



ΑΡΙΣΤΟΤΕΛΕΙΟ ΠΑΝΕΠΙΣΤΗΜΙΟ ΘΕΣΣΑΛΟΝΙΚΗΣ

ΠΡΟΣΑΡΜΟΓΗ ΤΗΣ ΕΚΤΙΜΗΣΗΣ ΖΗΜΙΑΣ
ΣΕ ΠΕΡΙΠΤΩΣΗ ΑΘΕΤΗΣΗΣ ΣΤΕΓΑΣΤΙΚΩΝ
ΔΑΝΕΙΩΝ ΣΤΟ ΠΛΑΙΣΙΟ ΚΛΙΜΑΤΙΚΩΝ ΚΑΙ
ΦΥΣΙΚΩΝ ΚΙΝΔΥΝΩΝ

ΔΙΠΛΩΜΑΤΙΚΗ ΕΡΓΑΣΙΑ

ΚΩΝΣΤΑΝΤΙΝΟΣ ΚΑΤΙΔΗΣ

AEM:9399

katidis98@gmail.com

ΕΠΙΒΛΕΠΩΝ: ΚΩΝΣΤΑΝΤΙΝΟΣ ΠΑΠΑΛΑΜΠΡΟΥ

ΕΠΙΚΟΥΡΟΣ ΚΑΘΗΓΗΤΗΣ ΑΠΘ

ΠΕΡΙΛΗΨΗ

Στην παρούσα διπλωματική εργασία, εξετάζουμε πώς μπορούμε να χρησιμοποιήσουμε τεχνικές προσομοίωσης και μηχανικής μάθησης για να εκτιμήσουμε τις ζημιές που μπορεί να προκύψουν σε στεγαστικά δάνεια λόγω φυσικών κινδύνων, με επίκεντρο την επίδραση της κλιματικής αλλαγής στην Ελλάδα. Για να το πετύχουμε αυτό, αναπτύξαμε ένα μοντέλο προσομοίωσης που καλύπτει έναν ορίζοντα 26 ετών (από το 2025 έως το 2050). Μέσω της προσομοίωσης, δημιουργούμε δεδομένα που περιλαμβάνουν παράγοντες κινδύνου όπως πλημμύρες, πυρκαγιές και σεισμούς, καθώς και επιδράσεις από οικονομικά σοκ και την ηλικία του ακινήτου, τα οποία συνδυάζονται για να υπολογίσουν πόση αξία μπορεί να χαθεί σε περίπτωση αθέτησης.

Το μοντέλο μας υπολογίζει την προσαρμοσμένη αξία των ακινήτων μετά από αυτές τις επιπτώσεις και, στη συνέχεια, υπολογίζει το LGD – δηλαδή το ποσοστό απώλειας σε σχέση με το δάνειο – με βάση τα δεδομένα αυτά. Έπειτα, χρησιμοποιούμε μοντέλα μηχανικής μάθησης, όπως το Random Forest και το XGBoost, για να μάθουν από τα προσομοιωμένα δεδομένα και να προβλέψουν το LGD για νέα παραδείγματα. Η εκπαίδευση και η αξιολόγηση των μοντέλων γίνεται με τεχνικές cross-validation, διασφαλίζοντας ότι οι προβλέψεις είναι ακριβείς και γενικεύσιμες.

Με αυτόν τον τρόπο, το έργο μας συνδυάζει θεωρητικά μοντέλα και τεχνικές μηχανικής μάθησης, παρέχοντας ένα ολοκληρωμένο πλαίσιο για την εκτίμηση των φυσικών κινδύνων στα στεγαστικά δάνεια στην Ελλάδα. Ο στόχος είναι να βοηθήσουμε στη διαμόρφωση αποτελεσματικών στρατηγικών διαχείρισης κινδύνου, λαμβάνοντας υπόψη τις μεταβαλλόμενες συνθήκες που επιβάλλει η κλιματική αλλαγή.

ABSTRACT

In this thesis, we examine how simulation techniques and machine learning can be used to estimate the damages that may occur in mortgage loans due to natural hazards, with a focus on the impact of climate change in Greece. To achieve this, we developed a simulation model that covers a 26-year horizon (from 2025 to 2050). Through simulation, we generate data that includes risk factors such as floods, fires, and earthquakes, as well as impacts from economic shocks and the age of the property, which are combined to calculate how much value might be lost in the event of default.

Our model computes the adjusted property value after these impacts and then calculates the LGD—the percentage loss relative to the loan—based on this data. Subsequently, we employ machine learning models, such as Random Forest and XGBoost, to learn from the simulated data and predict the LGD for new examples. The training and evaluation of the models are carried out using cross-validation techniques, ensuring that the predictions are accurate and generalizable.

In this way, our work combines theoretical models and machine learning techniques, providing a comprehensive framework for estimating the impact of natural hazards on mortgage loans in Greece. The goal is to assist in the formulation of effective risk management strategies, considering the evolving conditions imposed by climate change.

ΕΥΧΑΡΙΣΤΙΕΣ

Θα ήθελα να ευχαριστήσω τον επιβλέποντα καθηγητή κύριο Κωνσταντίνο Παπαλάμπρου για την ευκαιρία που μου έδωσε να συνεργαστώ μαζί του, την υπομονή και την καθοδήγηση που μου προσφέρθηκε. Θα ήθελα να ευχαριστήσω επίσης τους γονείς μου Γιώργο και Παρασκευή που ήταν δίπλα μου σε αυτό το ταξίδι ενηλικίωσης. Καμία επιτυχία μου δε θα ήταν εφικτή αν δεν υπήρχαν αυτοί οι άνθρωποι να μου δίνουν ελπίδα και έμπνευση να συνεχίσω το έργο μου. Θα ήθελα επίσης να ευχαριστήσω και τους φίλους μου Βασίλη, Γιώργο, Γιάννη και Λευτέρη που μου έδιναν κουράγιο και πίστη στις δυνατότητές μου να εκπληρώσω έναν από τους στόχους μου να γίνω μηχανικός. Ολοκληρώνοντας τη διπλωματική μου εργασία, δίνεται τέλος στον πρόλογο της καριέρας μου ως μηχανικός, ανοίγοντας τον δρόμο σε νέες ευκαιρίες, στόχους και γνωριμίες μέσα στον ωκεανό του αγνώστου.

Contents

1 Εισαγωγή	7
1.1 Δήλωση Προβλήματος(Problem Statement).....	7
1.2 Στόχοι και Ερευνητικά Ερωτήματα	8
2. Θεωρητικό Υπόβαθρο(Theoretical Background)	9
2.1 Πιστωτικός Κίνδυνος(Credit Risk)	9
2.2 ΖΗΜΙΑ ΚΑΤΑ ΑΘΕΤΗΣΗ (LGD).....	14
2.2.1 Παράγοντες που επηρεάζουν το LGD.....	14
Μαθηματική φόρμουλα.....	15
2.2.2 Σχέση LGD και Δείκτη LTV	16
2.2.3 Προβλέψεις LGD και Τεχνικές Μηχανικής Μάθησης	17
2.2.3.1 ΤΕΧΝΙΚΕΣ ΜΗΧΑΝΙΚΗΣ ΜΑΘΗΣΗΣ	18
2.2.3.2 Data Engineering.....	18
2.4 Προσομοιώσεις Monte Carlo	20
2.5 Κλιματική Άλλαγή και Επιπτώσεις στον Πιστωτικό Κίνδυνο	24
3 ΜΕΘΟΔΟΛΟΓΙΑ(METHODOLOGY)	25
3.1 ΣΥΛΛΟΓΗ ΚΑΙ ΠΡΟΕΠΕΞΕΡΓΑΣΙΑ ΔΕΔΟΜΕΝΩΝ.....	25
3.1.1 Δεδομένα ακινήτων και στεγαστικών δανείων	25
3.1.2 Κλιματικά και μετεωρολογικά δεδομένα.....	27
3.1.3 Δεδομένα Φυσικών Κινδύνων (σεισμοί, πυρκαγιές, πλημμύρες)	28
3.2 Υπολογισμός Κινδύνων και Παραμέτρων Δανείων	28
3.2.1 Σεισμικός Κίνδυνος (Earthquake Risk)	28
3.2.2 Κίνδυνος Πλημμύρας (Flood Risk)	30
3.2.3 Κίνδυνος Πυρκαγιάς (Wildfire Risk)	31
3.2.4 Loan Augmentation και Ολοκλήρωση Παραμέτρων Δανείων.....	32
3.3 Μοντέλα Μηχανικής Μάθησης για Εκτίμηση του LGD	33
3.3.1 Προετοιμασία Δεδομένων	33
3.3.2 Επιλογή και Εκπαίδευση Μοντέλων.....	34
3.3.3 Αξιολόγηση και Ανάλυση Αποτελεσμάτων	35
3.3.4 Λογική Επιλογής Μοντέλων.....	37
3.4 Προσομοιώσεις Monte Carlo	38
3.4.1 Σκοπός και Θεωρητική Βάση	38
3.4.2 Διαδικασία Προσομοίωσης.....	38
3.4.3 Λογική και Επιλογή των Παραμέτρων.....	40
3.4.4 Εφαρμογή και Χρήση των Αποτελεσμάτων	40
3.5 Απεικόνιση Αποτελεσμάτων και Διαγράμματα	41

3.5.1 Γραφικές Παραστάσεις και Διαγράμματα Χρόνου	41
3.5.2 Χαρτογράφηση	41
4. Αποτελέσματα	43
4.1 Αποτελέσματα Προσομοιώσεων Stress Testing	43
4.1.1 Ανάλυση Βασικού Σεναρίου	43
4.1.2 Ανάλυση Σεναρίου Υψηλού Κινδύνου Πλημμύρας.....	46
4.1.3 Ανάλυση Σεναρίου Υψηλού Σεισμικού Κινδύνου	49
4.1.4 Ανάλυση Σεναρίου Υψηλού Κινδύνου Πυρκαγιάς	50
4.1.5 Ανάλυση Συνδυασμένου Σεναρίου Υψηλού Κινδύνου Πλημμύρας και Πυρκαγιάς	53
4.2 Ανάλυση Μηχανικής Μάθησης (Machine Learning Analysis)	56
4.2.1 Αποτελέσματα Cross-Validation (CV) για το Random Forest.....	56
4.2.2 Αποτελέσματα Cross-Validation (CV) για το XGBoost.....	57
4.2.3. Σημασία Χαρακτηριστικών (Feature Importances)	57
5. Συμπεράσματα και Περιορισμοί	62
6. Προτάσεις για Μελλοντική Έρευνα	64
7. ΒΙΒΛΙΟΓΡΑΦΙΑ	65

1 Εισαγωγή

1.1 Δήλωση Προβλήματος(Problem Statement)

Η κλιματική αλλαγή αποτελεί μία από τις μεγαλύτερες παγκόσμιες προκλήσεις, με σημαντικές επιπτώσεις τόσο στο περιβάλλον όσο και στις οικονομικές και κοινωνικές δομές. Οι φυσικοί κίνδυνοι που συνδέονται με την κλιματική αλλαγή, όπως οι πλημμύρες, οι καύσωνες, οι ξηρασίες και τα ακραία καιρικά φαινόμενα, παρουσιάζουν αυξανόμενη συχνότητα και ένταση, επηρεάζοντας άμεσα τις ζωές των ανθρώπων και τις οικονομίες των περιοχών. Όταν μια φυσική καταστροφή πλήττει μια περιοχή, επηρεάζει τόσο τους ντόπιους κατοίκους όσο και τις γειτονικές περιοχές, προκαλώντας όχι μόνο υλικές ζημιές αλλά και σημαντικές ψυχολογικές επιπτώσεις.

Ένα χαρακτηριστικό παράδειγμα είναι η πρόσφατη πλημμύρα της Θεσσαλίας από την καταιγίδα Daniel, η οποία προκάλεσε ζημιές ύψους 2,14 δισεκατομμυρίων ευρώ (“Κακοκαιρία Ντάνιελ,” n.d.). Ένα άλλο αξιοσημείωτο παράδειγμα είναι ο σεισμός της Ιαπωνίας το 2011, που προκάλεσε οικονομικές απώλειες ύψους 25 τρισεκατομμυρίων γεν, ενώ οι ψυχολογικές επιπτώσεις στους κατοίκους παραμένουν εμφανείς μέχρι σήμερα. Οι φυσικές καταστροφές όχι μόνο αυξάνουν την αντίληψη του κινδύνου, αλλά επηρεάζουν και τις οικονομικές επενδύσεις. Επενδυτές συχνά επιδεικνύουν μεγαλύτερη επιφυλακτικότητα σε περιοχές που έχουν πληγεί από σημαντικές φυσικές καταστροφές, σε αντίθεση με περιοχές που έχουν υποστεί ελάχιστες ή καθόλου ζημιές. Αυτή η διαφοροποιημένη στάση υπογραμμίζει μια προκατάληψη στον τρόπο με τον οποίο οι διαχειριστές επενδυτικών κεφαλαίων ανταποκρίνονται σε τέτοια γεγονότα (Feng, Wang, and Lu 2022) .

Οι επιπτώσεις αυτών των φυσικών κινδύνων επεκτείνονται και στον χρηματοπιστωτικό τομέα, όπου οι στεγαστικές υποθήκες αποτελούν βασικό μέρος των δανειακών χαρτοφυλακίων των τραπεζών. Η πτώση της αξίας των ακινήτων λόγω πλημμυρών ή άλλων φυσικών φαινομένων οδηγεί σε αύξηση της αναλογίας Δανείου προς Αξία (Loan-to-Value, LTV) και συνεπώς σε αύξηση της Ζημίας Κατά την Αθέτηση (Loss Given Default, LGD). Επιπλέον, οι αυξημένες δαπάνες ανάκτησης και οι καθυστερήσεις στις διαδικασίες λόγω φυσικών καταστροφών επιβαρύνουν περαιτέρω το LGD καθιστώντας τις επιπτώσεις των φυσικών κινδύνων μια πολυδιάστατη πρόκληση για τον χρηματοπιστωτικό τομέα.

1.2 Στόχοι και Ερευνητικά Ερωτήματα

● Στόχοι Εργασίας

Η συγκεκριμένη εργασία επικεντρώνεται στην κατανόηση και την ποσοτικοποίηση της επίδρασης των φυσικών κινδύνων της κλιματικής αλλαγής στη Ζημία Κατά την Αθέτηση (Loss Given Default, LGD) για στεγαστικά δάνεια στον ελληνικό τραπεζικό τομέα. Οι βασικοί στόχοι της έρευνας περιλαμβάνουν:

1. Την αξιολόγηση της σχέσης μεταξύ φυσικών κινδύνων (π.χ. πλημμύρες, καύσωνες, σεισμοί) και της αξίας των εξασφαλίσεων (ακινήτων).
2. Την εφαρμογή τεχνικών μηχανικής μάθησης για την ακριβή πρόβλεψη του LGD σε σενάρια στρες που ενσωματώνουν τις επιπτώσεις των φυσικών καταστροφών.
3. Την ανάπτυξη ενός πλαισίου που να επιτρέπει την ενσωμάτωση κλιματικών κινδύνων στις πρακτικές διαχείρισης πιστωτικού κινδύνου των τραπεζών.

● Ερευνητικά Ερωτήματα

Η εργασία αποσκοπεί να απαντήσει στα ακόλουθα βασικά ερωτήματα:

1. Πώς επηρεάζουν οι φυσικοί κίνδυνοι της κλιματικής αλλαγής την αξία των ακινήτων και κατ' επέκταση την αναλογία Δανείου προς Αξία (Loan-to-Value, LTV);
2. Πώς οι μεταβολές στην αξία των ακινήτων επηρεάζουν τη Ζημία Κατά την Αθέτηση (LGD) για στεγαστικά δάνεια;
3. Πώς μπορούν οι ελληνικές τράπεζες να χρησιμοποιήσουν τα ευρήματα της έρευνας για να θωρακίσουν τα δανειακά τους χαρτοφυλάκια από φυσικούς κινδύνους;

● Διατύπωση Σχέσης

Στο πλαίσιο της εργασίας εξετάζεται η δυνατότητα εφαρμογής τεχνικών μηχανικής μάθησης για την πρόβλεψη του LGD υπό σενάρια κλιματικών καταστροφών. Η κατανόηση αυτών των παραγόντων είναι κρίσιμη για την ανάπτυξη αποτελεσματικών στρατηγικών διαχείρισης πιστωτικού κινδύνου, που θα θωρακίσουν τις τράπεζες απέναντι στις επιπτώσεις της κλιματικής αλλαγής.

2. Θεωρητικό Υπόβαθρο(Theoretical Background)

2.1 Πιστωτικός Κίνδυνος(Credit Risk)

Ορισμός: Ο πιστωτικός κίνδυνος είναι η πιθανότητα να μην ικανοποιηθεί η προσδοκία για ένα χρηματικό ποσό εντός ενός συγκεκριμένου αναμενόμενου χρονικού διαστήματος ή διαφορετικά είναι το ρίσκο ότι ένας δανειολήπτης δε θα πληρώσει το δάνειό του. (Caouette et al. 2008)

Ο πιστωτικός κίνδυνος είναι σημαντικός καθώς υπάρχει σε κάθε χρηματική συναλλαγή στην οποία η πληρωμή δεν είναι άμεση. Επηρεάζει μια ποικιλία οντοτήτων:

- Κυβερνήσεις οι οποίες δανείζονται για να χρηματοδοτήσουν έργα και να εκτελέσουν τις λειτουργίες ενός κράτους(δημόσιες υπηρεσίες όπως εκπαίδευση και σύστημα υγείας, αμυντική θωράκιση σε εθνικό επίπεδο ασφάλειας, κοινωνικά προγράμματα)
- Εταιρείες οι οποίες δανείζονται για τις εξαγορές τους (stock and assets) και την ανάπτυξή τους.
- Μικρομεσαίες επιχειρήσεις οι οποίες επιθυμούν να μεγαλώσουν και να αναβαθμιστούν.
- Ατομικά τους ανθρώπους, οι οποίοι δανείζονται σε προσωπικό επίπεδο για να αγοράσουν σπίτια, αμάξια και άλλα αγαθά.

Βασικές Μεταβλητές του Πιστωτικού Κινδύνου

Η κατανόηση του βάθους του πιστωτικού κινδύνου μπορεί να αποδοθεί από τους βασικούς δείκτες που συνδέονται μαζί του (“An Explanatory Note on the Basel II IRB Risk Weight Functions - July 2005,”):

- EAD (Exposure at Default) είναι η εκτίμηση του υπολειπόμενου ποσού ενός δανείου(των ήδη τραβηγμένων ποσών από τον δανειολήπτη αλλά και μελλοντικά ποσά που δεν έχει χρησιμοποιήσει ακόμη) από τη στιγμή που ο δανειολήπτης αθετήσει.
- LGD (Loss Given Default) είναι το ποσοστό της έκθεσης της τράπεζας που ενδέχεται να χάσει σε περίπτωση αθέτησης λαμβάνοντας υπόψη τις ανακτήσεις και τις εξασφαλίσεις. Οι απώλειες συνήθως εκφράζονται ως ποσοστό του EAD και εξαρτώνται από τον τύπο και το ποσό της εξασφάλισης καθώς επίσης και τον τύπο του δανειολήπτη και τα αναμενόμενα έσοδα από την εκκαθάριση του περιουσιακού στοιχείου που χρησιμοποιήθηκε ως εξασφάλιση του δανείου.

- PD (Probability of Default) ανά βαθμίδα αξιολόγησης δίνει το μέσο ποσοστό δανειοληπτών που αθετούν τις υποχρεώσεις τους σε αυτή τη βαθμίδα αξιολόγησης εντός ενός έτους.

Αυτές οι μεταβλητές απαρτίζουν τις αναμενόμενες απώλειες μιας τράπεζας (EL Expected Losses) οι οποίες είναι ένα στοιχείο της επιχειρηματικής δραστηριότητας των χρηματοπιστωτικών οργανισμών και τις διαχειρίζονται μέσω διαφόρων μέσων όπως η δημιουργία προβλέψεων.

Η φόρμουλα που ενώνει και υπολογίζει τις αναμενόμενες ζημιές (σαν καθαρό ποσό) είναι η εξής:

$$EL = PD * EAD * LGD$$

Ιστορική Αναδρομή

Η ιστορία του πιστωτικού κινδύνου χρονολογείται από το 1800 π.Χ., καθιστώντας τον την αρχαιότερη μορφή κινδύνου στις χρηματοπιστωτικές αγορές. Από την αρχαία Αίγυπτο μέχρι σήμερα, όπως αναφέρθηκε, παραμένει ουσιαστικά αμετάβλητος, καθώς πάντα υπήρχε ένα στοιχείο αβεβαιότητας σχετικά με το αν ένας δανειολήπτης θα αποπληρώσει το δάνειο του. Ο πιστωτικός κίνδυνος αποτέλεσε βασικό στοιχείο του δανεισμού από την εποχή που οργανώθηκαν οι πρώτες τράπεζες στη Φλωρεντία πριν από 700 χρόνια. Με την πάροδο του χρόνου, τα χρηματοπιστωτικά ιδρύματα έχουν αναπτύξει νέα εργαλεία και τεχνικές για τη διαχείριση, την τιμολόγηση και τη διανομή αυτής της αρχαίας μορφής χρηματοοικονομικού κινδύνου. (Caouette et al. 2008).

Η ιστορία των μοντέλων πιστωτικού κινδύνου εξελίχθηκε από υποκειμενικές αξιολογήσεις σε πολύπλοκα μαθηματικά μοντέλα, με στόχο την ακριβέστερη εκτίμηση και διαχείριση του πιστωτικού κινδύνου (Adamko, Klieštik, and Birtus, n.d.).

Εποχή χωρίς μοντέλα:

- Πριν την ανάπτυξη των μαθηματικών μοντέλων οι εταιρείες βασίζονταν σε υποκειμενικές αποφάσεις για την αξιολόγηση των δανειοληπτών. Οι αξιολογήσεις βασίζονταν σε παράγοντες όπως η φήμη, τα κεφάλαια και οι εξασφαλίσεις.
- Οι οίκοι αξιολόγησης διαδραμάτισαν σημαντικό ρόλο, παρέχοντας μια πιο αντικειμενική αξιολόγηση. Η John Moody and Company εισήγαγε το "Moody's Manual of Industrial and Miscellaneous Securities" το 1900, και ξεκίνησε τις αξιολογήσεις ομολόγων το 1909.
- Ο Henry Varnum Poor ήταν από τους πρώτους που δημοσίευσε μελέτες ασφαλείας το 1860.

Πρώτα Μαθηματικά Μοντέλα:

- Οι πρώτες προσπάθειες για μοντέλα πρόβλεψης αθέτησης έγιναν από τους Beaver (1967) και Altman (1968).
- Ο Altman πρότεινε scoring models χρησιμοποιώντας λογιστικούς και αγοραίους δείκτες. Το Z-score ήταν ένας γραμμικός συνδυασμός πέντε επιχειρηματικών δεικτών.
- Το Moody's KMV RiskCalc (2000) χρησιμοποιεί μια μεγαλύτερη βάση δεδομένων και μετασχηματισμένες τιμές δεικτών.

Μοντέλα Δομικής Μορφής (Structural-form models):

- Τα δομικά μοντέλα βασίζονται στην έννοια της "απόστασης προς την αθέτηση" (distance-to-default).
- Πρώτη γενιά: Το μοντέλο του Merton (1974) θεωρείται το πρώτο δομικό μοντέλο, βασισμένο στην αποτίμηση δικαιωμάτων προαιρεσης των Black και Scholes. Το μοντέλο θεωρεί ότι μια εταιρεία θα αθετήσει όταν η αγοραία αξία της είναι μικρότερη από τις υποχρεώσεις της.
- Δεύτερη γενιά: Τα μοντέλα αυτά επέτρεψαν την αθέτηση σε οποιαδήποτε στιγμή πριν τη λήξη του ομολόγου και χαλάρωσαν την υπόθεση σταθερής καμπύλης επιτοκίων, με συνεισφορές από τους Kim, Ramaswamy και Sundaresan (1993) και Longstaff και Schwartz (1995).
- Το κύριο πρόβλημα των δομικών μοντέλων είναι ότι δεν είναι συμβατά με την ουδέτερη ως προς τον κίνδυνο πιθανότητα αθέτησης.

Μοντέλα Μειωμένης Μορφής (Reduced-form models):

- Οι Jarrow, Lando και Turnbull (1997) και Duffie και Singleton (1999) παρουσίασαν σημαντικά μοντέλα μειωμένης μορφής.

- Τα μοντέλα μπορούν να χρησιμοποιηθούν για την αποτίμηση παραγώγων πίστωσης, λαμβάνοντας υπόψη αλλαγές στην πιστοληπτική ικανότητα.
- Αυτά τα μοντέλα αντιμετωπίζουν τα προβλήματα των δομικών, θεωρώντας ότι ο χρόνος αθέτησης επηρεάζεται από μια εξωγενή διαδικασία και όχι από τα χαρακτηριστικά της εταιρείας.
- Υπάρχουν δύο τύποι: Μοντέλα που βασίζονται στην ένταση και μοντέλα μετάβασης πιστοληπτικής ικανότητας.

Μοντέλα Credit Value-at-Risk (VaR):

- Οι τράπεζες ανέπτυξαν μοντέλα VaR στα τέλη της δεκαετίας του 1990, όπως τα μοντέλα αθέτησης (default mode) και τα μοντέλα mark-to-market.
- Τα σημαντικότερα μοντέλα VaR είναι τα CreditMetrics, KMV's Credit Portfolio Manager, Credit Risk+ και Credit Portfolio View.
- To CreditMetrics είναι ένα υβριδικό μοντέλο.
- To KMV CreditPortfolioManager εστιάζει στην πιθανότητα αθέτησης (EDF).
- To CreditRisk+ χρησιμοποιεί κατανομή Poisson και εστιάζει μόνο στην αθέτηση.
- To CreditPortfolioView συνδέει την πιθανότητα αθέτησης με την οικονομική κατάσταση.

Οι συμφωνίες της Βασιλείας

Το 1974 δημιουργήθηκε η επιτροπή της Βασιλείας η οποία ανέπτυξε τις συμφωνίες της Βασιλείας (Basel) για να διασφαλιστεί η παγκόσμια χρηματοπιστωτική σταθερότητα από τον πιστωτικό κίνδυνο, τη ρευστότητα και τον κίνδυνο της αγοράς. Οι συμφωνίες αυτές είναι σειρές διεθνών κανονισμών που καθορίζουν τα πρότυπα διαχείρισης του κινδύνου και της κεφαλαιακής επάρκειας στις τράπεζες. ("Bank for International Settlements," n.d.) Οι συμφωνίες είναι 3:

- Basel 1 (1988): Εισήγαγε την έννοια της RWA (risk-weighted assets) και απαιτούσε οι τράπεζες να κρατούν το κεφάλαιο στο 8% των RWA. Εστίαζε στον πιστωτικό κίνδυνο κατηγοριοποιώντας τα assets, δίνοντας αντίστοιχα βάρη κινδύνου στην κάθε κατηγορία και καθορίζοντας ένα ελάχιστο αναγκαίο κεφάλαιο. Στόχος ήταν οι τράπεζες να εξασφαλίζουν ένα επαρκές κεφάλαιο για να απορροφηθούν πιθανές απώλειες λόγω πιστωτικού κινδύνου.

- Basel 2 (2004): Επέκτεινε το Basel 1 δίνοντας μια πιο ολοκληρωμένη προσέγγιση επί της διαχείρισης των ρίσκων. Υπήρχαν 3 βασικές αλλαγές:
 1. Απαιτήσεις ελάχιστου κεφαλαίου: Ενίσχυσε το υπάρχων πλαίσιο για τον πιστωτικό κίνδυνο και εισήγαγε κεφαλαιακές απαιτήσεις για τον λειτουργικό κίνδυνο.
 2. Εποπτική Επισκόπηση: Τόνισε τη σημασία της ανασκόπησης των εσωτερικών εκτιμήσεων των τραπεζών για την κεφαλαιακή επάρκεια και τη διαχείριση κινδύνων.
 3. Πειθαρχία της Αγοράς: Ενθάρρυνε την πειθαρχία της αγοράς απαιτώντας από τις τράπεζες να αποκαλύπτουν περισσότερες πληροφορίες για τους κινδύνους τους και την κεφαλαιακή επάρκεια.
- Basel 3 (2010): Αντιδρώντας στην οικονομική κρίση του 2007-2009, η συγκεκριμένη συμφωνία εισήγαγε πιο αυστηρές απαιτήσεις κεφαλαίου και ρευστότητας με στόχο να ενισχυθεί η ανθεκτικότητα του τραπεζικού τομέα. Τα κύρια στοιχεία της ήταν τα εξής:
 1. Αυξημένες Απαιτήσεις Κεφαλαίου: Αυξήθηκε ο δείκτης κοινής κεφαλαιακής επάρκειας (CET1) στο 4.5% των RWA.
 2. Δείκτης Μόχλευσης: Είναι ένας δείκτης που χρησιμοποιείται για να αξιολογηθεί πόσο κεφάλαιο έρχεται σε μορφή δανείου ή προσδιορίζει την ικανότητα μιας εταιρείας να ανταποκριθεί στις οικονομικές της υποχρεώσεις. Εισήχθη για να υποστηριχθούν οι δείκτες κεφαλαίου που βασίζονται στον κίνδυνο, με στόχο τον περιορισμό της υπερβολικής μόχλευσης του τραπεζικού συστήματος.
 3. Πρότυπα Ρευστότητας: Καθιερώθηκαν ο Δείκτης Κάλυψης Ρευστότητας (LCR) και ο Δείκτης Σταθερής Χρηματοδότησης (NSFR) για να εξασφαλιστεί ότι οι τράπεζες διαθέτουν επαρκή ρευστότητα για να καλύψουν τις βραχυπρόθεσμες και μακροπρόθεσμες υποχρεώσεις τους αντίστοιχα.

2.2 ΖΗΜΙΑ ΚΑΤΑ ΑΘΕΤΗΣΗ (LGD)

2.2.1 Παράγοντες που επηρεάζουν το LGD

Το LGD είναι ο κύριος στόχος υπολογισμού της εργασίας. Πιο συγκεκριμένα θέλουμε να επικεντρωθούμε στις αλλαγές της τιμής του. Οι παράγοντες που επηρεάζουν το LGD μπορούν να κατηγοριοποιηθούν ως εξής:

- Τύπος και αξία των εξασφαλίσεων: Η αξία και ο τύπος των εξασφαλίσεων που υποστηρίζουν το δάνειο έχουν άμεση σχέση με το LGD. Όταν η αξία των εξασφαλίσεων μειώνεται, το LGD αυξάνεται. Ένας παράγοντας που εξετάζουμε είναι οι φυσικές καταστροφές. (“Integrating Climate Risk into Credit Risk Modelling,” n.d.) Η ύπαρξη stranded assets μπορεί να επηρεάσει το LGD όπως για παράδειγμα περιουσιακά στοιχεία που σχετίζονται με τον τομέα του πετρελαίου και του φυσικού αερίου μπορεί να χάσουν την αξία τους λόγω της μετάβασης σε μια οικονομία χαμηλών εκπομπών άνθρακα.
- Οικονομικές Συνθήκες: Οι μακροοικονομικές παράμετροι όπως το ΑΕΠ (Ακαθάριστο Εγχώριο Προϊόν) και το ποσοστό ανεργίας επηρεάζουν το LGD, η αξία των εξασφαλίσεων μειώνεται και οι απώλειες από αθετήσεις αυξάνονται, αυξάνοντας το LGD.
- Οι κινήσεις των επιτοκίων επηρεάζουν την ικανότητα των δανειοληπτών να αποπληρώσουν τα δάνεια τους, κάτι που με τη σειρά του επηρεάζει το LGD. Αυξήσεις των επιτοκίων μπορεί να οδηγήσουν σε αύξηση των αθετήσεων και του LGD.
- Ο πληθωρισμός μπορεί επίσης να επηρεάσει το LGD, καθώς μειώνει την πραγματική αξία των εξασφαλίσεων.
- Όροι του δανείου: Οι όροι του δανείου, όπως η διάρκεια και το επιτόκιο καθορίζουν την ικανότητα του δανειολήπτη να αποπληρώσει το δάνειο και κατά συνέπεια το LGD.
- Τύπος του δανείου: Ο κάθε τύπος του δανείου (όπως π.χ. το στεγαστικό που επικεντρωνόμαστε) έχει διαφορετικά χαρακτηριστικά και εξασφαλίσεις. Για παράδειγμα, τα στεγαστικά δάνεια συνήθως έχουν ως εξασφάλιση το ακίνητο, ενώ τα επιχειρηματικά δάνεια μπορεί να έχουν άλλα περιουσιακά στοιχεία ή εγγυήσεις. (Wong and Ho 2023)
- Σχέση δανείου-αξίας (LTV): Ένα υψηλότερο LTV υποδηλώνει υψηλότερο LGD, καθώς υπάρχει μικρότερο απόθεμα στην αξία του περιουσιακού στοιχείου για την κάλυψη της έκθεσης σε περίπτωση αθέτησης.

- Κλιματικοί κίνδυνοι: Οι ακραίες καιρικές συνθήκες μπορούν να οδηγήσουν σε άμεσες υλικές ζημιές σε σπίτια και εργοστάσια ή σε διαγραφές περιουσιακών στοιχείων που βρίσκονται σε περιοχές υψηλού κινδύνου, επηρεάζοντας τις αξίες των εξασφαλίσεων. (Wong and Ho 2023)

Μαθηματική φόρμουλα

Η μαθηματική φόρμουλα που θα ακολουθήσουμε για να συμπεριλάβουμε και τις ζημιές από τις καταστροφές είναι η αφαίρεση 1-Recovery Rate, η οποία μεταφράζεται ως εξής:

$$LGD_{it} = \max \left(\frac{EAD_{it} - (1 - \beta) \cdot Value\ underlying_{it}}{EAD_{it}}, 0 \right)$$

όπου

- EAD_{it} αντιπροσωπεύει το Exposure at Default του δανείου i στο χρόνο t. Δηλαδή, είναι το ποσό του δανείου που η τράπεζα έχει να χάσει σε περίπτωση που ο δανειολήπτης αθετήσει το δάνειο.
- $Value\ underlying_{it}$ είναι η τρέχουσα αξία του ακινήτου που χρησιμεύει ως εξασφάλιση.
- β αντιπροσωπεύει το «haircut» που εφαρμόζεται στο ακίνητο. Το haircut είναι ένα ποσοστό (συνήθως μεταξύ 0 και 1) το οποίο μειώνει την αξία που μπορεί να ανακτηθεί από το ακίνητο λόγω κόστους συναλλαγών, νομικών εξόδων, καθυστερήσεων στην πώληση και άλλων πρακτικών παραγόντων.

Υποσημείωση: Λόγω της ανάγκης να αξιολογηθεί η καθαρή επίπτωση των καταστροφών στην αξία της εξασφάλισής μας, θα θεωρήσουμε ότι το β μας είναι ίσο με το 0.

Αυτός ο υπολογισμός ποσοστού του LGD είναι κρίσιμος για να κατανοηθεί η καθαρή ζημιά (χωρίς άλλες επιδράσεις) στην αξία των εξασφαλισμένου περιουσιακού στοιχείου και το ποσό που μπορεί να ανακτηθεί σε περίπτωση αθέτησης του δανείου (Siegers, n.d.).

2.2.2 Σχέση LGD και Δείκτη LTV

Ορισμός δείκτη LTV(Loan to Value Ratio)

Ο Δείκτης Δανείου προς Αξία (LTV) ορίζεται ως η σχέση μεταξύ του εκκρεμούς ποσού του δανείου L_i και της αξίας του εξασφαλισμένου περιουσιακού στοιχείου C_i σε κάποια χρονική στιγμή. Ο τύπος είναι ο εξής:

$$LTV_i := \frac{L_i}{C_i} \in (0,1]$$

όπου $0 < L_i \leq C_i$ και $i=1,...n$ είναι το δάνειο i.

Ο δείκτης LTV εκφράζει το ποσοστό του περιουσιακού στοιχείου που χρηματοδοτείται μέσω του δανείου. Για παράδειγμα αν το LTV είναι ίσο με 70% σημαίνει ότι το 70% της αξίας του περιουσιακού στοιχείου καλύπτεται από δάνειο και το υπόλοιπο 30% καλύπτεται από προκαταβολές. (Greve and Hahnenstein, n.d.)

Επιπτώσεις του LTV στο LGD

Η κατανομή του λόγου δανείου προς αξία (LTV) επηρεάζει σημαντικά την απώλεια λόγω αθέτησης (LGD) ενός χαρτοφυλακίου δανείων:

- Γενική αρχή: Η κατανομή των λόγων LTV σε ένα χαρτοφυλάκιο περιέχει όλες τις πληροφορίες σχετικά με την πιθανή υπο- και υπερ-εξασφάλιση, οι οποίες είναι απαραίτητες για τον υπολογισμό των LGD υπό πίεση σε επίπεδο δανείου.
- Μη γραμμικότητα: Η μη γραμμικότητα της συνάρτησης LGD στο LTV τονίζει την εξάρτηση του μέσου LGD του χαρτοφυλακίου από το σχήμα της κατανομής LTV. Ειδικότερα, η ασύμμετρη μεταχείριση της υπερ- και υπο-εξασφάλισης στο LGD ενός δανείου, σημαίνει ότι η υπερ-εξασφάλιση δεν αντικατοπτρίζεται στην LGD, ενώ η υπό-εξασφάλιση αντικατοπτρίζεται.
- Διασπορά LTV: Η ευαισθησία του LGD του χαρτοφυλακίου στο στρες στα ποσοστά ανάκτησης μειώνεται με την αύξηση της διασποράς της κατανομής LTV. Δηλαδή, όταν υπάρχει μεγαλύτερη διακύμανση στα LTV των δανείων, το χαρτοφυλάκιο είναι λιγότερο ευαίσθητο στις αλλαγές στα ποσοστά ανάκτησης.

- Επίδραση της σοβαρής πίεσης: Όσο περισσότερη πίεση ασκείται στο χαρτοφυλάκιο μέσω χαμηλότερου ποσοστού ανάκτησης, τόσο λιγότερο σημαντική γίνεται η διασπορά των λόγων LTV. Όταν η υποτιθέμενη ανάκτηση είναι ίση με το ελάχιστο LTV όλων των δανείων υπό εξέταση, η επιρροή των διαφορών στις κατανομές LTV στο μέσο LGD του χαρτοφυλακίου εξαφανίζεται εντελώς.

Υποσημείωση: Στον υπολογισμό μας μπορούμε να χρησιμοποιήσουμε και τον αντίστροφο δείκτη του LTV, το VtL.

2.2.3 Προβλέψεις LGD και Τεχνικές Μηχανικής Μάθησης

Οι παραδοσιακές μέθοδοι πρόβλεψης έχουν αποδειχθεί ότι έχουν περιορισμούς στο βάθος αξιολόγησης του πιστωτικού κινδύνου και των επιπτώσεων του κλίματος στα χαρτοφυλάκια. Οι περιορισμοί είναι οι εξής (Nimmala 2024):

- Εξάρτηση από ιστορικά δεδομένα: Οι παραδοσιακές μέθοδοι βασίζονται κυρίως στα ιστορικά δεδομένα για την ανάλυση και πρόβλεψη μελλοντικών τάσεων και αποτελεσμάτων. Αυτό σημαίνει ότι η ικανότητά τους να προβλέψουν πρωτόγνωρα γεγονότα ή αλλαγές στις σχέσεις μεταξύ των μεταβλητών είναι περιορισμένη.
- Χρήση γραμμικών μοντέλων: Μοντέλα όπως η γραμμική παλινδρόμηση υποθέτουν γραμμικές συνδέσεις μεταξύ των μεταβλητών. Αυτή η υπόθεση ενδέχεται να μην είναι ακριβής για την αποτύπωση των περίπλοκων και μη γραμμικών αλληλεπιδράσεων μεταξύ των κλιματικών παραγόντων και των οικονομικών αποτελεσμάτων. Για παράδειγμα, η επίδραση ενός ακραίου καιρικού φαινομένου στην πιθανότητα αθέτησης ενός δανειολήπτη μπορεί να είναι μη γραμμική.
- Περιορισμένη ικανότητα ενσωμάτωσης μελλοντικών σεναρίων: Οι παραδοσιακές μέθοδοι έχουν περιορισμούς στην ικανότητα τους να ενσωματώνουν μελλοντικά κλιματικά σενάρια λόγω της αβεβαιότητας αυτών των κλιματικών εξελίξεων, καθιστώντας τες πιο αδύναμες στην προγνωστική τους συμπεριφορά στη σχέση πιστωτικού κινδύνου και κλίματος.
- Έλλειψη λεπτομερών πληροφοριών: Οι παραδοσιακές μέθοδοι συνήθως κάνουν ανάλυση σε συγκεντρωτικό επίπεδο με κίνδυνο να αγνοηθούν οι διάφορες ευπάθειες που αντιμετωπίζουν οι δανειολήπτες σε ατομικό επίπεδο.

Για αυτούς τους λόγους, υπάρχει μια αυξημένη ανάγκη να αξιοποιηθούν πιο προηγμένα και ευέλικτα εργαλεία, όπως η μηχανική μάθηση που γεμίζει τα κενά μεταξύ του πιστωτικού κινδύνου και τις επιπτώσεις της κλιματικής αλλαγής. Η μηχανική μάθηση δίνει τη δυνατότητα να αξιοποιήσει μεγάλα σύνολα δεδομένων και να ανακαλύψει σύνθετες μη γραμμικές σχέσεις ενσωματώνοντας μελλοντικά κλιματικά σενάρια και παρέχοντας πιο ολοκληρωμένες προβλέψεις.

2.2.3.1 ΤΕΧΝΙΚΕΣ ΜΗΧΑΝΙΚΗΣ ΜΑΘΗΣΗΣ

- Νευρωνικά Δίκτυα: Αυτά τα μοντέλα βαθιάς μάθησης είναι ισχυρά εργαλεία για τη μοντελοποίηση σύνθετων, μη γραμμικών σχέσεων. Τα Συνελικτικά Νευρωνικά Δίκτυα (CNNs) και τα Αναδρομικά Νευρωνικά Δίκτυα (RNNs) είναι ειδικές αρχιτεκτονικές που μπορούν να είναι αποτελεσματικές για την ανάλυση χωρικών και χρονικών πρότυπων, τα οποία είναι συχνά σημαντικά στη μοντελοποίηση του κλιματικού κινδύνου. (Nimmala 2024)
- Gradient Boosting Machines (GBMs): Αυτές οι μέθοδοι μάθησης συνόλων, όπως το XGBoost και το LightGBM, συνδυάζουν πολλούς "ασθενείς" εκπαιδευόμενους (συνήθως δέντρα αποφάσεων) για να δημιουργήσουν ένα ισχυρό προγνωστικό μοντέλο. Είναι γνωστά για την ικανότητά τους να χειρίζονται μεγάλα σύνολα δεδομένων, να αντιμετωπίζουν ελλιπείς τιμές και να καταγράφουν σύνθετες σχέσεις μεταξύ των μεταβλητών. Τα GBMs έχουν εφαρμοστεί με επιτυχία σε διάφορες εργασίες μοντελοποίησης πιστωτικού κινδύνου και μπορούν να προσαρμοστούν για την ενσωμάτωση παραγόντων κλιματικού κινδύνου. (Nimmala 2024)
- Random Forests: Αυτή είναι μια άλλη μέθοδος μάθησης συνόλων που κατασκευάζει πολλαπλά δέντρα αποφάσεων και συνδυάζει τις προβλέψεις τους για τη βελτίωση της ακρίβειας και τη μείωση της υπερπροσαρμογής. Μπορούν να επεξεργαστούν τόσο αριθμητικές όσο και κατηγορικές μεταβλητές, παρέχοντας παράλληλα ένα μέτρο της σημασίας των χαρακτηριστικών. Αυτό το μέτρο μπορεί να συμβάλει στην κατανόηση της σχετικής επίδρασης διαφόρων παραγόντων κλιματικού κινδύνου. (Nimmala 2024)

2.2.3.2 Data Engineering

Η αξιολόγηση του πιστωτικού κινδύνου με βάση την κλιματική αλλαγή έχει ως βασικό παράγοντα τη διαχείριση και την επεξεργασία των δεδομένων. Πιο συγκεκριμένα τα μοντέλα χρειάζονται μεγάλα και ποιοτικά δεδομένα προκειμένου να αναδειχθούν οι σύνθετες σχέσεις μεταξύ των κλιματικών μεταβλητών και των μετρήσεων πιστωτικού κινδύνου.

Για τη δημιουργία αυτών των δεδομένων χρειάζονται τρεις βασικές πηγές:

- Μεταβλητές κλιματικών σεναρίων: Τα δεδομένα αυτά είναι απαραίτητα στην καταγραφή πιθανών μελλοντικών επιπτώσεων της κλιματικής αλλαγής στον πιστωτικό κίνδυνο. Τα σενάρια αυτά πρέπει να καλύπτουν ένα ευρύ φάσμα πιθανών εξελίξεων, λαμβάνοντας υπόψη τόσο τους κινδύνους μετάβασης (π.χ. αλλαγές στην πολιτική και την τεχνολογία) όσο και τους φυσικούς κινδύνους (π.χ. πλημμύρες, ακραία καιρικά φαινόμενα). (Nimmala 2024)
- Οικονομικά στοιχεία δανειοληπτών: Για την αξιολόγηση της πιστοληπτικής ικανότητας των δανειοληπτών, είναι κρίσιμη η χρήση ιστορικών και προβλεπόμενων οικονομικών δεδομένων, όπως τα έσοδα, τα λειτουργικά έξοδα και οι ταμειακές ροές. Σε συνδυασμό με τις κλιματικές μεταβλητές, αυτά τα δεδομένα βοηθούν στο να κατανοήσουμε πώς μπορεί να επηρεαστεί η οικονομική κατάσταση των δανειοληπτών λόγω των αλλαγών που επιβάλλει το κλίμα.
- Χαρακτηριστικά περιουσιακών στοιχείων: Λεπτομερείς πληροφορίες για τα φυσικά χαρακτηριστικά και την τοποθεσία των ακινήτων που χρησιμοποιούνται ως ασφάλεια παιζουν επίσης σημαντικό ρόλο. Τέτοια παραδείγματα περιλαμβάνουν δεδομένα σχετικά με το υψόμετρο, τον κίνδυνο πλημμύρας και την ενεργειακή απόδοση του ακινήτου. Αυτές οι πληροφορίες συμβάλλουν στην εκτίμηση του βαθμού έκθεσης των περιουσιακών στοιχείων απέναντι στις επιπτώσεις της κλιματικής αλλαγής.

Η ποιότητα και η πληρότητα των δεδομένων είναι ουσιαστικές για την ανάπτυξη αξιόπιστων μοντέλων πιστωτικού κινδύνου. Αυτό σημαίνει ότι απαιτούνται διαδικασίες καθαρισμού, κανονικοποίησης και συμπλήρωσης των δεδομένων ώστε να αντιμετωπιστούν τυχόν ελλείψεις ή ασυνέπειες.

Οι προκλήσεις που δημιουργούνται από την αβεβαιότητα στα κλιματικά μοντέλα και τους περιορισμούς των ιστορικών δεδομένων, σε συνδυασμό με τις ασυνέπειες στη συλλογή των κλιματικών δεδομένων, απαιτούν συνεχόμενες προσπάθειες βελτίωσης της συλλογής, δημοσιοποίησης και τυποποίησής τους. Συνολικά, για να αξιοποιηθούν επιτυχώς οι τεχνικές μηχανικής μάθησης στην αξιολόγηση του πιστωτικού κινδύνου που συνδέεται με την κλιματική αλλαγή, είναι απαραίτητη η σωστή διαχείριση, επεξεργασία και διασφάλιση της ποιότητας των δεδομένων. Αυτό αποτελεί το θεμέλιο για την ανάπτυξη αξιόπιστων και αποτελεσματικών μοντέλων, τα οποία θα βοηθήσουν τους χρηματοπιστωτικούς φορείς να κατανοήσουν και να αντιμετωπίσουν καλύτερα τις επιπτώσεις των φυσικών κινδύνων.

2.4 Προσομοιώσεις Monte Carlo

Η μέθοδος Monte Carlo αποτελεί μια υπολογιστική τεχνική που βασίζεται στην τυχαία δειγματοληψία για την επίλυση πολύπλοκων προβλημάτων που συχνά δεν επιλύονται εύκολα με αναλυτικές μεθόδους. Με άλλα λόγια, αντί να αναζητήσουμε μια κλειστή λύση, η προσέγγιση αυτή χρησιμοποιεί επαναληπτικές προσομοιώσεις για να εξετάσει ένα ευρύ φάσμα πιθανοτήτων και να παράγει μια κατανομή πιθανών αποτελεσμάτων. (“Monte Carlo Risk Estimation,” n.d.)

Μια προσομοίωση Monte Carlo περιλαμβάνει ορισμένα θεμελιώδη στοιχεία που καθορίζουν το πρόβλημα και τη διαδικασία:

1. **Μεταβλητές Εισόδου (Τυχαίες Εισόδους):** Αυτές είναι οι αβέβαιες παράμετροι του μοντέλου, εκφρασμένες μέσω κατανομών πιθανότητας (π.χ. κανονική, ομοιόμορφη). Κάθε μεταβλητή μπορεί να λάβει μια σειρά από πιθανές τιμές σύμφωνα με την καθορισμένη κατανομή της. Κατά τη διάρκεια της προσομοίωσης, από αυτές τις κατανομές αντλούνται τυχαία δείγματα για να τροφοδοτηθεί το μοντέλο. Για παράδειγμα, σε μια προσομοίωση κόστους έργου, οι είσοδοι μπορεί να περιλαμβάνουν τυχαίες τιμές μονάδων κόστους, διάρκειες ή επίπεδα ζήτησης, βασισμένα σε ιστορικά δεδομένα. (“Monte Carlo Simulation Overview,” n.d.)
2. **Μαθηματικό Μοντέλο (Διαδικασία):** Αυτό αντιπροσωπεύει τη ντετερμινιστική λειτουργία ή τον αλγόριθμο που μετατρέπει τις εισόδους σε τελικό αποτέλεσμα. Το μοντέλο καταγράφει τη σχέση μεταξύ των μεταβλητών εισόδου και του αποτελέσματος. Μπορεί να είναι τόσο απλό όσο ένας τύπος (π.χ., Κέρδος = Έσοδα – Έξοδα) ή τόσο σύνθετο όσο μια λεπτομερής φυσική προσομοίωση. Σε κάθε δοκιμαστική εκτέλεση του μοντέλου, ένα σετ τυχαίων εισόδων οδηγεί στον υπολογισμό ενός αποτελέσματος. (“Monte Carlo Simulation Overview,” n.d.)
3. **Μεταβλητή Εξόδου (Αποτέλεσμα):** Η έξοδος είναι το μέγεθος ή ο δείκτης που μας ενδιαφέρει – για παράδειγμα, συνολικό κέρδος, χρονική ολοκλήρωση ενός έργου ή ποσοστό αστοχίας ενός συστήματος. Κάθε δοκιμή παράγει ένα αποτέλεσμα, και με την εκτέλεση πολυάριθμων δοκιμών, αυτά τα αποτελέσματα συνθέτουν μια κατανομή πιθανοτήτων. Συνήθως, οι αναλυτές χρησιμοποιούν ιστογράμματα ή συνοπτικές στατιστικές (μέσος όρος, διακύμανση, ποσοστιαίες τιμές) για να κατανοήσουν το εύρος και τις πιθανότητες διαφορετικών εκβάσεων. (“Monte Carlo Risk Estimation,” n.d.)

Συχνές Μεθοδολογίες στις Προσομοιώσεις Monte Carlo

Οι μέθοδοι Monte Carlo περιλαμβάνουν ποικίλες τεχνικές, οι οποίες επιλέγονται ανάλογα με το πρόβλημα που θέλουμε να λύσουμε και το επιθυμητό επίπεδο ακρίβειας:

1. Βασική Προσομοίωση Monte Carlo (Crude Monte Carlo)

Η απλούστερη προσέγγιση υιοθετεί τα εξής βήματα: παράγουμε ανεξάρτητα τυχαία δείγματα και υπολογίζουμε τα αποτελέσματα απευθείας. Αυτή η μέθοδος χρησιμοποιείται για την ολοκλήρωση τύπων Monte Carlo – όπου εκτιμούμε ένα ολοκλήρωμα ή μια αναμενόμενη τιμή υπολογίζοντας τον μέσο όρο των τυχαίων δειγμάτων – καθώς και για την προσομοίωση συστημάτων με σκοπό την πρόβλεψη αποτελεσμάτων (π.χ., οικονομικές προβλέψεις, εξελίξεις έργων). Παρά την απλότητά της, η βασική μέθοδος μπορεί να εφαρμοστεί σε ένα ευρύ φάσμα προβλημάτων (όπως βελτιστοποίηση, αριθμητική ολοκλήρωση και εκτίμηση πιθανοτήτων) μέσω της brute-force δειγματοληψίας. (“Monte Carlo Wiki,” n.d.)

2. Μαρκοβιανές Αλυσίδες Monte Carlo (MCMC)

Οι MCMC αποτελούν μια κατηγορία αλγορίθμων που επιτρέπουν τη δειγματοληψία από πολύπλοκες κατανομές πιθανοτήτων μέσω της δημιουργίας μιας αλληλουχίας εξαρτημένων δειγμάτων, γνωστής και ως Μαρκοβιανή αλυσίδα. Τεχνικές όπως το Metropolis–Hastings και το Gibbs sampling διευκολύνουν τη διεξαγωγή προσομοιώσεων σε υψηλών διαστάσεων χώρους (συχνά για την ανάλυση των Bayesian a posteriori κατανομών), όπου η άμεση δειγματοληψία είναι δύσκολη (“Monte Carlo Wiki,” n.d.). Οι μέθοδοι αυτές είναι ιδιαίτερα διαδεδομένες σε τομείς όπως η μηχανική μάθηση και η φυσική (π.χ., προσομοίωση συστημάτων σωματιδίων), επιτρέποντας την προσέγγιση κατανομών και αναμενόμενων τιμών όταν η ανεξάρτητη δειγματοληψία δεν είναι πρακτική.

3. Τεχνικές Μείωσης της Διακύμανσης

Αυτές οι τεχνικές αποσκοπούν στην αύξηση της ακρίβειας των εκτιμήσεων Monte Carlo χωρίς την απλή αύξηση του αριθμού διεξαγόμενων δοκιμών. Στόχος είναι να μειωθεί η διακύμανση των αποτελεσμάτων, προσφέροντας πιο σταθερές εκτιμήσεις για τον ίδιο αριθμό δειγμάτων. Σε αυτό το πλαίσιο, χρησιμοποιούνται μέθοδοι όπως:

- Αντεπαναλαμβανόμενες Μεταβλητές (Antithetic Variates)
- Μεταβλητές Ελέγχου (Control Variates)
- Δειγματοληψία Βάσει Σημαντικότητας (Importance Sampling)
- Στρωματοποιημένη Δειγματοληψία (Stratified Sampling, π.χ., Latin Hypercube Sampling)

Για παράδειγμα, η στρωματοποιημένη δειγματοληψία εξασφαλίζει την ομοιόμορφη κάλυψη του χώρου εισόδου και συχνά επιτυγχάνει την ίδια ακρίβεια με λιγότερα δείγματα. Αντίστοιχα, η δειγματοληψία βάσει σημαντικότητας δίνει μεγαλύτερη βαρύτητα σε δείγματα από κρίσιμες περιοχές (π.χ., στα άκρα μιας κατανομής για την ανάλυση κινδύνου), επιτρέποντας την καλύτερη σύλληψη σπάνιων γεγονότων. Με την εφαρμογή αυτών των τεχνικών, η προσομοίωση γίνεται πιο στατιστικά αποδοτική, δηλαδή παράγει μικρότερο σφάλμα για το ίδιο κόστος υπολογισμού.

4. Μέθοδοι Quasi-Monte Carlo

Σε αντίθεση με τη χρήση ψευδοτυχαίων αριθμών, οι μέθοδοι Quasi-Monte Carlo βασίζονται σε quasi-τυχαίες ακολουθίες (γνωστές και ως ακολουθίες χαμηλής διαφοράς), οι οποίες είναι προδιαγεγραμμένες και έχουν σχεδιαστεί ώστε να καλύπτουν ομοιόμορφα το χώρο εισόδου. Τέτοιες μέθοδοι, όπως αυτές που χρησιμοποιούν τις ακολουθίες Sobol' ή Halton, συχνά συγκλίνουν ταχύτερα από την παραδοσιακή τυχαία δειγματοληψία σε προβλήματα ολοκλήρωσης σε υψηλές διαστάσεις. Το μειονέκτημα είναι ότι ο υπολογισμός του σφάλματος γίνεται λιγότερο άμεσος, καθώς τα δείγματα δεν είναι πραγματικά τυχαία, αλλά στην πράξη οι μέθοδοι αυτές μπορούν να βελτιώσουν σημαντικά την ακρίβεια σε ορισμένες εφαρμογές, ιδιαίτερα στον τομέα των οικονομικών και του μηχανικού σχεδιασμού. ("Monte Carlo Wiki," n.d.)

5. Σειριακές Προσομοιώσεις Monte Carlo και Μέθοδοι Σωματιδίων

Σε δυναμικά προβλήματα ή σε εφαρμογές φιλτραρίσματος (όπως η επεξεργασία σήματος ή η παρακολούθηση), χρησιμοποιούνται σειριακές προσομοιώσεις Monte Carlo, γνωστές και ως φύλτρα σωματιδίων. Σε αυτές τις μεθόδους, ένα σύνολο τυχαίων δειγμάτων (ή "σωματίδια") μεταφέρεται διαδοχικά στο χρόνο, επιτρέποντας την εκτίμηση της εξελισσόμενης κατάστασης του συστήματος. Αυτή η εξειδικευμένη προσέγγιση είναι ιδιαίτερα χρήσιμη σε προβλήματα όπου η κατάσταση του συστήματος ενημερώνεται συνεχώς με νέα δεδομένα – ένα φαινόμενο που παρατηρείται συχνά στην Bayesian φιλτραρίσματος, στον εντοπισμό ρομποτικών συστημάτων και σε παρόμοιες εφαρμογές.

Κάθε μία από αυτές τις μεθοδολογίες έχει σχεδιαστεί για να αντιμετωπίζει συγκεκριμένες προκλήσεις: η MCMC είναι κατάλληλη για πολύπλοκες κατανομές, οι μέθοδοι Quasi-Monte Carlo προσφέρουν ταχύτερη σύγκλιση, ενώ οι τεχνικές μείωσης της διακύμανσης βελτιώνουν την αποτελεσματικότητα χωρίς να απαιτούν τεράστιο αριθμό δοκιμών. Στην πράξη, οι προσομοιώσεις Monte Carlo συχνά ενσωματώνουν πολλαπλές μεθόδους ταυτόχρονα – για παράδειγμα, συνδυάζοντας τεχνικές μείωσης διακύμανσης με μια βασική προσομοίωση ή χρησιμοποιώντας quasi-τυχαία δειγματοληψία σε μοντέλα οικονομικών κινδύνων.

Χρήση στη διαχείριση οικονομικού ρίσκου

Η διαχείριση κινδύνου περιλαμβάνει την κατανόηση ενός μεγάλου εύρους των πιθανών εκβάσεων σε αβέβαιες καταστάσεις, κατάσταση στην οποία οι Monte Carlo προσομοιώσεις είναι εξαιρετικά χρήσιμες. Γενικότερα η Monte Carlo ανάλυση στη διαχείριση κινδύνου έχει να κάνει με τη μετατροπή των what if σεναρίων σε ένα ποσοτικοποιημένο προφίλ κινδύνου. Συμβάλλει στην αξιολόγηση του πώς οι διαφορετικοί παράγοντες έχουν αντίκτυπο στο συνολικό αποτέλεσμα. Με την ταυτόχρονη αλλαγή πολλαπλών μεταβλητών εντός των ορίων της αβεβαιότητάς τους, η προσομοίωση αναπαράγει ένα ευρύ φάσμα από πιθανότητες και από εκβάσεις.

Σε επίπεδο ασφαλιστικού και επιχειρησιακού κινδύνου, οι εταιρείες χρησιμοποιούν τις μεθόδους Monte Carlo για να προσομοιώσουν ένα τεράστιο εύρος πιθανών συμβάντων (π.χ., φυσικές καταστροφές, ατυχήματα) και να διασφαλίσουν ότι διατηρούν επαρκές κεφάλαιο. (“Monte Carlo Simulation Overview,” n.d.)

Ένα βασικό προϊόν της ανάλυσης κινδύνου μέσω Monte Carlo είναι το προφίλ κινδύνου – μία καμπύλη πιθανοτήτων που απεικονίζει τις πιθανότητες για διάφορα επίπεδα ζημιών ή αποτυχιών. Αυτή η «εικόνα του κινδύνου» επιτρέπει σε όλους τους ενδιαφερόμενους όχι μόνο να κατανοήσουν το αναμενόμενο αποτέλεσμα, αλλά και να δουν τις ακραίες περιπτώσεις, παρέχοντας έτσι μια ολοκληρωμένη βάση για αποφάσεις που στηρίζονται στον κίνδυνο.

Συνολικά, οι προσομοιώσεις Monte Carlo αποτελούν θεμέλιο της ποσοτικής διαχείρισης κινδύνου, μετατρέποντας την αβεβαιότητα σε ένα χειροπιαστό σύνολο πιθανοτήτων που μπορούν να αναλυθούν και να διαχειριστούν με τρόπο που βελτιώνει την αξιοπιστία και την αποτελεσματικότητα στη λήψη αποφάσεων. (“Monte Carlo Simulation Overview,” n.d.)

Τέλος, στην κλιματική έρευνα οι προσομοιώσεις Monte Carlo αποτελούν ένα πολύτιμο εργαλείο για την υποστήριξη αναλύσεων πολιτικών και διαχείρισης κινδύνου. Οι ερευνητές προσομοιώνουν πληθώρα πιθανών μελλοντικών σεναρίων, λαμβάνοντας υπόψη μεταβαλλόμενες κοινωνικοοικονομικές υποθέσεις, διαφορετικές πορείες εκπομπών και την αντίδραση του κλίματος. Με αυτόν τον τρόπο, εξετάζουν πιθανούς κινδύνους όπως η πιθανότητα να ξεπεραστούν κρίσιμα όρια θερμότητας. Με άλλα λόγια, η προσομοίωση μπορεί να δείξει πόσο πιθανό είναι να φτάσουμε ή να υπερβούμε συγκεκριμένα όρια θερμοκρασίας. Ένας άλλος πιθανός κίνδυνος είναι η κατανομή των οικονομικών ζημιών, δηλαδή μέσω της ανάλυσης μπορεί να αποκαλυφθεί πόσο μεγάλη θα ήταν η συνολική οικονομική επίπτωση σε διαφορετικά σενάρια. Αυτές οι αναλύσεις βοηθούν τους υπεύθυνους λήψης αποφάσεων να διαμορφώσουν στρατηγικές ικανές να αντέξουν ένα ευρύ φάσμα μελλοντικών αβεβαιοτήτων. Έτσι, αντί να βασίζονται αποκλειστικά σε μία μόνο εκτίμηση, οι σχεδιαστές πολιτικής και οι διαχειριστές κινδύνου αποκτούν μια πιο ολοκληρωμένη εικόνα του μέλλοντος. (“Monte Carlo Wiki,” n.d.)

2.5 Κλιματική Αλλαγή και Επιπτώσεις στον Πιστωτικό Κίνδυνο

Η κλιματική αλλαγή αποτελεί μία από τις μεγαλύτερες προκλήσεις που αντιμετωπίζει σήμερα το παγκόσμιο οικονομικό σύστημα. Οι φυσικές επιπτώσεις της κλιματικής αλλαγής όπως οι πλημμύρες, οι πυρκαγιές και τα ακραία καιρικά φαινόμενα, επηρεάζουν άμεσα την αξία των ακινήτων, αποτελώντας κρίσιμο παράγοντα για τον πιστωτικό κίνδυνο των τραπεζών. Πιο συγκεκριμένα, οι τράπεζες αναγκάζονται όλο και περισσότερο να μελετούν τις επιπτώσεις της κλιματικής αλλαγής στα χαρτοφυλάκια δανείων τους, καθώς αυτά τα φαινόμενα μπορούν να υποτιμήσουν τα υποκείμενα περιουσιακά στοιχεία ή σε μείωση της πιστοληπτικής ικανότητας των δανειοληπτών.

Τα μοντέλα ενσωμάτωσης κλιματικών παραγόντων στον πιστωτικό κίνδυνο βασίζονται στην ιδέα ότι οι φυσικοί κίνδυνοι μπορούν να μετατραπούν σε ποσοτικοποιήσιμα μέτρα που αντικατοπτρίζουν την πιθανή ζημιά στα ακίνητα (Mbanyele and Muchenje 2022). Τα μοντέλα αυτά συχνά διαχωρίζουν τους κινδύνους σε δύο κατηγορίες:

- **Φυσικοί κίνδυνοι:** Οι οποίοι αφορούν ακραία καιρικά φαινόμενα όπως πλημμύρες, πυρκαγιές και σεισμοί. Αυτοί οι κίνδυνοι εκτιμώνται μέσω παραμέτρων όπως η συχνότητα και η ένταση του φαινομένου.
- **Μεταβατικοί κίνδυνοι:** Οι οποίοι συνδέονται με τη μετάβαση σε ένα περιβάλλον χαμηλών εκπομπών άνθρακα και περιλαμβάνουν αλλαγές στην πολιτική, την τεχνολογία και την καταναλωτική συμπεριφορά.

Στην παρούσα έρευνα, η έμφαση δίνεται κυρίως στους φυσικούς κινδύνους, δεδομένου ότι οι επιπτώσεις τους (π.χ. ζημιές στα ακίνητα) επηρεάζουν άμεσα τους δείκτες πιστωτικού κινδύνου όπως το Loss Given Default (LGD).

3 ΜΕΘΟΔΟΛΟΓΙΑ(ΜETHODOLOGY)

3.1 ΣΥΛΛΟΓΗ ΚΑΙ ΠΡΟΕΠΕΞΕΡΓΑΣΙΑ ΔΕΔΟΜΕΝΩΝ

Η πρώτη και θεμελιώδης φάση της μεθοδολογίας αφορά τη συλλογή, την προετοιμασία και την επεξεργασία των δεδομένων που απαιτούνται για την ανάπτυξη ολοκληρωμένων μοντέλων εκτίμησης πιστωτικού κινδύνου στεγαστικών δανείων στην Ελλάδα. Τα δεδομένα που χρησιμοποιήθηκαν χωρίζονται σε τρεις κατηγορίες:

- Δεδομένα ακινήτων από πόλεις της Ελλάδας
- Μετεωρολογικά και κλιματικά δεδομένα
- Δεδομένα φυσικών κινδύνων (σεισμοί, πυρκαγιές, πλημμύρες)

3.1.1 Δεδομένα ακινήτων και στεγαστικών δανείων

Αρχικά τα δεδομένα τα οποία χρησιμοποιήσαμε, αντλήθηκαν από το Kaggle και πιο συγκεκριμένα από το project “House Price Prediction from Listings Greece” (<https://www.kaggle.com/code/argyrisanastopoulos/house-price-prediction-from-listings-greece/input>) το οποίο περιέχει πληροφορίες σχετικά με τιμές, χαρακτηριστικά και γεωγραφικές πληροφορίες ακινήτων στις δύο μεγαλύτερες πόλεις της Ελλάδας, την Αθήνα και τη Θεσσαλονίκη.

Έτσι, προκειμένου να διευρύνουμε την έρευνά μας στην υπόλοιπη Ελλάδα και να επιτευχθεί μια πιο αντιπροσωπευτική γεωγραφική κάλυψη, προσθέσαμε ποικίλες πόλεις από όλη την Ελλάδα, όπως:

- Πελοπόννησος: Πάτρα, Αγρίνιο, Σπάρτη, Ναύπακτος, Τρίπολη, Πύργος
- Κρήτη: Ηράκλειο, Χανιά, Ρέθυμνο, Λασίθι
- Ήπειρος: Ιωάννινα, Άρτα, Πρέβεζα, Ηγουμενίτσα
- Θεσσαλία: Λάρισα, Βόλος, Τρίκαλα, Καρδίτσα
- Μακεδονία: Κιλκίς, Καβάλα, Βέροια, Κοζάνη, Σέρρες, Δράμα, Φλώρινα, Πολύγυρος
- Ιόνια Νησιά: Κέρκυρα, Λευκάδα, Κεφαλονιά, Ζάκυνθος, Ιθάκη
- Στερεά Ελλάδα: Λαμία, Χαλκίδα, Θήβα, Άμφισσα
- Θράκη: Αλεξανδρούπολη, Κομοτηνή, Ξάνθη
- Αιγαίο: Μύκονος, Σαντορίνη, Πάρος, Νάξος, Τήνος, Χίος, Λήμνος, Ικαρία, Σκύρος, Ρόδος, Κως, Σάμος, Λέσβος

Αυτό επιτεύχθηκε με τη μέθοδο της επέκτασης δεδομένων (data augmentation). Η διαδικασία αυτή είναι συνήθης στη μηχανική μάθηση για την αύξηση του μεγέθους και της ποικιλότητας των δεδομένων. Πιο συγκεκριμένα, πέρα από την επιλογή των περιοχών και των πόλεων κάναμε τα εξείς:

- **Γεωγραφική Αντιστοίχιση και Δημιουργία Συνθετικών Δεδομένων:** Για κάθε πόλη της εκτεταμένης λίστας, ανατέθηκαν οι αντίστοιχες γεωγραφικές συντεταγμένες (latitude και longitude) βάσει του OpenStreetMap (OSM), μια συνεργατική και ανοικτού κώδικα πλατφόρμα χαρτών. Στη συνέχεια, μέσω τυχαίας δειγματοληψίας και με την προσθήκη μικρού τυχαίου θορύβου στα βασικά χαρακτηριστικά (π.χ. τιμή ακινήτου, τετραγωνικά μέτρα, έτος κατασκευής), δημιουργήθηκαν συνθετικά δεδομένα που προσομοιώνουν την ποικιλία που υπάρχει στην ελληνική αγορά ακινήτων.
- **Ενοποίηση Δεδομένων:** Τα συνθετικά δεδομένα συγχωνεύτηκαν με τα αρχικά δεδομένα από Αθήνα και Θεσσαλονίκη, παράγοντας ένα ενιαίο dataset. Η διαδικασία αυτή όχι μόνο διευρύνει το μέγεθος του δείγματος, αλλά και ενισχύει την αντιπροσωπευτικότητα της ανάλυσης πιστωτικού κινδύνου, καθώς περιλαμβάνει στοιχεία από διάφορες περιοχές της χώρας.

Τα βασικά χαρακτηριστικά των ακινήτων περιγράφονται από τις στήλες των δεδομένων μας που προϋπήρχαν από το project του Kaggle είναι τα εξής:

- location_name: Το όνομα της πόλης όπου βρίσκεται το ακίνητο.
- location_region: Η περιοχή στην οποία ανήκει το ακίνητο (π.χ. Πελοπόννησος, Κρήτη κ.λπ.).
- res_type: Ο τύπος του ακινήτου (π.χ. διαμέρισμα, μονοκατοικία, μεζονέτα).
- res_address: Η διεύθυνση της τοποθεσίας του ακινήτου.
- res_price: Η τιμή πώλησης του ακινήτου.
- res_price_sqr: Η τιμή ανά τετραγωνικό μέτρο, που υπολογίζεται διαιρώντας την τιμή του ακινήτου με το εμβαδόν του.
- res_sqr: Το συνολικό εμβαδόν του ακινήτου σε τετραγωνικά μέτρα.
- construction_year: Το έτος κατασκευής του ακινήτου.
- levels: Ο συγκεκριμένος όροφος (π.χ. ισόγειο, 2ος όροφος) στο οποίο βρίσκεται το ακίνητο.
- bedrooms: Ο αριθμός των υπνοδωματίων.
- bathrooms: Ο αριθμός των μπάνιων.
- energyclass: Η ενεργειακή κλάση του ακινήτου, που αντανακλά την ενεργειακή απόδοση.

- auto_heating: Δυαδικός δείκτης (0 ή 1) που υποδηλώνει αν το ακίνητο διαθέτει αυτόματο σύστημα θέρμανσης.
- solar: Δυαδικός δείκτης που δείχνει αν υπάρχουν εγκαταστάσεις ηλιακής ενέργειας.
- cooling: Δυαδικός δείκτης για την ύπαρξη συστημάτων ψύξης.
- safe_door: Δυαδικός δείκτης που υποδηλώνει αν το ακίνητο είναι εξοπλισμένο με ασφαλή πόρτα.
- gas: Δυαδικός δείκτης που δείχνει αν το ακίνητο έχει πρόσβαση σε δίκτυο φυσικού αερίου.
- fireplace: Δυαδικός δείκτης που υποδηλώνει την ύπαρξη τζακιού.
- furniture: Δυαδικός δείκτης που δείχνει αν το ακίνητο είναι επιπλωμένο.
- student: Δυαδικός δείκτης που υποδηλώνει αν το ακίνητο απευθύνεται σε φοιτητές ή έχει χαρακτηριστικά που το καθιστούν ελκυστικό για τους φοιτητές.
- parking: Περιγραφή ή κατηγοριοποίηση της διαθέσιμης θέσης στάθμευσης (π.χ. ανοικτός χώρος, κλειστός χώρος, χωρίς θέση στάθμευσης).

3.1.2 Κλιματικά και μετεωρολογικά δεδομένα

Για την καταγραφή των κλιματικών και μετεωρολογικών συνθηκών που επηρεάζουν την αγορά ακινήτων και, κατ' επέκταση, τον πιστωτικό κίνδυνο στα στεγαστικά δάνεια, χρησιμοποιήσαμε δεδομένα από την πλατφόρμα CLIMPACT (<https://data.climpact.gr/en/dataset/497dc26d-45e0-4ad5-b8f3-5f8890f65129/resource/aef8997e-3b0f-4b02-9f59-0751a6093936>). Η CLIMPACT είναι το Εθνικό Δίκτυο Έρευνας για την Κλιματική Αλλαγή και τις Επιπτώσεις της, που εστιάζει στην Ελλάδα και την Κύπρο, μελετώντας τις κοινωνικές, οικονομικές και περιβαλλοντικές επιπτώσεις της κλιματικής αλλαγής.

Το σύνολο δεδομένων περιέχει ημερήσιες μετεωρολογικές παραμέτρους για τις πόλεις που χρησιμοποιήσαμε στα δεδομένα μας και για την περίοδο από το 2010 έως το 2023. Οι στήλες που χρησιμοποιήσαμε είναι οι εξής:

- T_mean: Αποτελεί τη μέση θερμοκρασία της ημέρας σε °C.
- T_max: Αποτελεί τη μέγιστη θερμοκρασία της ημέρας σε °C.
- T_min: Αποτελεί την ελάχιστη θερμοκρασία της ημέρας σε °C.
- Ac_R: Είναι η αθροιστική βροχόπτωση της ημέρας σε mm.
- WS_mean: Είναι η μέση ταχύτητα του ανέμου σε km/h.
- WG: Είναι η μέγιστη ριπή σε km/h.

Με αυτά τα δεδομένα ακολουθήσαμε τις εξής ενέργειες:

- Φιλτράρισμα για τους Καλοκαιρινές Μήνες: Εφαρμόστηκε φίλτρο για τους μήνες Ιούνιο έως Αύγουστο, ώστε να εξάγονται οι μέγιστες τιμές ταχύτητας ανέμου και ρυπίζις ανέμου κατά τη διάρκεια του καλοκαιριού λόγω της έντονης εμφάνισης πυρκαγιών τους καλοκαιρινούς μήνες.
- Εξαγωγή Συνοπτικών Στατιστικών: Υπολογίστηκαν μέσες τιμές θερμοκρασίας, μέγιστες και ελάχιστες τιμές θερμοκρασίας, καθώς και μέση βροχόπτωση. Αυτά τα συνοπτικά στοιχεία χρησιμεύουν ως βασικά input για την ανάλυση κινδύνου και για τη διαμόρφωση των κλιματικών σεναρίων στα επόμενα στάδια της μελέτης.

Συμπερασματικά, αυτή η διαδικασία επιτρέπει την εξαγωγή αξιόπιστων και συγκριτικών μετρήσεων για τις κλιματικές παραμέτρους, οι οποίες στη συνέχεια ενσωματώνονται στο κύριο dataset για την περαιτέρω ανάλυση του πιστωτικού κινδύνου.

3.1.3 Δεδομένα Φυσικών Κινδύνων (σεισμοί, πυρκαγιές, πλημμύρες)

Σε αυτήν την ενότητα, παρουσιάζουμε τα δεδομένα φυσικών κινδύνων που ενσωματώθηκαν στο συνολικό μας dataset, με στόχο την ακριβέστερη αξιολόγηση της έκθεσης των ακινήτων σε κινδύνους, όπως η σεισμική δραστηριότητα και οι πυρκαγιές.

Σεισμοί:

Τα δεδομένα σεισμών προέρχονται από το dataset “Greece's Earthquakes”, διαθέσιμο στο Kaggle (<https://www.kaggle.com/datasets/nickdoulos/greeces-earthquakes/data>). Το σύνολο αυτό περιέχει πληροφορίες όπως η ημερομηνία και ώρα εμφάνισης, το μέγεθος (magnitude), το βάθος και οι γεωγραφικές συντεταγμένες (latitude και longitude) των σεισμικών γεγονότων στην Ελλάδα καθημερινά από το 1965 έως και το 2023. Τα δεδομένα αυτά χρησιμοποιήθηκαν με σκοπό τον υπολογισμό του σεισμικού κινδύνου.

Πυρκαγιές:

Τα δεδομένα πυρκαγιών αντλήθηκαν από την επίσημη ιστοσελίδα της Πυροσβεστικής (<https://www.fireservice.gr/el/synola-dedomenon+temp>) με χρονολογία από το 2000 έως και το 2024. Το συγκεκριμένο σύνολο δεδομένων περιέχει πληροφορίες για την εμφάνιση, τη διάρκεια, την έκταση της φωτιάς καθώς επίσης και τον νομό με τις περιοχές που εμφανίστηκαν. Επιπλέον, οι πληροφορίες για τις ταχύτητες ανέμου, που αποτελούν σημαντικό παράγοντα στην εξάπλωση των πυρκαγιών, συλλέχθηκαν από το σύνολο δεδομένων του CLIMPACT. Η συνδυασμένη χρήση αυτών των πληροφοριών επιτρέπει τον υπολογισμό του δείκτη κινδύνου πυρκαγιάς (Wildfire Risk).

3.2 Υπολογισμός Κινδύνων και Παραμέτρων Δανείων

Σε αυτήν την ενότητα αναλύεται η σημασία των κινδύνων κλιματικής αλλαγής και φυσικών καταστροφών (ή διαφορετικά ρίσκων) που επηρεάζουν τα στεγαστικά δάνεια και παρουσιάζονται οι μαθηματικές φόρμουλες που χρησιμοποιήθηκαν για την εκτίμησή τους. Σε αυτό το πλαίσιο ο κίνδυνος ή αλλιώς ρίσκο (σεισμικό ρίσκο, ρίσκο πλημμύρας, κίνδυνος πυρκαγιάς) αναφέρεται στην πιθανότητα και το μέγεθος ζημιάς που μπορεί να προκύψει από φυσικές καταστροφές και επηρεάζει την ικανότητα ανάκτησης της αξίας του ακινήτου σε περίπτωση αθέτησης.

3.2.1 Σεισμικός Κίνδυνος (Earthquake Risk)

Ο σεισμικός κίνδυνος μετρά τη συχνότητα και την ένταση των σεισμών που έχουν συμβεί στην περιοχή ενός ακινήτου. Ο κίνδυνος αυτός επηρεάζει άμεσα την αξιοπιστία της αξίας του ακινήτου, καθώς οι σεισμοί μπορούν να προκαλέσουν σημαντικές ζημιές.

Μαθηματική Εξίσωση:

$$\text{Earthquake Risk} = \text{Normalized Earthquake Frequency} \times \text{Normalized Earthquake Magnitude}$$

Επεξήγηση:

- **Normalized Earthquake Frequency:** Κανονικοποιημένος δείκτης που δείχνει τη συχνότητα εμφάνισης σεισμών στην περιοχή. Διαιρείται η συχνότητα των σεισμών κάθε περίπτωσης με τη μέγιστη συχνότητα που παρατηρήθηκε στο σύνολο των δεδομένων. Αυτό εξασφαλίζει ότι οι τιμές μεταφέρονται στο διάστημα 0 έως 1, επιτρέποντας μια σύγκριση μεταξύ των διαφόρων περιοχών ανεξάρτητα από το απόλυτο μέγεθος της συχνότητας
- **Normalized Earthquake Magnitude:** Κανονικοποιημένη μέση τιμή του μεγέθους των σεισμών στην περιοχή, διαιρούμενη με μία μέγιστη τιμή. Κανονικοποιούμε τη μέση τιμή της μεγέθυνσης διαιρώντας την με το 10. Η επιλογή του 10 ως διαιρέτη βασίζεται στην παραδοχή ότι οι σεισμοί συνήθως παρουσιάζουν μεγέθη που κυμαίνονται μέχρι περίπου 10, οπότε αυτή η κλίμακα βοηθά να μεταφερθούν και αυτές οι τιμές στο ίδιο εύρος (0-1).

Με αυτήν την προσέγγιση, εξασφαλίζουμε ότι και οι δύο μεταβλητές έχουν

παρόμοια κλίμακα, διευκολύνοντας την ενσωμάτωση τους στα μοντέλα μας και την ανάλυση των αποτελεσμάτων.

Σε συνδυασμό με την τιμή του ακινήτου, αυτός ο δείκτης μπορεί να χρησιμοποιηθεί για την εκτίμηση της πιθανής ζημίας (Earthquake_Damage):

$$\text{Earthquake Damage} = \text{Earthquake Risk} \times \text{Property Price}$$

Αυτή η εξίσωση επιλέχθηκε επειδή ο σεισμικός κίνδυνος εξαρτάται τόσο από τη συχνότητα των σεισμών όσο και από τη μέση τους ένταση. Ο πολλαπλασιασμός των δύο κανονικοποιημένων παραμέτρων μας δίνει ένα ολοκληρωμένο μέτρο που αντικατοπτρίζει την πιθανότητα και την πιθανή σοβαρότητα των ζημιών, χωρίς να επιβαρύνει υπερβολικά το μοντέλο με επιπλέον πολυπλοκότητα.

3.2.2 Κίνδυνος Πλημμύρας (Flood Risk)

Ο κίνδυνος πλημμύρας υπολογίζεται με βάση τις βροχοπτώσεις, που είναι ένας κρίσιμος παράγοντας για την πρόκληση πλημμύρας. Οι δύο βασικές παράμετροι είναι η συχνότητα των βροχοπτώσεων και η ένταση τους. Η επίδρασή τους ορίζεται από σταθμίσεις που αντιπροσωπεύουν τη σχετική σημασία κάθε παραμέτρου.

Μαθηματική Εξίσωση:

$$\begin{aligned} \text{Flood Risk} = & (w_1 \times \text{Normalized Rainfall Frequency}) \\ & + (w_2 \times \text{Normalized Rainfall Intensity}) \end{aligned}$$

όπου:

- $w_1=0.6$ (βάρος για τη συχνότητα βροχοπτώσεων)
- $w_2=0.4$ (βάρος για την ένταση της βροχόπτωσης)

Επεξήγηση:

Για την κανονικοποίηση της συχνότητας και της έντασης της βροχόπτωσης, εφαρμόστηκε η τεχνική min-max normalization. Συγκεκριμένα, για κάθε μεταβλητή:

- **Συχνότητα Βροχόπτωσης:** Αφαιρείται η ελάχιστη τιμή της μεταβλητής από κάθε παρατήρηση και διαιρείται με τη διαφορά μεταξύ της μέγιστης και της ελάχιστης τιμής.

- **Ένταση Βροχόπτωσης:** Εφαρμόζουμε την ίδια διαδικασία, ώστε να φέρουμε και αυτή τη μεταβλητή στην ίδια κλίμακα.

Ο υπολογισμός της ζημίας από πλημμύρες (Flood_Damage) πραγματοποιείται ως εξής:

$$\text{Flood Damage} = \text{Flood Risk} \times \text{Property Price}$$

Η εξίσωση αυτή επιλέχθηκε διότι ο κίνδυνος πλημμύρας επηρεάζεται τόσο από τη συχνότητα της βροχής (Rainfall Frequency) όσο και από την έντασή της (Rainfall Intensity). Η χρήση σταθμίσεων ($w_1 = 0.6$ και $w_2 = 0.4$) επιλέχθηκε με τη θεώρηση ότι η συχνότητα έχει μεγαλύτερη επίδραση στην πιθανότητα πλημμύρας σε σύγκριση με την ένταση της βροχόπτωσης. Δηλαδή, μια υψηλή συχνότητα μηδενικών ή μικρών βροχοπτώσεων αυξάνει σημαντικά τον κίνδυνο συσσώρευσης νερού και συνεπώς τη δυνατότητα εμφάνισης πλημμύρας, ενώ η ένταση, αν και σημαντική, παίζει δευτερεύοντα ρόλο στην εκτίμηση της ζημιάς.

3.2.3 Κίνδυνος Πυρκαγιάς (Wildfire Risk)

Ο κίνδυνος πυρκαγιάς συνυπολογίζει διάφορα στοιχεία: τη συχνότητα εμφάνισης πυρκαγιών, τη διάρκεια τους, καθώς και μετεωρολογικούς παράγοντες όπως η υψηλή θερμοκρασία και η ταχύτητα ανέμου. Αυτοί οι παράγοντες συμβάλλουν στην εκτίμηση της πιθανότητας μιας πυρκαγιάς να προκαλέσει ζημιές στα ακίνητα.

Μαθηματική Εξίσωση:

$$\begin{aligned} \text{Wildfire Risk} &= \alpha \times \text{Normalized Frequency} + \beta \times \text{Normalized Duration} + \gamma \\ &\times \frac{\frac{\text{HIGH}_{TEMP}}{\text{HIGH}_{TEMPmax}} + \frac{\text{AVG}_{WINDSPEED}}{\text{AVG}_{WINDSPEEDmax}}}{2} \end{aligned}$$

όπου:

- $\alpha=0.5$ (βάρος για τη συχνότητα εμφάνισης πυρκαγιών)
- $\beta=0.3$ (βάρος για τη μέση διάρκεια της πυρκαγιάς)
- $\gamma=0.2$ (βάρος για τους μετεωρολογικούς παράγοντες)

Επεξήγηση:

- **Συχνότητα πυρκαγιών (Frequency):** Αυτή η παράμετρος καθορίζει τον αριθμό των πυρκαγιών που σημειώθηκαν σε κάθε περιοχή. Για την κανονικοποίησή της, διαιρείται η κάθε τιμή με τη μέγιστη τιμή της συχνότητας που παρατηρήθηκε στο σύνολο των δεδομένων.
- **Μέση διάρκεια πυρκαγιών (Avg_Fire_Duration):** Αυτή η παράμετρος αντιπροσωπεύει τη μέση διάρκεια των πυρκαγιών σε ώρες. Αντίστοιχα, διαιρείται η κάθε τιμή με τη μέγιστη μέση διάρκεια που παρατηρήθηκε, έτσι ώστε να μεταφερθούν όλες οι τιμές στο διάστημα 0 έως 1.
- **HIGH_TEMP και AVG_WIND_SPEED:** Αντιπροσωπεύουν αντίστοιχα τη μέγιστη θερμοκρασία και τη μέση ταχύτητα ανέμου, που κανονικοποιούνται ως προς τις μέγιστες τιμές τους για να δώσουν μια σχετική ένδειξη της επιρροής τους.

Η ζημιά από πυρκαγιά (wildfire damage) υπολογίζεται ως:

$$\text{Wildfire Damage} = \text{Wildfire Risk} \times \text{Property Price}$$

Στην εξίσωση του κινδύνου πυρκαγιάς, ορίσαμε τις σταθμίσεις με βάση την αντίληψη ότι η συχνότητα εμφάνισης πυρκαγιών έχει τη μεγαλύτερη επιρροή στον κίνδυνο, ακολουθούμενη από τη μέση διάρκεια των πυρκαγιών και τέλος τους μετεωρολογικούς παράγοντες (θερμοκρασία και ταχύτητα ανέμου). Αυτή η επιλογή αντικατοπτρίζει την κρίσιμη σημασία της συχνότητας, καθώς μια υψηλή συχνότητα πυρκαγιών αυξάνει σημαντικά την έκθεση του ακινήτου σε ζημιές, ενώ η διάρκεια και οι μετεωρολογικοί παράγοντες, αν και επίσης σημαντικοί, επηρεάζουν το μοντέλο σε μικρότερο βαθμό.

Παρά το γεγονός ότι οι μετεωρολογικοί παράγοντες (όπως η θερμοκρασία και η ταχύτητα ανέμου) είναι κρίσιμοι για την εξέλιξη μιας πυρκαγιάς, θεωρούνται περιστασιακοί παράγοντες που παρέχουν πρόσθετες πληροφορίες για την ένταση της επίπτωσης. Από την άλλη πλευρά, η συχνότητα και η διάρκεια αποτελούν βασικές μετρήσεις του γεγονότος, μετρώντας άμεσα την πιθανότητα εκδήλωσης και τη χρονική διάρκεια της πυρκαγιάς. Συνεπώς διαδραματίζουν καθοριστικό ρόλο στην εκτίμηση του σχετικού κινδύνου.

3.2.4 Loan Augmentation και Ολοκλήρωση Παραμέτρων Δανείων

Μετά τον υπολογισμό αυτών των κινδύνων, έγινε μετάβαση στην ενσωμάτωση αυτών των πληροφοριών για την εκτίμηση των παραμέτρων δανείων. Η διαδικασία αυτή περιλαμβάνει τα εξής βασικά βήματα:

- Υπολογισμός Σύνθετου Κινδύνου (Composite Risk):
Ο σύνθετος δείκτης κινδύνου ορίζεται ως το άθροισμα των επιμέρους δεικτών:

$$\text{Composite Risk} = \text{wildfire_risk} + \text{Earthquake_Risk} + \text{Flood_Risk}$$

Αυτός ο δείκτης παρέχει μια συνολική ένδειξη της έκθεσης ενός ακινήτου σε φυσικούς κινδύνους. Ο σύνθετος δείκτης χρησιμοποιείται για την ταξινόμηση των ακινήτων σε κατηγορίες κινδύνου.

- Λογική Καθορισμού του LTV

Η βασική λογική που ακολουθήσαμε είναι η αύξηση του LTV με βάση την αύξηση του σύνθετου κινδύνου. Στο μοντέλο μας, τα ακίνητα που παρουσιάζουν χαμηλό σύνθετο κίνδυνο ($\text{Composite Risk} < 0.5$) λαμβάνουν LTV 50%, ενώ εκείνα με μεσαίο σύνθετο κίνδυνο ($0.5 \leq \text{Composite Risk} \leq 0.7$) λαμβάνουν LTV 60% και τα ακίνητα με υψηλό σύνθετο κίνδυνο ($\text{Composite Risk} > 0.7$) λαμβάνουν LTV 70%. Δηλαδή, όσο μεγαλύτερη είναι η έκθεση σε φυσικούς κινδύνους, τόσο υψηλότερο ποσοστό χρηματοδότησης αποδίδεται.

Αιτιολόγηση της Επιλογής

Η εφαρμογή υψηλότερου LTV σε ακίνητα με υψηλό σύνθετο κίνδυνο επιτρέπει την προσομοίωση ακραίων σεναρίων, όπου οι ζημιές μπορούν να είναι σημαντικά μεγαλύτερες. Αυτό βοηθά στην αξιολόγηση της ευαισθησίας του χαρτοφυλακίου και στην εκτίμηση του απαιτούμενου κεφαλαίου για την κάλυψη πιθανών απωλειών.

Υπολογισμός του Ποσού Δανείου

Το ποσό του δανείου υπολογίζεται ως:

$$\text{Loan Amount} = \text{res_price} \times \text{LTV}$$

όπου:

- res_price : Η τιμή του ακινήτου.
- LTV: Το ποσοστό χρηματοδότησης που αποδίδεται βάσει του σύνθετου κινδύνου.

Για να ενισχύσουμε την ρεαλιστικότητα, προστίθεται μικρός τυχαίος θόρυβος (Gaussian noise) τόσο στο LTV όσο και στο ποσό του δανείου. Αυτό επιτρέπει την προσομοίωση των μικρών διακυμάνσεων που παρατηρούνται στην πραγματική αγορά.

To loan amount αντιπροσωπεύει το υπόλοιπο ποσό δανείου που παραχωρείται για ένα ακίνητο την χρονική στιγμή της ανάλυσης, δηλαδή το ποσό που εκτίθεται στον πιστωτικό κίνδυνο. Με άλλα λόγια, το loan amount αντιπροσωπεύει την "Exposure at Default" (EAD). Δηλαδή, αποτελεί το ποσό του δανείου που είναι εκτεθειμένο στον πιστωτικό κίνδυνο σε περίπτωση αθέτησης. Με αυτόν τον τρόπο, η EAD χρησιμεύει ως βασικό μέτρο για την εκτίμηση των πιθανών απωλειών σε ένα χαρτοφυλάκιο στεγαστικών δανείων.

3.3 Μοντέλα Μηχανικής Μάθησης για Εκτίμηση του LGD

Σε αυτή την ενότητα, παρουσιάζεται η ανάπτυξη και η αξιολόγηση μοντέλων μηχανικής μάθησης με σκοπό την εκτίμηση του Loss Given Default (LGD) για τα στεγαστικά δάνεια. Η χρήση των μοντέλων μηχανικής μάθησης παρέχει τη δυνατότητα ανάλυσης πολυδιάστατων δεδομένων και ανεύρεσης μη γραμμικών σχέσεων μεταξύ των μεταβλητών, συμβάλλοντας έτσι στη βελτίωση της ακρίβειας της πρόβλεψης.

3.3.1 Προετοιμασία Δεδομένων

Για την εκπαίδευση των μοντέλων χρησιμοποιήσαμε το συνδυασμένο dataset που προκύπτει από τις διαδικασίες συλλογής, επεξεργασίας και augmentation (βλ. ενότητες 3.1 και 3.2). Συγκεκριμένα, τα δεδομένα περιέχουν χαρακτηριστικά όπως:

- res_price: Η τιμή του ακινήτου.
- Flood_Risk, wildfire_risk, Earthquake_Risk: Δείκτες κινδύνου για πλημμύρες, πυρκαγιές και σεισμούς.
- Year: Το έτος παρατήρησης ή προσομοίωσης.
- loan_amount: Το υπολογισμένο ποσό δανείου (ή αλλιώς το Exposure At Default).

Για να αντιμετωπιστούν οι διαφορές μεταξύ των διαφορετικών σεναρίων προσομοίωσης (Baseline, High Flood Risk, High Wildfire Risk, High Earthquake Risk), εφαρμόστηκε one-hot encoding. Αυτή η τεχνική μετατρέπει κάθε κατηγορική τιμή σε μια δυαδική μεταβλητή, χωρίς να υποθέτει κάποια ιεραρχική διάταξη μεταξύ τους. Με αυτόν τον τρόπο, τα μοντέλα δεν αντιμετωπίζουν μεροληψία ή παραπλανητικές συσχετίσεις λόγω των κατηγορικών τιμών και μπορούν να ενσωματώσουν με μεγαλύτερη ακρίβεια τις διαφορές των σεναρίων ως ανεξάρτητες μεταβλητές.

Η χρήση της StandardScaler (ή άλλων τεχνικών κανονικοποίησης) εξασφαλίζει ότι όλες οι αριθμητικές μεταβλητές (ειδικά αυτές που αφορούν τους δείκτες κινδύνου) βρίσκονται σε μια κοινή κλίμακα. Αυτό είναι κρίσιμο για πολλά μοντέλα μηχανικής μάθησης, ιδιαίτερα εκείνα που βασίζονται σε αποστάσεις ή που χρησιμοποιούν βελτιστοποίηση με gradient descent, καθώς

αποτρέπει τις μεταβλητές με μεγαλύτερη κλίμακα να επιδρούν υπερβολικά στις τελικές προβλέψεις.

Όταν οι μεταβλητές βρίσκονται σε διαφορετικές κλίμακες, τα μοντέλα μπορούν να αντιμετωπίσουν προβλήματα συγκλίνουσας ορθότητας και να μάθουν λανθασμένες συσχετίσεις. Η κανονικοποίηση βοηθά να επιτευχθεί μια πιο σταθερή και γρήγορη σύγκλιση κατά τη διάρκεια της εκπαίδευσης, βελτιώνοντας την ακρίβεια των προβλέψεων.

Με την ενσωμάτωση των one-hot encoded μεταβλητών για τα σενάρια, το μοντέλο μπορεί να διαχωρίσει την επίδραση κάθε σεναρίου από τα υπόλοιπα χαρακτηριστικά. Έτσι, μπορούμε να αναλύσουμε πώς διαφοροποιούνται οι προβλέψεις ανάλογα με το σενάριο, επιτρέποντας μία πιο λεπτομερή ανάλυση ευαισθησίας.

Τέλος, για να αντικατοπτρίζεται η φυσική διακύμανση των δεδομένων και να αποφεύγεται η υπερπροσαρμογή (overfitting), εφαρμόστηκε ένας μικρός τυχαίος θόρυβος (Gaussian noise) σε επιλεγμένες μεταβλητές, συμπεριλαμβανομένων των LTV και άλλων οικονομικών χαρακτηριστικών.

3.3.2 Επιλογή και Εκπαίδευση Μοντέλων

Για την εκτίμηση του LGD χρησιμοποιήσαμε δύο κύρια μοντέλα μηχανικής μάθησης:

- Random Forest:

Το Random Forest είναι ένα ensemble μοντέλο που συνδυάζει πολλαπλά δέντρα αποφάσεων για να βελτιώσει την ακρίβεια και να μειώσει το overfitting. Είναι ιδιαίτερα κατάλληλο για δεδομένα που περιέχουν μη γραμμικές σχέσεις και αλληλεπιδράσεις μεταξύ των χαρακτηριστικών. Χωρίσαμε το dataset σε σύνολα εκπαίδευσης και δοκιμής και εφαρμόσαμε τεχνικές K-fold Cross-Validation για την αξιολόγηση της γενικευσιμότητας του μοντέλου. Η χρήση των μετρικών R^2 και Mean Squared Error (MSE) επέτρεψε την εκτίμηση της απόδοσης του μοντέλου. Μέσω του Random Forest μπορούμε να εντοπίσουμε τις μεταβλητές που έχουν τη μεγαλύτερη επίδραση στην πρόβλεψη του LGD, παρέχοντας έτσι πολύτιμα insights για τη διαχείριση κινδύνου.

- XGBoost:

Το XGBoost είναι ένα ισχυρό μοντέλο gradient boosting που επιδεικνύει υψηλή ακρίβεια σε προβλήματα με πολυδιάστατα και μη γραμμικά δεδομένα. Έχει αποδειχθεί ιδιαίτερα αποτελεσματικό σε προβλήματα πρόβλεψης και συχνά

υπερβαίνει τα παραδοσιακά μοντέλα σε πολλές εφαρμογές. Παρόμοια με το Random Forest, το XGBoost εκπαιδεύτηκε με χρήση train-test split και K-fold Cross-Validation. Τα αποτελέσματα αξιολογήθηκαν με τις ίδιες μετρικές (R^2 και MSE), επιτρέποντας τη σύγκριση της απόδοσης των δύο μοντέλων. Η διαδικασία ρύθμισης των υπερπαραμέτρων (π.χ. n_estimators, max_depth, learning_rate) για το XGBoost βελτιώνει την ακρίβεια του μοντέλου, επιτρέποντας την καλύτερη πρόβλεψη του LGD.

3.3.3 Αξιολόγηση και Ανάλυση Αποτελεσμάτων

Σε αυτήν την υποενότητα, αναλύονται λεπτομερώς τα αποτελέσματα των μοντέλων μηχανικής μάθησης που χρησιμοποιήθηκαν για την πρόβλεψη του LGD, καθώς και τα μετρικά και τις μεθοδολογίες που εφαρμόστηκαν για την αξιολόγησή τους.

1. Χωρισμός Δεδομένων και Διασταυρούμενη Επικύρωση:

- Train-Test Split:

Τα δεδομένα χωρίστηκαν σε σύνολα εκπαίδευσης και δοκιμής ώστε να εξασφαλιστεί ότι τα μοντέλα θα αξιολογηθούν σε ανεξάρτητα δεδομένα. Αυτό βοηθά στον έλεγχο της γενικευσιμότητας και της απόδοσης των μοντέλων σε νέες περιπτώσεις.

- K-fold Cross-Validation:

Εφαρμόστηκε η τεχνική K-fold (5-fold) Cross-Validation για την εκτίμηση της σταθερότητας και της γενικευσιμότητας των μοντέλων. Μέσω αυτής της μεθόδου, το dataset χωρίστηκε σε K υποσύνολα, και σε κάθε επανάληψη ένα υποσύνολο χρησιμοποιήθηκε για δοκιμή ενώ τα υπόλοιπα για εκπαίδευση. Αυτή η προσέγγιση βοηθά στην αποφυγή της υπερπροσαμογής (overfitting).

2. Μετρικές Αξιολόγησης:

Για την αξιολόγηση της απόδοσης των μοντέλων χρησιμοποιήθηκαν διάφορες μετρικές:

- R^2 (Coefficient of Determination):

Μετρά το ποσοστό της διακύμανσης του πραγματικού LGD που εξηγείται από το μοντέλο. Ένα υψηλό R^2 υποδεικνύει ότι το μοντέλο έχει μεγάλη εξηγητική δύναμη.

- Mean Squared Error (MSE):

Ο μέσος όρος των τετραγωνικών αποκλίσεων μεταξύ των προβλεπόμενων και των πραγματικών τιμών του LGD. Η MSE μας δίνει μια ένδειξη για το μέγεθος των σφαλμάτων πρόβλεψης.

3. Ανάλυση Feature Importance:

- Εξέταση της σημαντικότητας των χαρακτηριστικών:
Μέσω των μοντέλων Random Forest και XGBoost, αναλύσαμε τη σημαντικότητα των μεταβλητών (feature importance) για την πρόβλεψη του LGD. Αυτό μας επιτρέπει να κατανοήσουμε ποιοι παράγοντες (π.χ. τιμή ακινήτου, δείκτες κινδύνου, μεταβλητές σεναρίων) επηρεάζουν περισσότερο το αποτέλεσμα, δίνοντας πολύτιμα insights για τη διαχείριση κινδύνου.

4. Σύγκριση Μοντέλων και Επιλογή του Κατάλληλου:

- Σύγκριση
Τα μοντέλα Random Forest και XGBoost συγκρίθηκαν με βάση τις παραπάνω μετρικές, ώστε να εντοπιστεί ποιο παρέχει καλύτερη γενικευσιμότητα και ακρίβεια στην πρόβλεψη του LGD. Η σύγκριση αυτή μας επέτρεψε να επιλέξουμε το μοντέλο που καταδεικνύει τη μικρότερη απόκλιση (χαμηλότερο MSE) και υψηλότερο R^2 .
- Ρύθμιση Υπερπαραμέτρων (Hyperparameter Tuning):
Εφαρμόστηκαν τεχνικές ρύθμισης υπερπαραμέτρων για τη βελτιστοποίηση της απόδοσης των μοντέλων. Η ρύθμιση αυτή περιλαμβάνει παράγοντες όπως ο αριθμός των δέντρων, το βάθος των δέντρων, και ο ρυθμός μάθησης (learning rate) στο XGBoost.

5. Ανάλυση Ευαισθησίας και Σταθερότητας:

- Ευαισθησία σε Μεταβολές:
Διεξήχθη ανάλυση ευαισθησίας, όπου εξετάστηκαν οι επιπτώσεις μικρών μεταβολών στις εισόδους (π.χ., οι μεταβλητές κινδύνου) στα τελικά αποτελέσματα. Αυτό εξασφαλίζει ότι το μοντέλο είναι σταθερό και ότι οι προβλέψεις του δεν παρουσιάζουν μεγάλη διακύμανση με μικρές αλλαγές στα δεδομένα.
- Επικύρωση:
Επιπλέον, τα μοντέλα επικυρώθηκαν συγκρίνοντάς τα με ιστορικά δεδομένα και δημοσιευμένες μελέτες, επιβεβαιώνοντας την ακρίβεια των προβλέψεων και τη δυνατότητα εφαρμογής τους σε πραγματικά δεδομένα.

3.3.4 Λογική Επιλογής Μοντέλων

Η επιλογή των μοντέλων μηχανικής μάθησης για την εκτίμηση του LGD βασίστηκε σε μια σειρά από παραμέτρους και κριτήρια, με στόχο τη μεγιστοποίηση της ακρίβειας

και της γενικευσιμότητας των προβλέψεων. Συγκεκριμένα, επιλέχθηκαν τα μοντέλα Random Forest Regressor και XGBoost Regressor για τους εξής λόγους:

1. Αντιμετώπιση Μη Γραμμικών Σχέσεων:

Τα δεδομένα μας περιέχουν πολυδιάστατες και μη γραμμικές σχέσεις μεταξύ των χαρακτηριστικών (π.χ. οικονομικά χαρακτηριστικά, δείκτες κινδύνου, σενάρια προσομοίωσης). Τόσο το Random Forest όσο και το XGBoost είναι γνωστά για την ικανότητά τους να αποδίδουν καλά σε τέτοια δεδομένα, καθώς μπορούν να μοντελοποιήσουν πολύπλοκες αλληλεπιδράσεις χωρίς να απαιτούν προκαταρκτική μετασχηματισμό των μεταβλητών.

2. Ανθεκτικότητα στην Υπερπροσαρμογή (Overfitting):

Το Random Forest, μέσω της χρήσης πολλών δέντρων αποφάσεων και της τυχαίας δειγματοληψίας των χαρακτηριστικών, μειώνει σημαντικά την πιθανότητα υπερπροσαρμογής στα δεδομένα εκπαίδευσης. Αντίστοιχα, το XGBoost διαθέτει ενσωματωμένους μηχανισμούς regularization (όπως το L1 και L2 regularization) που συμβάλλουν στη σταθερότητα του μοντέλου.

3. Ευελιξία και Ρύθμιση Υπερπαραμέτρων:

Και τα δύο μοντέλα προσφέρουν μεγάλη ευελιξία μέσω της ρύθμισης υπερπαραμέτρων (π.χ., αριθμός δέντρων, βάθος δέντρων, learning rate στο XGBoost). Αυτό επιτρέπει την προσαρμογή του μοντέλου στις συγκεκριμένες απαιτήσεις του dataset μας, βελτιστοποιώντας την απόδοση με βάση μετρικές όπως το R2 και το MSE.

4. Δυνατότητα Εξαγωγής Feature Importance:

Η δυνατότητα εξαγωγής της σημαντικότητας των χαρακτηριστικών (feature importance) από το Random Forest και το XGBoost παρέχει πολύτιμα insights για τους παράγοντες που έχουν τη μεγαλύτερη επίδραση στην πρόβλεψη του LGD. Αυτή η πληροφορία είναι κρίσιμη για την ερμηνεία του μοντέλου και για τη λήψη αποφάσεων στη διαχείριση του κινδύνου.

5. Εφαρμογή σε Σεναριακές Προσομοιώσεις:

Η επιλογή των μοντέλων αυτών επιτρέπει την εύκολη ενσωμάτωση των διαφορετικών σεναρίων (μέσω one-hot encoding) και την ανάλυση της επίδρασης των μεταβλητών σε κάθε σενάριο. Αυτό είναι ιδιαίτερα χρήσιμο όταν επιδιώκουμε να αξιολογήσουμε την ευαισθησία του LGD σε διάφορες συνθήκες και να πραγματοποιήσουμε stress testing.

3.4 Προσομοιώσεις Monte Carlo

Σε αυτή την ενότητα, περιγράφεται η εφαρμογή της μεθοδολογίας Monte Carlo για την προσομοίωση των μελλοντικών συνθηκών και την εκτίμηση των επιπτώσεων των φυσικών και

οικονομικών κινδύνων στα στεγαστικά δάνεια. Η μέθοδος Monte Carlo αποτελεί μια στοχαστική προσέγγιση που χρησιμοποιείται για την προσομοίωση πολλαπλών σεναρίων, λαμβάνοντας υπόψη την αβεβαιότητα και τη διακύμανση των εισόδων.

3.4.1 Σκοπός και Θεωρητική Βάση

Η χρήση των προσομοιώσεων Monte Carlo επιτρέπει:

- Την εκτίμηση της διακύμανσης των μελλοντικών κινδύνων: Με την προσομοίωση πολλών πιθανών μελλοντικών καταστάσεων, το μοντέλο αποδίδει μια κατανομή πιθανών αποτελεσμάτων για παραμέτρους όπως το LGD και το Value-to-Loan (VtL).
- Την ανάλυση σεναρίων stress testing: Η μέθοδος επιτρέπει την εκτίμηση της επίδρασης διαφορετικών σεναρίων (π.χ. αύξηση κινδύνου πλημμύρας, πυρκαγιάς, σεισμικού κινδύνου) και την αξιολόγηση της ευαισθησίας του χαρτοφυλακίου.

3.4.2 Διαδικασία Προσομοίωσης

Η διαδικασία Monte Carlo υλοποιήθηκε με τα εξής βασικά βήματα:

Ορισμός Παραμέτρων Προσομοίωσης:

- Χρονικό Ορίζοντα: Οι προσομοιώσεις εκτελούνται για την περίοδο 2025 έως 2050.
- Αριθμός Επαναλήψεων (NUM_SIMULATIONS): Ορίζεται ο αριθμός των επαναλήψεων ανά έτος (100 προσομοιώσεις ανά έτος) για να αποτυπωθεί η αβεβαιότητα.
- Ρυθμοί Αύξησης και Noise Factors: Ορίστηκαν ετήσιοι ρυθμοί αύξησης για τους βασικούς δείκτες κινδύνου (πλημμύρες, πυρκαγιές, σεισμοί) με μικρή διακύμανση (με μέση τιμή και μικρό standard deviation), ώστε να προσομοιωθεί η σταδιακή μεταβολή των συνθηκών. Επιπλέον, προστέθηκε τυχαίος θόρυβος (Gaussian noise) για να αντικατοπτριστούν οι μικρές φυσικές διακυμάνσεις.

Εφαρμογή Σεναρίων και Παράγοντες Αβεβαιότητας:

Για την προσομοίωση ακραίων συνθηκών, καθορίσαμε διάφορα σενάρια κινδύνου τα οποία αντιπροσωπεύουν δυσμενείς αλλά ρεαλιστικές καταστάσεις. Τα σενάρια αυτά περιλαμβάνουν:

- Σενάριο Αύξησης Πλημμυρικού Κινδύνου: Σε αυτό το σενάριο, υποθέτουμε αύξηση της συχνότητας και της έντασης των βροχοπτώσεων, κάτι που οδηγεί σε μεγαλύτερη συσσώρευση νερού και αυξημένη πιθανότητα πλημμύρας.

- Σενάριο Αύξησης Κινδύνου Πυρκαγιάς: Σε αυτό το σενάριο, εξετάζουμε πώς η αύξηση της θερμοκρασίας και της συχνότητας των πυρκαγιών συμβάλλουν στην αλλαγή της αξίας των ακινήτων.
- Σενάριο Αύξησης Σεισμικού Κινδύνου: Υποθέτουμε ότι παρατηρείται αύξηση τόσο στη συχνότητα όσο και στο μέγεθος των σεισμών στην περιοχή, επηρεάζοντας σημαντικά την αξία των ακινήτων.
- Σενάριο Συνδυασμού Κινδύνων: Σε αυτό το σενάριο, συμβαίνουν ταυτόχρονα αυξημένα επίπεδα κινδύνου για πλημμύρες και πυρκαγιές, ενισχύοντας τη συνολική έκθεση και οδηγώντας σε πιο ακραίες εκτιμήσεις ζημιών.

Επίσης προσθέσαμε και τα εξής:

- Οικονομικές Μεταβολές: Εφαρμόζεται ένας τυχαίος οικονομικός θόρυβος (economic shock) για την προσομοίωση μεταβολών στις τιμές των ακινήτων. Ο οικονομικός θόρυβος αναφέρεται σε ξαφνικές και απρόβλεπτες μεταβολές στις οικονομικές συνθήκες που επηρεάζουν την αξία των περιουσιακών στοιχείων, όπως τα ακίνητα. Στο πλαίσιο της μεθοδολογίας μας, ο οικονομικός θόρυβος εισάγεται ως ένας παράγοντας πολλαπλασιαστής που τροποποιεί την τιμή του ακινήτου, αντανακλώντας τις επιπτώσεις απροσδόκητων μακροοικονομικών μεταβολών (π.χ. ύφεση, υψηλός πληθωρισμός ή άλλες κρίσεις). Συγκεκριμένα, μοντελοποιείται μέσω ενός τυχαίου παράγοντα που αντλείται από μια κανονική κατανομή με μέσο όρο 0 και μικρό standard deviation (0.02). Αυτός ο παράγοντας, όταν εφαρμοστεί ως πολλαπλασιαστής στην τιμή του ακινήτου, προσαρμόζει τις οικονομικές παραμέτρους σε κάθε επανάληψη της προσομοίωσης. Με αυτόν τον τρόπο, μπορούμε να προσομοιώσουμε το πώς οι ξαφνικές μεταβολές στην οικονομική κατάσταση επηρεάζουν το μέγεθος των ζημιών, τις παραμέτρους χρηματοδότησης (LTV) και κατ' επέκταση, τον πιστωτικό κίνδυνο του χαρτοφυλακίου. Η ενσωμάτωση του οικονομικού θορύβου στο μοντέλο μας έχει ως στόχο να αποδώσει την αβεβαιότητα που υπάρχει στην αγορά λόγω μακροοικονομικών παραγόντων. Επιπλέον, μπορεί να επιτρέψει τη διερεύνηση σεναριακών αποτελεσμάτων υπό διαφορετικές οικονομικές συνθήκες. Τέλος, μπορεί να ενισχύσει την ρεαλιστικότητα των προσομοιώσεων stress testing, δίνοντας τη δυνατότητα να αξιολογηθεί η ευαισθησία του χαρτοφυλακίου σε απροσδόκητες οικονομικές κρίσεις.
- Εισαγωγή Τυχαίου Θορύβου: Στο σύνολο των παραμέτρων εφαρμόζεται Gaussian noise για να αντικατοπτρίζεται η φυσική διακύμανση.
- Εκτέλεση των Επαναλήψεων: Για κάθε έτος και για κάθε σενάριο, τα δεδομένα ενημερώνονται με βάση τους ετήσιους ρυθμούς αύξησης και τους τυχαίους παράγοντες. Το μοντέλο επαναλαμβάνει

τη διαδικασία για κάθε μία από τις προσομοιώσεις, παράγοντας ένα πλήθος δυνατών μελλοντικών αποτελεσμάτων.

- **Συγκέντρωση Αποτελεσμάτων:**

Τα αποτελέσματα των προσομοιώσεων (π.χ., μεταβολές στα LGD, VaR) συγκεντρώνονται και αναλύονται στατιστικά, παράγοντας κατανομές (διαγράμματα, histograms) που επιτρέπουν την εκτίμηση της μέσης τιμής, της διακύμανσης και των ακραίων τιμών (VaR, Expected Shortfall).

3.4.3 Λογική και Επιλογή των Παραμέτρων

Η επιλογή των παραμέτρων για τις προσομοιώσεις βασίστηκε στα εξής κριτήρια:

- **Ρεαλισμός:** Οι ετήσιοι ρυθμοί αύξησης και οι παράγοντες θορύβου επιλέχθηκαν ώστε να αντικατοπτρίζουν πραγματικές μεταβολές στις κλιματικές συνθήκες και στην αγορά ακινήτων.
- **Ευαισθησία των Αποτελεσμάτων:** Με την εφαρμογή πολλαπλών επαναλήψεων, το μοντέλο επιτρέπει την ανάλυση της ευαισθησίας των αποτελεσμάτων σε μικρές μεταβολές των παραμέτρων.
- **Δυνατότητα Εφαρμογής σε Stress Testing:** Η προσέγγιση επιτρέπει την αξιολόγηση του κινδύνου σε ακραία σενάρια, δίνοντας τη δυνατότητα στους διαχειριστές κινδύνου να εντοπίσουν τα πιθανά ακραία αποτελέσματα και να προετοιμαστούν αναλόγως.

3.4.4 Εφαρμογή και Χρήση των Αποτελεσμάτων

Τα αποτελέσματα των προσομοιώσεων χρησιμοποιούνται για:

- **Αξιολόγηση της Συνολικής Έκθεσης:** Μέσω της ανάλυσης των κατανομών των προβλεπόμενων ζημιών (LGD), μπορεί να αξιολογηθεί η συνολική έκθεση του χαρτοφυλακίου.
- **Stress Testing και Διάγραμμα Προσομοιώσεων:** Τα αποτελέσματα εισάγονται σε απεικονίσεις (π.χ., heatmaps, γραφήματα μεταβολής LGD) για να παρουσιάσουν την επίδραση των διαφορετικών σεναρίων και την πιθανή κατανομή των ζημιών ανά έτος και περιοχή.
- **Λήψη Απόφασης:** Οι αναλύσεις αυτές παρέχουν κρίσιμα στοιχεία για την εκτίμηση του απαιτούμενου κεφαλαίου και την υιοθέτηση στρατηγικών διαχείρισης κινδύνου σε περιπτώσεις κρίσης.

3.5 Απεικόνιση Αποτελεσμάτων και Διαγράμματα

Η απεικόνιση αποτελεί κρίσιμο εργαλείο για την κατανόηση, επικύρωση και παρουσίαση των αποτελεσμάτων της μελέτης μας σχετικά με τον πιστωτικό κίνδυνο. Μέσω διαγραμμάτων και χαρτών, απεικονίζουμε τόσο τις χρονικές μεταβολές στους βασικούς δείκτες κινδύνου όσο και τη γεωγραφική κατανομή τους σε επίπεδο περιοχών.

3.5.1 Γραφικές Παραστάσεις και Διαγράμματα Χρόνου

- Στόχος και Σημασία:

Δημιουργήθηκαν γραμμικά διαγράμματα που απεικονίζουν την εξέλιξη των οικονομικών παραμέτρων με το πέρασμα του χρόνου για κάθε περιοχή ή σενάριο. Αυτά τα διαγράμματα επιτρέπουν τη σύγκριση των μεταβολών, καθώς και την ανάλυση της επίδρασης των διαφόρων σεναρίων στην εξέλιξη του πιστωτικού κινδύνου.

- Τεχνικές Λεπτομέρειες:

Χρησιμοποιήθηκαν κατάλληλες χρωματικές παλέτες και διαμορφώθηκαν οι άξονες (χρονική κλίμακα και τιμές των μεταβλητών) έτσι ώστε οι διαφορές να απεικονίζονται με σαφήνεια. Επίσης, περιλαμβάνονται annotations για να επισημανθούν οι τελικές τιμές ή κρίσιμα σημεία των γραφημάτων, διευκολύνοντας την ανάλυση των αποτελεσμάτων.

3.5.2 Χαρτογράφηση

Δημιουργήθηκαν γεωγραφικοί χάρτες που απεικονίζουν την κατανομή των κινδύνων σε επίπεδο περιφέρειας.

- Heatmaps Περιοχών: Οι χάρτες που δημιουργήθηκαν δείχνουν τη διακύμανση των οικονομικών και φυσικών παραμέτρων ανά περιοχή, όπως οι μέσες τιμές LGD. Χρησιμοποιήσαμε θερμοχάρτες (heatmaps) για να αναδειχθούν οι περιοχές με υψηλή ή χαμηλή έκθεση σε κινδύνους.
- Σύνδεση με Γεωγραφικά Δεδομένα: Μέσω της γεωγραφικής απεικόνισης, κάθε περιοχή χαρακτηρίζεται με βάση τα δικά της δεδομένα, επιτρέποντας την αναγνώριση των ευάλωτων περιοχών και την αξιολόγηση του αντίκτυπου των φυσικών καταστροφών σε γεωγραφικό επίπεδο.

4. Αποτελέσματα

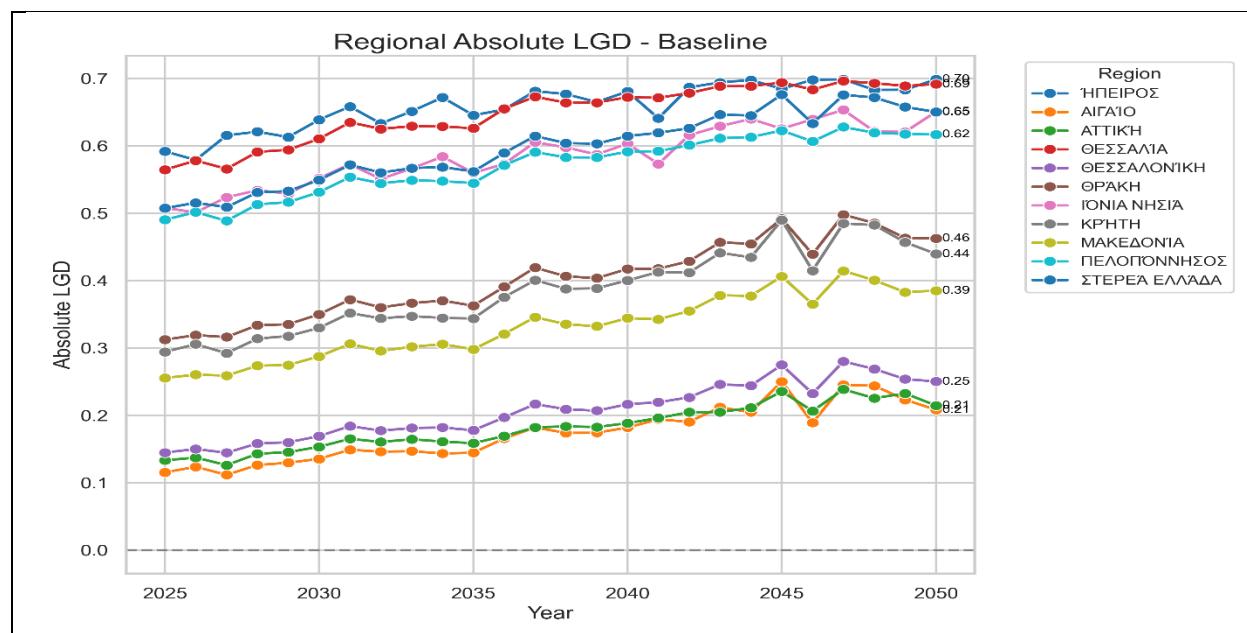
Σε αυτό το κεφάλαιο παρουσιάζονται τα αποτελέσματα της έρευνάς μας, τα οποία προέκυψαν από τις προσομοιώσεις stress testing και από την εφαρμογή μοντέλων μηχανικής μάθησης για την εκτίμηση του LGD στα στεγαστικά δάνεια. Μέσα από αυτή την ενότητα, επιδιώκουμε να δώσουμε μια ολοκληρωμένη εικόνα της επίδρασης των φυσικών κινδύνων και στο πιστωτικό χαρτοφυλάκιο, καθώς και να αναδείξουμε τις γεωγραφικές διαφορές και τις μεταβολές των βασικών οικονομικών παραμέτρων υπό διαφορετικά σενάρια.

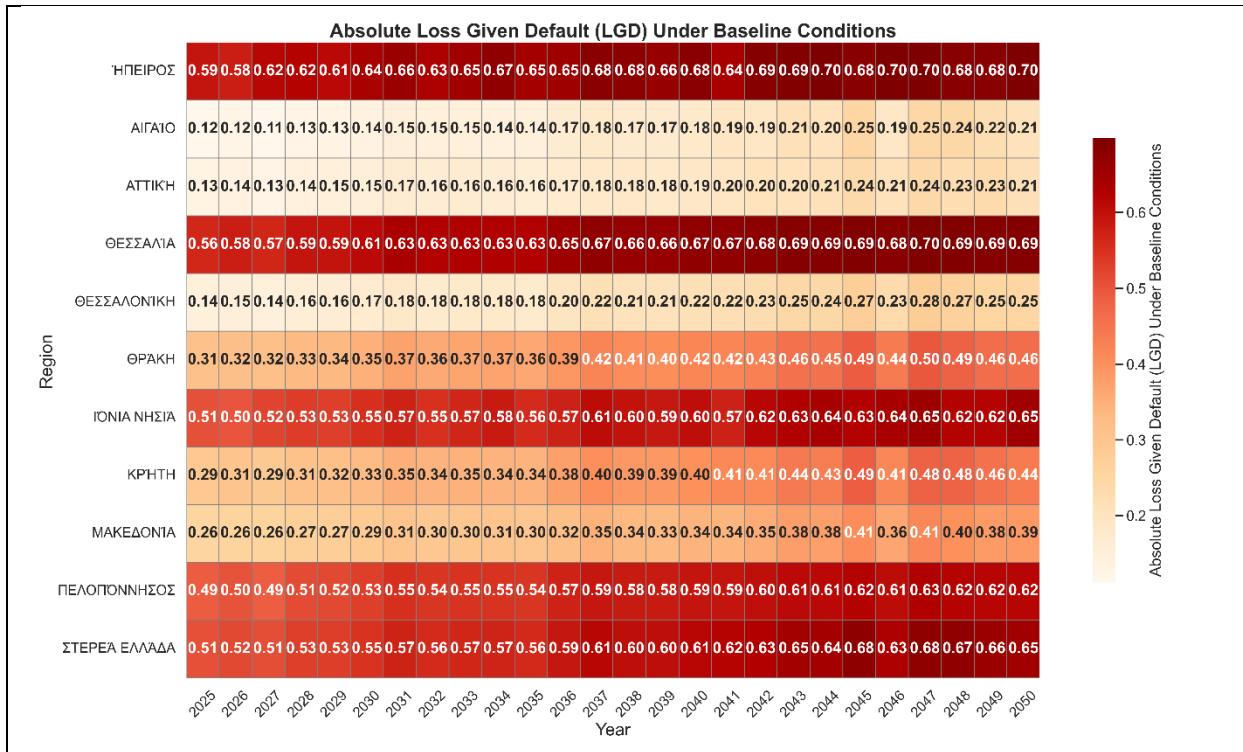
Παρακάτω θα αναφερθούν τα αποτελέσματα της έρευνάς που πραγματοποιήθηκε.

4.1 Αποτελέσματα Προσομοιώσεων Stress Testing

Στην ενότητα αυτή, παρουσιάζονται συνοπτικά τα βασικά αποτελέσματα των προσομοιώσεων stress testing που υλοποιήθηκαν. Μέσα από γραφήματα και heatmaps παρατηρείται πώς μεταβάλλονται οι δείκτες κινδύνου υπό διαφορετικά σενάρια – από τις τρέχουσες συνθήκες (Baseline) έως τις ακραίες συνθήκες (υψηλός κίνδυνος πλημμύρας, σεισμικός, πυρκαγιάς και συνδυαστικά σενάρια).

4.1.1 Ανάλυση Βασικού Σεναρίου



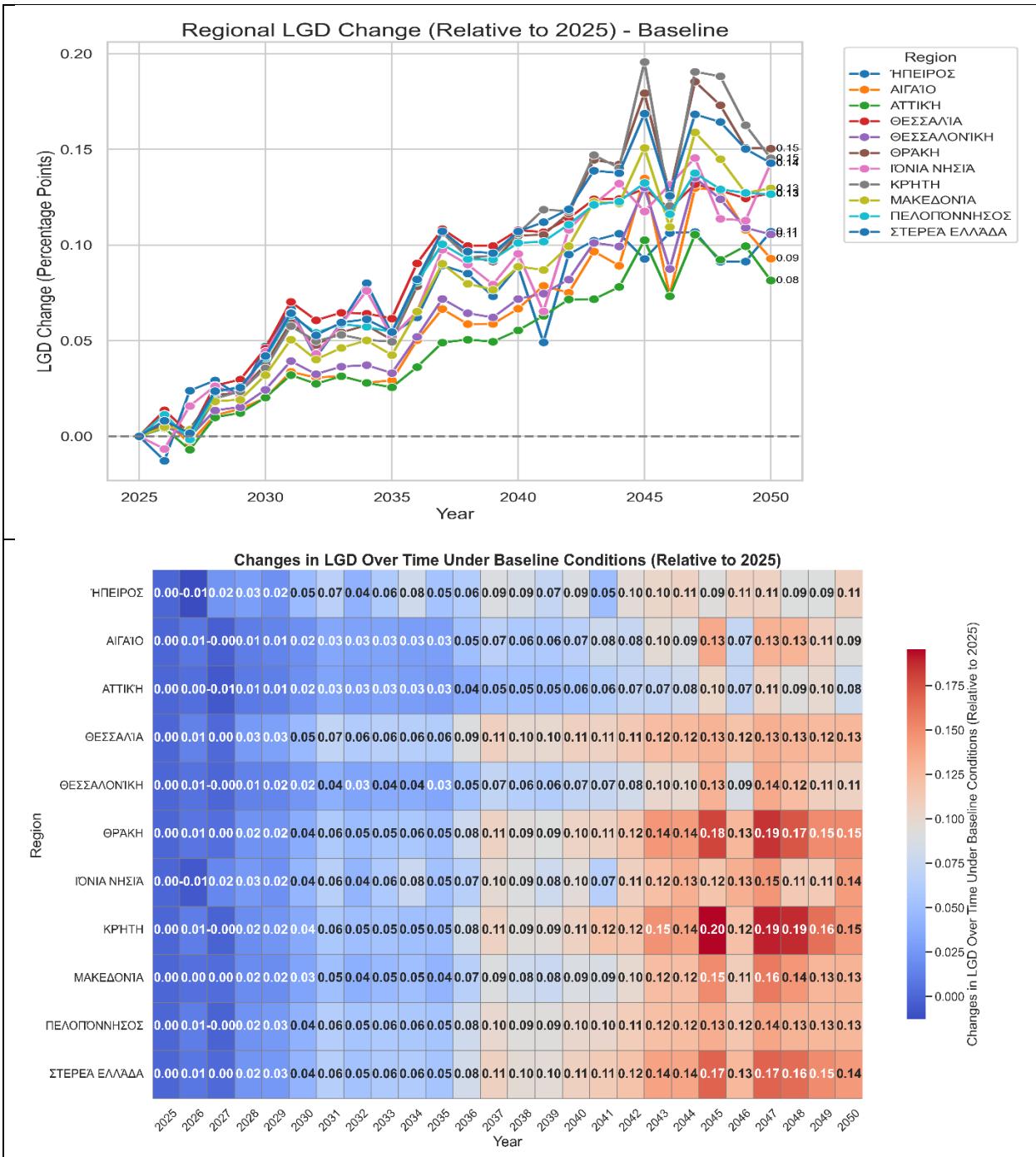


Εικόνα 4.1.1.1 Regional Absolute values of LGD for the baseline scenario

Η πρώτη γραφική παράσταση απεικονίζει την εξέλιξη της απόλυτης τιμής του LGD ανά περιοχή κατά την περίοδο 2025-2050. Παρατηρούμε ότι σε ορισμένες περιοχές, όπως η Στερεά Ελλάδα, η Θεσσαλία και τα Ιόνια Νησιά, το LGD παραμένει σταθερά υψηλό, ξεπερνώντας το 0.6 καθ' όλη τη διάρκεια της περιόδου. Αντίθετα, περιοχές όπως το Αιγαίο και η Αττική εμφανίζουν σημαντικά χαμηλότερες τιμές, με το LGD να παραμένει κοντά στο 0.2.

Μεταβολές του LGD σε Σχέση με το 2025

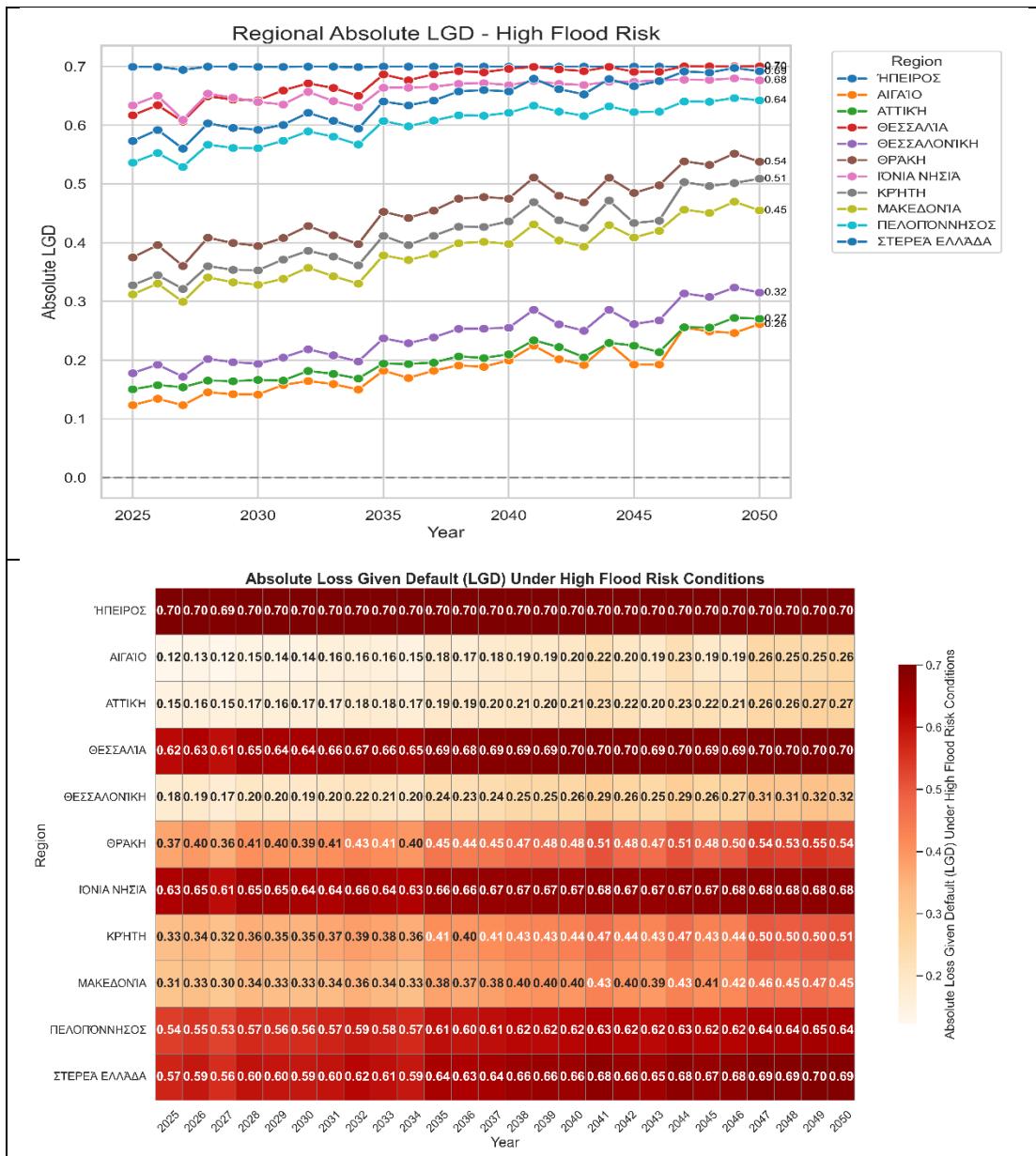
Οι παρακάτω απεικονίσεις δείχνουν τη μεταβολή του LGD σε σύγκριση με το 2025, επιτρέποντας τη σύγκριση των περιφερειακών αποκλίσεων. Εδώ φαίνεται ότι όλες οι περιοχές παρουσιάζουν μια ανοδική τάση στη μεταβολή του LGD, με κάποιες διακυμάνσεις γύρω από το 2045. Το φαινόμενο αυτό μπορεί να σχετίζεται με οικονομικές ή τραπεζικές κρίσεις που επηρεάζουν την αγορά ακινήτων και τη δυνατότητα αποπληρωμής των δανείων.



Εικόνα 4.1.1.2 Regional LGD changes for the baseline scenario

4.1.2 Ανάλυση Σεναρίου Υψηλού Κινδύνου Πλημμύρας

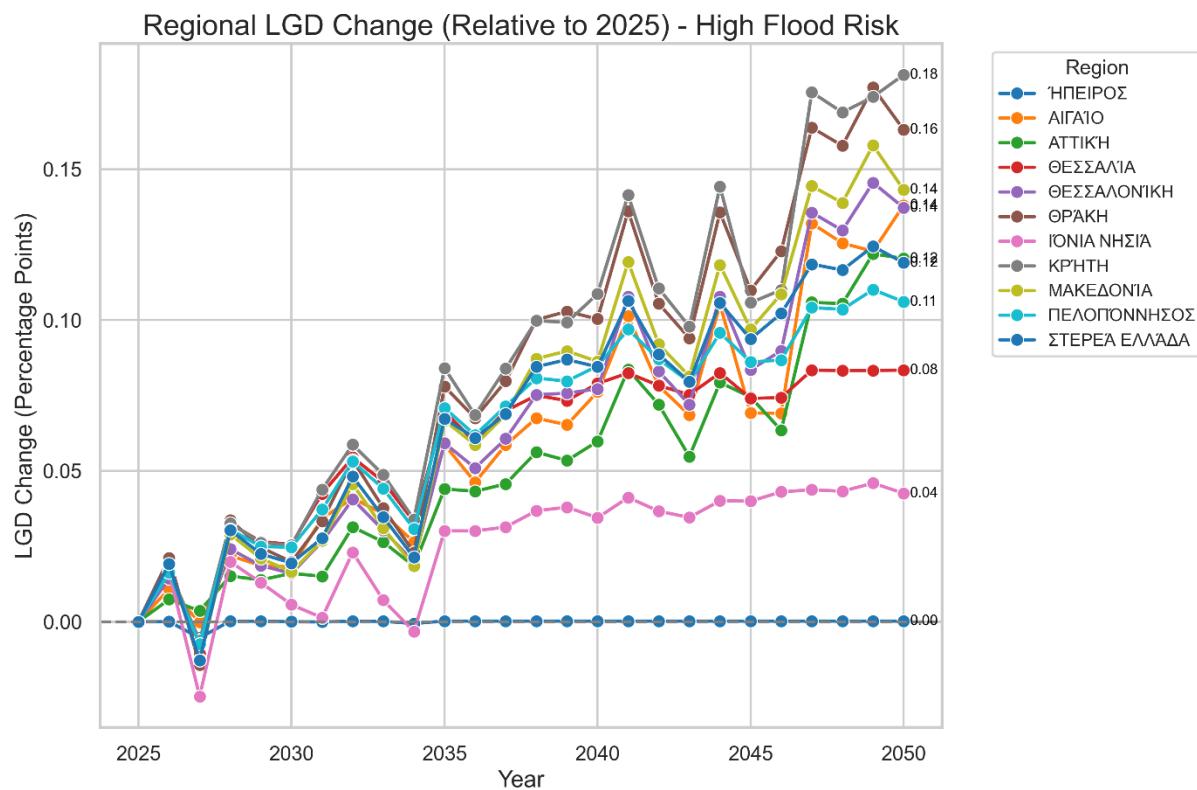
Σε αυτή την ενότητα εξετάζονται τα αποτελέσματα του σεναρίου υψηλού κινδύνου πλημμύρας, το οποίο προσομοιώνει συνθήκες όπου οι φυσικοί κίνδυνοι από πλημμύρες αυξάνονται σημαντικά σε σχέση με το βασικό σενάριο. Μέσω αυτής της προσομοίωσης, αξιολογούνται πώς οι επιπτώσεις των αυξημένων βροχοπτώσεων και της έντασής τους μεταβάλλουν το LGD και επηρεάζουν την πιστωτική έκθεση των στεγαστικών δανείων.

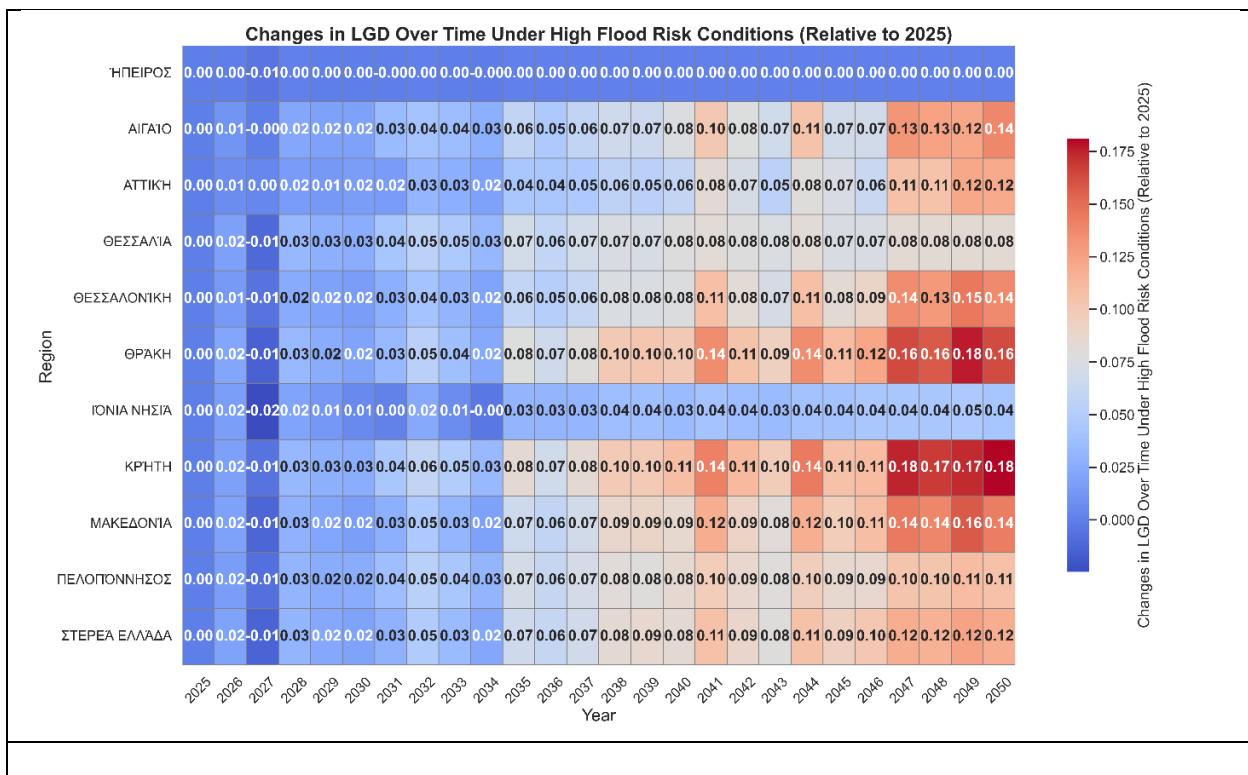


Εικόνα 4.1.2.1 Regional LGD changes for the high flood risk scenario

Στο πρώτο διάγραμμα γραμμών, όπου παρουσιάζεται η εξέλιξη του απόλυτου δείκτη LGD για το σενάριο υψηλού κινδύνου πλημμύρας (High Flood Risk), παρατηρούμε σαφή διαφοροποίηση μεταξύ των περιοχών. Συγκεκριμένα, η Ήπειρος, η Θεσσαλία, η Στερεά Ελλάδα και τα Ιόνια Νησιά παρουσιάζουν τις υψηλότερες απόλυτες τιμές LGD καθ' όλη τη διάρκεια της περιόδου (2025-2050), φτάνοντας ή και ξεπερνώντας το 0,65-0,7. Από την άλλη πλευρά, περιοχές όπως η Αττική και η Θεσσαλονίκη παραμένουν σε αισθητά χαμηλότερα επίπεδα LGD (0,27-0,32 περίπου), υποδεικνύοντας ότι, στο σενάριο πλημμύρας, αυτές οι αστικές περιοχές αντιμετωπίζουν χαμηλότερη απόλυτη έκθεση συγκριτικά με τις πιο εκτεθειμένες γεωγραφικά περιφέρειες.

Το δεύτερο διάγραμμα (heatmap) επιβεβαιώνει με σαφήνεια τις παραπάνω παρατηρήσεις. Οι περιοχές με υψηλότερο απόλυτο LGD (Ηπειρος, Θεσσαλία, Ιόνια Νησιά, Στερεά Ελλάδα) εμφανίζονται έντονα σκουρόχρωμες από την αρχή της περιόδου και σταθεροποιούνται σε ιδιαίτερα υψηλές τιμές, κοντά στο 0,65-0,70.





Εικόνα 4.1.2.2 Regional LGD changes for the high flood risk scenario

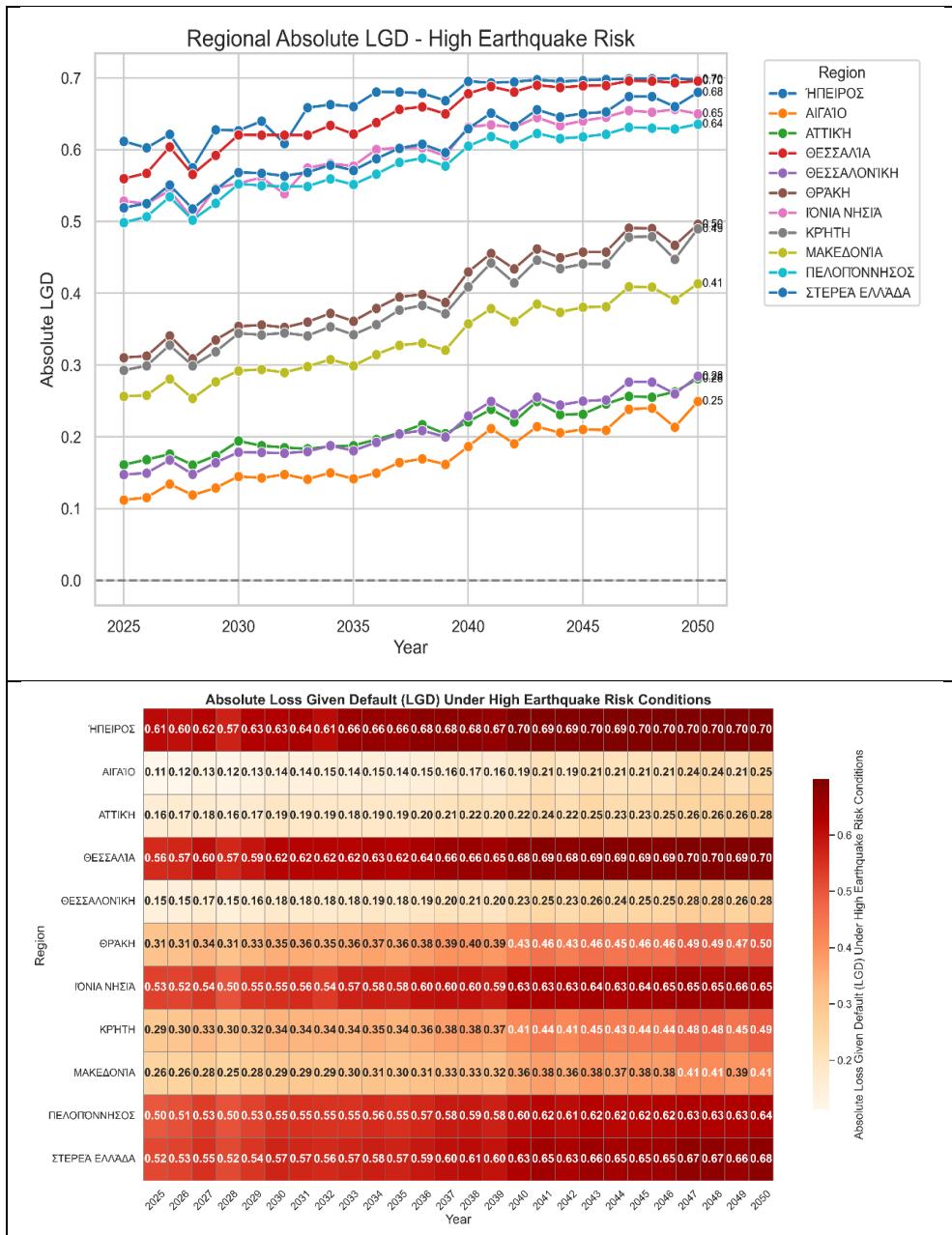
Από το διάγραμμα γραμμών, διαπιστώνουμε ότι οι περιοχές όπως η Ήπειρος, τα Ιόνια Νησιά, η Στερεά Ελλάδα η Θεσσαλία εμφανίζουν σχετικά απότομη αύξηση του LGD, ιδιαίτερα μετά το έτος 2035. Αυτές οι περιοχές, λόγω της γεωγραφικής τους θέσης και της μορφολογίας τους, παρουσιάζουν σημαντική ευαισθησία στις αυξανόμενες πλημμυρικές συνθήκες. Αντίθετα, περιοχές όπως η Αττική, το Αιγαίο και η Θεσσαλονίκη δείχνουν μικρότερες μεταβολές, γεγονός που επιβεβαιώνει τη σχετικά χαμηλότερη έκθεσή τους στον κίνδυνο πλημμύρας.

Στο heatmap, βλέπουμε μια πιο συνολική εικόνα της μεταβολής. Είναι εμφανές ότι περιοχές με εντονότερη αύξηση του LGD—όπως η Ήπειρος, η Θεσσαλία, τα Ιόνια Νησιά και η Στερεά Ελλάδα—μεταβαίνουν σε θερμές αποχρώσεις ήδη από τα πρώτα χρόνια των προσομοιώσεων, υπογραμμίζοντας την υψηλή ευαισθησία τους στις πλημμύρες. Αντιθέτως, η Αττική, η Θεσσαλονίκη και το Αιγαίο εμφανίζουν πιο σταδιακές και ήπιες μεταβολές, παραμένοντας σε ψυχρότερες αποχρώσεις για μεγαλύτερο χρονικό διάστημα.

Η παρουσία περιστασιακών αρνητικών διαφορών στα αρχικά χρόνια (ελαφρώς γαλάζια κελιά) υποδηλώνει τη φυσική διακύμανση της προσομοίωσης Monte Carlo, όπου ορισμένα σενάρια μπορούν να υποδείξουν προσωρινές μειώσεις στον πιστωτικό κίνδυνο, χωρίς ωστόσο να επηρεάζουν τη συνολική τάση αύξησης που είναι σαφής.

4.1.3 Ανάλυση Σεναρίου Υψηλού Σεισμικού Κινδύνου

Στην ενότητα αυτή εξετάζονται τα αποτελέσματα της προσομοίωσης που υλοποιήθηκαν στο πλαίσιο ενός σεναρίου αυξημένου σεισμικού κινδύνου (High Earthquake Risk). Σκοπός του σεναρίου αυτού είναι να αξιολογηθεί πώς μια πιθανή αύξηση στη συχνότητα και στην ένταση των σεισμικών γεγονότων μπορεί να επηρεάσει τον δείκτη LGD και συνεπώς, την αξιοπιστία του στεγαστικού χαρτοφυλακίου.



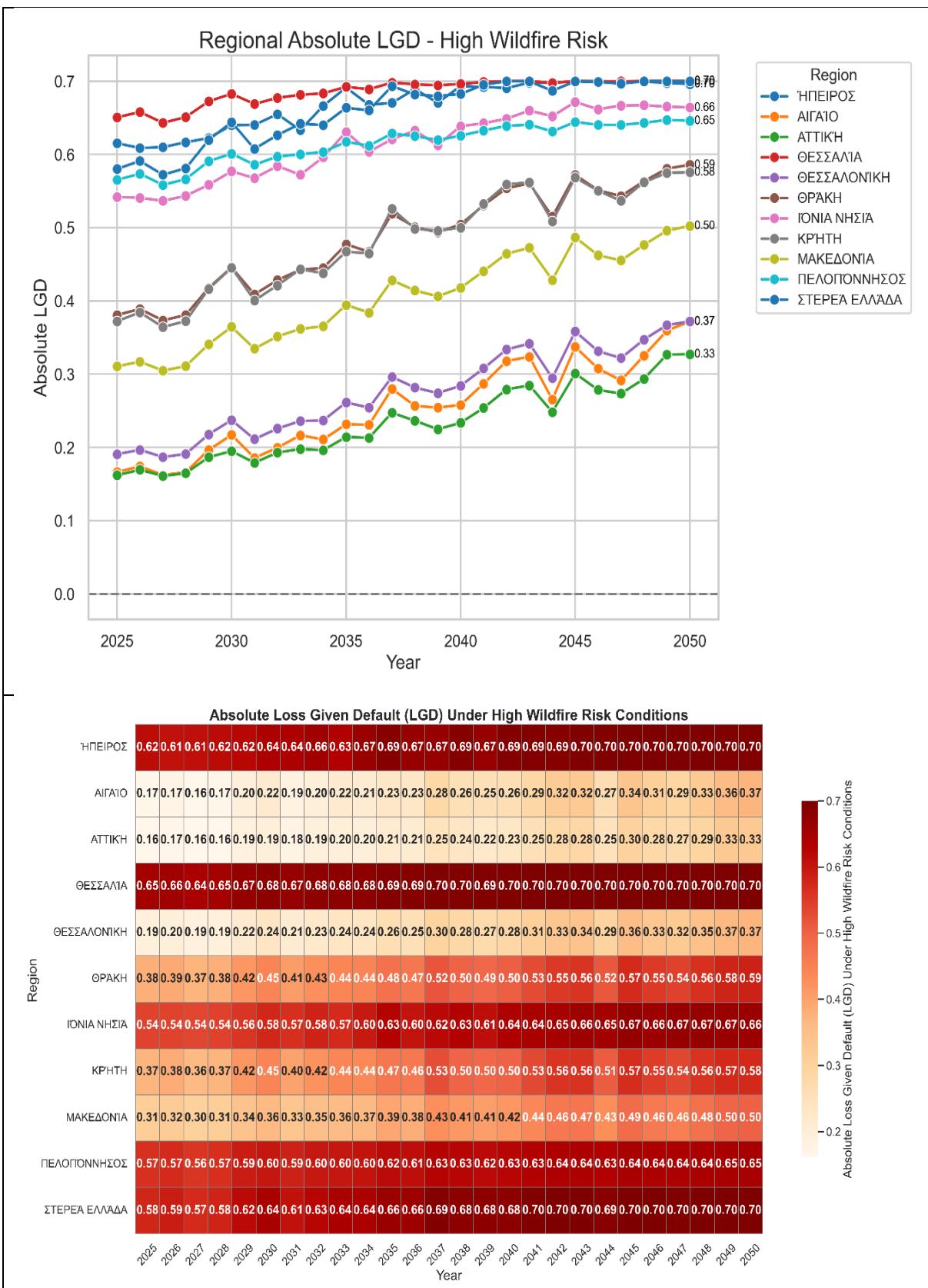
Εκόνα 4.1.3.1 Regional LGD changes for the high earthquake risk scenario

Στην ενότητα αυτή παρουσιάζονται αναλυτικά τα αποτελέσματα του σεναρίου αυξημένου σεισμικού κινδύνου. Εξετάζοντας αρχικά την εξέλιξη του απόλυτου δείκτη LGD μέχρι το 2050, παρατηρούμε ότι ορισμένες περιοχές, όπως η Ήπειρος, η Θεσσαλία και τα Ιόνια Νησιά, εμφανίζουν ιδιαίτερα υψηλές τιμές, που πλησιάζουν το 0,7.

Επιπλέον, μελετώντας τη μεταβολή του LGD συγκριτικά με το έτος βάσης (2025), γίνεται σαφές ότι από το 2035 και μετά κάποιες περιοχές, όπως η Ήπειρος και η Θεσσαλία, σημειώνουν σημαντική αύξηση. Αντιθέτως, περιοχές με σχετικά χαμηλότερη σεισμική δραστηριότητα, όπως η Αττική και το Αιγαίο, εμφανίζουν πιο ήπιες αυξήσεις, γεγονός που επιβεβαιώνει τη διαφοροποίηση του κινδύνου ανάλογα με τα τοπικά χαρακτηριστικά κάθε περιοχής.

4.1.4 Ανάλυση Σεναρίου Υψηλού Κινδύνου Πυρκαγιάς

Στην ενότητα αυτή παρουσιάζονται τα αποτελέσματα της ανάλυσης στο σενάριο υψηλού κινδύνου πυρκαγιάς.

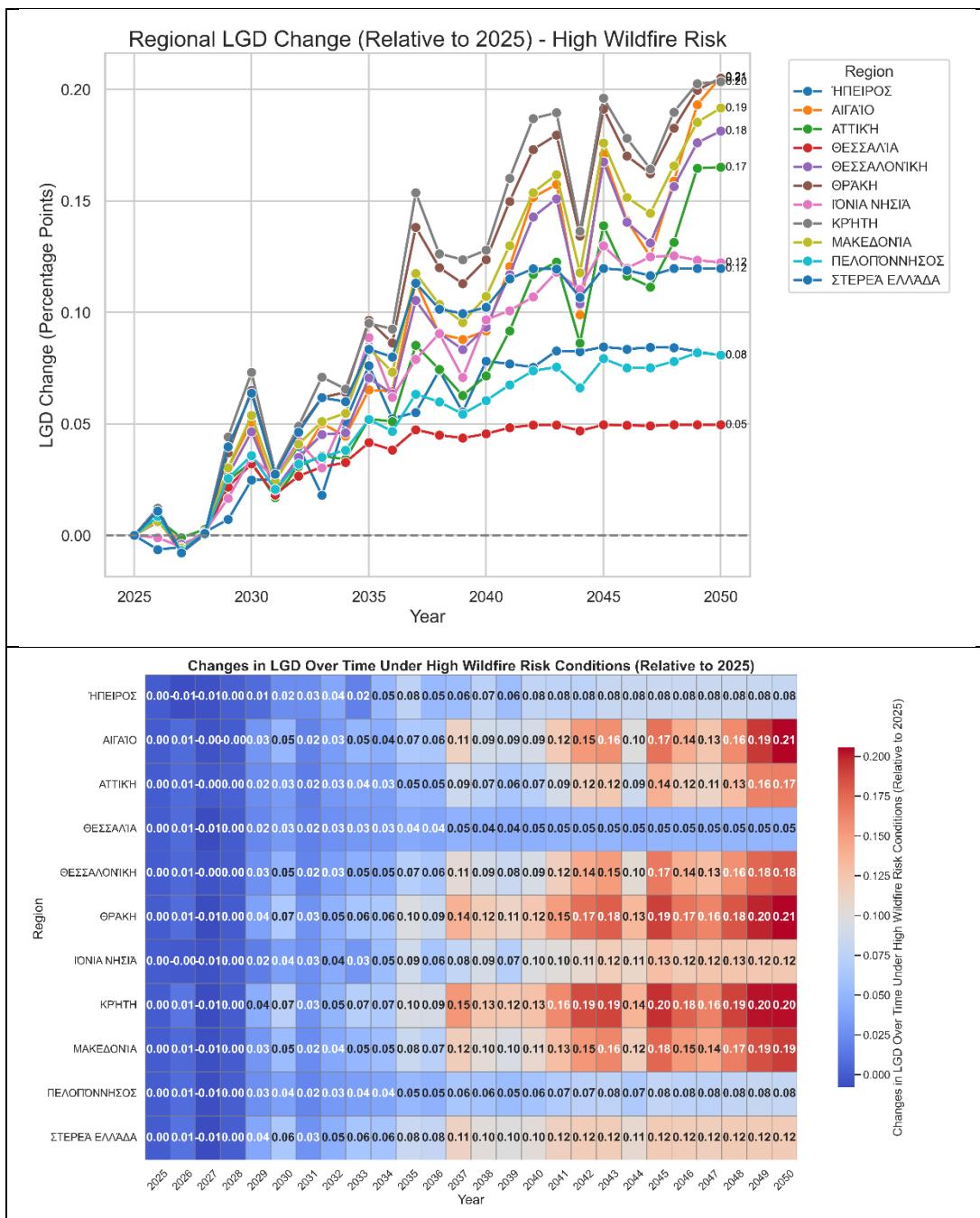


Εικόνα 4.1.4.1 Regional absolute LGD values for the high wildfire risk scenario

Στο σενάριο υψηλού κινδύνου πυρκαγιάς, τα αποτελέσματα παρουσιάζουν μια έντονη αυξητική τάση του δείκτη LGD στις περισσότερες γεωγραφικές περιοχές της Ελλάδας. Παρατηρούμε πως η Θεσσαλία, η Ήπειρος και η Στερεά Ελλάδα εμφανίζουν ιδιαίτερα υψηλές τιμές, φτάνοντας μέχρι και το 0.70 (70%) προς το τέλος της υπό εξέταση περιόδου (έτος 2050).

Εξετάζοντας πιο συγκεκριμένα τη σχετική μεταβολή του δείκτη LGD σε σχέση με το έτος βάσης (2025), διαπιστώνουμε ότι ιδιαίτερα μεγάλες μεταβολές παρουσιάζουν περιοχές όπως η Ήπειρος, η Μακεδονία, η Θράκη και η Κρήτη, με τις τιμές να αυξάνονται μέχρι και κατά 0.21 ποσοστιαίες μονάδες. Αντιθέτως, οι περιοχές του Αιγαίου και της Αττικής παρουσιάζουν πιο συγκρατημένη μεταβολή, καταδεικνύοντας ενδεχομένως μια χαμηλότερη ευαισθησία στον κίνδυνο πυρκαγιάς.

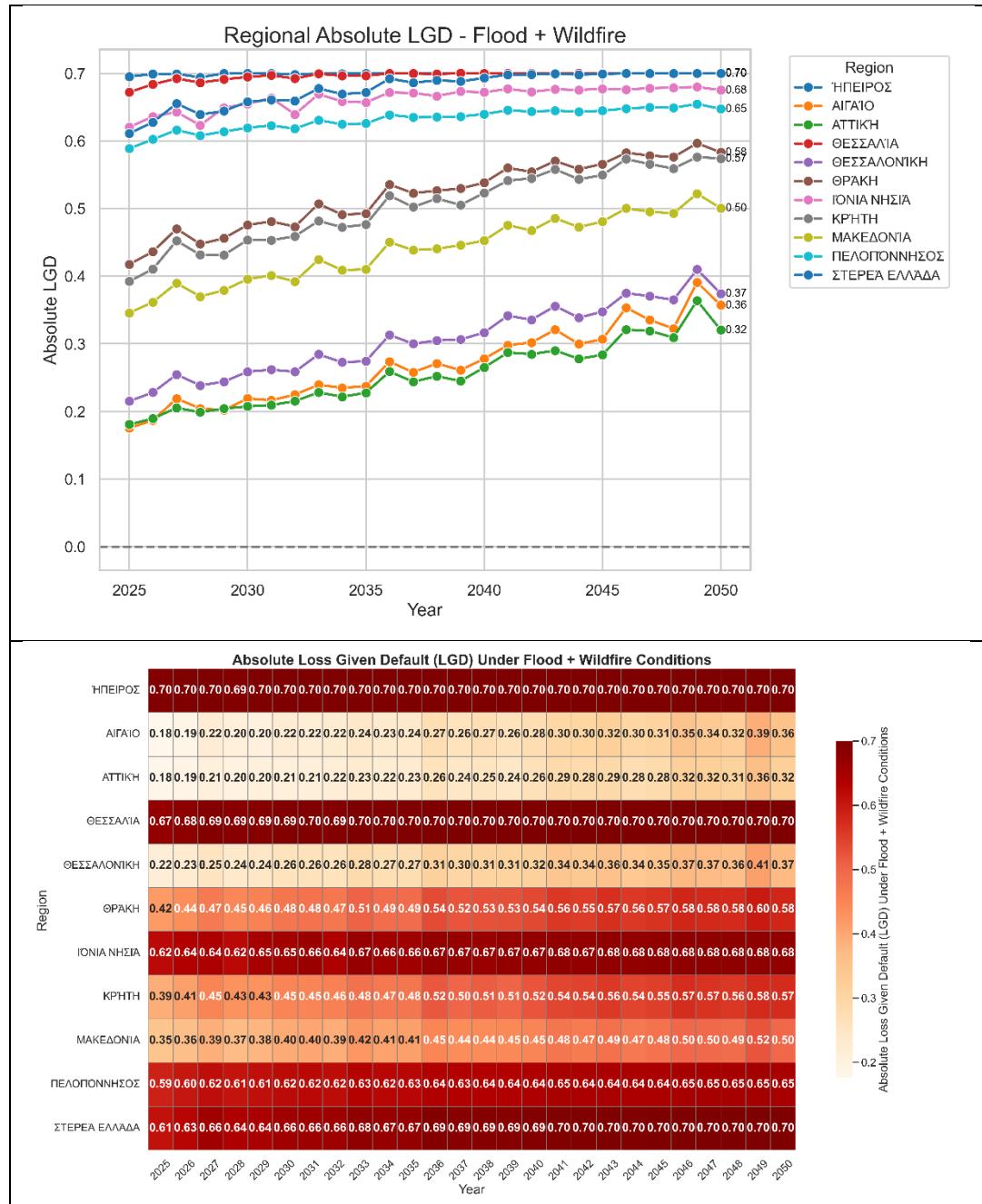
Οι απεικονίσεις heatmap αποτυπώνουν σαφέστερα την τάση αυτή, καθώς με την πάροδο του χρόνου τα χρώματα γίνονται όλο και πιο έντονα προς το κόκκινο, ειδικά για τις περιοχές της Θράκης και της Κρήτης. Από την άλλη, το Αιγαίο και η Αττική εμφανίζουν ηπιότερες αυξήσεις, κάτι που μπορεί να οφείλεται σε ιδιαίτερα γεωγραφικά και κλιματικά χαρακτηριστικά των συγκεκριμένων περιοχών.



Εικόνα 4.1.4.2 Regional LGD changes for the high wildfire risk scenario

4.1.5 Ανάλυση Συνδυασμένου Σεναρίου Υψηλού Κινδύνου Πλημμύρας και Πυρκαγιάς

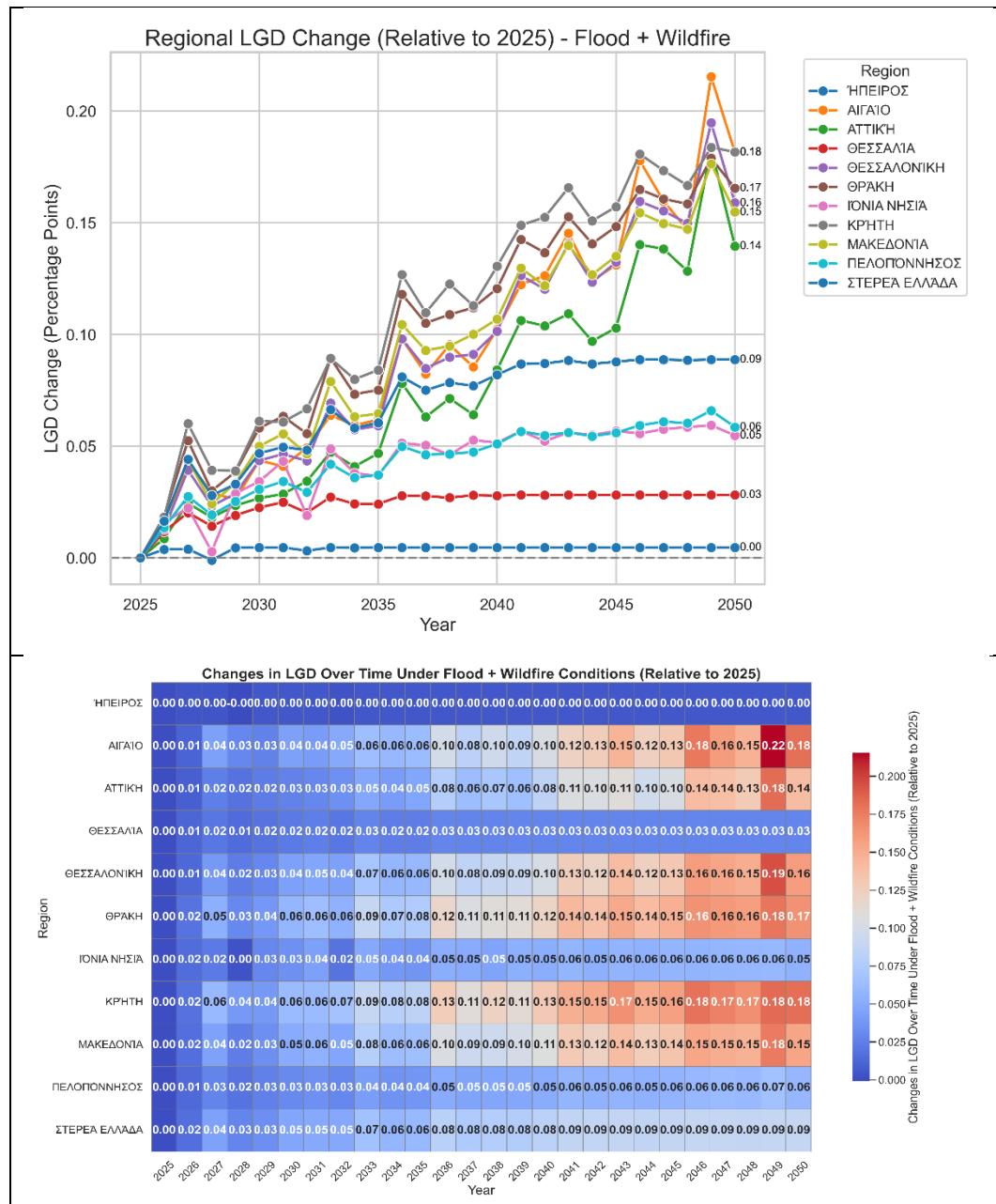
Στην ενότητα αυτή παρουσιάζεται η ανάλυση του σύνθετου σεναρίου υψηλού κινδύνου πλημμύρας και πυρκαγιάς, που στόχο έχει να αποτυπώσει την κοινή επίδραση δύο σημαντικών κλιματικών κινδύνων στον δείκτη LGD.



Eikóva 4.1.5.1 Regional absolute LGD values for the flood+wildfire risk scenario

Από τα αποτελέσματα που παρουσιάζονται, γίνεται εμφανές ότι περιοχές όπως η Ήπειρος και η Θεσσαλία αγγίζουν σχεδόν αμέσως το ανώτατο επίπεδο του απόλυτου LGD (0,70) ήδη από τα πρώτα έτη της προσομοίωσης. Παρόμοια υψηλή επίδραση παρουσιάζουν επίσης η Στερεά Ελλάδα, η Πελοπόννησος και τα Ιόνια Νησιά, τα οποία επίσης κινούνται προς πολύ υψηλές τιμές LGD, με ανοδική τάση που κορυφώνεται έως το 2050.

Η δυναμική αύξησης του LGD δείχνει ιδιαίτερα έντονη μεταβολή στις περιοχές της Θράκης, της Κρήτης και των Ιονίων Νήσων, όπου οι ποσοστιαίες μονάδες αυξάνονται σημαντικά, ειδικά μετά το έτος 2040.



Εικόνα 4.1.5.2 Regional LGD changes for the flood+wildfire risk scenario

4.2 Ανάλυση Μηχανικής Μάθησης (Machine Learning Analysis)

Στην ενότητα αυτή, παρουσιάζεται η εφαρμογή της μηχανικής μάθησης στην εκτίμηση του πιστωτικού κινδύνου των στεγαστικών δανείων. Ο βασικός στόχος της ανάλυσης μηχανικής μάθησης είναι η δημιουργία ενός μοντέλου που θα μπορεί να προβλέψει την εκτίμηση του δείκτη LGD (Loss Given Default).

Η διαδικασία που ακολουθήσαμε με τη χρήση του μοντέλου μηχανικής μάθησης (ML) για την ανάλυση των κινδύνων των ακινήτων περιλαμβάνει τη χρήση δύο κυριότερων αλγορίθμων: **Random Forest** και **XGBoost**. Παρακάτω παρατίθενται τα αποτελέσματα που προκύπτουν από τα μοντέλα αυτά και η ερμηνεία τους.

Τα αποτελέσματα της ανάλυσης μηχανικής μάθησης εξετάστηκαν μέσω διαφόρων μετρικών αξιολόγησης, όπως το **R²** για την αξιολόγηση της ικανότητας του μοντέλου να εξηγήσει τη μεταβλητότητα των δεδομένων, το **Mean Squared Error (MSE)** για την εκτίμηση του μέσου σφάλματος και το **Accuracy** για την αξιολόγηση της ποιότητας των προβλέψεων σε σχέση με τα πραγματικά δεδομένα.

Για την αξιολόγηση και εκπαίδευση των μοντέλων, διαιρέθηκε το σύνολο δεδομένων σε δύο υποσύνολα: το 70% των δεδομένων χρησιμοποιήθηκε για την εκπαίδευση και το υπόλοιπο 30% για τη δοκιμή του μοντέλου. Επιπλέον, εφαρμόστηκε στρατηγική 5-fold cross-validation χρησιμοποιώντας την προσέγγιση K-Fold με 5 διαιρέσεις.

Random Forest Regressor:

Χρησιμοποιήθηκαν 200 δέντρα (`n_estimators=200`), με μέγιστο βάθος 20 (`max_depth=20`) και ορίστηκε ότι για κάθε διαχωρισμό ενός εσωτερικού κόμβου απαιτούνται τουλάχιστον 10 δείγματα (`min_samples_split=10`).

XGBoost Regressor:

Σε αυτό το μοντέλο χρησιμοποιήσαμε 200 εκτιμητές, έναν ρυθμό μάθησης (learning rate) 0.2 και μέγιστο βάθος δέντρου (maximum tree depth) 10.

4.2.1 Αποτελέσματα Cross-Validation (CV) για το Random Forest

Το μοντέλο αυτό αξιολογήθηκε μέσω cross-validation, με τη μέθοδο R² και το MSE (Mean Squared Error) να χρησιμοποιούνται για τη μέτρηση της απόδοσης του μοντέλου.

- Αποτελέσματα R^2 για το Cross-Validation στα 5 σενάρια:
 - [0.95043106, 0.95027388, 0.95037601, 0.94950217, 0.94970113]

Η υψηλή τιμή του R^2 (περίπου 95%) υποδηλώνει ότι το μοντέλο Random Forest μπορεί να εξηγήσει περίπου το 95% της διακύμανσης της εξαρτημένης μεταβλητής LGD. Αυτό υποδηλώνει ότι το μοντέλο μπορεί να κάνει αξιόπιστες προβλέψεις για τις αλλαγές του LGD, βασισμένο στις εισόδους του.

- Μέσο R^2 : 0.9501

Η μέση τιμή του R^2 επιβεβαιώνει την ισχυρή ικανότητα του μοντέλου να αναπαράγει τα δεδομένα χωρίς σημαντικά σφάλματα.

- MSE (Μέσο Τετραγωνικό Σφάλμα): 0.0018964616813414111

Το μικρό MSE δείχνει ότι το μοντέλο έχει πολύ μικρές αποκλίσεις στις προβλέψεις του.

4.2.2 Αποτελέσματα Cross-Validation (CV) για το XGBoost

Το XGBoost είναι ένας άλλος ισχυρός αλγόριθμος που χρησιμοποιήθηκε για την ανάλυση του κινδύνου LGD. Όπως και το Random Forest, το XGBoost αξιολογήθηκε μέσω του cross-validation με τη χρήση της μετρικής R^2 .

- Αποτελέσματα R^2 για το Cross-Validation:
 - [0.96591853, 0.96625589, 0.96519239, 0.96531286, 0.96546773]

Όπως φαίνεται, το XGBoost αποδίδει καλύτερα σε σχέση με το Random Forest, με μέση τιμή $R^2 = 0.9656$. Αυτό σημαίνει ότι το μοντέλο XGBoost εξηγεί ακόμα καλύτερα τη διακύμανση των δεδομένων και έχει μεγαλύτερη ικανότητα πρόβλεψης.

- Μέσο R^2 : 0.9656

Αυτό δείχνει ότι το XGBoost αποδίδει με μεγαλύτερη ακρίβεια, εξηγώντας σχεδόν το 97% της διακύμανσης της εξαρτημένης μεταβλητής LGD.

- MSE: 0.0013175672644319424

Η τιμή του MSE στο XGBoost είναι χαμηλότερη από την τιμή του Random Forest, κάτι που δείχνει ότι το XGBoost κάνει ακόμη πιο ακριβείς προβλέψεις.

4.2.3. Σημασία Χαρακτηριστικών (Feature Importances)

Η σημασία των χαρακτηριστικών δείχνει πόσο σημαντικά είναι τα διάφορα χαρακτηριστικά για την πρόβλεψη του LGD. Από τα αποτελέσματα, μπορούμε να δούμε ότι ο κίνδυνος πλημμύρας (Flood Risk) είναι το πιο σημαντικό χαρακτηριστικό.

Random Forest:

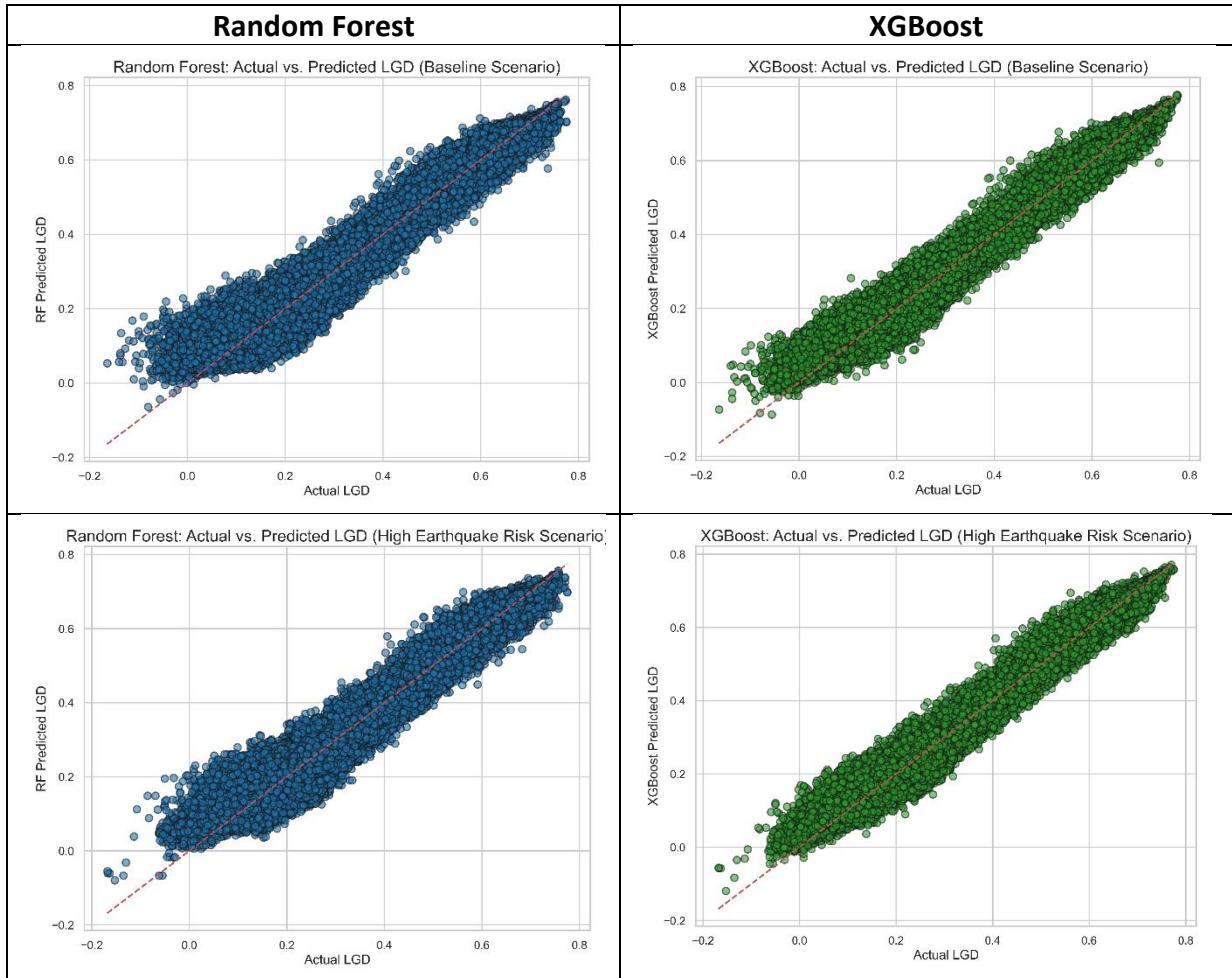
- Flood_Risk: 0.3780
- wildfire_risk: 0.2191
- Earthquake_Risk: 0.0974
- Year: 0.0534
- loan_amount: 0.0153

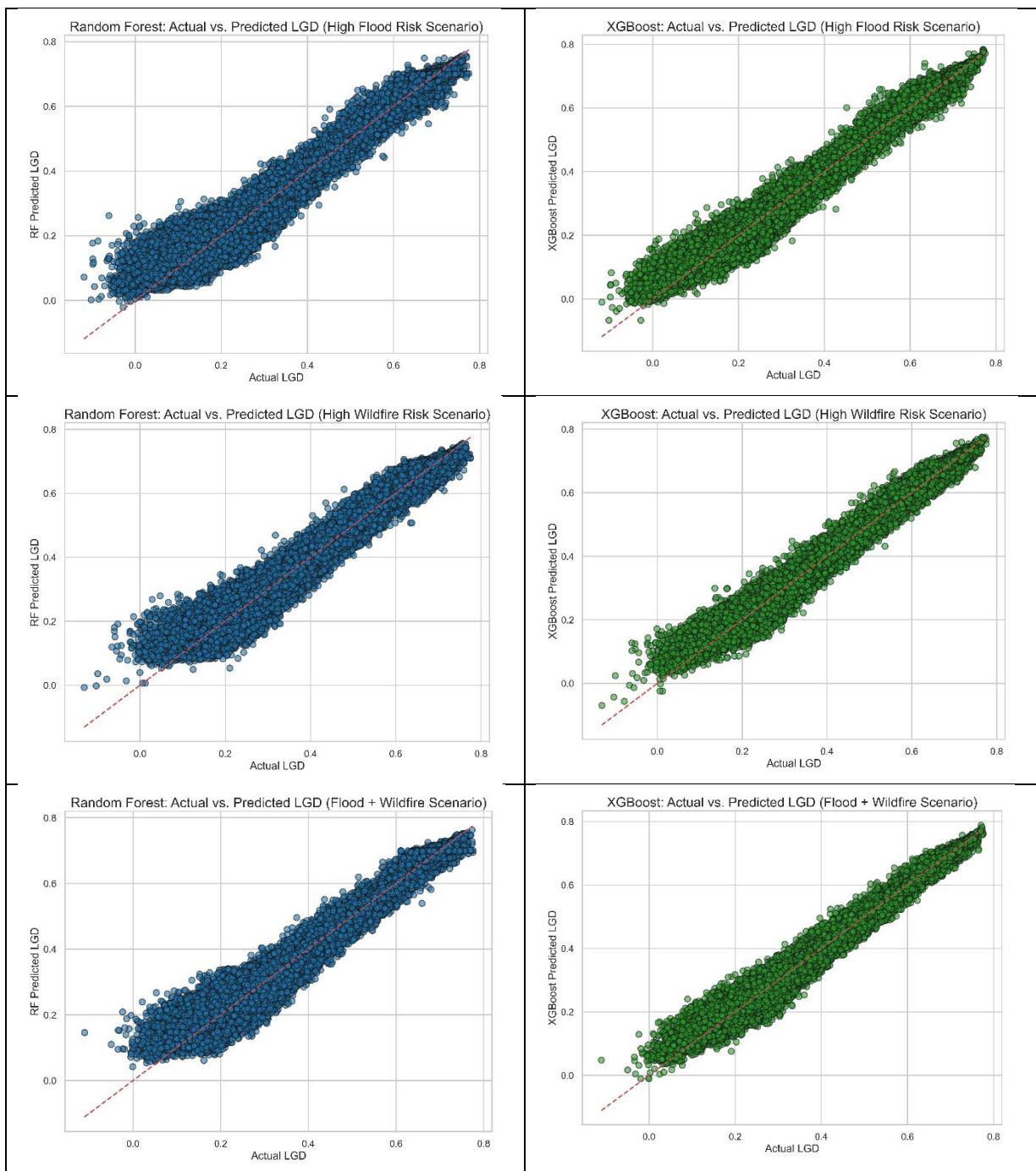
XGBoost:

- Flood_Risk: 0.2802
- Earthquake_Risk: 0.1553
- wildfire_risk: 0.1544
- Scenario_Flood + Wildfire: 0.0985
- Year: 0.1199

Ο κίνδυνος πλημμύρας (Flood Risk) είναι το πιο σημαντικό χαρακτηριστικό τόσο στο Random Forest όσο και στο XGBoost. Ενδιαφέρον παρουσιάζει το γεγονός ότι ο κίνδυνος σεισμού (Earthquake Risk) θεωρείται πιο σημαντικός στο XGBoost σε σύγκριση με το Random Forest.

Παρακάτω φαίνονται τα γραφήματα που συγκρίνουν τις Monte Carlo προσομοιώσεις με αυτές των μοντέλων προκειμένου να παρατηρηθεί πόσο κοντά είναι μεταξύ τους.





Εικόνα 4.2.1 Random Forest and XGBoost comparisons with LGD values of Monte Carlo simulations

Για το σενάριο **High Wildfire Risk**, τα διαγράμματα παρουσιάζουν μια ισχυρή γραμμική σχέση μεταξύ των προβλεπόμενων και των πραγματικών τιμών, με τα σημεία να συγκεντρώνονται κοντά στην ιδανική ευθεία (actual = predicted). Αυτή η στενή σύγκλιση

υποδηλώνει ότι και τα δύο μοντέλα, Random Forest και XGBoost, καταφέρνουν να αποδώσουν με αξιοπιστία τις επιπτώσεις του αυξημένου κινδύνου πυρκαγιών στο LGD.

Στο σενάριο **High Flood Risk**, τα διαγράμματα εμφανίζουν επίσης εξαιρετική απόδοση, όπου τα μοντέλα καταγράφουν υψηλή σύγκλιση μεταξύ προβλεπόμενων και πραγματικών τιμών. Ωστόσο, διαπιστώνεται ότι οι περιοχές που παρουσιάζουν υψηλότερη έκθεση σε πλημμύρες καταδεικνύουν ελαφρώς μεγαλύτερη διακύμανση στις προβλέψεις, κάτι που μπορεί να οφείλεται στην πολυπλοκότητα των επιπτώσεων των πλημμυρών στα ακίνητα.

Στο σενάριο **High Earthquake Risk**, παρατηρούμε ότι και πάλι οι προβλέψεις των μοντέλων συγκλίνουν πολύ καλά με τις πραγματικές τιμές, αν και μπορεί να υπάρχει κάποια μικρή διακύμανση στα άκρα των τιμών. Αυτό υποδηλώνει ότι το μοντέλο καταφέρνει να χειριστεί τις επιπτώσεις της αυξημένης σεισμικής δραστηριότητας, η οποία μπορεί να έχει διαφορετικό χαρακτήρα σε σχέση με τους άλλους φυσικούς κινδύνους.

Τέλος, στο συνδυαστικό σενάριο **Flood + Wildfire Risk**, τα διαγράμματα δείχνουν την επίδραση της ταυτόχρονης παρουσίας αυξημένου κινδύνου πλημμύρας και πυρκαγιών. Εδώ, οι προβλεπόμενες τιμές παρουσιάζουν επίσης υψηλή σύγκλιση με τις πραγματικές, υποδεικνύοντας ότι τα μοντέλα μπορούν να ενσωματώσουν τη σύνθετη αλληλεπίδραση μεταξύ διαφορετικών φυσικών κινδύνων. Παρά το γεγονός ότι το συνδυαστικό σενάριο προκαλεί μεγαλύτερη μεταβλητότητα στις τιμές, η γενική γραμμική σχέση παραμένει σταθερή, επιβεβαιώνοντας την ικανότητα των μοντέλων να διαχειρίζονται πολύπλοκα σενάρια κινδύνου.

Συμπερασματικά, τα διαγράμματα Actual vs. Predicted τα μοντέλα μηχανικής μάθησης δεν παρουσιάζουν υπερεκπαίδευση, καθώς οι προβλέψεις τους είναι σταθερές και γενικεύονται καλά στα δεδομένα δοκιμής.

5. Συμπεράσματα και Περιορισμοί

Συμπεράσματα

Η παρούσα εργασία εστιάζει στην ανάπτυξη μοντέλων εκτίμησης πιστωτικού κινδύνου για στεγαστικά δάνεια στην Ελλάδα, λαμβάνοντας υπόψη τους φυσικούς κινδύνους της κλιματικής αλλαγής και του σεισμογενούς χαρακτήρα της χώρας, όπως οι πλημμύρες, οι πυρκαγιές και οι σεισμοί. Χρησιμοποιήθηκαν το Random Forest και το XGBoost, για την πρόβλεψη του δείκτη Loss Given Default (LGD). Εκτός από τα δεδομένα των δανείων, σημαντικό ρόλο στην εκτίμηση του κινδύνου διαδραμάτισαν οι προσομοιώσεις φυσικών κινδύνων, οι οποίες παρείχαν κρίσιμες πληροφορίες για τα σενάρια κινδύνου που εξετάστηκαν.

Οι προσομοιώσεις των φυσικών κινδύνων (π.χ., σενάρια πλημμυρών, πυρκαγιών, σεισμών) έδωσαν τιμές για τη μεταβολή του LGD σε ακραίες καταστάσεις, οι οποίες στη συνέχεια χρησιμοποιήθηκαν ως πρόσθετα χαρακτηριστικά στα μοντέλα μηχανικής μάθησης. Τα αποτελέσματα από την εφαρμογή των μοντέλων αυτών έδειξαν ότι τόσο το **Random Forest** όσο και το **XGBoost** παρουσίασαν εξαιρετική απόδοση, με το **XGBoost** να υπερέχει λόγω της καλύτερης απόδοσης στην εκτίμηση του LGD.

Συνολικά, τα μοντέλα αυτά απέδωσαν υψηλούς δείκτες **R²** και χαμηλούς **MSE**, αποδεικνύοντας την αξία τους στην εκτίμηση του πιστωτικού κινδύνου με βάση τα φυσικά σενάρια κινδύνου. Οι υψηλοί δείκτες **R²** από τις διασταυρωμένες αξιολογήσεις (cross-validation) και οι χαμηλοί **MSE** υποδεικνύουν ότι τα μοντέλα δεν υπέφεραν από υπερεκπαίδευση (overfitting), καθώς παρουσίασαν σταθερή απόδοση τόσο στα δεδομένα εκπαίδευσης όσο και στις δοκιμές. Τα σενάρια πλημμύρας και πυρκαγιάς είχαν τον πιο σημαντικό αντίκτυπο στη μεταβολή του **LGD**, με τις περιοχές υψηλού κινδύνου να παρουσιάζουν μεγαλύτερες αυξήσεις στους δείκτες.

Επιπλέον, η ανάλυση της μεταβολής του LGD υπό διαφορετικά σενάρια κινδύνου (πλημμύρες, πυρκαγιές, σεισμοί) έδειξε ότι οι περιοχές με υψηλή φυσική επικινδυνότητα, όπως η Ήπειρος, η Θεσσαλία και τα Ιόνια Νησιά, αντιμετωπίζουν αυξημένο πιστωτικό κίνδυνο, κάτι που επιβεβαιώνεται από τις αυξήσεις του LGD σε αυτά τα σενάρια. Αυτά τα ευρήματα υπογραμμίζουν τη σημασία της ενσωμάτωσης των φυσικών κινδύνων στις στρατηγικές διαχείρισης πιστωτικού κινδύνου των τραπεζών και άλλων χρηματοπιστωτικών οργανισμών.

Περιορισμοί

Παρά τα θετικά αποτελέσματα και την αξία της προσέγγισης, υπάρχουν ορισμένοι περιορισμοί που επηρεάζουν τα συμπεράσματα της έρευνας:

- Περιορισμένα δεδομένα:** Η διαθεσιμότητα των δεδομένων ήταν περιορισμένη, καθώς χρησιμοποιήθηκαν ανοιχτές πηγές δεδομένων που παρέχουν περιορισμένες πληροφορίες για συγκεκριμένες περιοχές της Ελλάδας. Η έλλειψη πλήρων δεδομένων για ακίνητες περιουσίες στην Ελλάδα και η ανεπάρκεια δεδομένων για ακραίες καταστάσεις περιορίζουν την ακριβή εκτίμηση του κινδύνου σε διάφορες περιοχές.
- Περιορισμένοι υπολογιστικοί πόροι:** Η υπολογιστική ισχύς του υπολογιστή που χρησιμοποιήθηκε για την εκπαίδευση των μοντέλων ήταν περιορισμένη. Παρά τη χρήση εξελιγμένων αλγορίθμων, η επεξεργασία μεγάλων όγκων δεδομένων και η εκπαίδευση μοντέλων όπως το **XGBoost** απαιτεί μεγαλύτερη υπολογιστική δύναμη, κάτι που ενδέχεται να περιορίσει τις δυνατότητες βελτιστοποίησης των μοντέλων για ακόμα πιο σύνθετα σενάρια. Πιο συγκεκριμένα, οι προσομοιώσεις έφταναν τις 100m και εμείς μπορούσαμε να χρησιμοποιήσουμε ένα μικρό δείγμα των 5m.
- Σενάρια προσομοίωσης:** Αν και οι προσομοιώσεις για τα φυσικά σενάρια (πλημμύρες, σεισμοί και πυρκαγιές) προσδιόρισαν τις μεταβολές του LGD, οι παραδοχές που χρησιμοποιήθηκαν για την κατασκευή αυτών των σεναρίων ενδέχεται να μην καλύπτουν πλήρως όλες τις πιθανές μεταβλητές που θα μπορούσαν να επηρεάσουν τις συνέπειες αυτών των κινδύνων στο μέλλον. Για παράδειγμα, οι παράγοντες όπως η ανακαίνιση των κτιρίων δεν ενσωματώθηκαν πλήρως στις προσομοιώσεις, γεγονός που ενδέχεται να περιορίσει την ακρίβεια των εκτιμήσεων.
- Γενίκευση των αποτελεσμάτων:** Τα αποτελέσματα της μελέτης είναι επικεντρωμένα στην Ελλάδα και βασίζονται σε δεδομένα για συγκεκριμένες περιοχές και φυσικούς κινδύνους. Η προσαρμογή αυτών των μοντέλων σε άλλες χώρες ή περιοχές με διαφορετικά χαρακτηριστικά κινδύνου ή οικονομικές συνθήκες απαιτεί πρόσθετες έρευνες και βελτιώσεις.
- Υποθέσεις και απλοποιήσεις:** Για την ανάπτυξη των μοντέλων, έγιναν ορισμένες υποθέσεις που ενδέχεται να μην αντανακλούν πλήρως την πραγματικότητα. Για παράδειγμα, η προσέγγιση για τον υπολογισμό των κινδύνων βασίστηκε σε δεδομένα προσομοίωσης, τα οποία ενδέχεται να μην καταγράψουν όλες τις πραγματικές μεταβλητές που επηρεάζουν τον πιστωτικό κίνδυνο.

6. Προτάσεις για Μελλοντική Έρευνα

Η παρούσα εργασία ανέπτυξε μοντέλα μηχανικής μάθησης για την εκτίμηση του πιστωτικού κινδύνου που προκύπτει από φυσικές καταστροφές, ωστόσο υπάρχουν αρκετοί τομείς στους οποίους μπορεί να επεκταθεί και να βελτιωθεί η έρευνα στο μέλλον. Παρακάτω παρατίθενται ορισμένες προτάσεις για μελλοντική έρευνα που θα μπορούσαν να βελτιώσουν την ακρίβεια των μοντέλων και να διευρύνουν την εφαρμογή τους.

- **Επέκταση των Δεδομένων:** Ένας από τους κύριους περιορισμούς της τρέχουσας μελέτης ήταν η διαθεσιμότητα δεδομένων. Η συλλογή περισσότερων και πιο ποιοτικών δεδομένων από διαφορετικές γεωγραφικές περιοχές της Ελλάδας ή ακόμα και από άλλες χώρες θα μπορούσε να βοηθήσει στην καλύτερη εκτίμηση των κινδύνων. Επιπλέον, η ενσωμάτωση νέων παραμέτρων, όπως οι τοπικές οικονομικές συνθήκες και τα χαρακτηριστικά των ακινήτων (π.χ. έτος κατασκευής, υλικό κατασκευής κ.λπ.), θα μπορούσε να βελτιώσει την ακρίβεια των προβλέψεων.
- **Υπολογιστικοί Πόροι και Βελτιστοποίηση Μοντέλων:** Για την εκπαίδευση πιο προηγμένων μοντέλων μηχανικής μάθησης και την καλύτερη αξιολόγηση των δεδομένων, η έρευνα αυτή θα μπορούσε να επωφεληθεί από τη χρήση ισχυρότερων υπολογιστικών πόρων. Η βελτιστοποίηση των αλγορίθμων και η χρήση τεχνικών όπως το **ensemble learning** ή το **deep learning** μπορεί να βελτιώσει την απόδοση των μοντέλων και να προσφέρει ακόμα ακριβέστερες προβλέψεις για το πιστωτικό ρίσκο.
- **Περισσότερες προσομοιώσεις σε πραγματικές συνθήκες:** Η δημιουργία πιο ρεαλιστικών προσομοιώσεων που να αντανακλούν τις μελλοντικές συνθήκες κινδύνου, όπως η αύξηση της συχνότητας φυσικών καταστροφών λόγω της κλιματικής αλλαγής, θα μπορούσε να προσφέρει πιο ακριβείς προβλέψεις για το μέλλον.

7. ΒΙΒΛΙΟΓΡΑΦΙΑ

Adamopoulos, I. P., Syrou, N. F., Adamopoulou, J. P., & Mijwil, M. M. (2024). Southeast Mediterranean and Middle Eastern countries are experiencing impacts from the climate crisis, extreme weather events, and the conventional method of water use: A comprehensive scoping study. SSRN.

Schuermann, T. (2004). What do we know about loss given default? SSRN.

Author Unknown. (2023, September 7). Extreme flooding caused by Storm Daniel devastates Greece. *Financial Times*. <https://www.ft.com/content/b46d34c0-fef8-45f2-8930-a01c10b4d13a>

Feng, Z., Wang, C., & Lu, Y. (2022). The impact of climatic disaster on corporate investment policy. *Journal of Multinational Financial Management*, 66, 100773. <https://doi.org/10.1016/j.mulfin.2022.100773>

Caouette, J., Altman, E., Narayanan, P., & Nimmo, R. (2008). Managing Credit Risk: The Great Challenge for the Global Financial Markets. Wiley.

Bank for International Settlements. (n.d.). History of the Basel Committee. *Bank for International Settlements*. Retrieved March 18, 2025, from <https://www.bis.org/bcbs/history.htm>

Adamko, P., Klieštik, T., & Birtus, M. (2014). History of credit risk models. In *2014 2nd International Conference on Economics and Social Science (ICESS 2014)*. University of Žilina, Slovakia.

Valášková, K., & Donegov, V. (2014). Assessing credit risk by Moody's KMV model. ResearchGate. Retrieved from <https://www.researchgate.net/publication/289204220>

Basel Committee on Banking Supervision. (2005, July). *An Explanatory Note on the Basel II IRB Risk Weight Functions*. Bank for International Settlements.

Adamko, P., Klieštik, T., & Birtus, M. (n.d.). History of Credit Risk Models.

Wong, M., & Ho, H. (2023). A Framework for Integrating Extreme Weather Risk, Probability of Default, and Loss Given Default for Residential Mortgage Loans. *Sustainability*, 15(15), 11808. <https://doi.org/10.3390/su151511808>

Greve, C., & Hahnenstein, L. (n.d.). Stress Testing the Credit Risk of Mortgage Loans: The Relationship between Portfolio-LGD and the Loan-to-Value Distribution.

Siegers, E. H. (2023). Assessing the economic implications of physical climate risks on Dutch banks: A case study [Master's thesis, University of Twente]. Faculty of Behavioural Management and Social Sciences.

Wikipedia Contributors. (n.d.). Monte Carlo method. *Wikipedia*.

https://en.wikipedia.org/wiki/Monte_Carlo_method

Amazon Web Services. (n.d.). What is Monte Carlo simulation? AWS.

<https://aws.amazon.com/what-is/monte-carlo-simulation/>

Investopedia. (n.d.). Monte Carlo simulation basics. *Investopedia*.

<https://www.investopedia.com/articles/financial-theory/08/monte-carlo-multivariate-model.asp>

Nimmala, R. (2024). Enhancing Credit Risk Assessment in the Era of Climate Change: A Machine Learning Approach for Loan and Lease Portfolios. *Journal of Artificial Intelligence, Machine Learning and Data Science*, 2(2), 594-601. <https://doi.org/10.51219/JAIMLD/rohit-nimmala/155>

Mbanyele, W., & Muchenje, L. (2022). Climate change exposure, risk management and corporate social responsibility: Cross-country evidence. *Journal of Multinational Financial Management*, 66, 100771. <https://doi.org/10.1016/j.mulfin.2022.100771>

CLIMPACT. (n.d.). Climatic Indices for Greece [Data set]. *CLIMPACT*.

<https://data.climpact.gr/en/dataset/497dc26d-45e0-4ad5-b8f3-5f8890f65129/resource/aef8997e-3b0f-4b02-9f59-0751a6093936>

Doulos, N. (n.d.). Greece's Earthquakes [Data set]. *Kaggle*.

<https://www.kaggle.com/datasets/nickdoulos/greeces-earthquakes?resource=download>

OpenStreetMap.(n.d.)

<https://www.openstreetmap.org/#map=15/37.67499/21.43437&layers=T>

Hellenic Fire Service. (n.d.). Σύνολα δεδομένων (Datasets). *Hellenic Fire Service*.

<https://www.fireservice.gr/el/synola-dedomenon>

Anastopoulos, A. (n.d.). House price prediction from listings in Greece [Kaggle Notebook]. *Kaggle*. <https://www.kaggle.com/code/argyrisanastopoulos/house-price-prediction-from-listings-greece/input>