nature15725. URL https://www.nature.com/articles/nature15725. Number: 7577 Publisher: Nature Publishing Group.

[11] M. Dell, B. F. Jones, and B. A. Olken. Temperature Shocks and Economic Growth: Evidence from the Last Half Century. *American Economic Journal: Macroeconomics*, 4(3):66–95, July 2012. ISSN 1945-7707. doi: 10.1257/mac.4.3.66. URL https://www.aeaweb.org/articles?id=10.1257/mac.4.3.66.

[12] M. Dell, B. F. Jones, and B. A. Olken. What Do We Learn from the Weather? The New Climate-Economy Literature. SSRN Scholarly Paper 2345072, Social Science Research Network, Rochester, NY, Oct. 2013. URL https://papers.ssrn.com/abstract=2345072.

[13] S. Dietz, A. Bowen, C. Dixon, and P. Gradwell. 'Climate value at risk' of global financial assets. *Nature Climate Change*, 6(7):676–679, July 2016. ISSN 1758-6798. doi: 10.1038/nclimate2972. URL https://www.nature.com/articles/nclimate2972/. Number: 7 Publisher: Nature Publishing Group.

[14] R. R. Fernández, I. Martín de Diego, V. Aceña, A. Fernández-Isabel, and J. M. Moguerza. Random forest explainability using counterfactual sets. *Information Fusion*, 63:196–207, Nov. 2020. ISSN 1566-2535. doi: 10.1016/j.inffus.2020.07.001. URL https://www.sciencedirect.com/science/article/pii/S1566253520303134.

[15] A. Gande and D. C. Parsley. News spillovers in the sovereign debt market. *Journal of Financial Economics*, 75(3):691–734, Mar. 2005. ISSN 0304-405X. doi: 10.1016/j.jfineco.2003.11.003. URL https://www.sciencedirect.com/science/article/pii/S0304405X04001576.

[16] P. Gazzotti. RICE50+: DICE model at country and regional level. *Socio-Environmental Systems Modelling*, 4:18038–18038, Apr. 2022. ISSN 2663-3027. doi: 10.18174/sesmo.18038. URL https://sesmo.org/article/view/18038.

[17] M. E. Kahn, K. Mohaddes, R. N. C. Ng, M. H. Pesaran, M. Raissi, and J.-C.

Yang. Long-term macroeconomic effects of climate change: A cross-country analysis. *Energy Economics*, 104:105624, Dec. 2021. ISSN 0140-9883. doi: 10.1016/j.eneco.2021.105624. URL https://www.sciencedirect.com/science/article/pii/S0140988321004898.

[18] M. Kalkuhl and L. Wenz. The impact of climate conditions on economic production. Evidence from a global panel of regions. *Journal of Environmental Economics and Management*, 103:102360, Sept. 2020. ISSN 0095-0696. doi: 10.1016/j.jeem.2020.102360. URL https://www.sciencedirect.com/science/article/pii/S0095069620300838.

[19] P. Klusak, M. Agarwala, M. Burke, M. Kraemer, and K. Mohaddes. Rising temperatures, falling ratings: The effect of climate change on sovereign creditworthiness. SSRN Scholarly Paper 3811958, Social Science Research Network, Rochester, NY, Mar. 2021. URL https://papers.ssrn.com/abstract=3811958.

[20] M. Kraemer, M. Mrsnik, and A. Petrov. Storm Alert: Natural Disasters Can Damage Sovereign Creditworthiness. page 32.

[21] T. M. Mitchell. *Machine Learning*. McGraw-Hill, 1997. ISBN 978-0-07-115467-3. Google-Books-ID: EoYBngEACAAJ.

[22] N. D. Rao, P. Sauer, M. Gidden, and K. Riahi. Income inequality projections for the Shared Socioeconomic Pathways (SSPs). *Futures*, 105:27–39, Jan. 2019. ISSN 0016-3287. doi: 10.1016/j.futures.2018.07.001. URL https://www.sciencedirect.com/science/article/pii/S001632871730349X.

[23] O. Sagi and L. Rokach. Explainable decision forest: Transforming a decision forest into an interpretable tree. *Information Fusion*, 61:124–138, Sept. 2020. ISSN 1566-2535. doi: 10.1016/j.inffus.2020.03.013. URL https://www.sciencedirect.com/science/article/pii/S1566253519307869.

[24] O. Sagi and L. Rokach. Approximating XGBoost with an interpretable decision tree. *Information Sciences*, 572:522–542, Sept. 2021. ISSN 0020-0255. doi: 10.1016/j.ins.2021.05.055. URL https://www.sciencedirect.com/science/article/pii/S0020025521005272.

[25] T. Vidal, T. Pacheco, and M. Schiffer. Born-Again Tree Ensembles. *arXiv:2003.11132 [cs, stat]*, Aug. 2020. URL http://arxiv.org/abs/2003.11132. arXiv: 2003.11132.

A | Appendix A

Long-term foreign currency issuer rating symbol	Numerical rating	Rating grade	
S&P			
AAA	20	Prime high grade	Investment grade
AA+	19	High grade	
AA	18		
AA-	17		
A+	16	Upper medium grade	
A	15		
A-	14		
BBB+	13	Lower medium grade	
BBB	12		
BBB-	11		
BB+	10	Speculative	Non-investment grade
BB	9		
BB-	8		
B+	7	Highly speculative	
B	6		
B-	5		
CCC+	4	Substantial risks	
CCC	3		
CCC-	2		
CC	1	Extremely speculative	
C	1		
D/SD	1	In default	

Figure A.1: Conversion of S&P grade into a 20-notch numerical scale

The scale on the figure A.1 is the scale used in Klusak et al. from S&P, (it is a picture from their article).

B | Appendix B

Appendix where you find more data visualisation corresponding to the Chapter 2.

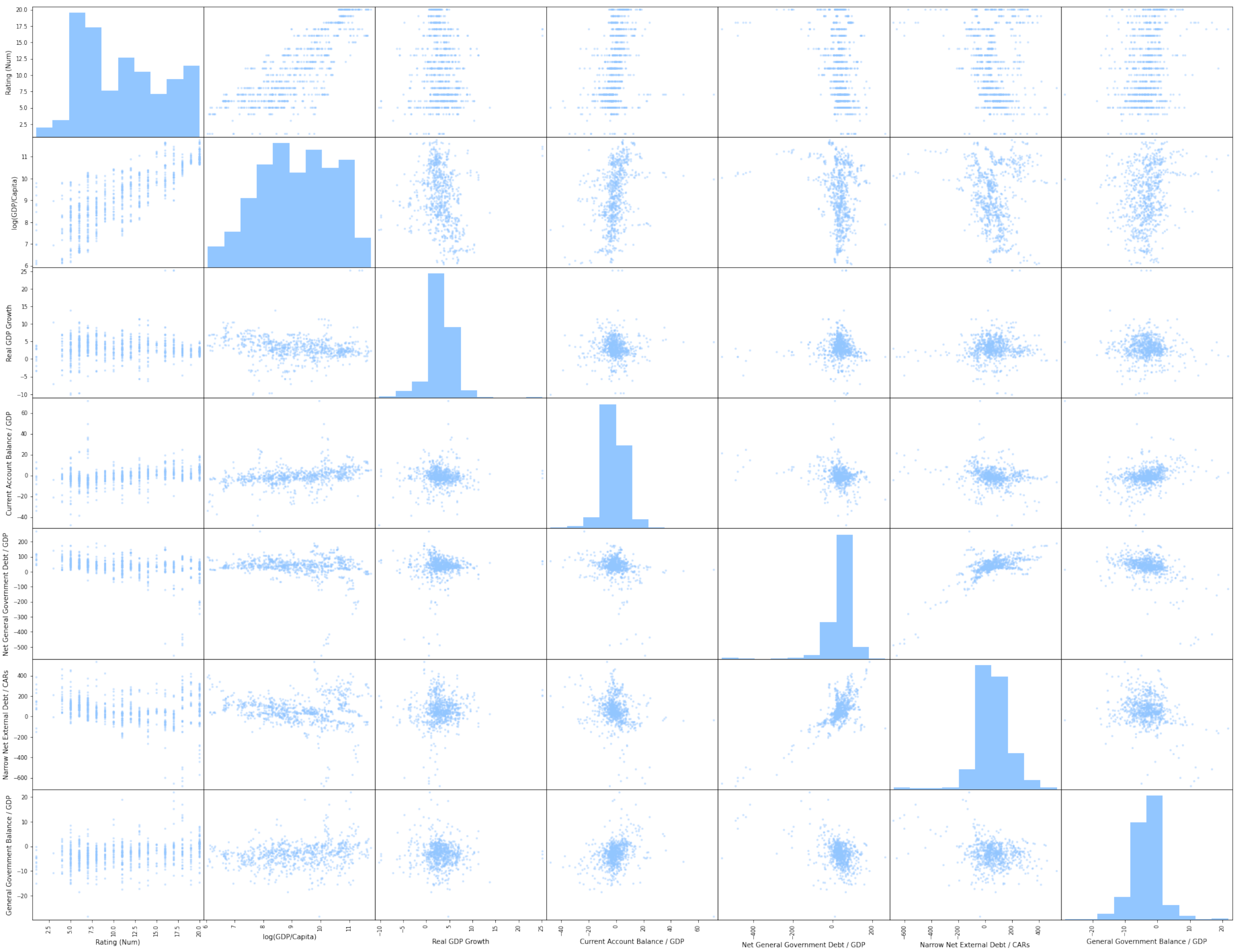


Figure B.1: Data Visualisation

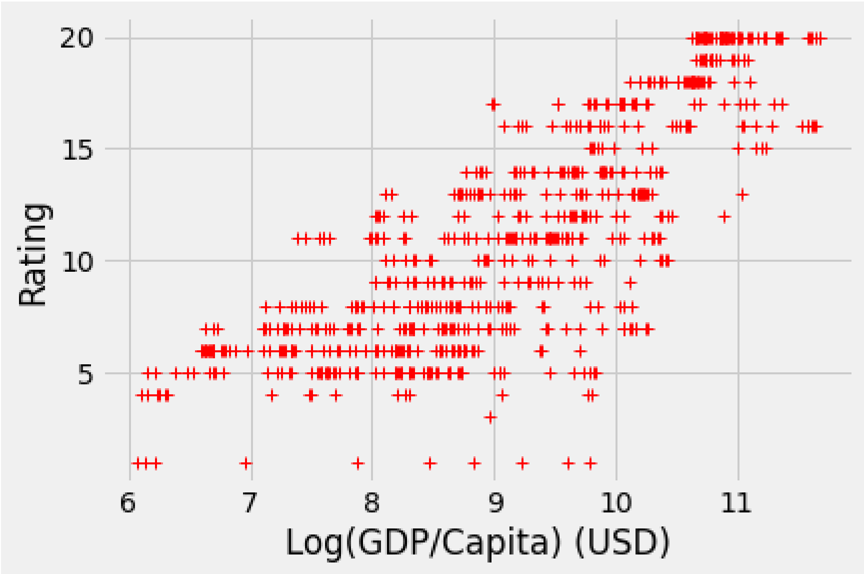
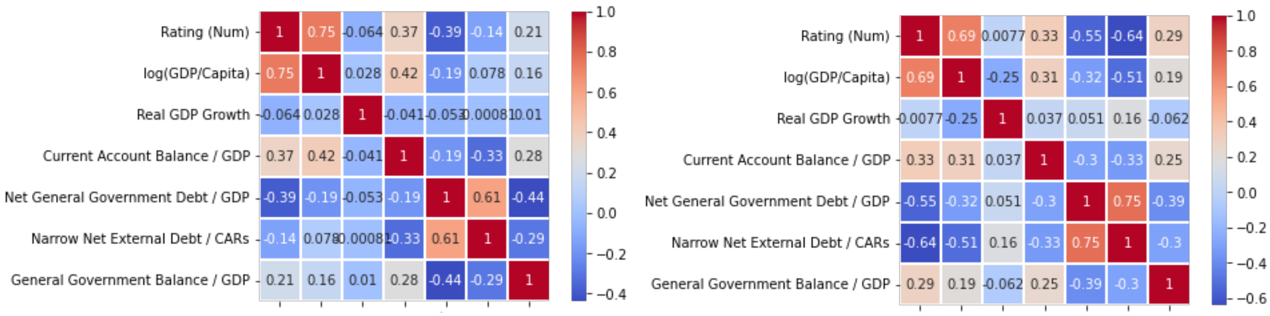


Figure B.2: Relation between log(GDP/Capita) and Rating



(a) (b)

Figure B.3: Correlations Matrices for OECD (a) and Non OECD countries (b)

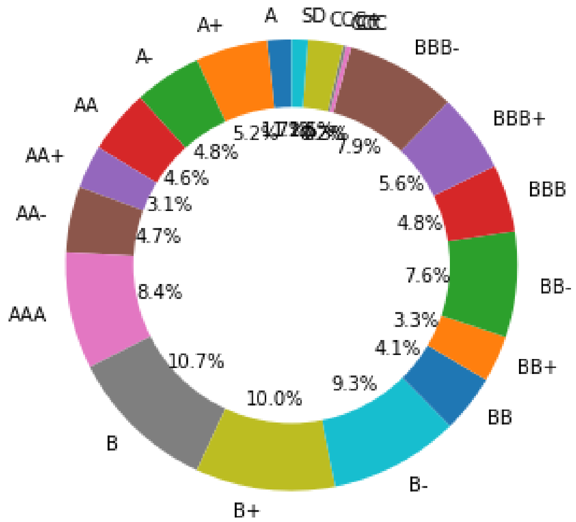


Figure B.4: Ratings Repartition in our data

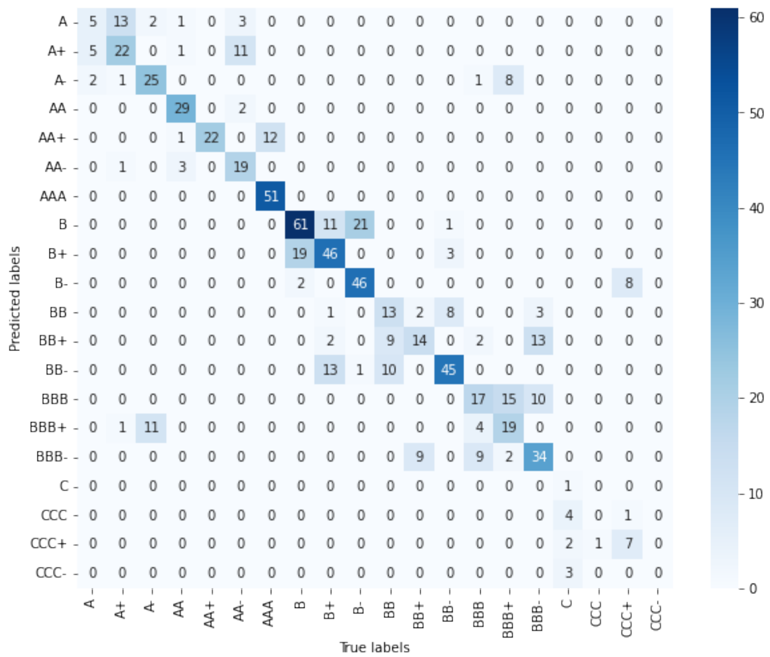


Figure B.5: Random Forest confusion matrix on the training

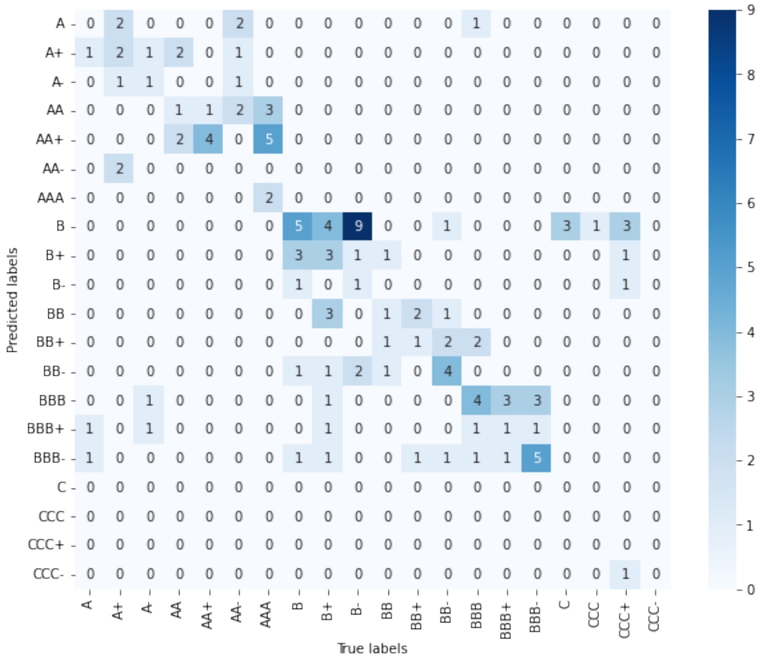


Figure B.6: Random Forest confusion matrix on the test

C | Appendix C

Chapter with example of decision tree and some visualization.

```
|--- log(GDP/Capita) <= 9.86
|   |--- Narrow Net External Debt / CARs <= 57.86
|   |   |--- log(GDP/Capita) <= 8.76
|   |   |   |--- Narrow Net External Debt / CARs <= -16.88
|   |   |   |   |--- Net General Government Debt / GDP <= 179.26
|   |   |   |   |   |--- value: [10.75]
|   |   |   |   |--- Net General Government Debt / GDP >  179.26
|   |   |   |   |   |--- value: [1.00]
|   |   |   |--- Narrow Net External Debt / CARs >  -16.88
|   |   |   |   |--- log(GDP/Capita) <= 6.89
|   |   |   |   |   |--- value: [4.57]
|   |   |   |   |--- log(GDP/Capita) >  6.89
|   |   |   |   |   |--- value: [8.00]
|   |   |--- log(GDP/Capita) >  8.76
|   |   |   |--- Net General Government Debt / GDP <= 94.33
|   |   |   |   |--- General Government Balance / GDP <= -3.09
|   |   |   |   |   |--- value: [11.13]
|   |   |   |   |--- General Government Balance / GDP >  -3.09
|   |   |   |   |   |--- value: [13.12]
|   |   |   |--- Net General Government Debt / GDP >  94.33
|   |   |   |   |--- value: [4.60]
|   |--- Narrow Net External Debt / CARs >  57.86
|   |   |--- Net General Government Debt / GDP <= 65.77
|   |   |   |--- log(GDP/Capita) <= 8.07
|   |   |   |   |--- value: [6.11]
|   |   |   |--- log(GDP/Capita) >  8.07
|   |   |   |   |--- Real GDP Growth <= 1.32
```

```
|   |   |   |   |   |--- value: [5.64]
|   |   |   |   |--- Real GDP Growth >  1.32
|   |   |   |   |   |--- value: [8.58]
|   |   |--- Net General Government Debt / GDP >  65.77
|   |   |   |--- log(GDP/Capita) <= 7.05
|   |   |   |   |--- value: [2.43]
|   |   |   |--- log(GDP/Capita) >  7.05
|   |   |   |   |--- value: [5.38]
|--- log(GDP/Capita) >  9.86
|   |--- log(GDP/Capita) <= 10.47
|   |   |--- Net General Government Debt / GDP <= 63.56
|   |   |   |--- Current Account Balance / GDP <= 29.35
|   |   |   |   |--- Current Account Balance / GDP <= -2.32
|   |   |   |   |   |--- value: [12.86]
|   |   |   |   |--- Current Account Balance / GDP >  -2.32
|   |   |   |   |   |--- value: [15.49]
|   |   |   |--- Current Account Balance / GDP >  29.35
|   |   |   |   |--- value: [7.00]
|   |   |--- Net General Government Debt / GDP >  63.56
|   |   |   |--- Current Account Balance / GDP <= -0.06
|   |   |   |   |--- value: [8.85]
|   |   |   |--- Current Account Balance / GDP >  -0.06
|   |   |   |   |--- value: [12.31]
|   |--- log(GDP/Capita) >  10.47
|   |   |--- Real GDP Growth <= 3.57
|   |   |   |--- General Government Balance / GDP <= -0.39
|   |   |   |   |--- value: [18.28]
|   |   |   |--- General Government Balance / GDP >  -0.39
|   |   |   |   |--- value: [19.78]
|   |   |--- Real GDP Growth >  3.57
|   |   |   |--- value: [16.92]
```

```
findfont: Font family ['Arial'] not found. Falling back to DejaVu Sans.
```

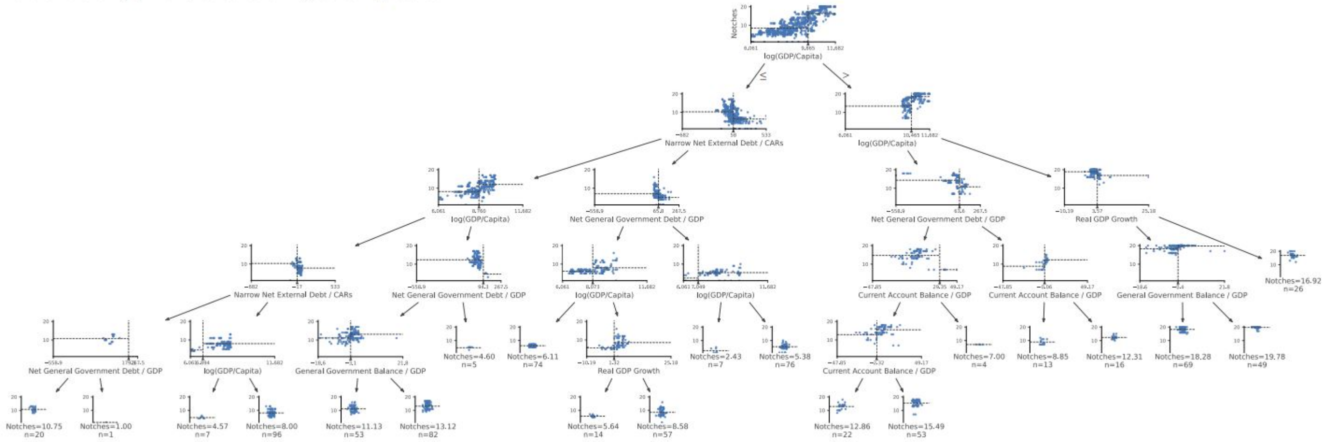


Figure C.1: Visualisation of the mechanism of the tree

	Mean Error	Exact	Within 1	Within 2	Within 3
Random Forest	*1.16*	*36.26*	*58.71*	*82.18*	*92.10*
Whole Sample	1.63	24.75	45.23	69.84	83.15
OECD Countries	1.14	47.30	68.37	84.78	90.26
Non OECD Countries	1.52	26.38	48.40	73.81	85.93
Investment Grade Countries	1.21	36.5	58.43	80.04	90.96
Non Investment Grade Countries	1.05	33.03	61.57	88.26	95.45

Table C.1: Decision Trees results with Max Depth = 5

D | Appendix D

Appendix with the other example of fitting between the macroeconomic variable and GDP Loss as described and explained in section 2.3

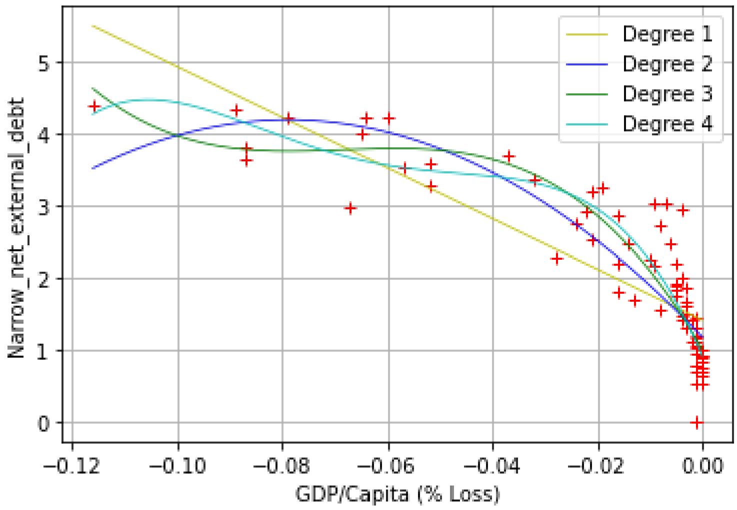


Figure D.1: Fitting models of the effect of GDP loss on Narrow Net External Debt

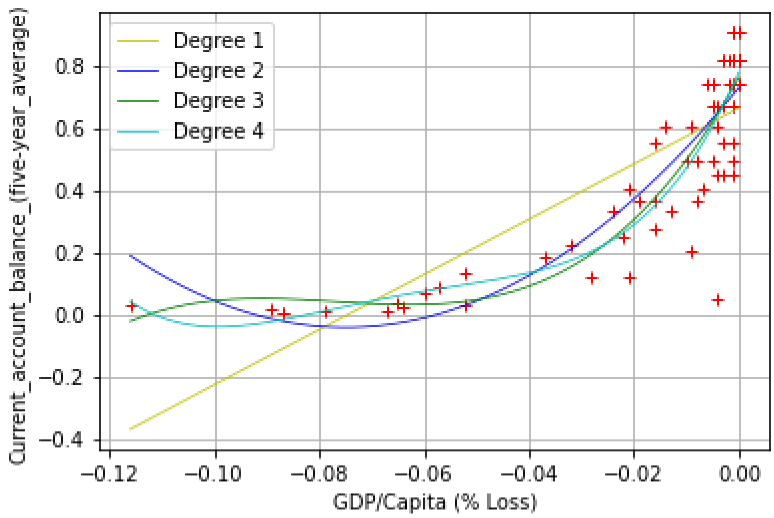


Figure D.2: Fitting models of the effect of GDP loss on Current Account Balance

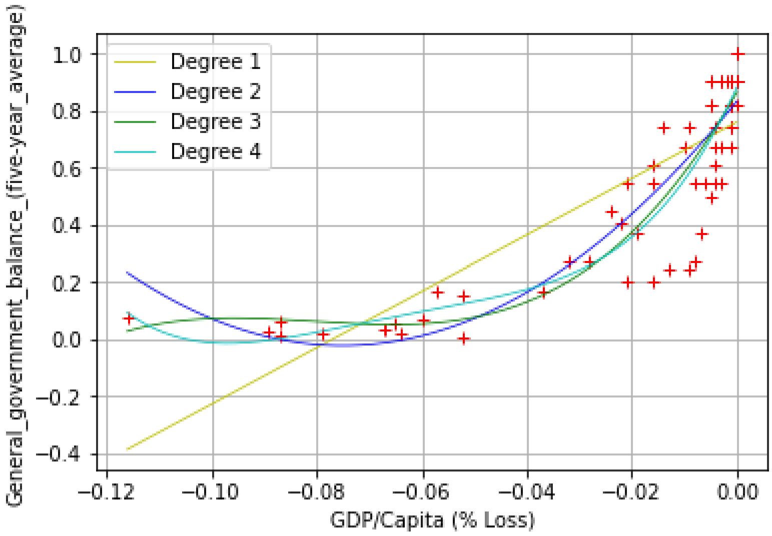


Figure D.3: Fitting models of the effect of GDP loss on General Government Balance

E | Appendix E

Appendix where you find more data visualisation and results corresponding to the Chapter 3.

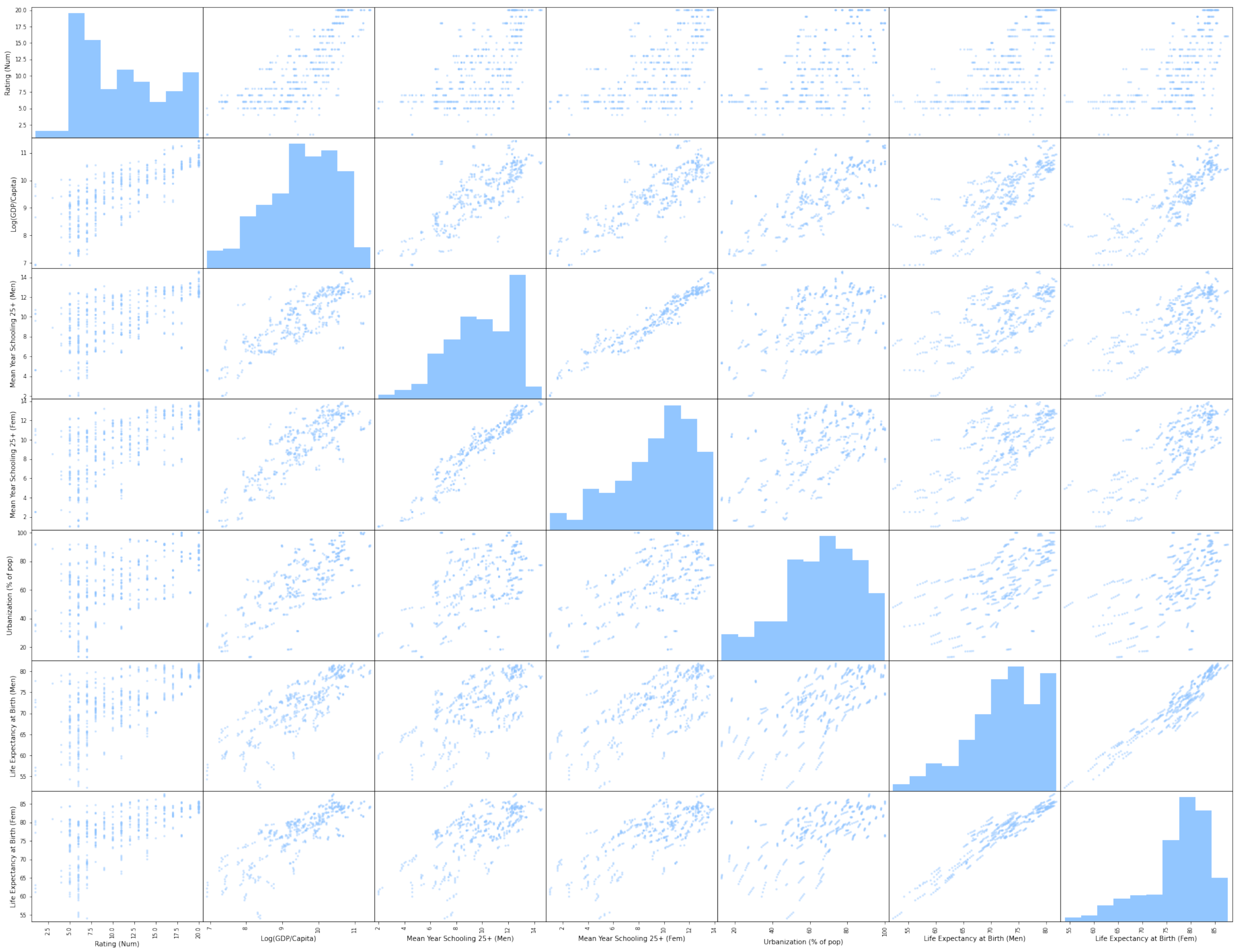


Figure E.1: New Data Visualisation

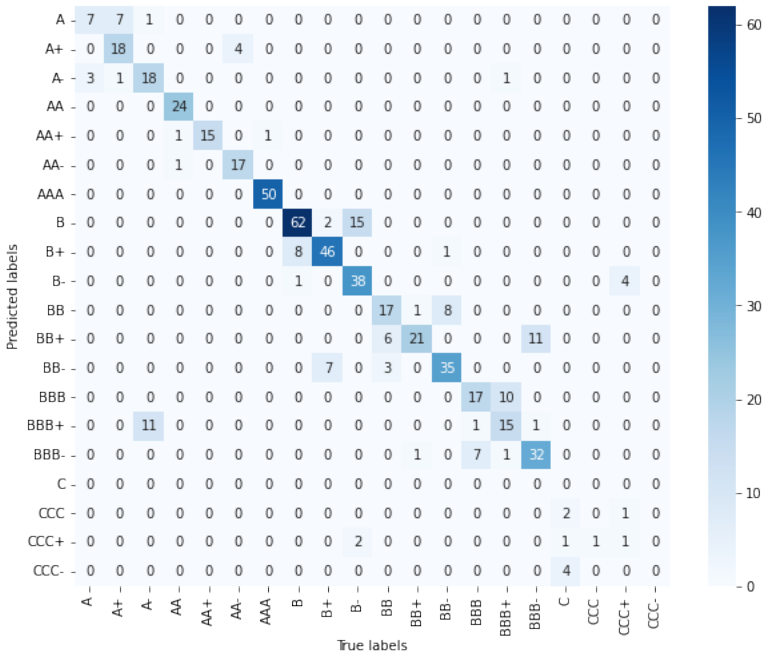


Figure E.2: Random Forest confusion matrix with Model 1

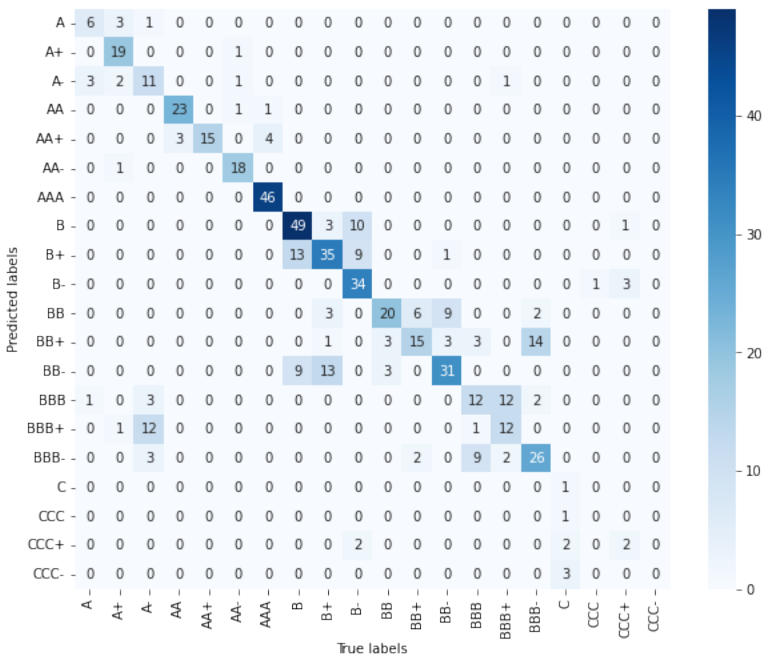


Figure E.3: Random Forest confusion matrix with Model 2

List of Figures

List of Tables

Acronyms

BIS Bank for International Settlement. 2

CARs Current Account Receipts. 18

CRA Credit Rating Agencies. 8, 19, 22, 31

ECB European Central Bank. 1

ESG Environmental, Social, and Governance. 1

GDP Gross Domestic Product. 6

GRP Gross Regional Product. 8

IAM Integrated Assessment Models. 5, 30

IPCC Intergovernmental Panel on Climate Change. 1

PPP Purchasing Power Parity. 29

RCP Representative Concentration Pathway. 6

SSP Shared Socioeconomic Pathways. 5, 29, 31, 34

Acknowledgements

Thanks to Prof. Tavoni for giving me the opportunity to make this thesis. Thanks to Matteo Calcaterra for his support and guidance.

Thanks to my family for his support during my studies.

Thanks to Centrale Nantes for giving me the opportunity to go in double degree and thanks to Politecnico di Milano for welcoming me.

The mistake you make, don't you see, is in thinking one can live in a corrupt society without being corrupt oneself. After all, what do you achieve by refusing to make money? You're trying to behave as though one could stand right outside our economic system. But one can't. One's got to change the system, or one changes nothing. One can't put things right in a hole-and-corner way, if you take my meaning.

Keep the Aspidistra Flying, George Orwell