



УНИВЕРЗИТЕТ “СВ. КИРИЛ И МЕТОДИЈ” ВО
СКОПЈЕ



ГРАДЕЖЕН ФАКУЛТЕТ СКОПЈЕ

Ристе Здравко Ристов

ВЛИЈАНИЕ НА КАРАКТЕРИСТИКИТЕ НА ПАТОТ И ОКОЛИНАТА ВРЗ
БРОЈОТ НА СООБРАЌАЈНИТЕ НЕЗГОДИ

Докторски труд

Скопје, 2026

Докторанд:
РИСТЕ ЗДРАВКО РИСТОВ

Тема:
ВЛИЈАНИЕ НА КАРАКТЕРИСТИКИТЕ НА ПАТОТ И ОКОЛИНАТА ВРЗ БРОЈОТ НА
СООБРАЌАЈНИТЕ НЕЗГОДИ

Ментор:
Проф. д-р ЗЛАТКО ЗАФИРОВСКИ,
Универзитет „Св. Кирил и Методиј“ во Скопје, Градежен факултет Скопје

Комисија за одбрана:
Проф. д-р СЛОБОДАН ОГЊЕНОВИЌ (претседател)
Универзитет „Св. Кирил и Методиј“ во Скопје, Градежен факултет Скопје

Проф. д-р ЗЛАТКО ЗАФИРОВСКИ,
Универзитет „Св. Кирил и Методиј“ во Скопје, Градежен факултет Скопје

Проф. д-р РАДОЈКА ДОНЧЕВА,
Универзитет „Св. Кирил и Методиј“ во Скопје, Градежен факултет Скопје

Проф. д-р МАРИЈА МАЛЕНКОВСКА ТОДОРОВА,
Универзитет „Св. Климент Охридски“ – Битола, Технички Факултет Битола

Вонр. проф. д-р МАРКО РЕНЧЕЉ,
Универзитет во Марибор, Словенија, Факултет за градежништво, транспортно
инженерство и архитектура Марибор

Научна област:
ТРАНСПОРТНА ИНФРАСТРУКТУРА

Датум на одбрана: март 2026

БЛАГОДАРНОСТ

Изразувам искрена благодарност до мојот ментор, проф. д-р Злајко Зафировски, за конитинуираната стручна поддршка, јасните насоки, конструктивните забелешки и довербата во мене на целиот процес на изработка на оваа докторска дисертација. Неговото искуство, посветеност и критички пристап имаа суштинско значење за мојот и унапредувањето на истражувањето.

Посебна благодарност укажувам до претседателот на комисијата за одбрана, проф. д-р Слободан Огњеновиќ, за вниманието, деталните согледувања и стручните сугестии, кои значително придонесоа за подобрување на квалитетот и јасноста на трудот.

Благодарност укажувам и до членовите на комисијата, проф. д-р Радојка Дончева, проф. д-р Марија Маленковска Тодорова и вонр. проф. д-р Марко Ренчел, за нивниот професионален ангажман, конструктивните забелешки и вредните насоки дадени во мене на оценувањето на трудот.

Благодарност изразувам и до колегите, соработниците и пријателите кои на различни начини придонесоа во мене на истражувањето, преку размена на идеи, стручни дискусии и несебична поддршка.

На крај, укажувам најдлабока благодарност до моето семејство, пред сè до сопругата Гордана и децата Марика, Илина и Дамјан, за неизмерната поддршка, приемието и разбирањето во мене на целиот период на изработка на докторската дисертација.

Изјава за оригиналност

Изјавувам дека докторскиот труд е оригинален труд што го имам изработено самостојно.

Своерачен потпис на докторандот,

м-р Ристе Ристов дипл.град.инж.

Изјава за идентичност на верзии

Изјавувам дека електронската верзија на докторскиот труд е идентична со отпечатениот докторски труд.

Потпис на авторот,

м-р Ристе Ристов дипл.град.инж.

АПСТРАКТ

Патната безбедност останува критично прашање на јавниот интерес и инженерството, со значајни здравствени, социоекономски и етички импликации. И покрај континуираните напори на меѓународно и национално ниво, бројот на сообраќајните незгоди и натаму е висок. На светско ниво, се губат повеќе од еден милион животи годишно. Во Европа бројот на загинали изнесува околу 45.5 на милион жители, односно 4.6 жртви на 1 000 km патишта. Статистиката во Македонија е значително поизразена во однос на европскиот просек, со 69.5 на милион жители односно 8.8 жртви на 1 000 km. Оваа разлика ја потврдува потребата од систематски, мерлив и практично применлив инженерски пристап. Оттука, следи насочена анализа на достапните податоци од нашата патна мрежа. Трудот ја анализира врската меѓу карактеристиките на патот, околината и бројот на сообраќајни незгоди на главната државна мрежа, при што рамката се темели на реални инфраструктурни, сообраќајни и климатски податоци од јавни институции и терен. Целта е да се постави едноставен и повторлив процес за процена на безбедноста на делниците, кој ги поврзува институционалните бази и теренските анализи, при што овозможува споредливост преку нормализиран тежински индекс на незгоди (W_i). На оваа основа, се применуваат три методи: интегративна формула, модели на машинско учење и Баесови мрежи со три структури. За оперативна употреба се поставуваат граници на W_i (0-5, 5-10, 10-30) и се обезбедува валидација без протекување на информации меѓу поставките. Резултатите покажуваат дека интегративната формула со седум параметри постигнува коефициент на детерминација од $R^2=0.72$ и овозможува континуирано рангирање на делниците по безбедност со прегледен увид во ефектите. Во машинското учење, Extra Trees со девет параметри има $R^2=0.63$ и се издвојува како најстабилен избор, со мала променливост во тестирања со преклоп. Кај Баесовите мрежи, носечки структури се GTT со точност 0.83 и PC со 0.85, со конкретни граници и точна поделба во три категории. Во овој распоред, интегративната формула нуди најдобра читливост и едноставна примена, Extra Trees обезбедува стабилно предвидување, а Баесовите мрежи даваат прегледно категоризирање и објаснување. Од комплетната анализа, може да се увиди дека сообраќајните параметри имаат најголемо влијание, при што ПГДС и брзината се најизразени. Патната инфраструктура и геометријата се втори по влијателност, како конзистентна група која ја опишува состојбата на мрежата и карактеристиките на трасата. Факторите на околината во целина имаат послаб ефект, но стануваат значајни во специфични модели и услови. Анализата со Баесови мрежи го потврдува овој распоред преку директни јазли на W_i , каде најчесто се појавуваат сообраќајните и патните параметри. Практичната примена се гледа во тоа што, трите пристапи даваат слични рангирања и стабилно ги разграничуваат делниците во три категории според оперативните граници. Стабилноста на резултатите потврдена е преку два пристапи: со наизменично изоставување по еден метод и со тестирање на дополнителни делници кои не се дел од главната анализа. Просторните прикази во ГИС јасно ги истакнуваат зоните со повисок ризик и служат за планирање на увид и инженерски одлуки, од брзи мерки како обновување на обележување и сигнализација и подобрување на одводнување, до поголеми интервенции таму каде профилите на ризик го оправдуваат вложувањето. Во целина, постапката е јасна и повторлива, дава споредливи резултати и може да се пренесе на други категории патишта при соодветни податоци.

Клучни зборови: безбедност, сообраќај, делници, W_i , ПГДС, коловоз, прегледност, сигнализација, ExtraTrees, ГИС.

ABSTRACT

Road safety remains a critical issue of public interest and in engineering, with significant health, socioeconomic, and ethical implications. Despite ongoing efforts at the international and national levels, the number of traffic crashes remains high. Globally, more than one million lives are lost each year, which confirms the need for a systematic, measurable, and practically applicable engineering approach. In Europe, the average is about 45.5 fatalities per million inhabitants, i.e., 4.6 deaths per 1,000 km of road network. In Macedonia, the rate is 69.5 per million inhabitants, i.e., 8.8 deaths per 1,000 km, which indicates a higher risk than the European average. This difference motivates a focused analysis of the available data from the national road network. The study analyzes the relationship between road and environmental characteristics and the number of traffic crashes on the main state road network, with the framework based on real infrastructure, traffic, and climate data from public institutions and field observations. The aim is to establish a simple and repeatable process for assessing the safety of road sections that links institutional databases with field analyses and enables comparability through a normalized weighted crash index (W_i). On this basis, three methods are applied: an integrative formula, machine learning models, and Bayesian networks with three structures. For operational use, thresholds are set for W_i (0–5, 5–10, 10–30), and validation is ensured without information leakage across configurations. The results show that the integrative formula with seven parameters achieves a coefficient of determination of $R^2 = 0.72$ and enables continuous ranking of sections by safety, with a clear view of parameter effects. In the machine learning framework, Extra Trees with nine parameters has $R^2 = 0.63$ and stands out as the most stable option, with low variability in cross-validation tests. For the Bayesian networks, the leading structures are GTT with an accuracy of 0.83 and PC with an accuracy of 0.85, with clear thresholds and accurate classification into three categories. In this arrangement, the integrative formula offers the best readability and ease of application, Extra Trees provides stable prediction, and the Bayesian networks provide clear categorization and interpretable explanation. Overall analysis indicates that traffic parameters have the greatest influence, with AADT (PGDS) and speed being the most prominent. Road infrastructure and geometry rank second in influence as a consistent group describing the condition of the network and route characteristics. Environmental factors generally have a weaker effect but become significant in specific models and conditions. The Bayesian network analysis confirms this ordering through the direct parent nodes of W_i , where traffic and road parameters most frequently appear. In practice, the three approaches yield similar rankings and consistently separate sections into three categories according to the operational thresholds. The stability of the results is confirmed through two approaches: a leave-one-method-out procedure and evaluation on additional sections that were not part of the main analysis. GIS spatial displays clearly highlight higher-risk zones and support inspection planning and engineering decisions, from quick measures such as renewing markings and signage and improving drainage to larger interventions where risk profiles justify the investment. Overall, the procedure is clear and repeatable, produces comparable results, and can be transferred to other road categories given the availability of appropriate data.

Keywords: safety, traffic, sections, W_i , PGDS, pavement, visibility, signage, ExtraTrees, GIS

СОДРЖИНА

Апстракт	IV
Abstract	V
Содржина.....	VI
Список на слики	IX
Список на табели.....	XI
Листа на симболи.....	XII
1. Вовед.....	1
1.1 Дефинирање на проблем.....	2
1.2 Предмет и опфат на истражувањето	3
1.3 Цели на истражувањето.....	4
1.4 Причини за избор на темата.....	5
1.5 Очекуван научен придонес.....	6
1.6 Практична вредност на истражувањето.....	6
1.7 Содржина на докторската дисертација	7
2. Теоретска рамка и клучни концепти	9
2.1 Принципи на современата безбедност на патиштата	9
2.2 Епидемиолошки пристап и Хадонова матрица	11
2.3 Клучни поими за безбедност на патиштата	12
2.4 Индикатори и мерење на безбедноста на патиштата.....	13
2.5 Патна инфраструктура и безбедност.....	14
2.6 Управување со брзини	15
2.7 Човечки фактор во сообраќајот	17
2.8 Возила и безбедносни технологии.....	17
2.9 Околина и услови на движење	18
2.10 Квалитет на податоци и споредливост.....	19
2.11 Концепциска шема на системот на патна безбедност	19
2.12 Кратка синтеза на теоретската рамка	20
3. Преглед на достигнувањата во научната област	21
3.1 Карактеристики на патот и прегледност	21
3.2 Карактеристики на сообраќајот.....	23
3.3 Околина и оперативни услови	25
3.4 Просторна категоризација на ризик.....	27
3.5 Влијателни параметри и објаснивост	28
3.6 Аналитички методи	30
3.7 Метрики и валидација	33
3.8 Синтеза и импликации	35
4. Образложение на работните хипотези и тези.....	37
4.1 Главна теза	37
4.2 Главни хипотези.....	37
4.3 Работни хипотези.....	38
4.4 Критериуми за проверка и мерни показатели	38
5. Опис на податоците и истражувачкиот опфат	40
5.1 Општи податоци за патната мрежа	40
5.2 Технички карактеристики на јавните патишта	42
5.2.1 Автопати.....	42
5.2.2 Експресни патишта.....	44

5.2.3	Магистрални патишта со две сообраќајни ленти	45
5.3	Начин на собирање на податоци.....	45
5.4	Категоризација на податоци	46
5.4.1	Категоризација според динамиката на податоците.....	47
5.4.2	Категоризација според основните карактеристики на податоците	47
5.4.3	Опис на податоците и нивна обработка.....	47
5.5	Подготовка и почетна анализа на податоците	62
5.5.1	Распределби по параметар	62
5.5.2	Интервал и атипични вредности	64
5.5.3	Атипични податоци според IQR	64
5.5.4	Зависности меѓу параметрите (Спирманова корелација).....	65
5.5.5	Мултиколинearност (VIF)	66
5.5.6	Клучни резултати и насоки	67
6.	Методологија.....	68
6.1	Општ аналитички пристап	68
6.2	Кластер анализа	68
6.2.1	Податоци и променливи	69
6.2.2	Пристапи и избор на број на кластери	69
6.2.3	Стабилност, надворешна проверка и изводи	70
6.2.4	Кластеризирање на параметри (feature clustering)	70
6.3	Иницијална проверка на техники	71
6.3.1	Дрво на одлука (Decision Tree)	72
6.3.2	Наизменично генерирани дрва на одлуки (Random Forest).....	72
6.3.3	Дополнителни дрва на одлуки (Extra Trees)	73
6.3.4	Комбинирани подмножества (Bagging Regressor)	74
6.3.5	Адаптивно засилување (AdaBoost)	75
6.3.6	Итеративно засилување (Gradient Boosting).....	75
6.3.7	Екстремно засилување (XGBoost).....	76
6.3.8	Лесно адаптивно засилување (LightGBM).....	78
6.3.9	Категориско засилување (CatBoost).....	78
6.3.10	Итеративно засилување со хистограмски прагови (HGBoosting).....	79
6.3.11	Објаслива машина за засилување (Explainable Boosting Machine, EBM)	80
6.3.12	Потпорна векторска регресија (Support Vector Regression, SVR)	81
6.3.13	Повеќеслоен перцептрон (Multilayer Perceptron, MLP).....	82
6.4	Дефинирање на влијателни фактори	83
6.4.1	SHAP анализа	84
6.4.2	Пермутацииска анализа	85
6.4.3	Пресметка на средна вредност	86
6.5	Интегративна формула – конструкција и калибрација	87
6.5.1	Полиномна регресија	88
6.5.2	Lasso регресија (L1-регуларизација).....	89
6.5.3	Ridge регресија (L2-регуларизација)	90
6.5.4	Тестирање и валидација.....	91
6.6	Машинско учење – конфигурација на моделите и избор.....	92
6.6.1	Хиперпараметри и тренинг-поставки	92
6.6.2	Тестирање и валидација.....	93
6.7	Општи карактеристики на анализата со Баесови мрежи.....	94

6.7.1	Дискретизација	95
6.7.2	Greedy Thick Thinning (GTT)	96
6.7.3	PC (Peter-Clark).....	97
6.7.4	TAN (Tree-Augmented Naive Bayes).....	99
6.7.5	Тестирање и валидација.....	99
7.	Резултати и валидација	101
7.1	Кластер анализа	101
7.1.1	Поставки и параметри.....	101
7.1.2	Избор на број на кластери.....	101
7.1.3	Резултати по метод и проверка со W_i	103
7.1.4	Профили на кластери	105
7.1.5	Кластеризирање на параметри (feature clustering)	106
7.1.6	Работен избор за понатамошни анализи	107
7.2	Иницијална проверка на техники	108
7.3	Влијателни параметри	108
7.3.1	Поединечни резултати по модел	109
7.3.2	Средна вредност на влијанијата	110
7.3.3	Чувствителност на обединувањето на влијанијата	111
7.4	Интегративна формула.....	112
7.4.1	Евалуација на модели	113
7.4.2	Избран модел и предиктивна формула за W_i	114
7.4.3	Валидација на предиктивна формула	115
7.5	Избор на оптимален модел од техники на машинско учење	117
7.5.1	Резултатна анализа и избор на модел според тест перформанси.....	117
7.5.2	Избор на стабилен носечки модел	119
7.5.3	Тестирање и валидација.....	121
7.6	Баесови мрежи	123
7.6.1	GTT (Greedy Thick Thinning)	123
7.6.2	PC (Peter-Clark).....	124
7.6.3	TAN (Tree-Augmented Naive Bayes).....	124
7.6.4	Споредбени дијаграми и калибрација.....	125
7.6.5	Избор на водечки модел.....	127
7.7	Категоризација на безбедноста на делниците.....	127
7.8	Надворешна валидација	128
8.	Начин на имплементација.....	130
9.	Дискусија	135
9.1	Синтеза на главните резултати	135
9.2	Толкување на механизмите	137
9.3	Усогласеност со светската литература	138
9.4	Стабилност на резултатите и ограничувања	139
10.	Научен придонес и примена на резултатите од истражувањето.....	140
10.1	Научен придонес	140
10.2	Практична примена	141
11.	Идни истражувања.....	143
12.	Заклучоци	146
13.	Список на литература	149
14.	Прилози	157

14.1	Опис на параметри.....	157
14.2	Збирна табела на сите податоци.....	158
14.3	Дескриптивна статистика на сите податоци	166
14.4	Дескриптивна статистика на тренинг податоци (80%).....	166
14.5	Дескриптивна статистика на тест податоци (20%).....	167
14.6	Кластер профили по податоци (GMM m=6)	168
14.7	Кластер профили по параметри (m=8).....	169
14.8	Целосна интегративна формула (OLS k=7).....	170
14.9	Оптимална детерминација	171
14.10	Квантилна дискредитација на сите податоци	172

СПИСОК НА СЛИКИ

Слика 2.1	Експоненцијална врска меѓу брзина и стапка на незгоди (Ji, 2020)	10
Слика 2.2	Систем возач, возило, пат и околина (Postransky, 2020).....	11
Слика 2.3	Концептуална шема на системот на патна безбедност (ISO 39001:2012).....	20
Слика 3.1	Зависност меѓу радиус на кривина и зачестеност на незгоди (Othman, 2009)...	22
Слика 3.2	Зависност меѓу сообраќаен обем и зачестеност на незгоди на крстосници (Retallack, 2020).....	24
Слика 3.3	Релативен ризик по температура и климатска зона (Malin, 2017)	26
Слика 3.4	Мапирање на јавнобезбедносен ризик по МВР-сектори (Babanoski, 2016).....	27
Слика 3.5	SHAR-рангирање на влијателни карактеристики во MLP-модел (Yan, 2023)	29
Слика 5.1	Прегледна карта на државни патишта (ЈПДП, 2024)	40
Слика 5.2	Стандарден профил на автопат.....	42
Слика 5.3	Стандарден профил на автопат со дополнителна возна лента.....	43
Слика 5.4	Стандарден профил на автопат без дополнителна лента за застанување	43
Слика 5.5	Профил на автопат без дополнителни ленти за застанување.....	43
Слика 5.6	Стандарден профил во една насока и магистрален профил во друга насока ...	44
Слика 5.7	Профил со стеснета разделна лента и лента за застанување.....	44
Слика 5.8	Редуциран автопатски профил	44
Слика 5.9	Експресен пат	45
Слика 5.10	Пат со две сообраќајни ленти (лево V=80-100km/h, десно V=50-80km/h).....	45
Слика 5.11	Геометриска интерпретација на кривинска карактеристика	48
Слика 5.12	Наклон со спротивна насоченост на отклонувањето	49
Слика 5.13	Карактеристични случаи на проверка на прегледност на застанување	51
Слика 5.14	Должина на застанување за патишта од техничка група А (МТВ, 2009).....	52
Слика 5.15	Мултифункционално возило Dynatest (Dynatest International, 2009)	53
Слика 5.16	Опаѓање на РСІ во зависност од староста на коловозот (C&CPIC 2021)	55
Слика 5.17	Просечни годишни врнежи од дожд за период 2014-2023 (mm).....	58
Слика 5.18	Просечни годишни врнежи од снег за период 2014-2023 (mm)	59
Слика 5.19	Средни годишни температури за период 2014-2023 (°C)	59
Слика 5.20	Максимални годишни температури за период 2014-2023 (°C).....	60
Слика 5.21	Минимални годишни температури за период 2014-2023 (°C).....	61
Слика 5.22	Хистограми на параметрите	63
Слика 5.23	Опсег и атипични вредности	64
Слика 5.24	Атипични вредности по параметар според IQR.....	65

Слика 5.25 Спирманова матрица на корелации (ρ) за сите параметри.....	66
Слика 5.26 VIF по параметар со референтна линија на VIF = 10	67
Слика 6.1 Методолошка шема на кластер анализа	68
Слика 6.2 Методолошка шема за кластеризирање на параметри.....	71
Слика 6.3 Дрво на одлука (Decision Tree).....	72
Слика 6.4 Наизменично генерирани дрва на одлуки (Random Forest)	73
Слика 6.5 Дополнителни дрва на одлуки (ExtraTrees).....	74
Слика 6.6 Комбинирани подмножества (Bagging Regressor).....	74
Слика 6.7 Адаптивно засилување (AdaBoost).....	75
Слика 6.8 Итеративно засилување (Gradient Boosting).....	76
Слика 6.9 Екстремно засилување (XGBoost)	77
Слика 6.10 Лесно адаптивно засилување (LightGBM).....	78
Слика 6.11 Категориско засилување (CatBoost)	79
Слика 6.12 Итеративно засилување со хистограмски прагови (HistGradientBoosting).....	80
Слика 6.13 Објаслива машина за засилување (EBM).....	81
Слика 6.14 Потпорна векторска регресија (SVR)	82
Слика 6.15 Повеќеслоен перцептрон (MLP)	83
Слика 6.16 Шематски приказ на SHAP алгоритмот.....	84
Слика 6.17 Шематски приказ на пермутациска анализа (Permutation Importance)	85
Слика 6.18 Полиномна регресија ($\text{deg} = 2$) (Hastie, 2009).....	88
Слика 6.19 Lasso со полиномска основа ($\text{deg} = 2$) (Kitali, 2019).....	89
Слика 6.20 Ridge со полиномска основа ($\text{deg} = 2$) (Garefalakis, 2022)	90
Слика 6.21 Основен пример за Баесова мрежа (Carrodano, 2024).....	94
Слика 6.22 Нормализиран боксплот со средни вредности	95
Слика 6.23 Распределба на записите по три состојби по дискретизација	96
Слика 6.24 Структура на Баесова мрежа добиена со алгоритмот GTT	97
Слика 6.25 Структура на Баесова мрежа добиена со алгоритмот PC.....	98
Слика 6.26 Структура на Баесова мрежа добиена со алгоритмот TAN.....	99
Слика 7.1 ВИС по k и коваријансни форми (GMM)	102
Слика 7.2 Просечна силуета по k (Ward).....	103
Слика 7.3 PCA проекција на делниците – GMM ($k = 6$).....	104
Слика 7.4 W_i по кластери – GMM ($k = 6$).....	104
Слика 7.5 Дендрограм на параметри	107
Слика 7.6 R^2 (%) и RMSE по модел на 80/20 поделба.....	108
Слика 7.7 Влијание на параметри по модел (80/20).....	109
Слика 7.8 Усреднет ранг на параметрите според усреднетото влијание	111
Слика 7.9 Чувствителност на усреднети важности по прагови	112
Слика 7.10 Спирманова корелација на рангирањата меѓу прагови на R^2_{Ok}	112
Слика 7.11 Test R^2 во зависност од бројот на параметри k (OLS, Ridge, Lasso).....	114
Слика 7.12 Кумулативна крива на добивки во зависност од покриеноста на сегменти....	115
Слика 7.13 Крива на зголемување во зависност од покриеноста.....	116
Слика 7.14 R^2 според број на параметри	117
Слика 7.15 Насочено прецизирање.....	118
Слика 7.16 Точност и стабилност по модел (5-fold)	120
Слика 7.17 Точност и стабилност по модел (5-fold).....	120
Слика 7.18 Распределба на R^2 по преклопи (5-fold)	121
Слика 7.19 Крива на добивка (Gains крива) за ET	122

Слика 7.20 Крива на зголемување (Lift крива) за ЕТ	122
Слика 7.21 ROC крива со референтна дијагонала	123
Слика 7.22 Споредба на AUC и вкупна точност	125
Слика 7.23 ROC-криви за класа $W_i = \text{Optimal}$	126
Слика 7.24 Калибрациски (reliability) криви и HL-тест	126
Слика 7.25 Карта на безбедност на делниците според оперативните граници (W_i).....	128
Слика 7.26 Реално и пресметано W_i по метод.....	129

СПИСОК НА ТАБЕЛИ

Табела 2.1 Хадонова матрица за патна безбедност (Sleet, 2011)	12
Табела 2.2 Геометрија и очекуван ефект врз реалните брзини (World Bank, 2024)	15
Табела 6.1 Техники и клучни хиперпараметри за SHAP-анализата ($R^2 \geq 0.45$)	85
Табела 6.2 Техники и клучни хиперпараметри за пермутациска анализа.....	86
Табела 6.3 Хиперпараметри и тренинг-поставки по модел	92
Табела 7.1 Клучни поставки за кластер-анализа	101
Табела 7.2 Големини на кластери – GMM ($k = 6$).....	103
Табела 7.3 Профили на кластери – медијани на избрани параметри (GMM, $k = 6$).....	105
Табела 7.4 Работен избор и употреба.....	107
Табела 7.5 Тежини по модел според точност на 80/20 поделба	110
Табела 7.6 Прецизност, покриеност и лифт на 10, 20 и 30 проценти покриеност	116
Табела 7.7 Максимално R^2 (базно) и m^* по техника	118
Табела 7.8 Корективен преглед.....	119
Табела 7.9 Матрица на забуни – GTT	124
Табела 7.10 Матрица на забуни – PC.....	124
Табела 7.11 Матрица на забуни – TAN.....	125
Табела 7.12 Категоризација во три оперативни граници	127
Табела 7.13 Надворешна валидација за три регионални делници.....	128

ЛИСТА НА СИМБОЛИ

Симбол	Единица	Значење
Wi	поени	тежински индекс на сообраќајни незгоди
PGDS	возила/ден	просечен годишен дневен сообраќај
LIMIT	km/h	ограничена/референтна брзина
V_sign	години	старост на вертикална сигнализација
H_sign	1/година	фреквенција на обновување на хоризонтална сигнализација
Kr.Kar.	°/km	кривинска карактеристика/коэффициент на закривеност
Ave_Rad	m	среден радиус на хоризонтални кривини
Ave_Inc	%	просечен надолжен наклон
KON.NAK.	1/km	индекс на контранаклони
Lat_Force	m/s ²	бочен потисок
SSD_H	1/km	индикатор за непрегледност – хоризонтала
SSD_V	1/km	индикатор за непрегледност – вертикала
DSD	m	растојание за одлучување
PCI	поени	индекс на состојба на коловоз
IRI	m/km	индекс на надолжна нерамност
RDB_RUT	mm	длабочина на колотрази
SKIDRES	–	коэффициент на триење/склоност кон пролизгување
K.Int. P.T.	1/km	густина на јазли/крстосници/наплатни по km
K.Bridges	1/km	густина на мостови и вијадукти по km
Ave_Rain	mm	просечни годишни врнежи од дожд
Max_Snow	mm	максимални годишни врнежи од снег (воден еквивалент)
Ave_Temp	°C	средна годишна температура
Max_Temp	°C	максимална годишна температура
Min_Temp	°C	минимална годишна температура
Ave_Height	m	средна надморска висина на делницата
R ²	–	коэффициент на детерминација
RMSE	Ед. на Y	корен од средноквадратна грешка
MAE	Ед. на Y	просечна апсолутна грешка
AUROC	–	подрачје под ROC крива
F1	–	хармоничен просек на прецизност и опфат
УХМР	–	управа за хидрометеоролошки работи
ГИС	–	географски информациски систем
ЈПДП	–	јавно претпријатие за државни патишта

1. ВОВЕД

Безбедноста на патиштата е едно од суштинските прашања на модерните општества, затоа што влијае врз здравјето, економската продуктивност и квалитетот на живот на милиони луѓе. Иако во последната деценија се вложуваат значителни напори на глобално и европско ниво, бројот на жртви и понатаму останува неприфатливо висок. Секоја година, приближно 1.19 милиони лица го губат животот во сообраќајни незгоди а дополнителни 20–50 милиони се евидентираат како повредени, често со долготрајни последици по здравјето и работната способност. Сообраќајните незгоди се и водечка причина за смрт кај децата и младите од 5 до 29 години [1]. Овие податоци ја нагласуваат важноста да се премине од општи насоки кон систематски и докажани ефикасни пристапи со мерливи резултати на терен. Во Македонија, бројот на загинати изнесува во просек околу 135 лица годишно, што одговара на приближно 69.5 на милион жители (ЕУ≈45.5) и околу 8.8 загинати на 1 000 km патишта (ЕУ≈4.6). Изразено по 1 милион возила е околу 196.6 (ЕУ≈52.4), што е значително над европскиот просек.

Според европските индикатори, напредокот не е ист насекаде. Некои држави постојано го намалуваат бројот на жртви, кај други тоа оди побавно. Обично ризикот е десетици загинати на милион жители годишно. Позициите се менуваат бавно и зависат од јавните политики, квалитетот на патната инфраструктура и државниот систем за управување со патна безбедност (институции, закони, контрола, образование и комуникации). Во оваа насока, споредбата не служи само за поставување цели, туку и за препознавање на мерките што носат најголем ефект (како управување со брзини, подобрување на прегледноста, квалитет на коловозот и заштита на ранливи учесници) [2].

Причините за настанување и тежината на последиците се резултат на интеракција на повеќе групи фактори и нивната важност се менува по региони. Во врнежливи области, највлијателни се одводнувањето и површинската текстура на коловозот за пневматикот да има доволно прилепување. На места со снег и изразити ниски температури, клучно е навремено зимско одржување и мерки за спречување на пролизгување. Во случај на густа и честа магла, од големо значење се добра прегледност, соодветно осветлување и прегледна сигнализација. На патишта со голем сообраќаен обем, главно влијаат ПГДС (просечен годишен дневен сообраќај) и контролата на брзините. Во региони со силни ветрови, ризикот расте поради странични налети и намалена стабилност на возилата. А во средини со послаба економска моќ, поголем дел од влијанието се префрла на состојбата на коловозната површина и на недостатокот или дотраеноста на сообраќајната сигнализација. Овие примери посочуваат дека факторите треба да се разгледуваат заеднички со фокус на локалниот контекст, наместо изолирано и универзално.

Современите истражувања спојуваат два пристапи кои взаемно се надополнуваат. Првиот користи статистички модели за да утврди колку често се случуваат незгоди и колкава е нивната тежина. Вториот се ослонува на податоци за да предвиди ризик и навреме да ги рангира делниците со помош на просторна (ГИС) анализа. Предвидувањата се оценуваат со конкретни критериуми, за да се види кои фактори најмногу влијаат и кои мерки носат најголем ефект. По потреба, делниците се делат на групи, се одбираат најважните параметри и се применуваат едноставни модели за анализа.

Во оваа насока се оформува прецизна постапка што води од собирање на податоци до донесување одлука. Ризикот на делниците не се гледа како случаен настан, туку како

резултат на повеќе услови што влијаат заедно. Во анализата се земаат предвид карактеристиките на патот (фактори на кои може директно да се влијае), потоа условите во сообраќајот (фактори кои се менуваат индиректно преку управување) и на крај условите на околината, (фактори кои не можат да се променат). Наместо да се следат само општи правила или прописи, овој пристап овозможува за секоја делница посебно да се направи процена на ризикот и да се одредат приоритети врз основа на докази.

Споредено со европската пракса, податоците за нашите патишта често покажуваат повисок ризик од просекот. Дел од разликите се должат на обемот на сообраќај, а дел на квалитетот и одржувањето на патната инфраструктура и на начинот на сообраќајно уредување. Важно е да се истакне дека има реален простор за значајни подобрувања преку системски и технички мерки, без да се навлегува подлабоко во детални податочни анализи [2,3].

Анализата ги опфаќа делниците од главната патна мрежа во претежно вонградски области, каде има мал пешачки сообраќај и мерките можат брзо да дадат резултат. Податоците се обработени по единствен протокол и се поврзани со точни локации, за моделите и проверките да бидат споредливи и да може да се повторуваат.

Со овој пристап се земаат проверени меѓународни практики и се приспособуваат за локална употреба, со јасни процедури што можат да се применат на терен. Целта е да се изберат вистинските приоритети и мерки за намалување на ризикот, редовно да се проверува состојбата на патиштата и да се следи напредокот.

1.1 Дефинирање на проблем

Патната безбедност претставува сложен општествен и инженерски предизвик, при што состојбата на мрежата произлегува од повеќе меѓусебно поврзани области. За да се утврди каде и зошто ризикот е зголемен, потребен е заеднички преглед со сите показатели. Тука се нагласуваат главните пречки што покажуваат зошто е потребна заедничка, национална анализа на патната безбедност. Фокусот е на поврзани проблеми кои ја намалуваат ефикасноста на мерките и го отежнуваат донесувањето конкретни одлуки [4].

Досегашните анализи опфаќаат многу мал дел од параметрите кои влијаат врз безбедноста. Во научната и применетата пракса често се обработува само поединечен фактор, како мерка на изложеност (обем на сообраќај и брзини), човечки фактори (употреба на појаси, алкохол, замор и одвлекување на вниманието), состојба на коловоз и одводнување (триење и аквапланинг), временски и светлосни услови (дожд, снег, магла и ноќно возење), геометрија и прегледност (наклони, радиуси и крстосници), како и безбедносни карактеристики на возилата и спроведување на правилата во сообраќајот. Постојат и истражувања што комбинираат повеќе фактори, поврзуваат сообраќајно оптоварување и брзина со бројот на незгоди, ја спојуваат состојбата на коловозот со врнежи и температура и разгледуваат врски меѓу геометрија, осветлување, состав на сообраќајниот тек и заштитни елементи покрај патот. Иако овие пристапи даваат корисен увид, нерамномерната примена ја отежнува споредбата меѓу различни средини а комбинирањето без јасни правила ретко дава стабилни приоритети [5]. Потребен е заеднички пристап што ги поврзува факторите од патот, сообраќајот и околината во една анализа. Таквата анализа овозможува да се види како овие фактори меѓусебно делуваат и

влијаат еден на друг, како и кога повеќе услови заедно создаваат заеднички ризик на одредени делници.

За да се овозможи ваква анализа, потребно да се соберат податоци од повеќе институции. Во моментов, овие податоци кај нас се од затворен тип и не се достапни за поширока употреба или поврзување со други бази, што ја ограничува можноста за нивна анализа и примена во систем за следење на патната безбедност. Податоците за сообраќајните незгоди често не се доволно точни, не се геореференцирани на ист начин и постојат разлики во описите и категориите на последиците. Поради тоа прво е потребно усогласување и нормализација на податоците, а потоа да се изврши заедничка и релевантна анализа [9]. Неопходно е воспоставување институција (агенција) што ќе ги поврзува базите на податоци и ќе дефинира единствени правила за нивно усогласување. Овие податоци треба да се надополнуваат со минимални теренски снимања/мерења. Такво тело рутински би препознавало небезбедни делници, би следело како од одреден податок се стигнува до препорака за мерка и би ги проверувало ефектите во текот на времето [6].

Механизмот за управување со безбедносен ризик на патиштата не е добро поставен и ретко се користи континуирано. Институциите не прават редовни, проактивни проверки на мрежата со објективни мерки, па одлуките најчесто се носат по одреден настан и се насочуваат кон неколку познати точки (локации на патот), наместо кон делници [5]. Анализите често само го дефинираат бројот на незгоди и неколку основни податоци, без посредничка врска со причините. Потребен е заеднички преглед што ги спојува можните причинители во една целина, за да се види каде и зошто ризикот од незгоди се зголемува.

Главниот проблем е недостигот од единствен, применлив систем што навреме ги открива и категоризира небезбедните делници. Ваквиот систем треба да ги поврзува податоците за патот, сообраќајот и околината во еднакво мерливи показатели, па преку анализа да посочува кон инженерски решенија за подобрување на безбедноста на патот. Начинот треба да биде доволно едноставен за редовна употреба и доволно прецизен за избор на мерки со најголем ефект [4].

1.2 Предмет и опфат на истражувањето

Предмет на истражувањето е поврзување на карактеристиките на патот, сообраќајот и околината со појавата и тежината на сообраќајните незгоди на делници од магистралната патна мрежа во Македонија. Истражувањето ја разгледува врската меѓу одделни и комбинирани фактори како и нивниот придонес кон зголемување на ризикот од сообраќајни незгоди на ниво на делница. Со други зборови, целта е едноставно да се утврди како овие фактори поединечно и заеднички влијаат на ризикот по делници.

Опфатот на истражувањето вклучува 161 делница со вкупна должина од околу 1 300 km, вонградски патишта, претежно наменети за моторен сообраќај. Единица на набљудување е делница, континуиран дел од пат меѓу две јасно определени точки, како што се крстосници, населени места или граници на одржување. Кога податоците тоа го налагаат, делницата се дели на подделници, пократки делови со хомогени услови во поглед на геометрија, состојба на коловоз, ограничувања на брзината и слично. Покрај линискиот аспект, се разгледува и напречниот опфат (патен коридор до 100 m ширина) околу патот, кој се користи за процена на влијанијата од околината, врз основа на достапните геопросторни податоци, со цел да се уочат локации со намалена прегледност.

Групите на променливи се организирани во три области, со цел интегрирано оценување:

- **Карактеристики на патот:** геометриски карактеристики (во ситуација надолжен и напречен профил), прегледност, состојба на коловоз, густина на патни јазли, број на објекти и сообраќајна сигнализација.
- **Сообраќајни карактеристики:** ПГДС (просечен годишен дневне сообраќај) и брзини.
- **Карактеристики на околината:** временски услови со врнежи од дожд и снег, средни, максимални и минимални температури, осветленост и видливост, како и релјефни влијанија (надморска висина).

Показателите за безбедност се мерат преку бројот на незгоди и тежината на последиците. Податоците се прибираат од институции и од терен, потоа се прочистуваат, усогласуваат и се геореференцираат, за да се обезбеди прецизна анализа и меѓусебна поврзаност на делниците. Кога е потребно, резултатите се усогласуваат на еднаква основа, односно се пресметуваат по километар или според сообраќајно оптоварување, за делниците да можат реално да се споредуваат меѓу себе.

Опфатот ги исклучува урбаните делници со изразен пешачки сообраќај и специфични локални уредувања што бараат поинаков аналитички пристап. Ова јасно ги поставува границите на анализата и објаснува на кои податоци и променливи се фокусира истражувањето.

1.3 Цели на истражувањето

Поаѓајќи од утврдените потреби и опфатот, може јасно да се издвојат целите на истражувањето. Најпрво се поставува општата цел, а потоа се дефинираат специфични цели кои ја насочуваат методологијата, обработката на податоците и начинот на прикажување на резултатите.

Општа цел е, да се развие интегрирана и контролирана методологија за рано откривање и рангирање на небезбедни делници во Македонија, заснована на податоци за пат, сообраќај и околина.

Специфични цели:

- **спојување и геореференцирање на база на податоци:** креирање на единствена база на податоци од релевантни институции со усогласена терминологија и конкретни правила за собирање и нормализација, со соодветни проверки на квалитет;
- **параметри по три области:** избор и опишување на параметри од патот, сообраќајот и околината, со утврдени начини на пресметување;
- **влијанија на фактори (поединечни/комбинирани):** испитување на поединечни и комбинирани влијанија на параметрите врз бројот и тежината на незгодите на ниво на делница;
- **модели за процена на безбедносен ризик:** изработка на модели за одредување на небезбедни делници со задоволителна точност и едноставна примена во пракса;

- **категоризација и граници за приоритети:** дефинирање категории и прагови за избор на приоритетни делници (за подобрување) и утврдување мерки со најголем очекуван ефект;
- **валидација и ретроспективна проверка:** тестирање на применетиот аналитички пристап со соодветни методи на валидација и споредба на резултатите во текот на времето;
- **оперативен протокол за редовна проверка:** дефинирање чекори за спроведување на проверките, распоред и показатели за систематско следење на резултатите и примената во институционална пракса.

Остварувањето на овие цели ќе обезбеди практична и корисна основа за донесување одлуки. На овој начин резултатите ќе покажат кои делници се најризични и каде треба прво да се преземат мерки за намалување на бројот на незгоди.

1.4 Причини за избор на темата

Изборот на темата произлегува од потребата, безбедносниот ризик на патната мрежа да се следи и управува поединечно за секоја делница. Потребен е поширок пристап што ги обединува сите фактори кои влијаат на појавата на незгоди во единствена рамка, за подобро да се согледа каде и зошто ризикот расте и каде треба да се насочат мерките за најголем ефект. На овој начин одлуките се носат врз основа на конкретни показатели, а врз општи проценки.

Во пракса најчесто се прават одделни анализи за различни делови од патната мрежа, со податоци кои не се усогласени и тешко се споредуваат. Затоа е отежнато следењето на промените во текот на времето, како и одредувањето на приоритетите. Недостасува единствен пристап каде, со користење на соодветни податоци, ќе се препознаваат небезбедните делници и ќе се покажува каде треба да се преземат мерки. Потребно е податоците, мерките и нивните резултати да бидат меѓусебно поврзани, за да се утврди дали преземените активности навистина го намалуваат ризикот. Токму овој недостаток е главна причина за изборот на темата.

Постојат и реални предуслови за спроведување на ваков пристап. Во Македонија се достапни административни податоци за патната мрежа, нејзините карактеристики, сообраќајот и дел од условите на околината. Кога овие податоци ќе се поврзат, усогласат и геореференцираат, и ќе се дополнат со теренско снимање каде недостасуваат податоци, може да се изгради стабилна база на ниво на делница без потреба од долги и скапи мерења. Потоа, со ГИС-алатки, резултатите може лесно да се прикажат преку соодветни наменски карти.

Изборот на темата е поттикнат и од нејзината непосредна примена. Резултатите од анализата служат како алатка која ги групира делниците според безбедносен ризик, поставува граници за приоритет на подобрување, насочува кон конкретни мерки и овозможува редовно следење на постигнатото. Со ваков пристап процената на ризикот води до одлука која може да се спроведе на терен, потоа да се следи нејзиниот ефект и, доколку е потребно, мерката да се корегира. На овој начин се создава континуиран циклус на подобрување и се намалува веројатноста за незгоди, со повисоки стандарди на патната безбедност.

1.5 Очекуван научен придонес

Овој труд, има цел да понуди практични и применливи решенија за проценка на патната безбедност во Македонија. Во продолжение следуваат клучните придонеси:

- **интегрирана рамка за процена на безбедносен ризик по делници:** ги поврзува сите релевантни податоци за секоја делница во една целина, со усогласени термини и правила за пресметување и нормализација по километар, за да се добијат споредливи резултати и полесно да се одредат приоритетите за подобрување;
- **геореференцирана национална база на податоци на ниво на делница:** која ќе содржи поврзани податоци од различни институции, усогласени по термини и формати, со протокол за спојување, проверка и редовно ажурирање во текот на времето. На овој начин секоја делница може да се анализира и спореди низ различни години и локации, без потреба од нови мерења;
- **модел за утврдување на безбедност по делници:** во кои се вклучени методи од машинско учење за да се одреди нивото на безбедност, да се откријат клучните фактори на ризик и да се објасни нивното дејство врз резултатите;
- **категоризација и граници за дефинирање приоритети по делници:** каде се утврдуваат категории и прагови по делници, со едноставни критериуми и мерни показатели, за да се издвојат најризичните делници и да се насочат мерки со најголем очекуван ефект;
- **кластер профилирање на делници:** или групирање на слични делници според нивните карактеристики, како што се геометријата, состојбата на коловозот и сообраќајното оптоварување. На овој начин, делниците со слични услови се анализираат заедно, што овозможува полесно планирање на интервенции и прилагодување на мерки според нивниот тип;
- **валидација со тестирање на независен сет и проверки на стабилност во текот на времето:** каде се проверува точноста и стабилноста на резултатите, за да се потврди нивната сигурност и практична применливост;
- **оперативен протокол за редовна проверка и следење:** содржи конкретни чекори, распоред и показатели за проценка на состојбата и ефектите од применетите мерки, кој може директно да се примени во институциите за континуирано следење и изработка на извештаи за безбедност.

Со наведените придонеси, се очекува трудот да даде практична и корисна основа за донесување одлуки. Резултатите ќе помогнат полесно да се одредат приоритетите за подобрување на безбедноста и да се намали ризикот од сообраќајни незгоди на национално ниво.

1.6 Практична вредност на истражувањето

Овој пристап директно се применува во планирање, управување и одржување на патштгата. Резултатите се мерливи и лесно разбирливи. Ризикот по делници се рангира и се прикажува на карта во ГИС, што овозможува побрзо и подобро донесување на одлуки.

Ризикот зависи од состојбата на коловозот, прегледноста, сигнализацијата и сообраќајните услови, па затоа предност имаат мерки со најголем очекуван ефект. На овој начин се добива основа преку која, со конкретни параметри може да се планираат идните инвестиции и да се контролира, што е добиено после нив.

Користејќи ги истите правила и чекори, може да се воспостави повторлива постапка за редовна проверка и следење. Ова обезбедува споредливи резултати во текот на времето и навремено ажурирање на приоритетите за подобрување на безбедноста. Истата логика може да се примени и на други категории патишта, каде што има соодветни податоци, со цел намалување на ризикот и на остатокот од мрежата.

1.7 Содржина на докторската дисертација

Содржината на докторскиот труд се состои од следните поглавја:

1. Вовед

- Контекст и значење на темата
- Проблем и мотивација
- Цели, предмет и опфат

2. Преглед на релевантни истражувања

- Карактеристики на патот и прегледност
- Сообраќајни карактеристики
- Околина и оперативни услови
- Просторна категоризација на ризик
- Влијателни параметри и нивно дејство
- Аналитички методи
- Показатели и валидација

3. Тези и работни хипотези

- Главна теза
- Хипотези по области (пат, сообраќај и околина)

4. Податоци за анализираните делници

- Извори на податоци, период, единица на набљудување и геореференцирање
- Показатели по области и подготовка и чистење на базата
- Нормализација (изложеност) и проверка на конзистентност

5. Методологија на истражувањето

- Истражувачка анализа и преглед на интеракции
- Модели за утврдување на безбедност и критериуми за оценка
- Рангирање и граници, ГИС приказ на резултати

6. Резултати од анализите

- Описни статистики и клучни односи
- Оценка на моделите и важност на параметрите
- Ранг-листи на делници и примери

7. Дискусија

- Толкување на резултатите и споредба со литературата
- Примена во практиката
- Ограничувања и преносливост на резултатите

8. Научен придонес и примена на резултатите од истражувањето

- Новини и методолошки придобивки
- Институционална примена и следење на резултатите

9. Препораки за идни истражувања

- Проширување на податочната основа и области
- Оценка “пред и потоа“ на мерки и економска анализа
- Прилагодување на пристапот на други мрежи

10. Заклучок

- Синтеза на исполнетите цели
- Најважни резултати и значење
- Главни насоки за примена

11. Список на користена литература

12. Прилози

Оваа содржина претставува основа за насочување кон натамошниот истражувачки процес.

2. ТЕОРЕТСКА РАМКА И КЛУЧНИ КОНЦЕПТИ

Безбедноста на патиштата е карактеристика на целата патна мрежа и зависи од интеракцијата меѓу возачот, возилото, патната инфраструктура и околината [7]. Целта е да се намали ризикот и последиците да се ограничат преку повеќе усогласени мерки, наместо вината да се бара само кај поединецот. Вообичаено, полициските извештаи како причина ја наведуваат грешката на возачот, но треба да се разгледа и поширокиот контекст, односно како условите на патот, организацијата на сообраќајот и околината придонеле за таква грешка.

Епидемиолошкиот пристап ја разгледува патната безбедност во три временски моменти: пред настан, за време на настан и по настан, и ги опфаќа четирите делови од системот: возач, возило, пат и околина. Мерките се распоредуваат за да спречи незгода, да го намалат ризикот и последиците доколку се случи незгода и да обезбедат брз и координиран одзив по незгодата.

Од брзината, зависи количината на енергија што се ослободува при судир со друго возило или објект, што директно влијае врз тежината на последиците. Колку е поголема брзината, толку се посериозни повредите и материјалната штета. Поради тоа, управувањето со брзината е клучен дел од секое решение насочено кон ублажување на последиците од незгодите [8].

Следењето и управувањето со безбедноста се потпираат на две групи информации: мерливи показатели за состојбите во системот (возач, возило, пат и околина) и податоци за незгоди и последици (настан, локација, тежина на последиците и сл.). Во европската пракса се користи основен сет на мерливи показатели за брзина, корисници, возила, патна инфраструктура и одзив по незгода, што овозможува споредливост и одлуки засновани на податоци [9].

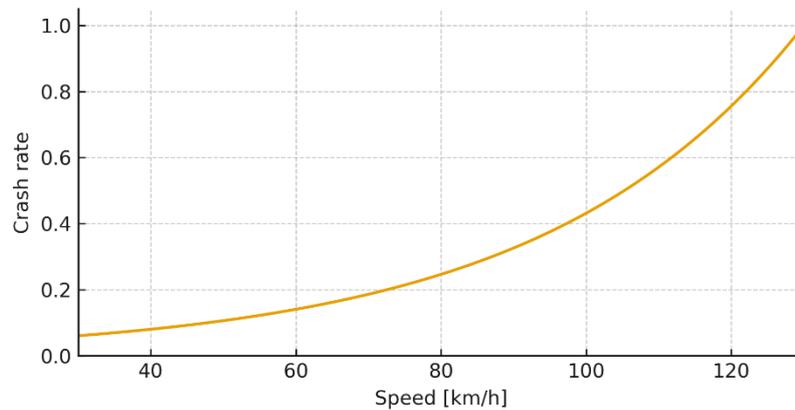
2.1 Принципи на современата безбедност на патиштата

Безбедноста претставува едно од клучните карактеристики на целиот сообраќаен систем. При нормални сообраќајни услови може да се појават неправилности и грешки. Затоа, системот се организира така што, тие грешки да не доведат до сообраќајна незгода. Доколку сепак дојде до незгода, последиците од неа да бидат што помали. [10]. Овој пристап се применува преку повеќе мерки што заедно ја зголемуваат безбедноста. Во него се вклучуваат подобрувања на патиштата, управување со брзината, користење на побезбедни возила, одговорно однесување на учесниците во сообраќајот и добра организација за брза реакција кога ќе се случи незгода.

Патиштата ја обликуваат насоката и начинот на движење на возилата, па нивните карактеристики, како прегледноста, ширината на лентите и состојбата на коловозот, имаат директно влијание врз безбедноста и ризикот од незгоди. Прегледна траса овозможува предвидливи линии на движење и го намалува непотребното менување на лента. Прегледноста обезбедува време за забележување и маневар. Ширините на сообраќајните ленти и банкините обезбедуваат простор за корекција на мала грешка, а системот за одводнување ја одведува водата од коловозот и го намалува ризикот од аквапланинг. Добро поставена и усогласена сигнализација му помага на возачот брзо и точно да одлучи, како и соодветно да ја усогласи брзината [10]. Кога овие елементи се

усогласени, патот овозможува континуиран проток и ја намалува веројатноста, случајна грешка да се претвори во сообраќајна незгода.

Врската меѓу брзината и безбедносниот ризик е меѓу најистражуваните во патната безбедност. Во пракса се забележува изразен тренд на зголемување на незгодите со раст на брзината [11]. На слика 2.1 е прикажана стапката на незгоди (број на незгоди по километар во даден период), која се зголемува експоненцијално со зголемување на брзината.



Слика 2.1 Експоненцијална врска меѓу брзина и стапка на незгоди (Ji, 2020)

Според дијаграмот, ограничувањата на брзината и мерките за нивно почитување треба да се прилагодат на геометријата, сообраќајниот режим и локалните услови. Со ова усогласување се намалува веројатноста за настанување незгода, се ублажуваат последиците, а мерките полесно се планираат и применуваат во пракса.

Возилата придонесуваат кон повисока сообраќајна безбедност на два начина:

- **активните системи**, како сопирачки, стабилизација и асистенции за возачот, го намалуваат ризикот за судир со подобра контрола при движење;
- **пасивните системи**, како безбедносни појаси, воздушни перничувања и конструкција на возилото, ги ублажуваат повредите ако настане судир.

Со дизајнирање на возила со понови технологии, се зголемува ефективноста на мерките за патна и сообраќајна безбедност и се намалува зачестеноста на незгодите [12].

Човечкиот фактор останува важен, бидејќи вниманието, заморот, перцепцијата и донесувањето одлуки влијаат врз ризикот. Обуката и едукацијата на возачите помагаат, но системот не смее да се потпира на совршено однесување. Потребен е пристап со толеранција на мала грешка и континуиран надзор што го држи ризикот од незгоди во прифатливи граници, со правила и спроведување што лесно се препознаваат на терен. Дополнително, користење безбедносен појас во автомобил и заштитна кацига на мотоцикл или велосипед, воздржување од консумирање алкохол и избегнување на одвлекување на вниманието, го намалуваат ризикот на ниво на возач [12].

Проценката на безбедноста се врши систематски. Се користат показатели што опишуваат ризични состојби, како усогласеност со ограничувањата на брзината, употреба на безбедносна опрема и податоци за незгоди и последици. Во европската пракса постои основен сет на показатели за брзина, безбедносна опрема, консумирање алкохол, одвлекување на вниманието, безбедност на возилата, безбедност на патната инфраструктура и реакција по незгода, со што се овозможува споредливост и преимена на стратегии засновани на податоци [12]. За правилна процена на безбедносниот ризик, треба

да се земе предвид и сообраќајното оптоварување, односно да се пресметува стапка на незгоди според обемот на сообраќајот и неговата структура, за пореално да се утврди дали и како ризикот се менува [11].

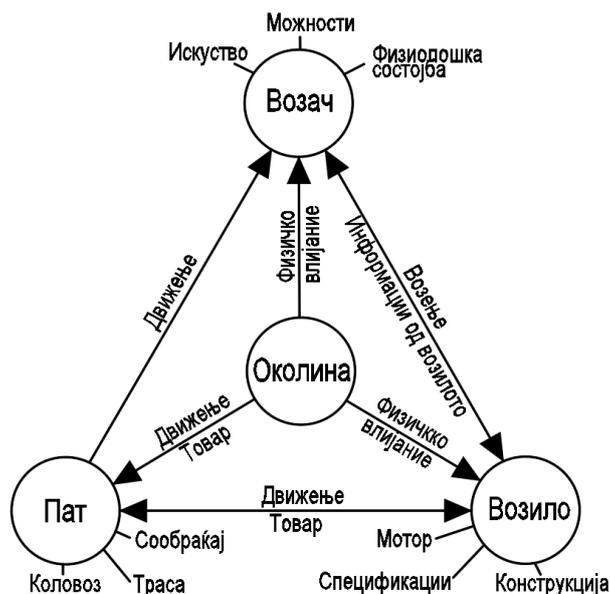
2.2 Епидемиолошки пристап и Хадонова матрица

Епидемиолошкиот пристап ја објаснува безбедноста како резултат на заедничко дејство на возачот, возилото, патот и околината низ фазите пред, за време и по настанот. Овој начин на разгледување помага мерките да се насочат таму каде што најмногу го намалуваат ризикот и ги ублажуваат последиците [13].

Возач, возило, пат и околина. Во реални услови, возачот може да направи грешка, техничките системи во возилата имаат свои ограничувања, а патот и околината ги дефинираат условите за движење. Кога сите овие елементи се усогласени со јасни правила, обука на возачи, современи системи во возилата и уреден пат со прегледна траса и правилна сигнализација, се намалува можноста една мала грешка да доведе до сообраќајна незгода. Ако сепак се случи незгода, последиците полесно се контролираат и се помали [14].

Фази на делување. Целта е мерките да делуваат во сите три фази на настанот. Пред незгодата се преземаат мерки што ја намалуваат веројатноста да се случи. За време на незгодата се намалува силата на ударот и се штитат учесниците. По незгодата се реагира брзо и координирано, за повредите да бидат помали и опоравувањето побрзо [15].

На слика 2.2 прикажан е односот меѓу возачот, возилото, патот и околината, со стрелки што го означуваат правецот на влијанијата. Горниот дел го поврзува возачот со неговите способности, искуство и моментална состојба. Долниот дел го прикажува односот меѓу патот и возилото преку начинот на движење и нивната техничка состојба, додека околината влијае врз сите три елементи. Овој приказ укажува дека мерките ретко делуваат изолирано и дека подобрувањето во еден дел има поголем ефект, кога е системски поддржано од останатите елементи на системот [16].



Слика 2.2 Систем возач, возило, пат и околина (Postransky, 2020)

Дијаграмот покажува усогласени правила за движење и сигнализација, прегледна и уредена траса, соодветни технички системи во возилото и контрола на влијанијата од околината кои заедно го намалуваат ризикот од незгода и го стабилизираат целиот систем возач-возило-пат-околина.

Во продолжение, табела 2.1 го прикажува распоредот на мерки според Хадоновата матрица со кратки и општи формулации што се користат за планирање на мерки за безбедност.

Табела 2.1 Хадонова матрица за патишна безбедност (Sleet, 2011)

Фаза \ Дел	Човек	Возило	Пат и околина
Пред настан	Обука, информирање, контрола	Асистенции за возач, системи за предупредување	Прегледна траса, јасни ограничувања, јасна сигнализација
Настан	Користење на задолжителна опрема	Автоматско сопирање, стабилизација	Уредна коловозна површина, заштитни огради, безбедни завршетоци
По настан	Брзо пријавување и насочување на лице место	Автоматско повикување помош	Пристапност за итни служби, јасни локациски ознаки

Табелата прикажува како мерките за безбедност се распределуваат низ сите составни делови на системот и три фази. На овој начин се посочува кој е одговорен за секоја мерка и каков резултат се очекува, што овозможува полесна поделба на задачите, поставување на рокови и следење на напредокот.

Со ваков распоред, однесувањето на возачите, техничките својства на возилата и карактеристиките на патот и околината се поврзуваат со конкретни интервенции и мерење на ефектите, што создава координиран начин за управување со ризикот на патиштата.

2.3 Клучни поими за безбедност на патиштата

Безбедноста на патиштата се објаснува преку неколку поими кои овозможуваат лесно разбирање и споредливост. Овој дел ги сумира основните дефиниции и ги поврзува со практичната примена во мерење и управување:

- **фактори на ризик и елементи на системот:** нивото на ризик зависи од односот меѓу човекот, возилото, патот и околината, каде секоја промена во еден дел може да влијае на безбедноста на целиот систем;
- **изложеност, ризик и стапки:** изложеноста го претставува сообраќајното оптоварување или извозените километри на трасата, што се анализира. Ризикот ја опишува веројатноста за настанување на незгода при дадено ниво на изложеност. Стапката ги поврзува незгодите со изложеноста, најчесто изразена како број на незгоди на милион извозени километри. Со користење на стапки, може да се утврди колку реално се менува самиот ризик со текот на времето [17];
- **нивоа на тежина на последици:** последиците најчесто се групираат во категории по тежина на исходи (без повреда, лесна повреда, тешка повреда и смртен случај). Оваа поделба овозможува следење на промените со текот на времето и избор на мерки според сериозноста на последиците. Кога е можно, класификацијата се усогласува

со здравствените дефиниции за степенот на повреди, со цел да се обезбеди поголема споредливост меѓу различни извори на податоци;

- **показатели за перформанси и број на незгоди и последици:** показателите ги опишуваат ризичните состојби во системот на патна безбедност и служат за процена на слабостите и потребата од превентивни мерки. Податоците за незгоди и последици ги опишуваат самите настани, повредите и смртните случаи. Двете групи се користат заедно, при што показателите укажуваат каде треба да се насочи делувањето, а податоците покажуваат каков ефект имаат преземените мерки [18];
- **управување со брзини:** од брзината зависи енергијата при незгода, а влијае и на зголемување на веројатноста од судир и на тежината на последиците. Затоа ограничувањата на брзината треба да се усогласат со карактеристиките на трасата и околината. Соодветната сигнализација и контролата на сообраќајот придонесуваат реалните брзини да одговараат на условите на патот и неговата околина. Така мерките се насочуваат таму каде што најмногу го намалуваат ризикот [19];
- **активна и пасивна безбедност на возилата:** активната безбедност опфаќа системи што помагаат за време на движењето (сопирање, стабилизација, системи за асистенција на возачот) и ја намалуваат веројатноста за судир. Пасивната безбедност ги намалува повредите при судир (појаси, воздушни перничкиња, конструкциска заштита). Понови технологии кај возилата придонесуваат за намалување на бројот и сериозноста на незгодите;
- **мерење и толкување:** мерењето се планира однапред со прецизно дефинирање кое се мери, каде, во кој период и со кои средства и методи. Вклучувањето на податоците за изложеност овозможува следење на промените со текот на времето. Кога терминологијата за поимите што се мерат и методите за прибирање и обработка на податоците се усогласени, резултатите се споредливи во време и простор и овозможуваат донесување одлуки засновани на податоци [18,19].

Овие поими обезбедуваат заеднички начин за опишување на состојбите и резултатите. Со нив мерките на терен директно се поврзуваат со индикаторите што се следат, за реализирање на целите за помалку незгоди и помалку повреди.

2.4 Индикатори и мерење на безбедноста на патиштата

Мерењето го претвора поимот за безбедноста на патиштата во збир од разбирливи и споредливи податоци. Се користат две главни групи информации. Првата група се податоци (индикатори) што опишуваат ризични состојби во системот. Втората група се податоци за незгоди и последици. Кога двете групи се анализираат заедно, може појасно да се утврди каде треба да се делува и каков ефект имаат преземените мерки [20].

Вообичаено, се користи основен сет индикатори што се едноставни за толкување и погодни за споредба меѓу различни места и временски периоди. Во нив спаѓаат: усогласеност со ограничувањата на брзината, употреба на безбедносен појас и заштитна кацига (кај мотоциклисти), консумирање алкохол кај возачите, користење мобилен уред при возење, безбедност на возилата, состојба на патната инфраструктура и организирана помош по незгода. Брзината особено е важна, бидејќи ја зголемува и веројатноста за појава и тежината на последиците, па усогласеноста со ограничувањата на брзината претставува

клучен и прегледен индикатор [21]. Овие индикатори се усогласени со европската пракса и претставуваат основа за креирање политики засновани на податоци [22].

Изворите на податоци се разновидни и меѓусебно се дополнуваат. Во полициските извештаи се бележат настаните и околностите на незгодите, здравствените служби го опишуваат степенот на повредите, а инженерите за управување со патишта и општинските надзори даваат податоци за состојбата на патот, сигнализацијата и физичките промени (нови приклучоци, промена на ленти или воведување кружни текови). Дополнително, се користат теренски проспекции, автоматски мерења и анкетирање. За правилно следење на состојбите се применуваат мерки на изложеност, како должина на трасата или сообраќајно оптоварување.

Квалитетното мерење бара прецизно дефинирани постапки и јасно утврдени правила. Потребно е да се знае што точно се брои, каде, во кој период и на кој начин се прибираат податоците, како и кои извори се користат. Се проверува репрезентативноста и стабилноста на податоците со текот на времето. Кога е можно, се поврзуваат полициските и болничките податоци за попрецизно утврдување на тежината на повредите кај учесниците во незгодите.

За поцелосна слика се користат стапки, тековни просеци и споредби според видот на местото и времето на денот. На овој начин се создаваат валидни показатели што помагаат при избор на соодветни мерки за подобрување.

За систематско работење, за секој индикатор се наведува: што значи и во која мерна единица се мери, каков резултат се очекува, од кој извор доаѓаат податоците, колку често се мери и кој е надлежен. Потоа се поставуваат почетна и целна вредност, со конкретни чекори како да се стигне до целта. Со ваков пристап се обезбедува континуитет во следењето и транспарентност во оценувањето на резултатите.

2.5 Патна инфраструктура и безбедност

Патната инфраструктура има големо влијание, врз тоа, дали ќе се случи сообраќајна незгода и колку ќе бидат сериозни последиците. За да се намали ризикот, трасата треба да биде таква што возачите ќе имаат добра прегледност, доволна видливост за навремено да реагираат, а сигнализацијата и ознаките да бидат јасни и усогласени. Во европските земји ова се постигнува со организиран систем на постапки, што вклучува безбедносни проверки при проектирање, редовни инспекции на патната мрежа, управување со опасни места и класификација на ризик. Овие активности помагаат, инженерите својата работа да ја насочат кон цели што може да се измерат и оценат, согласно резултатите [23].

Геометриските елементи го насочуваат движењето и ги намалуваат конфликтите. Усогласени радиуси на хоризонтални кривини и соодветно витоперење со обезбедување континуитет на трасата го намалуваат изненадувањето и потребата за брзи маневри. Ширините на сообраќајните ленти и банкните овозможуваат простор за корекција на мала грешка. Физичкото разделување на спротивни насоки го намалува фронталниот ризик (челен судир). На крстосници се применуваат решенија што ја поедноставуваат комуникацијата меѓу движењата и ја намалуваат брзината. Коловозот обезбедува доволно триење и рамномерност при движење. Системот за одводнување ја отстранува водата за да се намали ризикот од пролизгување. Страничните зони (непосредната околина на патот) се уредуваат со огради и безбедни завршетоци на објекти за да се намалат последиците при излетување.

Планирањето интервенции се базира на процена на ризикот. Со безбедносни функции и корективни фактори се предвидува промената на бројот на незгоди при промени на геометријата, крстосниците, ограничувањата на брзина или контролата на пристап. На тој начин се споредуваат сценарија и се избираат решенија со најголем очекуван ефект, потврдени со податоци [24].

Кога се планираат инвестиции за подобрување на патиштата, се користат алатки што ја поврзуваат проценката на безбедноста на конкретни делници и точки со мерки што треба да се преземат, а со тоа за колку ќе се намали ризикот. Потоа се прави листа по редослед: кои мерки даваат најголем ефект со најмалку вложени финансиски средства. Таа листа станува програма на интервенции што води до мерливо намалување на незгодите и повредите. Напредокот се следи преку индикатори и редовни инспекции [25].

2.6 Управување со брзини

Брзината влијае и на тоа колку често се случуваат незгоди и колку се сериозни нивните последици. Целта на управувањето со брзината е возачите да се движат со реални брзини што одговараат на условите на местото и на функцијата на трасата. Ова се постигнува преку добро усогласени ограничувања, уреден изглед на патот што го насочува вниманието на возачите и преку постојано следење, дали тие ги почитуваат ограничувањата [26].

Најсигурното и најграјно управување со брзините се постигнува со геометриски решенија во трасата и профилот. Геометријата ја обликува очекуваната брзина преку хоризонтални и вертикални кривини, должини на правци, надолжни наклони, витоперење и прегледност. Ширината на сообраќајните ленти и напречниот профил влијаат на изборот на брзина и на стабилноста на движењето. Кога геометриските карактеристики на патот, прегледноста и сообраќајната сигнализација се усогласени, возачите природно се приближуваат до целните брзини без потреба од постојана интервенција. [27]. Во табела 2.2 се дадени клучни геометриски елементи со очекуван ефект врз реалните брзини, со кратки напомени за безбедност.

Табела 2.2 Геометрија и очекуван ефект врз реалните брзини (World Bank, 2024)

Геометриски елемент	Што се менува	Ефект врз реални брзини	Коментар за безбедност
Хоризонтални кривини и должини на правци	Радиус на кривина и должина на правец	Помал радиус и пократок правец ја намалува брзината. Голем радиус и долг правец ја зголемува брзината	Избегнување на нагли премини меѓу одделни делови од патот со различни карактеристики
Вертикални кривини и надолжни наклони	Радиус на вертикална кривина и надолжен наклон	Поголем надолжен наклон и конвексна кривина ја намалуваат брзината. Мал надолжен наклон и конкавна кривина дозволуваат повисока брзина	Проверка на должина на кочење и прегледност
Витоперење	Големина на напречен наклон во кривина	Соодветното витоперење го стабилизира движењето и ја ограничува брзината	Усогласеност со радиусот на хоризонталната кривина и очекуваните брзини

Прегледност (вклучувајќи и преттигнување)	Должина на прегледност	Поголема прегледност поттикнува повисока брзина. Ограничена прегледност ја намалува брзината	Обезбедување растојание за запирање и безбедно преттигнување
Ширина на сообраќајни ленти	Стандардна ширина	Пошироки ленти поттикнуваат повисоки брзини. Потесни намалуваат брзина	Рамнотежа меѓу стабилноста на возилото и контролата на движењето
Напречен профил и ширина на банкени	Ширина на банкени и уреденост на рабовите на коловозот	Широки и чисти банкени поттикнуваат повисоки брзини	Комбинација со геометриски мерки за поефикасно ограничување на брзината
Физичко разделување на насоки	Разделен појас и огради	Разделувањето често ја зголемува избраната брзина	Намалување на ризик од фронтални судири
Уредување на крстосници	Тип, геометрија и влезно-излезни ленти	Кружни текови и стеснети влезови ја намалуваат брзината	Обезбедување прегледност и предвидливи траектории
Рабно уредување и обрабување	Пешачки острови, рабници паркирање во профилот на улицата	Обрабувањето и паркирањето во профилот на улицата ја намалуваат брзината	Поголема внимателност кон велосипедисти и пешаци

Табелата служи како брз преглед при оценување и избор на решенија. Со неа геометријата, прегледноста и напречниот профил се поврзуваат со очекувани промени на реалните брзини и со јасни напомени за безбедност, што помага мерките да се насочат таму каде што најмногу го намалуваат ризикот.

Ограничувањата на брзината имаат смисла само кога се прилагодени на локацијата и нејзината намена. На делници со многу приклучоци, пешачки премини и содржини покрај патот се применуваат пониски ограничувања на брзината. На делници со разделени насоки на движење, добра прегледност и уредени странични површини, ограничувањата може да бидат повисоки, но само ако вкупниот ризик реално се намалува. Тоа се постигнува со усогласени линии на движење, соодветни геометриски решенија и мерки за смирување на сообраќајот, за реалните брзини природно да се приближат до целните.

Инженерските решенија и контролите треба да бидат комплементарни за возачите постојано да ги почитуваат ограничувањата. Примена на камери, мобилни и стационарни контроли, јасна сигнализација и информации во реално време го зголемуваат почитувањето на ограничувањата на брзината. Системите во возилата, како интелегентно помошно ограничување на брзината и автоматско сопирање, дополнително го намалуваат ризикот, особено во градски услови [28].

Следењето на брзината се врши со показатели преку кои се анализира распределбата на брзините, средната вредност и распонот на конкретната локација. Се пресметува колку возила се движат во рамките на ограничувањето и колку го надминуваат. Се споредуваат мерења пред и по спроведување на мерки за безбедност на истите места и во исти временски периоди. За правилно следење на промените, со текот на времето се користат

мерки на изложеност, како извозени километри или сообраќајно оптоварување, со што може да се провери како реално се менува безбедносниот ризик.

2.7 Човечки фактор во сообраќајот

Влијанието на човечкиот фактор е најголемо врз ризикот од сообраќајни незгоди. Вниманието, оптовареноста со задачи, заморот, консумирањето алкохол и други супстанции, ја менуваат способноста за перцепција и донесување одлука. Затоа, управувањето со ризикот не се потпира на совршено однесување, туку на услови што го намалуваат ризикот и ги ограничуваат последиците [29].

Одлуките на возачот често се резултат од состојбата на патот и околината. Начинот на обликување на трасата, степенот на прегледност, јасноста на сигнализацијата и количината на информации влијаат врз тоа, како се разбира ситуацијата на лице место и како се избира правилен маневар. Проектните решенија треба да го водат возачот кон помалку грешки и кон понизок ризик, бидејќи не секоја грешка произлегува од лични физиолошки слабости, туку понекогаш од нејасни услови на патот [30].

Клучни однесувања со најголем ефект се изборот на брзина според условите, задржување на вниманието при возење, употребата на безбедносна опрема, воздржувањето од консумирање алкохол и супстанции што влијаат на возењето и навремен одмор. Особено, користењето мобилни уреди при управување со возилото го одвлекува вниманието и го забавува времето на реакција. Возачот потешко ја забележува ситуацијата во сообраќајот и побавно реагира на ненадејни промени [31].

Подобрувањето на однесувањето се постигнува преку доследна примена на сообраќајните прописи, со обука и информирање, и со поддржувачка технологија во возилото. Системите за асистенција како предупредување за напуштање лента, автоматско сопирање и препознавање знаци, помагаат да се спречат грешки и ја намалуваат веројатноста за незгода.

Следењето на однесувањето се прави со едноставни и повторливи мерења на повеќе места и во подолг период, како на пример колкав е делот од возачи што користат мобилен уред додека управуваат, колкав е делот под влијание на алкохол и колку правилно се користат појаси и заштитни кациги (кај мотоциклистите).

2.8 Возила и безбедносни технологии

Возилото влијае на ризикот преку две групи технологии: активната безбедност ја намалува веројатноста за судир, а пасивната безбедност ги ублажува последиците ако настане судир. Постепеното обновување на возниот парк со возила со понови технологии го зголемува ефектот на мерките на патиштата и води кон понизок ризик од појава на незгоди.

Активна безбедност. Електронската контрола на стабилноста (ESC) статистички значително ги намалува судирите поврзани со загуба на контрола врз возилото и излетувања од коловозот, со што опаѓа и веројатноста за тешки последици [32]. Автоматското сопирање при опасност од судир (АЕВ) ја намалува зачестеноста на судири одзади и ги ублажува последиците доколку сепак настане незгода, особено при ниски и средни брзини [33,34]. Системите за задржување во лента и одржување на безбедно

растојание ја поддржуваат стабилноста на движењето и изборот на соодветна брзина, што индиректно го намалува ризикот и го олеснува управувањето со возилото.

Пасивна безбедност. Безбедносните појаси, воздушните перничкиња и конструктивните елементи со контролирана деформација ја ограничуваат енергијата што се пренесува кон патниците и ја намалуваат сериозноста на повредите. Современите каросерии ја распределуваат силата на ударот низ повеќе носечки делови, со што се подобрува заштитата на патничкиот простор [35].

Следење на напредокот. Напредокот може да се следи преку едноставни индикатори, како што се уделот на возила опремени со ESC, АЕВ и системи за поддршка на возачот, просечната старост на возилата, и уделот на возила со високи безбедносни оценки. Кога овие индикатори растат, се очекува соодветно намалување на бројот и сериозноста на незгодите при споредливи сообраќајни услови.

2.9 Околина и услови на движење

Околината влијае на безбедноста преку временските услови, прегледноста и рељефот. Дождот, снегот, мразот и појавата на магла го намалуваат триењето меѓу пневматикот и коловозот, ја намалуваат должината на прегледност и го зголемуваат растојанието потребно за запирање. Целта е возачите да возат со брзина и стабилност што одговараат на условите на дадена локација на патот.

Дождот ја намалува адхезијата меѓу пневматикот и коловозот и ја зголемува должината за запирање. Прскањето со вода ја намалува прегледноста, а на повисоки брзини се зголемува можноста возилото да пролизга по воден слој (аквапланинг). Ова најчесто се случува кога коловозот е истрошен, има слаба површинска текстура или недоволно одводнување [36,38]. Снегот и мразот ја намалуваат контролата врз насоката и управувањето со брзината. Силниот страничен ветер ја намалува стабилноста кај високи и лесни возила. Ризикот се намалува со функционален систем за одводнување, навремено зимско одржување, соодветна површинска текстура на коловозот и со прилагодување на брзината кон моменталните услови [36,38].

Прегледноста се менува со појава на магла, врнежи (дожд и снег) и слаба природна светлина (особено во зори и самрак), што го скратува времето за забележување и го зголемува потребното растојание за реакција. При вакви услови потребни се урамнотежени геометриски решенија во трасата и профилот и разбирлива и усогласена сигнализација што ја поддржува правилната проценка на возачот [36,37].

Теренот влијае на начинот на движење и на техничките барања на трасата. На планински делници со големи надолжни наклони и кратки растојанија меѓу кривините се потребни континуирана прегледност, функционално одводнување и геометрија што овозможува возење со брзина без нагли промени. На патиштата со раздвоени коловози и безбедно уредена непосредна околина, ризикот е понизок при исти метеоролошки услови. Контролата на брзините и редовното одржување остануваат клучни за стабилно и безбедно движење [37, 39].

Следењето на влијанијата од околината се врши со јасни показатели и усогласен метод на сите локации. Се користи и мерка на изложеност, како број на возила и поминати километри, за подобро следење на промените со текот на времето. Врз основа на резултатите, мерките за безбедност се прилагодуваат навремено и проактивно, со цел да се одржи континуирана безбедност на сообраќајот.

2.10 Квалитет на податоци и споредливост

Квалитетните податоци се основа за точно толкување на трендовите и избор на мерки што даваат најголем ефект. Потребни се јасни описи, усогласени методи и стабилни постапки за прибирање и обработка на информациите. Податоците од полициските евиденции треба внимателно да се анализираат, бидејќи не сите настани и повреди се пријавени, па затоа заклучоците се потврдуваат преку проверка на повеќе извори [40].

Споредливоста е можна кога термините, класификациите и мерните единици се исти на различни места во текот на времето. Потребна е јасна и достапна документација за методите. Поголема точност се постигнува со поврзување на полициски и болнички записи или користење на податоци за сериозни повреди со правила за претворање на медицинските класификации. Со тоа се добива поцелосна слика и се намалува можноста за потценување на повредите [41].

Набљудувањата на терен и автоматските мерења се спроведуваат по однапред утврден план, при што јасно се утврдува кои податоци се бројат, каде, во кој период и со кои инструменти. Се проверува застапеноста и повторливоста на измерените податоци со текот на времето. Индикаторите се усогласуваат со сообраќајното оптоварување, преку број на возила или поминати километри, за да се утврди реалната промена на ризикот [42].

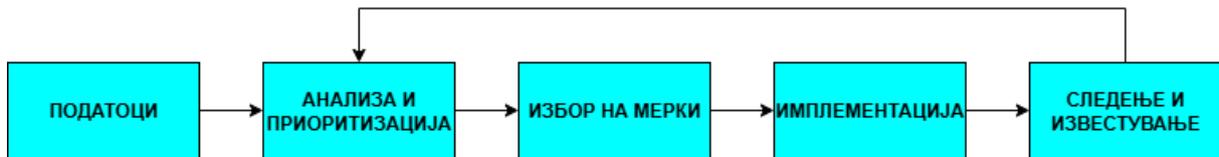
За меѓународни споредби најпрво се вршат контроли на валидност и усогласеност, по што податоците се вклучуваат во заеднички анализи. Кога дефинициите и постапките се усогласени, индикаторите стануваат споредливи на меѓународно ниво низ времето, а процената на ефектите од мерките станува поточна и поверодостојна [43].

2.11 Концепциска шема на системот на патна безбедност

Концепциската шема покажува како теоријата се претвора во практична работа, со јасен редослед на чекори што почнуваат со податоци, продолжуваат со анализа и одредување приоритети, а потоа следи: избор на мерки, имплементација и следење на резултатите. Влезните податоци се групираат во индикатори што опишуваат ризични состојби и на податоци за незгоди и повреди, а анализата ги прикажува различните делници во различни временски интервали, пред и потоа, за да се види како се менува состојбата. Од тие споредби се поставуваат цели и се одредуваат приоритети, со што мерките се насочуваат таму каде што се очекува најголем ефект [44].

На слика 2.3 прикажана е шемата со пет чекори и повратна врска. Прво се прибираат податоци со усогласени правила, потоа следи анализа и дефинирање приоритети според очекуван ефект и изводливост, потоа следува избор и концепт на мерки по делови од системот, потоа имплементација со јасни задачи и рокови. На крајот се врши следење преку истите индикатори за да се оцени постигнатиот резултат. Истите правила важат на ниво на делница и на ниво на мрежа, што овозможува резултатите да се споредуваат во текот на времето и меѓу дефинирани локации.

Изборот на мерки се одвива според системот возач, возило, пат и околина, при што се обрнува внимание на управувањето со брзини, геометриските карактеристики на патот и прегледноста, безбедносните системи во возилата и организацијата на реакцијата по незгодата. Имплементацијата се запишува, така што секоја активност има назначен извршител, рок и очекуван резултат. Следењето се врши со исти правила и методи за да се добијат споредливи резултати кои може со сигурност да се потврдат.



Слика 2.3 Концептуална шема на системот на илмина безбедност (ISO 39001:2012)

Прикажаната шема служи како практичен водич за редослед на активности и распределба на одговорности. Овој целосен преглед ја нагласува улогата на информациите и повратната врска во намалување на ризикот, во согласност со современи инженерски пристапи и управување со перформанси [45]. Со ваков распоред одлуките се базираат на податоци, мерките се прилагодуваат според резултатите, а споредливоста и насоката се задржуваат.

2.12 Кратка синтеза на теоретската рамка

Безбедноста може да се разгледува како карактеристика на системот во кој учествуваат возачот, возилото, патот и околината. Наместо да се очекува безгрешно однесување, решенијата се насочуваат кон намалување на веројатноста за незгода и кон ограничување на последиците ако истата се случи. Важна е улогата на управувањето со брзини и усогласувањето на геометриските карактеристики на патот, сигнализацијата и безбедносните системи во возилата со условите на дадена локација.

Епидемиолошкиот пристап и Хадоновата матрица ги поврзуваат фазите пред незгода, за време на незгода и по незгода со четирите делови од системот. На овој начин секоја мерка добива свое место и цел, од обука и контрола, преку уредување на трасата и безбедносните системи во возилата, до брза реакција по незгодата.

Индикаторите за перформанси и податоците за незгоди се користат заедно. Индикаторите ја опишуваат состојбата што создава ризик, податоците за незгоди го опишуваат исходот. Кога мерењата се усогласени и се изразуваат преку сообраќајно оптоварување, се добиваат споредливи промени и ефекти што може сигурно да се проверат.

Целиот процес се спроведува по редослед: прибирање податоци, анализа и дефинирање на приоритетите, избор на мерки, нивна имплементација и следење. Повратната врска овозможува резултатите повторно да се вклучат во анализата за подобрување на приоритетите и концептите на преземените мерки [42,44].

Во оваа насока, истражувањето ги стандардизира поимите, ги класифицира индикаторите од патот и околината, и воспоставува аналитички пристап за проверка на безбедноста.

3. ПРЕГЛЕД НА ДОСТИГНУВАЊАТА ВО НАУЧНАТА ОБЛАСТ

Прегледот ги поврзува трите групи кои влијаат на безбедноста и на бројот на сообраќајни незгоди. На почеток се опишуваат карактеристиките на патот: правци и кривини, надолжни наклони, витоперење, состојба на коловоз, сигнализација и прегледност. Потоа следат сообраќајните карактеристики како ПГДС, структура на сообраќај и брзини. На крај се додаваат условите на околината и оперативните услови како што се временските услови, делот од денот и теренот. На овој начин се гради целосна слика за ризикот, на секоја делница посебно.

Покрај описот на податоците, се користат ГИС-алатки за приказ на ризик на карта, постапки за издвојување на влијателни параметри и начини за толкување на резултатите. Дополнително се користат методи за анализа, мерки и правила за проверка и споредба на модели, во насока на тоа резултатите да бидат сигурни и лесни за споредување. Клучната идеја е дека, незгодите произлегуваат од заедничко дејство на патни и сообраќајни фактори и фактори на околината. Затоа се следи нивната зачестеност и се користи тежинскиот индекс на сообраќајни незгоди (W_i) за оценување и рангирање на ризик по делница. Резултатите се презентираат по ист редослед: најпрво локално, потоа регионално и на крај глобално, со насока за практична примена.

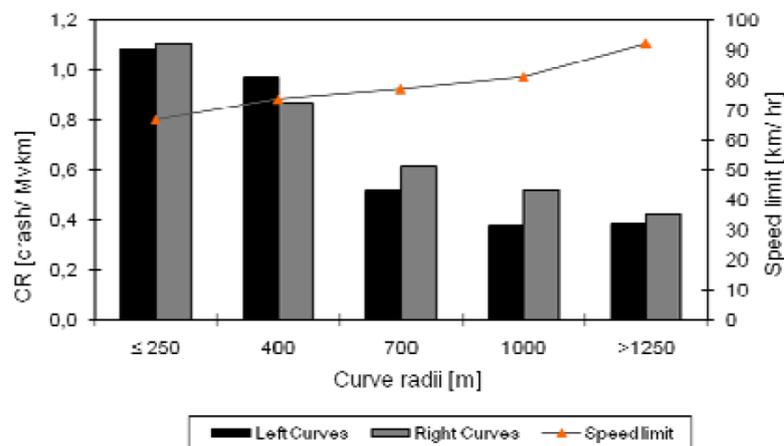
3.1 Карактеристики на патот и прегледност

Карактеристиките на патот се клучен фактор за ризикот по делница, бидејќи ја определуваат траекторијата на движење, прегледноста и реалните брзини. На локално ниво, трудот „Влијание на елементите на патот врз бројот на сообраќајните незгоди“ [46] опфаќа петгодишни податоци за делница на А4 кај нас, со анализа по делници на правци, кривини, надолжни наклони, витоперење, состојба на коловоз и прегледност. Методот се базира на идентификација и споредба на ризични точки по делница. Резултатите потврдуваат дека усогласената геометрија и точната сигнализација ја намалуваат зачестеноста на незгоди. Ова е директно релевантно, бидејќи истите параметри се главни фактори што го одредуваат бројот на незгоди при моделирањето по делница.

Во оваа насока, „Impact of Road and Traffic Characteristics on the Traffic Accidents“ [47] додава регионален опфат со регресиона анализа на ниво на делници според радиусот на хоризонталните кривини, надолжниот наклон и густината на пристапните точки. Резултатите издвојуваат критични граници, каде наклон над 3 проценти и хоризонтални кривини со мал радиус се поврзани со значително повисока зачестеност на незгоди, што дава прецизни инженерски насоки за реконструкција и управување со пристапите на магистрални делници. Ова ја поткрепува идејата дека корекциите на геометриските елементи треба да бидат приоритет, таму каде што отстапуваат од препорачаните граници.

„Identifying Critical Road Geometry Parameters Affecting Crash Rate and Crash Type“ [48] прецизира кои комбинации на геометриски елементи се најпроблематични. Опфатот вклучува патишта со средна разделна лента со детален опис на банкини и состојба на коловоз. Анализата на стапки по геометриски класи покажува дека, комбинацијата мал хоризонтален радиус, тесни банкини или пат без банкини и деградиран коловоз, значајно ја зголемува зачестеноста и ја менува структурата на типовите на незгоди. На слика 3.1 е прикажана зависноста меѓу незгоди и радиус на хоризонтална кривина, при што критични се хоризонтални кривини со радиуси под 700 m, што укажува на давање приоритет на

реконструкција на делници со такви вредности на хоризонталните кривини. Овој графички приказ ја олеснува проверката на делниците што отстапуваат од безбедните граници и служи како основа за планирање на мерки.



Слика 3.1 Зависносии меѓу радиус на кривина и зачестиеносии на незгоди (Othman, 2009)

Состојбата на коловозот директно ја менува безбедноста на делницата, бидејќи со стареење се влошуваат адхезијата, рамноста и одводнувањето. Се појавуваат микроштетувања, полирање на агрегатот и ситни нерамнини што го намалуваат триењето, особено во влажни услови.

Трудот „Effect of Road Pavement Types and Ages on Traffic Accident Risks“ [49] покажува јасен временски образец, ризикот расте со поминати години од последната санација и ефектот е поизразен кога на коловозот има присуство на вода. Тоа важи и за порозните асфалти, каде пропустливоста со време опаѓа, водата побавно се одведува, се зголемува веројатноста за пролизгување и се продолжува должината на сопирање. Во пракса, приоритетите за одржување не треба да се темелат само на визуелен впечаток, туку на систематско следење на два едноставни индикатори: датумот на последната рехабилитација и поминатото време од неа, во комбинација со податоци за условите при незгоди во влажни услови. Согласно тоа, навреме се препознаваат делници со растечки ризик и се планираат рехабилитации и мерки за подобро одводнување.

Во урбани услови, „The Impact of Road Geometric Formation on Traffic Crash and Its Severity Level“ [50] анализира 17 006 незгоди (Будимпешта, 2017–2021) со мултиномна логистичка регресија и MLP-ANN и покажува дека хоризонталните кривини и делниците со повеќе од 3 сообраќајни ленти се поврзани со повисока сериозност на последниците од незгоди. За да се споредат групите, авторите воведуваат индекс на сериозност (SI) како удел на исходи во дадена група, кој се движи од 0 до 1, при што збирот на уделите по категории (лесни, тешки, смрт) изнесува приближно 1. Во таа рамка, за групата со ≥ 3 ленти, $SI(\text{смрт}) \approx 0.017$ произлегува од 48 смртни исходи на 2 791 незгода. Овој резултат упатува на потреба од подобра прегледност и внимателно поставени ограничувања на брзина на локации со комплексна геометрија.

Непосредната околина на патот често го дефинира степенот на последици кога ќе се случи грешка. „Influence of Road Infrastructure Design over the Traffic Accidents: A Simulated Case Study“ [51] користи симулации за да испита што се случува кога возило излегува во отворен одводен канал покрај пат и моделира две типични ситуации. Се мерат брзината и оптоварувањата на возачот и патниците при удар и движење по каналот, при што се добиваат брзини околу 79 km/h и 65 km/h при забрзувања од приближно 50 m/s² во

надолжна насока и 30 m/s^2 во попречна насока. Овие вредности укажуваат на големи оптоварувања и висока веројатност за тешки повреди. Поради тоа, отворените канали покрај патот претставуваат сериозна опасност при излегување и треба да се заштитат со покривање или поставување заштитни огради.

Трудот „The Role of Attention in Horizontal Curves: A Comparison of Advance Warning, Delineation, and Road Marking Treatments“ [52] експериментално споредува три решенија за кривини: знаци за предупредување пред кривина, обележување на самата кривина и патни ознаки на коловозот, во и пред кривината. Мерењата се прават во контролирани услови со следење на брзината (пред и во кривината), позиција на возилото (држење во лента) и вниманието на возачот. Резултатите покажуваат дека решенијата што визуелно ја истакнуваат кривината и нејзината закривеност (шеврони, јасни рабни и разделни линии, водилки) поефикасно водат до намалување на брзината во кривината и постабилно држење на лентата отколку само навремени предупредувања со класични знаци. Главниот заклучок е дека, геометријата добива реален безбедносен ефект кога е придружена со јасно и континуирано обележување, што ја прави кривината читлива за возачот и ја намалува веројатноста за грешки.

Литературата укажува на повеќе клучни фактори што влијаат на ризикот. Најизразени се радисот на хоризонталната кривина, надолжниот наклон и витоперењето. Со тек на време прегледноста и состојбата на коловозот се менуваат и го зголемуваат влијанието врз безбедноста. Кога ќе настане незгода, елементите на непосредна околина на патот и сигнализацијата често ја одредуваат тежината на последиците. Во пракса потребно е брзо препознавање и санација на делници со мали радиуси на хоризонталните кривини и несоодветни банкини. Во плановите за одржување треба да се следат староста на коловозот и датумот на последната интервенција. Ограничувањата на брзината треба да се усогласат со пресметаната прегледност. Со таков пристап мерките се насочуваат таму каде што најмногу го намалуваат ризикот.

3.2 Карактеристики на сообраќајот

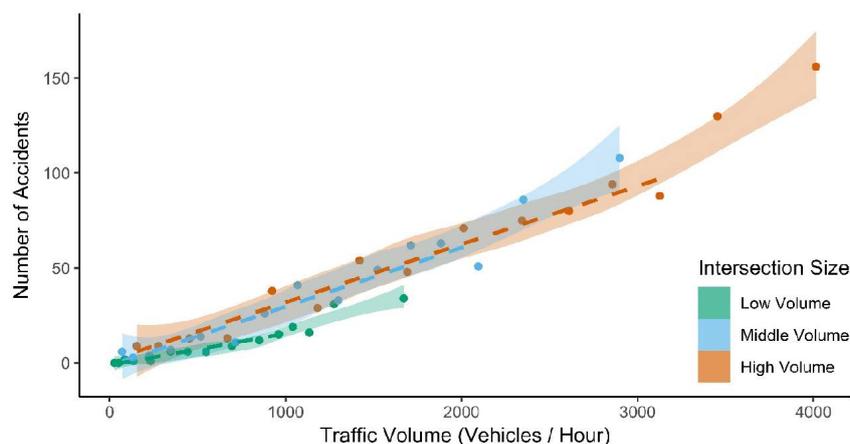
Карактеристиките на сообраќајот ја определуваат изложеноста и динамиката на движењето по делници, при што ПГДС, структурата на сообраќајниот тек и управувањето со брзините најчесто се поврзуваат со зачестеноста на незгоди. На локално ниво, „Систематизација и оценување на мерки за смирување на сообраќајот на секундарна патна мрежа“ [53] покажува дека по воведување мерки за смирување на сообраќајот, средните брзини се намалуваат за 10–15 km/h, брзината под која се движат 85% од возилата опаѓа за 22–29%, а незгодите се намалуваат за 15–30% во споредливи периоди. Ова е значаен показател дека, нормализирањето на дијаграмот на брзина носи мерливо намалување на ризикот.

Регионалното истражување „Predicting Road Traffic Accidents–Artificial Neural Network Approach“ [54] собира податоци по делници од Србија и Република Српска за да предвиди каде ќе има повеќе незгоди и какви ќе бидат последиците. На почеток се подготвуваат влезовите, како ПГДС, должина, основни геометриски податоци и услови на околината. Потоа се градат два модели со вештачки невронски мрежи (ANN). Првиот предвидува вкупен број незгоди, додека вториот класифицира какви ќе бидат последиците. По обучувањето се проверува кои променливи најмногу придонесуваат. Како најсилни сигнали се издвојуваат ПГДС и должината на делницата. Должината

учествува со околу 40% релативна важност во моделот за вкупни незгоди. Моделот за последици постигнува приближно 75% точност. Заклучокот е дека овие резултати се доволни за оперативно рангирање на делници, насочување контроли и планирање мерки таму каде што ефектот се очекува да биде најголем.

Меѓународните истражувања потврдуваат дека повисока просечна брзина значи поголем ризик. „The Relation Between Speed and Crashes“ [55] анализира податоци од повеќе патишта и ја поврзува просечната брзина со бројот и тежината на незгодите. Прво се споредуваат делници со различни профили на брзина. Потоа се моделира врската меѓу просечната брзина и ризикот, одделно за незгоди со повредени и за незгоди со фатални последици. Резултатите покажуваат дека, кога просечната брзина расте, расте и ризикот, при што ефектот е посилен кај најтешките последици. Зголемување од 10% во просечната брзина поврзано е со околу 21% повеќе незгоди со повредени и околу 46% повеќе незгоди со фатален исход, додека при намалување важи обратен ефект. Оттука произлегува дека усогласувањето на брзините по делници претставува ефикасен начин за намалување на ризикот, особено кога се работи за најтешките последици.

Во рамки на обемот и конфликтите на крстосници, трудот „Relationship Between Traffic Volume and Accident Frequency at Intersections“ [56] анализира 120 крстосници во Аделаида, Австралија, со 1 629 незгоди и повеќе од пет милиони часовни мерења на обем. На мал обем, врската со незгодите е приближно линеарна. Кога обемот е многу висок, незгодите растат побрзо од линеарно. Тоа се опишува со квадратен член во Поасонов и негативно-биномни модели. Дополнително, во анализата е вклучен и ефектот од дожд, кој го зголемува ризикот. На слика 3.2 прикажан е изразен раст на бројот на незгоди со зголемување на обемот на сообраќај, особено кога крстосницата е преоптоварена. Овие констатации помагаат мерките да се насочат кон најоптоварените пристапи и кон критичните периоди.



Слика 3.2 Зависност меѓу сообраќаен обем и зачестеност на незгоди на крстосници (Retallack, 2020)

Овој приказ е важен бидејќи покажува дека обемот на сообраќај треба да се моделира нелинеарно и со вклучување на временските услови (особено дожд), за да се согледа брзиот раст на ризикот и да се утврдат приоритетите за подобрување на ризичните делници и крстосници во понатамошната анализа.

Истражувањата кај нас, во регионот и во меѓународната литература се усогласени. ПГДС и управувањето со брзина добро ја објаснуваат промената на ризикот по делници. Кај крстосници, големината на сообраќајот и временските услови делуваат заедно, па ризикот расте. Затоа најдобро функционираат мерки што комбинираат управување со брзина, подобрување на прегледноста и јасно обележување на пристапите.

3.3 Околина и оперативни услови

Факторите на околината и оперативните услови го менуваат ризикот на делниците преку прегледност, пролизгување и начин на возење. На локално ниво, „Analysis of Traffic Accidents in North Macedonia from 2010 to 2019“ ги разгледува сообраќајните незгоди во Македонија од 2010 до 2019 година. Се прави споредба по години и по главни причински фактори. Пребрзото возење се јавува во приближно секоја шеста незгода. Алкохолот учествува во околу секоја десетта незгода. Бројот на незгоди поврзани со алкохол се зголемил приближно 2.7 пати во периодот 2010–2019. Со тоа се покажува дека ризикот се менува со времето и со условите на возење. Ова укажува дека моделите за процена на ризик треба да го следат распоредот на незгодите по периоди и по услови (време, прегледност, пролизгување и начин на возење), за навремено да се препознаат зголемувањата на ризик и да се планираат соодветни мерки [57].

Во регионален контекст, „SARIMA Modelling Approach for Forecasting of Traffic Accidents“ анализира месечен број на сообраќајни незгоди во Белград (2016–2019) за да утврди дали постои повторливост во текот на годината и да изгради краткорочни прогнози. Анализата користи SARIMA (Seasonal AutoRegressive Integrated Moving Average) модел, а изборот на моделите се прави преку ACF [Autocorrelation Function] и PACF [Partial Autocorrelation Function], со проверка на точноста. Резултатите покажуваат изразена сезонска нерамномерност: во последниот квартал (октомври–декември) има најмногу незгоди. Моделот постигнува MAPE [Mean Absolute Percentage Error] = 5.22%, што значи дека прогнозата во просек отстапува околу 5 на секои 100 случаи, што е доволно добро за практично планирање.

Вакви сезонски модели можат навреме да посочат кои месеци имаат повисок очекуван ризик, за службите да планираат мерки по делници, токму во тие периоди [58].

На глобално ниво, „Effects of Adverse Weather on Traffic Crashes: Systematic Review and Meta-Analysis“ прикажува систематски преглед на истражувања за влијанието на неповолни временски услови врз сообраќајните незгоди. Во трудот се опфатени повеќе студии од различни земји и години. Откритијата покажуваат дека при врнежи, бројот на незгоди се зголемува. Снегот има посилен ефект од дождот и може да го зголеми бројот на незгоди за околу 84% а повредите за околу 75%. Резултатите упатуваат, во идното моделирање и планирање да се вклучат периодите со дожд и снег, со акцент на прегледноста, пролизгувањето и прилагодувањето на брзината. Мерките треба да се распределат по делници, во периодите кога ризикот очекувано расте [59].

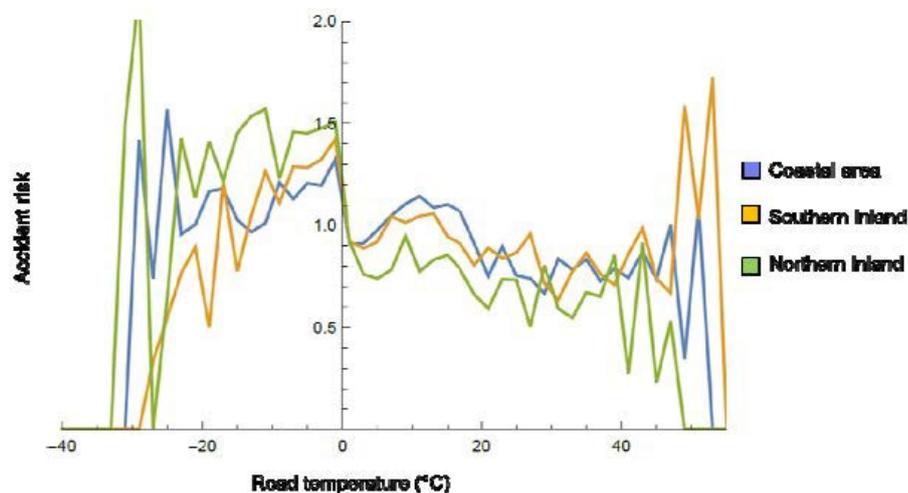
Како дополнување, „A Review of the Effect of Road Weather Conditions on Road Safety“ сумира преглед на истражувања од повеќе земји и континенти, со опфат на урбани сообраќајници, автопати и регионални делници. Се разгледуваат дожд, снег, магла, ниски температури и ветер, како и нивното дејство преку намалена прегледност, послаб контакт на пневматиците со коловозот и подолг пат на сопирање. Резултатите покажуваат зголемување на незгодите за време на врнежи, при што снегот обично има посилен ефект од дождот, а влијанието се менува според местото, категорија на пат и периодот од годината. Временските услови треба јасно да се внесат во моделите и во планот на мерки по делници. Фокусот треба да биде во периодите и местата каде податоците покажуваат зголемен број незгоди [60].

„Road Accidents and Rainfall in a Large Australian City“ анализира дневни сообраќајни незгоди во Мелбурн (1987–2002), со нормализација на бројот на незгоди според обемот на

сообраќај. Во дождливи денови ризикот е повисок отколку во суви денови. Зголемувањето е поизразено кога дождот следи по подолг сув период, а влијанието варира по сезони, со поистакнат ефект во есен отколку во пролет. Овие резултати покажуваат дека моделите и плановите треба јасно да ги вклучат временските услови, со акцент на времетраењето на сувите денови и на сезоната [61].

Трудот „Weather-Related Road Accidents in England and Wales: A Spatial Analysis“ ги разгледува временски предизвиканите сообраќајни незгоди во Англија и Велс и ја анализира нивната просторна распределба, користејќи полициски записници сумирани на ниво на локални власти. Целта е да се утврди каде и колку временските појави како дожд, магла и силен ветер се поврзани со поголем број незгоди. Резултатите покажуваат значајни регионални разлики и групирање на незгоди поврзани со временски услови во одредени подрачја, како и јасни разлики меѓу периодите од годината. Овие резултати упатуваат дека, при процена на ризик по делници треба да се вклучат просторни показатели и климатски одлики, за мерките да се приспособат на местата каде временските услови најмногу придонесуваат за незгоди [62].

Во студени услови, трудот „Identification of Accident Risk on the Finnish Main Road Network in Different Weather Conditions“ ја анализира главната патна мрежа во Финска и врската меѓу временските услови и ризикот од сообраќајни незгоди. Користени се обемни сетови на податоци од сензори и полициски записи, со поделба на мрежата по климатски зони и по температури на коловозот. Следени се денови со снег, мраз, дожд и суви услови, како и интервали со температура над и под 0°C. На слика 3.3 прикажана е врската меѓу температурата на коловозот и релативниот ризик по зони. На температури под 0°C ризикот брзо се зголемува, при што линиите за крајбрежната, јужната и северната внатрешност се разликуваат, што покажува дека секоја зона реагира поинаку на студени временски услови.



Слика 3.3 Релативен ризик по температура и климатска зона (Malin, 2017)

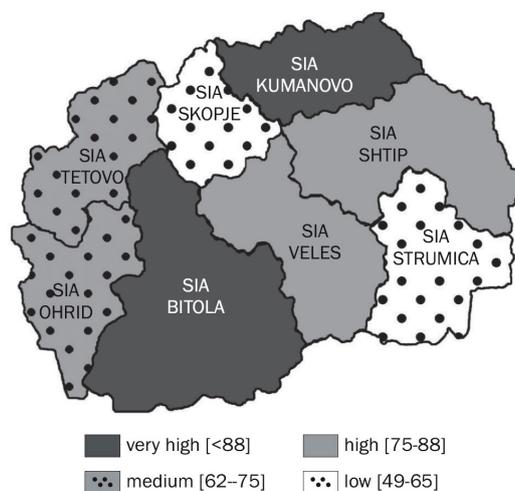
Резултатите укажуваат дека ризикот е најголем при мраз и насобран снег. Профилите на ризик се разликуваат по зона и по период од годината, со поизразен ризик во зимските месеци и навечер. Оттука, управувањето со брзина и предупредувањата, треба да се прилагодат по зона и по температура на коловоз, со јасни граници за активирање во студени периоди [63].

Сумирано, зимските услови со снег и мраз најмногу го зголемуваат безбедносниот ризик, а дождот ја зголемува зачестеноста на незгоди (особено при мал сообраќаен обем).

Во ноќните часови и зимските месеци се препорачува засилена контрола и поинтензивно информирање на возачите.

3.4 Просторна категоризација на ризик

Просторната категоризација го прави ризикот видлив по делници и по подрачја, со мапи и мерни показатели што овозможуваат споредба и приоритет. Во трудот „Analysis of Road Traffic Safety through Direct Relative Indicators for Traffic Accidents Fatality: Case of Republic of Macedonia“ се пресметува јавен ризик на 100 000 жители, сообраќаен ризик на 10 000 возила и динамички ризик по патник-километри за период 2005–2014, а потоа се прави картографска поделба по сектори. Вредностите за јавен ризик се движат од околу 6.28 до 9.58 (загинати на 100 000 жители), во четири нивоа, што овозможува просторна приоритизација на мерките [64].



Слика 3.4 Мапирање на јавнобезбедносен ризик по МВП-сектори (Babanoski, 2016)

На слика 3.5 прикажана е поделба на ризикот по сектори со четири нивоа. Овој приказ го насочува изборот на мерки кон подрачјата со највисок ризик и помага да се одреди редоследот на интервенции по делници.

„A Comparative Analysis of Hotspot Identification Methods“ споредува седум методи за идентификација на ризични точки (Hotspot Identification Methods, HSID) на автопатот A16 во Италија, користејќи пет годишни податоци за сообраќајни незгоди. Разгледани се: зачестеност на незгоди, стапка на незгоди, еквивалент на незгоди со само материјална штета, пропорција на незгоди, емпириски Баес (Empirical Bayes, статистички пристап што комбинира историски податоци и податоци од конкретна локација за попрецизна процена на ризикот) за вкупни и за тешки незгоди и потенцијал за подобрување. Точноста и стабилноста се оценуваат преку четири тестови: конзистентност по локација, конзистентност по метод, вкупна разлика во рангови и вкупен комбиниран скор. Емпирискиот Баес (ЕВ) излегува најстабилен и најсигурен, дури и кога се менува прагот за избор на критични точки (најгорни 1%, 5% или 10% локации според рангирањето). Едноставните мерки, како зачестеност и стапка, даваат понестабилни и помалку усогласени листи, па во практиката е оправдано предност да имаат ЕВ-пристапите за поверојатно утврдување на приоритетните делници за мерки [65].

Трудот „A Machine Learning Approach for Classifying Road Accident Hotspots“ анализира 10 годишни полициски податоци за незгоди на бразилските федерални

автопати, околу 1.8 милиони записи, и класифицира делници со висок ризик за тешки незгоди според период во денот, ден во неделата, тип и траса на патот, насока, временски услови и вид на незгода. Најдобар резултат дава невронска мрежа со точност 83% (Accuracy), прецизност 84% (Precision), одзив 83% (Recall) и F1 - score 82% (F1 - score, мерка што ја спојува прецизноста и откриените случаи во една вредност) на урамнотежен сет на податоци, а XGBoost (Extreme Gradient Boosting, метод на екстремно засилување што ја подобрува точноста на предвидувањата) постигнува F1-скор 71% во балансирана поставка со атрибут за зачестеност. Клучниот наод е дека, балансирањето на класите е неопходно бидејќи без него метриците за откривање тешки незгоди опаѓаат иако вкупната точност изгледа висока, што укажува дека комбинирањето на оперативни и услови од околина со машинско учење овозможува попрецизно и постабилно издвојување на приоритетни делници за преземање мерки [66].

Методологијата „iRAP Methodology and Star Ratings (incl. SR4D)“ објаснува како се оценува безбедноста на патиштата со ѕвезди од 1 до 5, при што 1 ѕвезда значи најнизок степен на безбедност, а 5 ѕвезди највисок. Методологијата користи податоци за патот, околината и ограничувања на брзина. Врз основа на овие елементи се пресметува ризик од тешки последици при незгода за различни учесници во сообраќајот, одделно за возачи, мотоциклисти, пешаци и велосипедисти, па се доделува оценка со ѕвезди по корисник. Покрај оценувањето, системот предлага и мерки за подобрување, како уредување на раскрсници, подобро обележување, управување со брзини, нови пешачки премини или заштитни огради, со проценка на ефект и трошок. SR4D (Star Rating for Designs) е примена на истата логика врз проектни решенија пред патот да се изгради. Наместо готова делница, се оценува проектот и се проверува колку ѕвезди би добиле различните учесници. Потоа, проектот може да се коригира уште во фаза на проектирање, за да се постигне повисока безбедносна оценка со мали промени, а подоцна и полесно да се одредат приоритетни мерки по делници [67].

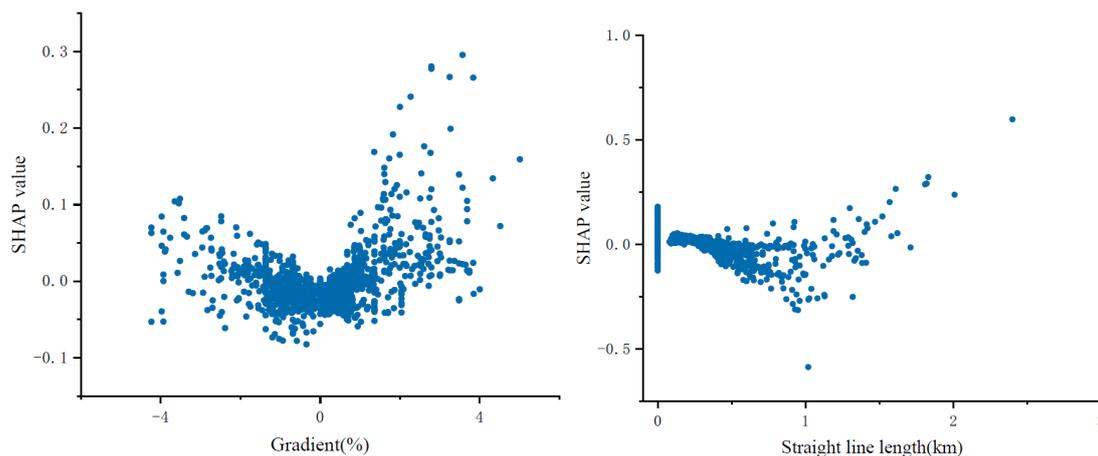
Картите со индикатори, статистичките методи за наоѓање критични места, класификацијата со машинско учење и системот за рангирање (iRAP) работат заедно. Со нив се прави рангирање по простор каде што ризикот е најголем и се планираат идни интервенции. Со оваа рамка делниците со најголем ризик се откриваат навреме, се споредуваат по подрачја и се поврзуваат со мерки што носат најголема корист.

3.5 Влијателни параметри и објаснивост

Овој преглед објаснува кои параметри најмногу влијаат на ризикот по делници и како тие се рангираат во анализите. Во врска со на автопатиштата, „A Crash-Prediction Model for Multilane Roads“ се анализираат пет годишни податоци (1999–2003) од патишта со четири сообраќајни ленти со разделна лента на автопатска делница во Италија. Целта е да се предвиди бројот на незгоди и да се утврдат највлијателните фактори, со одделни модели за правци и за кривини. Во анализата влегуваат должината на делницата, закривеноста на трасата, ПГДС, прегледноста, радијална компонента на коефициентот на отпор од триење, надолжниот наклон и присуството на крстосници. Резултатите покажуваат дека кај кривини најизразено делуваат должината, закривеноста и ПГДС, а кај правци, должината, ПГДС и крстосниците. Влажниот коловоз дополнително го зголемува бројот на незгоди. Заклучно, моделите овозможуваат рано препознавање на делници чувствителни на промени во сообраќајното оптоварување, геометриските елементи и во

состојбата на коловозот, што помага мерките да се насочат таму каде што ефектот ќе биде најголем [68].

Трудот „Relationship between Highway Geometric Characteristics and Accident Risk: A Multilayer Perceptron Model (MLP) Approach“ истражува како геометриските одлики на патиштата се поврзани со ризикот од незгоди и дали моделот со повеќеслоен перцептрон (MLP, невронска мрежа со повеќе слоеви што учи врски меѓу влезните параметри и резултатот) може подобро да го предвиди тој ризик од класични статистички пристапи [69]. Користени се податоци за делници со мерки како радиус на кривини, надолжен наклон, ширина на ленти и банкини, број и растојание на крстосници, разделен појас, прегледност и ограничувања на брзина, заедно со евиденција на незгоди по делница. Резултатите покажуваат дека MLP дава повисока точност при разликување делници со повисок и понизок ризик во споредба со едноставни референтни модели. Највлијателни променливи се закривеноста (помал радиус - поголем ризик), поголем надолжен наклон, тесни ленти и банкини, висока густина на крстосници и недоволна прегледност, додека разделен појас и пошироки банкини се поврзани со понизок ризик. Дијаграмите на слика 3.5 прикажуваат, како придонесот на моделот расте при поголеми надолжни наклони и при подолги правци, што упатува на мерки за управување со брзината и предупредувања на делници каде вредностите значително отстапуваат од нула.



Слика 3.5 SHAP-рангирање на влијателни карактеристики во MLP-модел (Yan, 2023)

Резултатите покажуваат дека моделите од типот MLP можат рано да идентификуваат делници чувствителни на геометриски карактеристики и услови на патот. Затоа, мерките како корекции на кривини и надолжни наклони, проширување на ленти и банкини и управување со крстосници и брзини се насочуваат таму каде што ќе донесат најголемо намалување на ризикот.

„Effect of Road Pavement Types and Ages on Traffic Accident Risks“ анализира како типот и староста на коловозот се поврзани со ризикот од незгоди, со спојување на евиденција за незгоди и база за состојба на коловоз по делници [70]. Се следат материјалот на коловозот, изминато време од последната рехабилитација, рамноста и фатливост на пневматиците со коловоз, одделно на сув и влажен коловоз и преку ден и во текот на ноќта. Резултатите покажуваат дека, на влажен коловоз ризикот е околу 60% повисок од сув коловоз, што значи дека истата делница при исти услови на сообраќај има поголема веројатност за незгода кога е влажна. Староста на коловозот додава околу 4% дополнителен ризик на секоја година. На кривини при дожд ова зголемување е уште посилено, околу 10–16% по година, што во пракса значи дека стар и истрошен коловоз на кривина побрзо станува ризичен. Овие резултати упатуваат, во моделите да се внесат

староста и типот на коловоз, а во планирањето да се даде приоритет на рехабилитации на постари и лизгави делници, особено на кривини и во дождливи периоди.

Во зимски услови, „Risk Prediction for Winter Road Accidents on Expressways“ се анализираат сообраќајни незгоди на автопати и експресни патишта со средна разделна лента во Кореја, за периодот декември-февруари во 2017-2019. Податоците ги спојуваат полициските записи, метеоролошките мерења и геометриските карактеристики на трасата. Поради мал број зимски незгоди поврзани со време, податоците најпрво се нормализираат, па потоа се тренираат модели на машинско учење како логистичка регресија, невронска мрежа, XGBoost и Random Forest (метод што користи повеќе случајни дрва на одлука за да постигне поголема точност и стабилност). Random Forest постигнува Accuracy = 0.984 и AUC = 0.907 (Area Under the Curve, мерка што ја покажува способноста на моделот да разликува различни нивоа на ризик), што укажува на добра способност да се разликува ризикот по делници во зима. Најизразени фактори од временските услови се релативната влажност и температурата, поврзани со мраз и магла. Од геометриските карактеристики најмногу влијаат малите радиуси на хоризонталните кривини и големиот прекршен агол, а значајни се и надморската височина и промените во непосредната околина на патот. Комбинирањето на временските услови и геометријата овозможува навремено означување на делници за управување со брзини, предупредувања и зимско одржување [71].

Заклучно, највлијателни параметри се геометриските карактеристики (должина, радиус на хоризонтална кривина, ширина на сообраќајна лента), сообраќајно оптоварување и услови на коловозот (влажност, снег, староста на коловозот). SHAP дава јасно рангирање на влијанијата по карактеристики и овозможува резултатите практично да се применат во моделирање по делници.

3.6 Аналитички методи

Аналитичките пристапи за моделирање на сообраќајни незгоди се движат од класични нумерички модели кон податочни ориентиран модели, со едноставни правила за избор на модел (критериуми за соодветност, вкргрстена валидација, споредба на метрики), така што толкувањето и предвидувањето се спојуваат.

„The Statistical Analysis of Crash-Frequency Data: A Review and Assessment of Methodological Alternatives“ е преглед на методи за моделирање на број на незгоди по делници. Трудот објаснува кога и кој пристап треба да се користи, и како едноставно да се толкуваат резултатите. Во трудот се издвојуваат два клучни резултати корисни за примена. Кога бројот на незгоди силно варира меѓу делници, соодветен избор е Негативно-биномниот модел, бидејќи подобро ја опишува променливоста и дава постабилни процени. Кога одреден фактор е во изразен раст, корисни се модели кај кои ефектите јасно се читаат како промени на ризик (Поасонов модел или Негативно-биномен модел со лог врска), за полесен увид во промената на очекуваниот број на незгоди. Овој пристап помага, резултатите да се претворат во јасно рангирање и во план на мерки што е лесен за примена [72].

Продолжувајќи со избор на распределба, „The Relationship Between Truck Accidents and Geometric Design of Road Sections: Poisson Versus Negative Binomial“ ги споредува моделите за врската меѓу незгоди на товарни возила и геометриски карактеристики на патот. Анализата е спроведена на државната патна мрежа во САД, со делници од автопати

и главни сообраќајници. Опфатени се податоци за ПГДС по лента, закривеност на трасата, радиус на хоризонтална кривина, надолжен наклон, ширини на ленти и банкини, како и отстапувања во нивната ширина. Покажано е дека Поасоновиот модел често не е доволен, бидејќи не ја опфаќа целата варијабилност. Негативно-биномната регресија дава подобри резултати кога има повеќе возила и изразена закривеност. Резултатите покажуваат дека највлијателни параметри за незгоди на товарни возила се, поголем ПГДС по лента, поголема закривеност, помал радиус на хоризонтална кривина, поголем надолжен наклон и недоволни ширини или отстапувања на ленти и банкини. Ова насочува кон примена на Негативно-биномна регресија кај посложени услови и делници со големи разлики во сообраќајното оптоварување и геометријата [73].

Во трудот „Application of Finite Mixture of Negative Binomial Regression Models With Varying Weight Parameters“ се испитува како да се моделираат незгоди кога делниците изгледаат слично по број на незгоди, но се разликуваат по карактеристики како радиус на хоризонтална кривина и надолжен наклон. Како замена за еден модел, се користи мешавина од повеќе негативно-биномни модели за подобро да се откријат скриените разлики меѓу делниците. Овој пристап дава пониски вредности на AIC (Akaike Information Criterion, мерка што го оценува квалитетот на статистички модел со земање предвид на неговата сложеност) и поголема точност, па може да ги разликува групите на патишта со различен ризик, иако на прв поглед изгледаат слично. Ваквото моделирање овозможува појасна и попрacticна поделба на делниците, според нивната опасност [74].

За истовремено моделирање на повеќе типови незгоди, „A Multivariate Random Parameters Zero-Inflated Negative Binomial Modeling Approach for Traffic Crash Data“ опишува модел применет на делници од автопати и главни сообраќајници, каде често има многу делници без ниту една незгода и повеќе видови незгоди што се јавуваат паралелно. Анализата ги вклучува ПГДС, ограничување на брзина, број и ширина на ленти, банкини, разделен појас, закривеност, надолжен наклон и околина. Пристапот комбинира Негативно-биомнен модел и случајни параметри, за истовремено да се објаснат и нулите и разликите меѓу делници, каде во еден систем се поврзуваат повеќе типови незгоди. Резултатите покажуваат подобро вклопување од класичните модели и дека ефектите од обемот на сообраќај, брзините и геометријата не се исти на секоја делница. Ова овозможува попрецизно рангирање и избор на специфични мерки во зависност од локацијата како управување со брзини, проширување на ленти или банкини и корекции на траса [75].

Трудот „A Multivariate Poisson-Lognormal Regression Model for Prediction of Crash Counts by Severity, Using Bayesian Methods“ користи мултиваријантен Поасон-логнормален (MVPLN) модел за истовремено моделирање на бројот на незгоди по сериозност (материјална штета, повреда, тешка, фатална) на рурални патишта со две ленти. Во анализата наместо да се градат одделни модели, се вметнуваат заеднички случајни ефекти што дозволуваат поврзаност меѓу исходите и објаснуваат поголема варијабилност од очекуваното. Проценката е Баесова мрежа со MCMC (Markov Chain Monte Carlo, метод за статистичка симулација), со алгоритмите Gibbs и Metropolis-Hastings, при што како влезни променливи се ПГДС, ограничување на брзина, ширина на коловоз/банкини, закривеност и наклон. Во споредба со едноваријантни модели, MVPLN покажува подобро вклопување и помали граници на сигурност, бидејќи ефикасно ги користи врските меѓу нивоата на сериозност, при што практично, поголем ПГДС и повисоки брзини се поврзани со повеќе незгоди, додека пошироки ленти/банкини и поквалитетна геометрија со помал ризик.

Методот овозможува пореално рангирање на делници и дава очекуван број незгоди по тежина на последиците, за полесно планирање мерки [76].

„Hybrid Feature Selection Based Machine Learning Classification System for the Prediction of Injury Severity in Single and Multiple Vehicle Accidents“ опфаќа податоци за незгоди со едно и повеќе возила, со информации за патот, возилото, возачот и условите на околината, за да се предвиди тежината на последици и да се утврди кои фактори најмногу влијаат. На почеток се применува Boruta филтер (за отстранување на непотребни податоци) врз Random Forest кој ги споредува променливите со нивни случајни копии и ги задржува само важните фактори. Потоа се тренираат модели како XGBoost, логистичка регресија, kNN (k-Nearest Neighbors, метод што класифицира според најблиските примери) и Naive Bayes (метод за брза и сигурна класификација заснован на Баесовата теорема) на поделба тренинг-тест 80/20 и се оценуваат со AUC, F1 и Accuracy на посебен тест дел. Резултатите покажуваат дека XGBoost најчесто има најдобри предвидувања, а рангирањето на највлијателните фактори помага да се разбере, зошто моделите донесуваат такви одлуки [77].

Трудот, „Utilizing Support Vector Machine in Real-Time Crash Risk Evaluation“ објаснува како во реално време да се процени ризик од незгода на автопати и експресни патишта. Се користат кратки интервали пред незгоди и податоци од патни детектори и метеоролошки станици. Во моделот влегуваат брзина, обем на сообраќај, капацитет на сообраќај, променливост на брзината, период во денот и дожд. Алгоритмот SVM (Support Vector Machine, метод за препознавање шеми и класификација на податоци) учи од овие податоци и препознава шеми што најавуваат повисок ризик. Резултатите покажуваат дека ризикот расте кога брзината опаѓа, капацитетот на сообраќај се заситува, варијабилноста на брзината е зголемена и има појава на врнежи. Ова овозможува навремени предупредувања и прилагодување на брзините во моментите кога ризикот почнува да расте [78].

Надоврзано на тоа, „Real-Time Crash Risk Prediction on Arterials Based on LSTM-CNN“ објаснува процена на ризик од незгода во реално време на урбани сообраќајници со семафори, користејќи кратки временски интервали пред незгода и податоци од патни детектори, временски услови и фази со семафори. Во моделот влегуваат брзина, обем на сообраќај, задоцнување на раскрсници, должина на редици и информација за активната фаза со семафори. Методот спојува повратна невронска мрежа (LSTM) за временските шеми и конволуциона невронска мрежа (CNN) за просторните шеми меѓу ленти и соседни раскрсници. Резултатите покажуваат повисока точност од класични пристапи и подобра рана детекција на ризични моменти, при што ризикот расте кога брзината опаѓа, кога промените на брзината се поголеми, редиците се издолжуваат и има неповолни комбинации на фази на семафорите и врнежи. Овој пристап овозможува навремени предупредувања, прилагодување на брзините и динамично управување со семафори на делници и крстосници, кога ризикот почнува да расте [79].

Класичните модели се корисни бидејќи се лесни за разбирање и даваат стабилна основа, а кога се комбинираат повеќе пристапи или се користат повеќе променливи полесно се гледаат скриените разлики меѓу делниците. Временските модели помагаат да се предвидат незгодите по месеци и да се планираат мерки навреме, додека машинското и длабокото учење нудат предупредувања во реално време.

3.7 Метрики и валидација

Аналитичките пристапи за моделирање на сообраќајни незгоди се движат од класични нумерички модели, кон податочни ориентиран модели. За избор на соодветен модел се користи едноставен сет правила, со проверка на соодветност, вклучена валидација и споредба на метрики за точност. Во пракса потребно е тренинг и тест податоците да бидат строго одвоени за да се избегне преголемо прилагодување. Со ваков пристап, толкувањето и предвидувањето одат заедно и резултатите се поупотребливи.

Трудот, „The Statistical Analysis of Crash-Injury Severities: A Review of Analytical Methods“ објаснува како да се изберат и проверат метрики за модели што ја опишуваат сериозноста на повреди. Опфатени се мерки за вклопување на моделот, како лог-веројатност, мерка за отстапување, AIC (Akaike Information Criterion – мерка што спречува прекумерно вклопување) и BIC (Bayesian Information Criterion – мерка што спречува прекумерно вклопување, фокусирана на моќ на предвидување), кои помагаат да се избере соодветна структура. Потоа, следи проверка на предвидувањето на одвоени податоци (тренинг-тест), каде покрај вкупна точност се гледа и колку добро моделот ги препознава тешките и фаталните случаи (прецизност, одзив, F1). Кога тешките случаи се ретки, Precision-Recall дава појасна слика од ROC (Receiver Operating Characteristic), а корисна е и проверка на калибрацијата за да се види дали предвидените веројатности се совпаѓаат со реалноста. Со оваа комбинација на метрики се добива и јасна споредба меѓу модели и понадежна основа за практична примена [80].

„Goodness-of-Fit Testing for Accident Models With Low Mean“ опфаќа проверка на модели што објаснуваат број на незгоди по делници со реални податоци каде просечниот број е мал и отстапувањето често е поголемо од средна вредност, што е типично за патни мрежи со различна должина на делници и различен сообраќај. Анализата проверува дали Поасоновиот или Негативно биномен модел се вклопува подобро, со користење на мерка за отстапување и Пирсоново χ^2 во однос на степените на слобода. Изборот на модел се потврдува со AIC и BIC за да се добие едноставно и прецизно решение. Проблемот посочува дека класичните тестови можат да дадат погрешна процена кога средината е мала, па доколку се читаат директно без уреден пристап може да се добие неточна проценка. Резултатите покажуваат дека при преголемо отстапување Негативно биномен модел најчесто дава порамномерни разлики во резултатите и пониски AIC и BIC од Поасоновиот модел. Ова овозможува поусогласена споредба на делници и посигурни процени [81].

Во анализи низ време „On the Use of Cross-Validation for Time Series Predictor Evaluation“ објаснува како најправилно да се оценат модели што предвидуваат месечен број на незгоди во текот на времето. Опфатени се симулации и реални серии со тренд и сезонски промени со цел да се покаже зошто случајната крос валидација не е соодветна за вакви податоци и зошто треба да се почитува редоследот на времето. Проблемот е дека мешањето на минати и идни податоци носи протекување на информации (data leakage) и дава премногу оптимистички резултати. Наместо тоа, се користи динамичка историја при што моделот се учи на постари месеци, а секогаш се предвидува следниот месец и чекор по чекор се поместува временскиот интервал. Резултатите покажуваат дека оваа постапка дава пореална и постабилна слика за точност кога сериите се зависни во текот на времето и дека споредбата на модели е пообјективна, ако за секој чекор се пресметаат грешките RMSE (Root Mean Square Error) и MAPE (Mean Absolute Percentage Error) а потоа се прикажат просеците и отстапувањето [82].

За класификација со ретки настани, „Precision-Recall Plot Is More Informative Than the ROC Plot When Evaluating Binary Classifiers on Imbalanced Datasets“ објаснува како да се оценат бинарни модели, кога позитивниот настан е редок. Опфатот е споредба на две мерки за валидација, ROC и Precision-Recall (крива која балансира помеѓу прецизност и реакција за ретки настани во неурамнотежени податоци), на неурамнотежени сетови и повеќе модели. Резултатите покажуваат дека ROC може да изгледа добро и кога моделот слабо ги препознава ретките позитивни случаи, бидејќи големиот број негативни примери ја прикрива слабата прецизност. Precision-Recall дава пореална слика во вакви услови, затоа што директно ги мери прецизноста и одзивот за позитивната класа. Во пракса е корисно, покрај ROC-AUC, да се прикажат PR-крива и PR-AUC, како и прецизност, чувствителност и F1 на избрани граници. Потоа, точноста се проверува на одвоен тест сет или со крос-валидација со рамномерна поделба по класи, за споредбата на модели да биде појасна и корисна [83].

„The Relationship Between Precision-Recall and ROC Curves“ дополнително го покажува односот меѓу ROC и Precision-Recall при оценување на бинарни модели, со фокус на тоа, како се менуваат метриците кога класите се нерамномерни. Опфатот вклучува теоретска врска меѓу паровите точки на двете криви и улогата на застапеноста на позитивната класа во вредностите на прецизноста и одзивот. Резултатите покажуваат дека ист ROC резултат може да одговара на различни PR-вредности зависно од уделот на позитивните случаи, па затоа PR е почувствителна на нерамнотежа и подобро ја покажува корисноста на моделот за ретката класа. За валидирање најдобро е да се прикажат и ROC и PR криви (со AUC), да се изберат неколку граници со табели за прецизност, чувствителност и F1 и секогаш да се наведе застапеноста на позитивната класа, за споредбата меѓу модели да биде појасна и корисна [84].

Во оценување на места со зголемена зачестеност, трудот „Experimental Evaluation of Hotspot Identification Methods“ споредува методи за наоѓање ризични места на патишта и покажува како да се провери нивната вредност со метрики. Опфатени се зачестеност, стапка, еквивалент на незгоди само со материјална штета (EPDO), емпириски Баес и потенцијал за подобрување, применети на реални делници. Валидацијата се прави со стабилност на рангирањето при различни граници (врвни 1, 5, 10 проценти), преклоп на листи меѓу методи и удел на ризични точки во врвот (capture rate). Резултатите покажуваат дека емпирискиот Баес има постабилни рангови и поголем удел на вистински ризични локации, додека едноставните мерки почесто го менуваат редоследот и пропуштаат критични точки. Практично, изборот на метод треба да се темели на овие валидациони метрики, а не само на сурови броеви [85].

За преносливост, трудот „International Transferability of Accident Modification Functions for Horizontal Curves“ препорачува како да се провери дали функциите за модификација на незгоди важат и во друга земја. Факторите за модификација на незгоди (CMF-Crash Modification Factor, AMF-Accident Modification Factor) се множители што покажуваат колку се менува очекуваниот број незгоди по воведена мерка: вредност 1 значи без промена. На пример: 0.8 значи намалување од 20%, 1.3 значи зголемување од 30%. Опфатот вклучува тестирање на постојни CMF/AMF на нова патна мрежа и споредба на предвиденото, со набљудуваното по категорија на пат и услови. Резултатите покажуваат дека без локална рекалибрација точноста опаѓа, па се додава калибрационен фактор и повторно се проверува со метрики како MAPE и RMSE, како и со проверка на пристрасност и покриеност на доверливите интервали. По рекалибрација, резултатите

стануваат поусогласени со реалните податоци, а CMF/AMF може да се користат за валидни процени и споредливи одлуки на новата мрежа [86].

„Estimating Safety by the Empirical Bayes Method: A Tutorial“ објаснува како без експеримент да се процени ефектот од мерки за безбедност со споредба „пред-потоа“ на истите места. Опфатени се чекори од емпириски Баес со користење подолгорочни податоци и функции за безбедност за да се пресмета колку незгоди би се очекувале и без мерката. Анализата се прави поради едноставните споредби кои можат да ги проценат придобивките при враќање кон средина и промени во сообраќајот што не се должат на мерката. Проблемот е дека ако овие фактори се игнорираат, сликата не е реална. Резултатите покажуваат дека емпирискиот Баес ја намалува оваа пристрасност и дава поцврсти процени. Поради ова, се посочува дека подобро е резултатите да се прикажат како процентуална промена со граници на сигурност [87].

Сумирано, изборот на метрики се усогласува со задачата. При споредба на модели се користат AIC, BIC и мерка за отстапување, а при прогнози се следат RMSE и MAPE во подвижни прозорци во текот на времето. Во класификација се прикажуваат PR-AUC и ROC-AUC со табели со избрани граници, а при рангирање на делници и оценка на мерки се применуваат Баесови методи, проверка на преносливост на функции и споредби „пред и потоа“ со граници на сигурност.

3.8 Синтеза и импликации

Патните карактеристики, во заеднички склоп со сообраќајот и околината ја одредуваат зачестеноста и сериозноста на незгодите, при што се издвојуваат неколку приближни шеми. Малите радиуси на хоризонтални кривини, надолжни наклони над околу 3%, долги правци без смирувања (со вештачки закривувања), како и ограничена прегледност, се поврзуваат со повисок ризик на ниво на делници. Состојбата и староста на коловозот се особено изразени во врнежливи услови, каде што пролизгувањето ја засилува можноста од излегување. Дополнително на ова, присуството на небезбедни странични елементи (банкини, риголи, берми, канавки) ја зголемува изложеноста на ризик кон судир со неподвижни препреки. Овие резултати укажуваат на мерки насочени кон геометриски корекции, подобрување на сообраќајната сигнализација, целно одржување и расчистување на страничните појаси.

Од аспект на сообраќајните карактеристики, кога обемот на сообраќај и брзината се во прираст, незгодите се зголемуваат рамномерно. Поголем удел на тешки возила ја намалува просечната брзина, но судирите со нив се со потешки последици. На крстосници со мал сообраќај и изразени врнежи од дожд, почесто се грешат проценката на растојанието и времето за премин. Затоа се потребни прилагодени ограничувања на брзина, појасно обележување и добро одводнување.

Временските услови и околината дополнително го поттикнуваат ризикот од сообраќајни незгоди. Дождот ја намалува прегледноста и коефициентот на триење, па сопирањето се продолжува и полесно настанува пролизгување. Снегот делува во две насоки, го намалува триењето како и дождот, а страничниот насобран снег ја стеснува оперативната ширина и ги прикрива рабовите на лентите, што ја отежнува ориентацијата. Мразот, исто така, го намалува триењето и ја зголемува веројатноста за губење контрола. Овие влијанија се поизразени на делници со неповолна геометрија и деградиран коловоз, што упатува на комбинирани инженерски и оперативни мерки и навремено планирање.

Просторно гледано, рангирањето по делници е поедноставно за примена, бидејќи не се врзува за поединечни вредности. Наместо да се прикажуваат конкретни податоци, се применуваат граници што ги класифицираат вредностите во нивоа, па се дефинираат рангови на безбедноста по делници. Мапите со ризик и приказ на индикатори, заедно со редовно следење по делници, овозможуваат јасни приоритети и транспарентно планирање, од забележан проблем до спроведена мерка.

Податоците покажуваат кои фактори најмногу влијаат на ризикот по делница. Најчесто тоа се должината на делницата, ПГДС, радиусот на хоризонтални кривини, надолжниот наклон, староста на коловозот, периодот во денот и врнежите. Кога овие влијанија делуваат заедно или се комбинираат со други фактори, ризикот може брзо да порасне на исклучително високо ниво. Затоа насоките за планирање и одржување треба да ги земаат предвид и поединечните и здружените ефекти.

Практичен редослед за моделирање предвидува прво да се избере модел што поднесува големи разлики меѓу делници и многу нули во податоците, при што се проверува дали има делници без регистрирани незгоди и се дозволуваат нелинеарни односи за одредени податоци. Потоа, резултатите се споредуваат со алтернативни модели за да се провери доследноста на заклучоците. Стабилноста се оценува со вкрстена валидација и со податоци од друг период. На крај, просторното насочување се потврдува со техники што ги ублажуваат случајните скокови, со проверка на друга мрежа за потврда на применливоста.

Сумирано, ризикот расте кога неповолна геометрија и деградиран коловоз се спојуваат со голем обем на сообраќај и неповолни временски услови. Во пракса најефективно подобрување се прави преку корекција на геометријата на слабите делници, управување со брзини, јасна сигнализација и насочено одржување со приоритет во врнежливи периоди.

4. ОБРАЗЛОЖЕНИЕ НА РАБОТНИТЕ ХИПОТЕЗИ И ТЕЗИ

Во ова истражување, хипотезите и тезите имаат цел да објаснат и да утврдат начин за одредување на безбедни делници. Претходните истражувања покажуваат дека ризикот од сообраќајни незгоди е поврзан со повеќе фактори, како геометриски карактеристики на патиштата, прегледноста, интензитетот на сообраќајот и условите од околината, потоа е потребно нумеричко проучување на секој од нив [88]. Избраните хипотези и тези овозможуваат детално да се испитаат главните фактори што го зголемуваат ризикот и да се развие метод за негова нумеричка проценка по делница врз основа на тие фактори. Хипотезите се основни претпоставки за насоката и големината на влијанието на факторите, а тезите ја насочуваат нивната практична примена за креирање препораки и конкретни интервенции. Примената на статистички техники и методи од машинско учење овозможува класификација на делници на побезбедни и поризични, обработка на голем број податоци и издвојување на критичните делници врз основа на резултатите и категориите на ризик [89].

4.1 Главна теза

Главната теза наведува дека, анализата со машинско учење врз јасно дефинирани показатели (параметри) овозможува прецизна проценка и рангирање на нивото на ризик на патните делници во Македонија. Пристапот го користи машинското учење како алатка за откривање врски меѓу параметрите и нивото на безбедност на начин што останува разбирлив за практична примена, а оттука резултатите се насочени кон точност и објаснивост. Анализата има за цел, да покаже кои фактори најмногу придонесуваат, во која насока делуваат и колкаво е нивното дејство, како и да обезбеди картографски прикази што помагаат при дефинирање на приоритени мерки на делници со повисок ризик. Оваа теза служи како рамка која подоцна се проверува врз збир на податоци со стандардни чекори на валидација.

4.2 Главни хипотези

Главните хипотези ја поставуваат основата за процена на нивото на безбедност на патните делници со користење на мерливи инженерски параметри и постапки од машинско учење. Тие се формулирани општо и се проверуваат на независен дел од податоците, за да се добијат резултати што се разбирливи и применливи, а тие се:

- **Интегрирани параметри.** Комбинирањето на патни, сообраќајни и параметри од околина, овозможува објективна процена на безбедноста по делници. На овој начин, споредбата меѓу делниците е полесна и помалку зависна од локални услови или од отстапувања во податоците. На крајот се добива појасна слика каде ризикот е повисок и каде прво треба да се делува.
- **Фактори што може да се корегираат.** Интервенциите во патни и сообраќајни карактеристики најчесто носат поголемо намалување на ризикот од условите од околината кои не можат да се контролираат. Ова значи дека мерките насочени кон овие области имаат поголема практична корист и полесно се планираат според расположливите ресурси.

- **Сложени врски.** Влијанието на параметрите не расте рамномерно и често не делува заедно. Методите на машинско учење полесно ги препознаваат овие врски отколку едноставни линеарни модели. Со тоа се добива попрецизна процена на ризикот и поразбирливо толкување на придонесот на секој параметар.
- **Стабилност на резултатите.** Рангирањата на параметрите и процените на ризик треба да останат слични и при умерени промени на подесувањата или поделбите на податоците. Кога е присутна ваква стабилност, резултатите се соодветни за сигурна примена во реални услови.
- **Оперативна применливост.** Картографското прикажување и рангирање според ниво на ризик, помагаат да се одредат приоритетни делници кога ресурсите се ограничени. Ваквиот приказ ја олеснува координацијата меѓу надлежни служби и планирањето на фазни интервенции, што го забрзува преминот од анализа кон конкретни мерки.

Овие хипотези овозможуваат прегледен пристап за процена и управување со ризикот на делниците, поставуваат конкретни услови за анализа и насочуваат кон резултати што директно се користат во практична примена.

4.3 Работни хипотези

Работните хипотези се поставени за процена на нивото на безбедност на патните делници врз основа на параметри од патот, сообраќајот и околината, при што како работна претпоставка се смета дека неповолни вредности во овие параметри ги зголемуваат ризиците по делница. Тврдењата се проверуваат на независен дел од збир на податоци, со исти мерни правила на сите делници, како на пример:

- **Патни фактори:** ја одредуваат поврзаноста меѓу геометријата на трасата, прегледноста и состојбата на коловозот со одводнување. Овие елементи заедно го обликуваат нивото на ризик на делниците и овозможуваат споредливи процени на влијанијата низ целата мрежа.
- **Сообраќајни фактори:** објаснуваат како обемот на сообраќајот, брзината и организацијата на раскрсници и патни јазли влијаат на ризикот. Овие фактори ја покажуваат динамиката на движењето и помагаат подобро да се види колку е ризична секоја делница.
- **Фактори од околина:** ги опфаќаат временските услови, релјефот и појавите на самата локација. Тие делуваат како надворешни услови што ја менуваат безбедноста и го зајакнуваат дејството на патните и сообраќајните фактори.

Овој пристап овозможува да се утврди како секоја категорија придонесува и во која насока делува. Добиените мерења по делница се доволни за сигурна класификација и за избор на следни чекори во реални услови.

4.4 Критериуми за проверка и мерни показатели

Ова поглавје објаснува како се проверуваат резултатите и кои показатели се користат, за да бидат споредливи и применливи во реални услови. Правилата се применуваат на

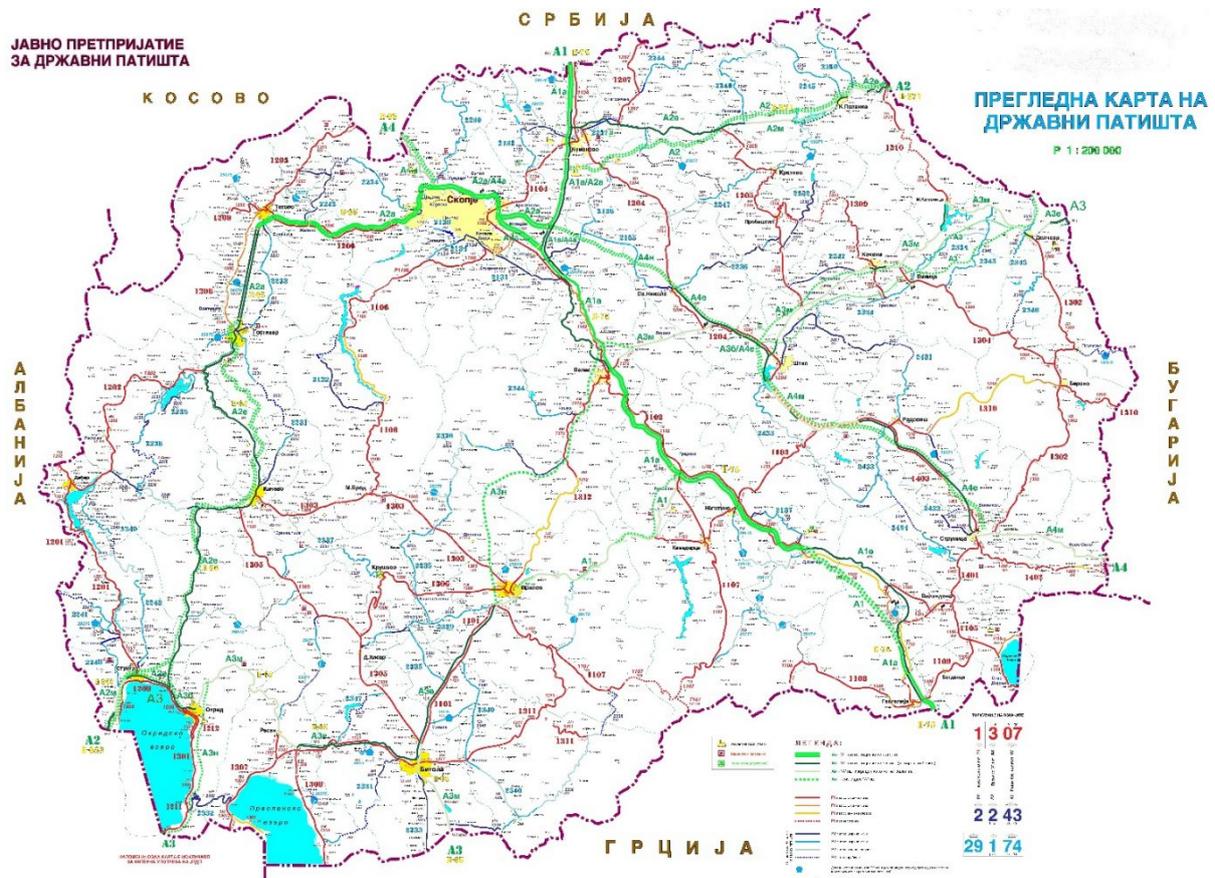
поставените хипотези и на главната теза, за заклучоците да бидат потврдени со податоци корисни за пракса.

Проверката се изведува на независен тест-дел со поделба 80/20. Сите подготовки и подесувања се прават само на тренинг-делот, за да нема протекување на информации. Точноста се оценува со R^2 , MAE и RMSE на тест-делот. Дополнително, се анализира распределбата на грешките по групи делници со слични својства, за да се утврдат можни систематски отстапувања поврзани со тип на пат или со конкретни услови. Стабилноста се проценува со повеќе почетни поделби и минимални промени на подесувањата. Се следи дали рангирањата на параметрите и вредностите на мерните показатели остануваат во сличен опсег при различни поделби на податоците. Влијанието на секој параметар се оценува со SHAP (Shapley Additive Explanations, метод што покажува колку секој параметар придонесува во пораст или пад на резултатот) и пермутациски метод (постапка за проверка на важноста на секој параметар преку менување на неговите вредности). За секој параметар се прикажува насока и приближна јачина на влијанието, усогласени со практичната логика. Оперативната употреба се оценува преку картографски прикази. На карта се прикажува рангирањето на делниците според проценетиот ризик, се поставува праг за приоритет и се проверува делот од вкупниот ризик што го опфаќа приоритетната група. Контролата на квалитет опфаќа проверка на усогласеност на податоците, третман на вредности кои недостасуваат (missing values) и екстреми, како и проверка на просторни несовпаѓања. Овие насоки овозможуваат резултатите лесно да се споредуваат, да останат стабилни и да можат да се користат во пракса. Проценката на нивото на безбедност се заснова на поставени тези и работни хипотези, поддржани со критериуми за проверка и тестирање.

5. ОПИС НА ПОДАТОЦИТЕ И ИСТРАЖУВАЧКИОТ ОПФАТ

5.1 Општи податоци за патната мрежа

Патната мрежа во Северна Македонија опфаќа 14 475 km патишта, поделени на магистрални, регионални и локални. Од нив 897 km се магистрални, 3 790 km се регионални, а останатите 9 788 km се локални патишта. Од магистралните патишта 330 km се автопати, 90 km експресни патишта и 477 km патишта со две ленти [90, 91]. Според националната класификација околу 558 km од магистралните патишта се дел од европската патна мрежа (Е-патишта), од кои 239 km се автопат и 319 km пат со две ленти. Во поглед на автопатските решенија патната мрежа во Македонија не е на нивото на европските држави (ЕУ), бидејќи таму должината на автопати изнесува околу 28 km/1000 km², додека кај нас е околу 14 km/1000 km².



Слика 5.1 Прегледна карта на државни патишта (ЈПДП, 2024)

На слика 5.1 прикажан е распоредот на државните патишта, што овозможува брз преглед на главните коридори и нивната меѓусебна поврзаност.

Главната структура на мрежата ја даваат четири магистрални коридори (A1, A2, A3 и A4), кои се главните артерии на сообраќајот во Македонија. Преку нив, транспортот се пренесува од делници со најголем капацитет кон делници со помал дневен проток.

Подетално, главните магистрални коридори се:

- **Магистрален пат А1:** Граница со Србија (ГП Табановце) – Куманово – Миладиновци – Велес – Неготино – Демир Капија – Гевгелија – граница со Р. Грција (ГП Богородица). Овој патен правец е дел од европскиот коридор (Е-75) во најголем дел

автопатски профил, со варијации по локации. Вкупната должина на овој магистрален коридор изнесува 177 km.

- **Крак на магистрален пат А1:** Патен јазол Градско - крстосница Прилеп. Геометрискиот напречен профил на овој патен правец главно е со две ленти, со дополнителна лента за бавни возила на пострмни делови. Должината на овој коридор изнесува 50 km.
- **Магистрален пат А2:** Граница со Бугарија (ГП Деве Баир) - Крива Паланка - Страцин - Романовце (Куманово) - Миладиновци - Обиколница Скопје - Тетово - Гостивар - Кичево - Требениште - Струга - граница со Албанија (ГП Кафасан). Овој патен правец опфаќа делови од три европски коридори (E65=149 km, E852=20 km, E871=74 km) и има делови со автопат и делови со две сообраќајни ленти. Автопатски профил има на делниците Гостивар - Тетово и Тетово - Скопје, како и на делницата која е заедничка со Коридорот 10. Останатите делници се со две ленти, и со лента за бавни возила, каде е тоа неопходно. Должината на овој магистрален коридор изнесува 308 km.
- **Магистрален пат А3:** Крстосница Требениште (врска со А2) - крстосница Подмоље - Охрид - Косел - Ресен - Битола - Прилеп - Велес - Штип - Кочани - Делчево - граница со Бугарија (ГП Рамна Нива), (E65=103 km), делница: Битола - крстосница Кукуречани - граница со Грција (ГП Медитлија) (E65=21 km) и делница: Косел (врска со А3) - Охрид - граница со Р. Албанија (ГП Љубаниште). По целата должина се издвојуваат две делници со експресен пат: Велес - Кадрифаково и Штип - Кочани, а останатите се патишта со две ленти, со лента за бавни возила на пострмни места. Вкупната изградена должина на овој магистрален коридор изнесува 228 km (133 km европски патишта), а должината на останатиот неизграден дел (Велес - Прилеп, Требениште - Косел - граница со Албанија) ќе се прецизира со идната проектна документација.
- **Магистрален пат А4:** Граница со Косово (ГП Блаце) - крстосница Стенковец - Обиколница Скопје - Петровец - Миладиновци - Свети Николе - Штип - Радовиш - Струмица - граница со Бугарија (ГП Ново Село) (E65=14 km). Овој коридор има три типа напречни профили:
 - Автопат: Крстосница Стенковец - Обиколница Скопје - Петровец - Миладиновци - Свети Николе - Штип,
 - Експресен пат: Штип - Радовиш,
 - Пат со две сообраќајни ленти: Граница со Косово (ГП Блаце) - крстосница Стенковец и Радовиш - Струмица - граница со Бугарија (ГП Ново Село).

Вкупната должина на овој магистрален коридор изнесува 192 km (14 km европски патишта).

Магистралните и регионалните патишта се во надлежност на ЈПДП, а локалните патишта на локалната самоуправа (ЛС).

Одржувањето на магистралните и регионалните патишта го спроведува Јавното претпријатие за одржување и заштита на магистрални и регионални патишта Скопје (ЈП Македонијапат). Одржувањето се врши според финансискиот план на ЈПДП, а договорите се склучуваат секоја година во согласност со планот.

Оваа распределба и надлежност создава појасна слика за носечките правци и поддржува планирање на одржување и унапредување на капацитетите.

5.2 Технички карактеристики на јавните патишта

5.2.1 Авиопаати

Автопатиштата, кои се во употреба во моментот, имаат вкупна должина од 330 km и се дел од три различни магистрални коридори А1 (К-10), А2 (К-8) и А4. На овие делници се применуваат шест типизирани геометриски напречни профили, при што дозволената брзина варира од 50 km/h до 130 km/h.

• Стандарден профил

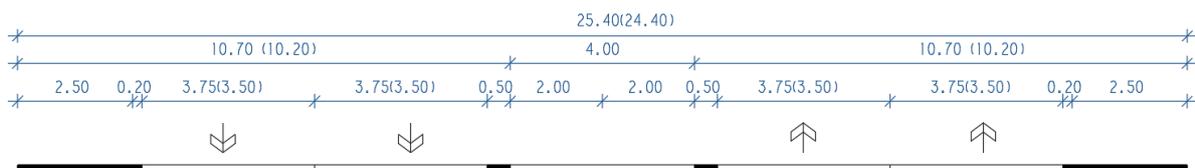
Овој профил дефиниран е со претходниот ЈУС стандард (сега МКС) и е типично решение за поголем дел од автопатиштата кај нас.

Главните карактеристики на овој профил се:

- Сообраќајни ленти	$4 \times 3.75(3.50) = 15.00 (14.00) \text{ m}$
- Рабни ленти (до лента за претекнување)	$2 \times 0.50 = 1.00 \text{ m}$
- Рабни ленти (пред лента за застанување)	$2 \times 0.20 = 0.40 \text{ m}$
- Ленти за застанување	$2 \times 2.50 = 5.00 \text{ m}$
- Средна разделна лента	min 4.00 m
- Вкупно:	25.40(24.40) m

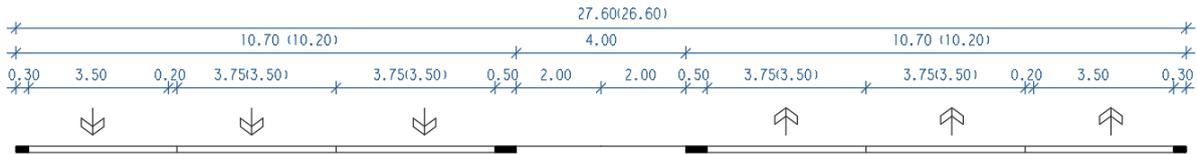
Профилот се протега на следниве делници:

- ГП Табановце – ПЈ Куманово,
- Миладиновци – Катланово,
- Велес – Демир Капија,
- Смоквица – ГП Богородица,
- Миладиновци – Штип,
- Миладиновци – Хиподром – Тетово,
- Хиподром – Петровец.



Слика 5.2 Стандарден профил на авиопаати

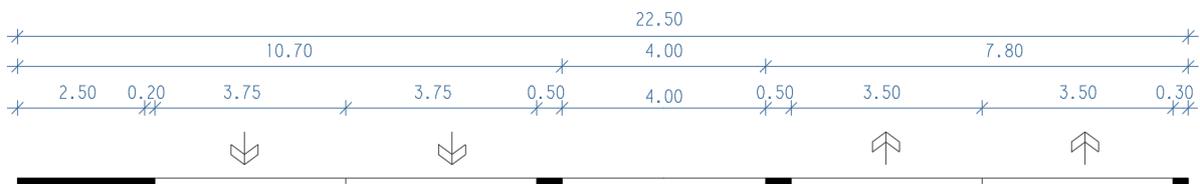
Експлоатационата брзина на овие делници варира од 80 km/h до 130 km/h. Во зависност од локалните услови (лента за бавни возила или ленти за приклучување и исклучување), профилот се модифицира на кратки делови со проширување на лентата за застанување на 3.50 m и рабна лента од 0.30 m.



Слика 5.3 Стандарден профил на авиоајџ со дојолништелна возна лениа

- **Стандарден профил без дојолништелна лениа за застанување**

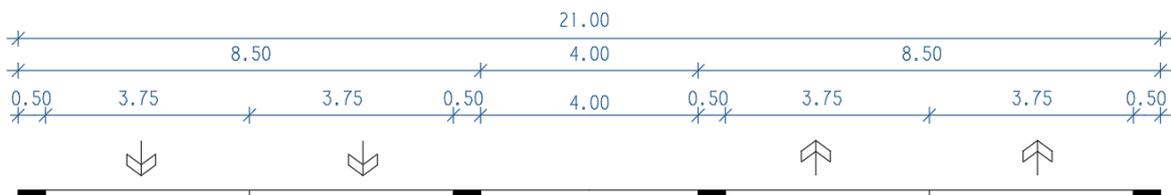
Карактеристичниот пресек се однесува на делницата од ПЈ Окта (рафинерија Окта) до ПЈ Куманово и се јавува на растојание од приближно 20 km. Отстапките од типизираното решение важат само за насока Југ-Север, каде на десната коловозна лента се применети помали ширини и не е изведена лента за застанување. Принудното застанување овозможено е со дополнително изведени ниши и одморалишта, а дозволената брзина изнесува 100–120 km/h.



Слика 5.4 Стандарден профил на авиоајџ без дојолништелна лениа за застанување

- **Напречен профил без дојолништелни лениа за застанување од две страни**

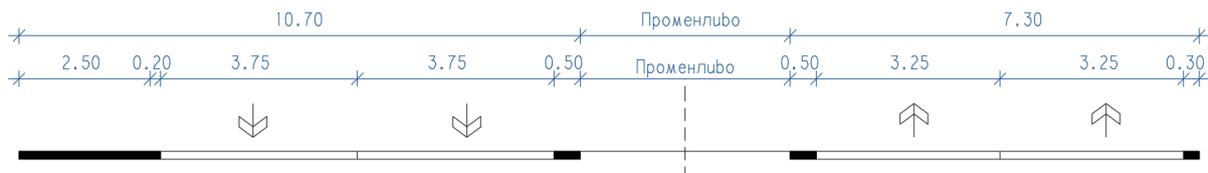
Овој геометриски напречен профил е применет по целата должина на делницата од ПЈ Миладиновци до ПЈ Петровец. Отстапките од стандардниот автопатски профил, се во тоа што, се изоставени лентите за застанување во двете насоки за возење. Дозволената брзина на оваа делница е 120 km/h на отворените делници и 100 km/h на патните јазли.



Слика 5.5 Профил на авиоајџ без дојолништелни лениа за застанување

- **Стандарден авиоајџски профил во една насока и задржан магистрален профил во сировивна насока**

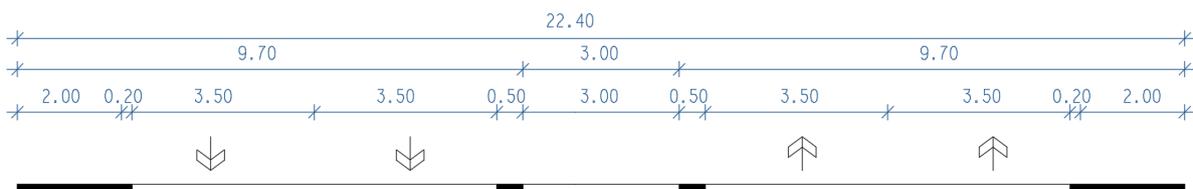
Овој профил се наоѓа на делницата од Катланово до Велес. Отстапките на типизираното решение се однесуваат само за насока Север-Југ, каде од десната страна е задржан постојниот коловоз од поранешниот регионален/магистрален пат Велес – Скопје. Принудното застанување на оваа страна овозможено е со дополнително изведени ниши и одморалишта. Брзината на оваа делница варира од 50 km/h до 80 km/h.



Слика 5.6 Стандарден профил во една насока и магистрален профил во друга насока

- **Профил со општа кај разделна лента и лента за застанување**

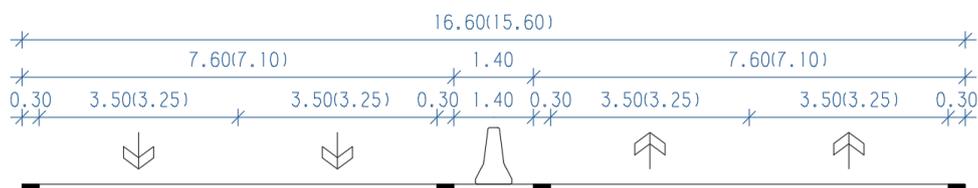
Изведениот напречен пресек се наоѓа на делницата од Демир Капија до Смоквица. Измените од типските карактеристики се однесуваат на ширините на средната разделна лента (3.0 m) и лентата за застанување (2.0 m). Дозволената брзина варира од 80 km/h до 120 km/h.



Слика 5.7 Профил со стеснета разделна лента и лента за застанување

- **Редуциран автострајски профил**

Овој профил се наоѓа на делницата од Тетово до Гостивар. Отстапките од стандардниот профил се однесуваат на значително намалена ширина на средната разделна лента, тесни рабни ленти кај лентите за претекување и изоставена лента за застанување. Брзината на оваа делница е ограничена на 80 km/h. Според најновиот Закон за јавни патишта и Законот за безбедност на патиштата, овој напречен профил се класифицира како експресен пат.



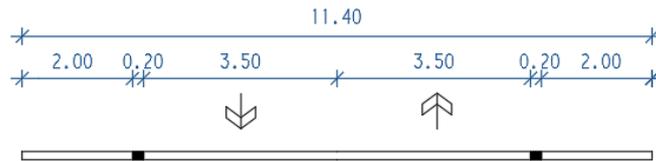
Слика 5.8 Редуциран автострајски профил

5.2.2 Експресни патишта

Геометрискиот напречен профил на експресните патишта вклучува по една сообраќајна лента во насока и лента за застанување со вклучена рабна лента. Вкупната изградена должина на експресни патишта во моментот е 90 km и се состои од следниве делници:

- Велес – Кадрифаково А3 (Р1312)
- Штип – Радовиш А4
- Три Чешми – Кочани А3

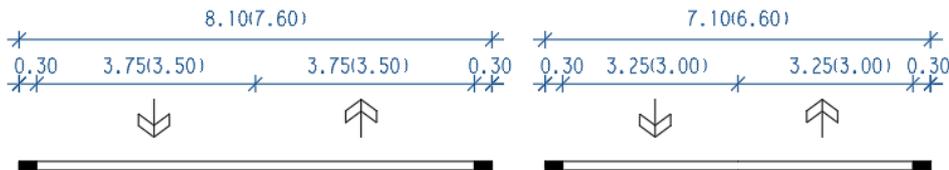
Дозволената брзина на овие делници е во опсег од 80 km/h до 110 km/h.



Слика 5.9 Експресен пати

5.2.3 Магистрални патишта со две сообраќајни ленти

Останатиот дел од магистралните патишта има геометриски напречен профил со по една сообраќајна лента во насока и нивната должина изнесува 330 km. Во поглед на брзинските карактеристики, грубо се издвојуваат делници кои овозможуваат брзини од 80 km/h до 100 km/h и делници каде брзините се во граница од 50 до 80 km/h.



Слика 5.10 Пати со две сообраќајни ленти (лево $V=80-100\text{km/h}$, десно $V=50-80\text{km/h}$)

Генерално, се издвојуваат две делници каде е дозволена брзина од 100 km/h, и тоа:

- Прилеп – Битола (A3)
- Радовиш – Струмица (A4)

Останатите делници од магистралната мрежа се класифицираат како подгрупа со помали возно-динамички можности и брзини под 80 km/h. Според податоци на Државниот завод за статистика (2022), приближно 75 km магистрални патишта се водат како земјани или идни (непробиени) делници.

5.3 Начин на собирање на податоци

Во ова истражување, податоците за карактеристиките на патната мрежа се собрани преку соработка со релевантни институции и преку теренски снимања. Ова овозможи сеопфатна анализа и систематско дефинирање на параметрите кои ја опишуваат состојбата на патиштата.

• Податоци од Јавното истражување за државни патишта (ЈИДП)

ЈИДП располага со база на податоци и веб ГИС платформа преку која се ажурираат собраните податоци. Од нивната база преземен е ГИС модел, кој служи како подлога за ова истражување. Тие податоци се:

- точна должина на секоја делница,
- ограничување на брзината,
- коефициент на отпорот при триење (Skid Resistance),
- колотрази (Rutting),
- надолжна рамност (IRI – International Roughness Index),
- број, локација и состојба на мостови,
- висински коти,
- датум на поставување на вертикалната сигнализација,
- просечен годишен дневен сообраќај (ПГДС).

- **Податоци од Управата за хидрометеоролошки работи (УХМР)**

Управата за хидрометеоролошки работи води евиденција и анализира податоци за климатските и временските услови кои можат да влијаат на состојбата на патиштата, нивната проодност и возната динамика. Овие податоци се користат за проценка на надворешните фактори што можат да влијаат и на одржливоста и функционалноста на патната инфраструктура, како и на условите што придонесуваат за зголемен ризик од сообраќајни незгоди. Од нивната архива, добиени се податоци за:

- просечните годишни врнежи од дожд,
- просечните максимални годишни дебелини на снежна покривка,
- максималните, минималните и средните температури.

- **Податоци од Јавното претпријатие за одржување и заштита на магистрални и регионални патишта (ЈП Македонија пат)**

Јавното претпријатие за одржување и заштита на магистрални и регионални патишта, води евиденција за редовното и периодичното одржување на патиштата, од каде се добиени следните податоци:

- видот и фреквенцијата на одржување спроведено на секоја делница,
- временскиот период за обнова на хоризонталната сигнализација.

- **Податоци од Министерството за внатрешни работи (МВР)**

Министерството за внатрешни работи прибира и обработува податоци за сообраќајните незгоди врз основа на службени записници и теренски истраги, кои се искористени за:

- бројот на сообраќајни незгоди,
- сериозноста на сообраќајните незгоди,
- локацијата на сообраќајните незгоди.

- **Видео снимки со ГПС камера (DOD LS460W, IS420W)**

Снимањето овозможува 3Д визуелизација на конкретната траса со координати (OpenStreetMap) и прикажување на моменталната брзина на секоја делница (минимална, максимална и просечна). Овие снимки се искористени за:

- дефинирање на бројот на контранаклоните по делница,
- утврдување на индексот на површинска оштетеност (PCI – Pavement Condition Index),
- анализа на должината на застанување (SSD – Stopping Sight Distance).

5.4 Категоризација на податоци

Анализата на податоците што влијаат на зачестеноста на сообраќајните незгоди е клучна за подобро разбирање на функционирањето на патниот систем и за прецизно идентификување на факторите што придонесуваат за нивната појава. Истражувањето опфаќа широк спектар на податоци што ја отсликуваат патната мрежа, сообраќајните

услови и влијанија од околина, со цел да се постави стабилна основа за натамошна обработка и анализа. Притоа, податоците се групираат во две основни категории:

- категоризација според динамика на податоците,
- категоризација според основните карактеристики.

Првата поделба ги разгледува податоците според нивната временска природа и динамика на промените, додека втората ги групира според нивните основни карактеристики и улоги во рамките на истражувањето.

5.4.1 Категоризација според динамиката на податоците

Анализата опфаќа податоци собрани во период од 10 години (2014–2023), при што тие се групирани според нивната динамика и значење:

- Променливи со значителни годишни варијации: податоците за обем на сообраќај и зачестеноста на сообраќајните незгоди покажуваат изразени промени од година во година и затоа се анализираат како годишни вредности.
- Параметри со мала процентуална годишна промена: климатските услови (температури, врнежи, снежни наноси) и состојбата на коловозот се обработуваат како просечни вредности за целиот десетгодишен период.
- Фиксни параметри: геометријата на патот и прегледноста остануваат непроменливи во времето, па затоа се разгледуваат без временска компонента.

5.4.2 Категоризација според основните карактеристики на податоците

За подетално структурирање, податоците се разгледуваат според нивните основни карактеристики. Оваа поделба ги објаснува природата и улогата на различните аспекти на патната мрежа, сообраќајот, околината и сообраќајните незгоди:

- Карактеристики на патот врз кои може директно да се влијае: геометриски параметри, брзина, состојба на коловозот и сигнализација.
- Карактеристики на сообраќајот врз кои директно не може да се влијае: просечен годишен дневен сообраќај (ПГДС).
- Карактеристики на околината врз кои не може да се влијае: климатски услови (температури, врнежи и снежни наноси).
- Податоци за сообраќајните незгоди: зачестеност на незгодите и нивната просторна распределба по делници.

5.4.3 Опис на податоците и нивна обработка

Комбинирањето на двете категоризации овозможува целосен и систематски пристап за обработката на податоците. Динамичната поделба ги посочува временските промени, додека категоризацијата според карактеристиките ја појаснува улогата и влијанието на секој фактор. Податоците се обработени со примена на ГИС-алатки и статистички методи, со мапирање и согледување на клучните просторни и временски промени.

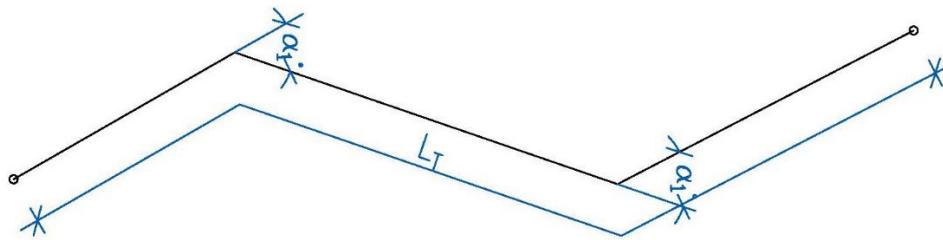
- **Просечна ограничена брзина (LIMIT)**

Просечната ограничена брзина е средна вредност на реалните теренски ограничувања по должина на анализираните делници, изразена во km/h. Вредностите се земаат од

податоците на ЈПДП поврзани со вертикална сигнализација. Кога на една делница има повеќе ограничувања на брзината, се пресметува просек по должина за да се рефлектира уделот на секое ограничување во трасата. Просечната ограничена брзина служи како референтна мерка за дозволеното движење во однос на геометријата и условите на трасата и влегува во пресметки на бочниот потисок, што е директно поврзано со стабилноста при движење на кривини.

- **Кривинска карактеристика (Kr.Kar.)**

Кривинската карактеристика претставува однос меѓу збирот на сите прекршни агли изразени во степени и вкупната должина на трасата изразена во километри ($^{\circ}/\text{km}$). Овој параметар ја оценува хоризонталната геометрија на трасата [92].



Слика 5.11 Геометриска интерпретација на кривинска карактеристика

Формулата за пресметување е:

$$K_k = \frac{\sum \alpha}{L_T} \quad 5.1$$

α - прекршен агол (степени)

L_T - вкупна должина на трасата (km)

Поголемата вредност укажува на траса со остри кривини и нагласени завои, додека помалата вредност дефинира траса со поблаги кривини, што придонесува за зголемена безбедност и удобност при возење.

- **Среден хоризонтален радиус (Ave.Rad)**

Средниот хоризонтален радиус е важен параметар кој го опишува квалитетот на геометриската закривеност на патот преку анализа на сите хоризонтални кривини на делницата. Тој се пресметува како средна вредност на радиусите на сите кривини, што овозможува проценка на хоризонталната закривеност на патот.

Формулата за пресметка е:

$$R_a = \frac{\sum_{i=1}^n R_i}{n} \quad 5.2$$

R_a - среден хоризонтален радиус (m)

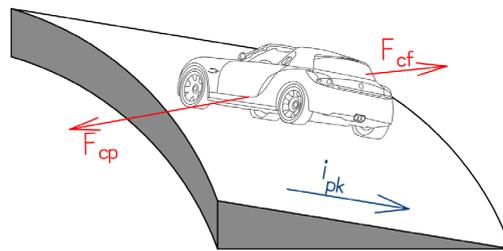
R_i - радиусот на i -тата хоризонтална кривина (m)

n - вкупен број на хоризонтални кривини на делницата

- **Коефициент на напречен наклон (KON.NAK.)**

Контранаклоните, односно наклони со спротивна насоченост на напречниот пад, се делови од патот каде напречниот наклон е насочен кон надворешната, наместо кон

внатрешната страна на кривината. Иако, според постојните регулативи овие контранаклони се дозволени, нивното присуство може да влијае на управувањето и стабилноста на возилата.



Слика 5.12 Наклон со спротивна насоченост на отклонувањето

Психолошки, контранаклоните можат да предизвикаат чувство на несигурност и замор кај возачите, што може да доведе до намалена концентрација и зголемен ризик од сообраќајни незгоди. Ова е особено изразено кај неискусни или повозрасни возачи.

Од физички аспект, контранаклоните влијаат на динамиката на возилото при движење низ хоризонтални кривини. Центрифугалната сила, која дејствува кон надворешноста на кривината, во комбинација со контранаклонот на патот, резултира со зголемен бочен потисок врз возилото. Ова може да доведе до намалена стабилност и контрола, особено при поголеми брзини или неповолни временски услови.

За квантитативна оценка на присуството на контранаклони на одредена делница, се користи коефициент кој се пресметува како однос на збирот на сите попречни наклони со спротивна насоченост и вкупната должина на делницата:

$$K = \frac{\sum_{i=1}^n n_i}{L} \quad 5.3$$

K - коефициент на контранаклони

n_i - број на попречни наклони со спротивна насоченост

L - вкупна должина на делницата (m)

Овој коефициент овозможува објективна проценка на присуството и интензитетот на контранаклоните на анализираната делница.

- **Просечен надолжен наклон (Ave_Inc)**

Просечните надолжни наклони за секоја делница се определени преку ГИС анализа со помош на софтверот ArcMap. Методот се базира на пресметка на наклоните меѓу последователни темиња (врски помеѓу секое прекршување) по должина на патната траса, користејќи дигитален теренски модел (DTM) за добивање на висинските точки. За да се пресметаат наклоните за дадена делница, потребни се најмалку три висински точки, бидејќи секој сегмент се формира меѓу две последователни точки. Ако делницата има n висински точки, тогаш можат да се дефинираат n-1 сегменти.

За секој сегмент меѓу две точки i и i+1, надолжниот наклон i_{ni} се пресметува како однос на висинската разлика и хоризонталното растојание помеѓу нив:

$$i_{ni} = \frac{Z_{i+1} - Z_i}{D_{i,i+1}} \times 100 \quad 5.4$$

i_{ni} – наклонот меѓу точките i и $i+1$ (%)

Z_i и Z_{i+1} – висински точки (mNv)

$D_{i,i+1}$ – хоризонталното растојание меѓу точките (m)

Откако се пресметуваат вредностите за сите сегменти на одредена делница, просечниот надолжен наклон i_{nave} се добива како аритметичка средина на сите сегменти со различен надолжен наклон:

$$i_{nave} = \frac{\sum_{i=1}^{n-1} i_{ni}}{n-1} \quad 5.5$$

i_{nave} – просечен надолжен наклон за дадена делница (%)

$n-1$ – број на сегменти по должина на анализиран пат

Овој метод овозможува прецизна вертикална анализа на сите делници, како и идентификација на критични сегменти како оние со стрмни наклони, кои можат да влијаат на стабилноста и кочењето, така и на оние со наклони блиски до хоризонтала, каде што постои зголемен ризик од аквапланинг или појава на голомразица во зимски услови.

- **Средна надморска висина (Ave_Height)**

Средната надморска висина на секоја делница се определува преку ГИС анализа, користејќи ги висинските точки (темиња) по целата должина на патот. Процедурата е идентична со методот за пресметка на надолжниот наклон, каде од дигиталниот теренски модел (DTM) се извлекуваат висинските вредности за секое теме, а потоа се одредува нивната средна вредност.

Средната надморска висина H_{avg} за дадена делница, се пресметува како:

$$H_{avg} = \frac{\sum_{i=1}^n H_i}{n} \quad 5.6$$

H_{avg} – средна надморска висина на делницата (mNv)

H_i – висинска точка на теме i (mNv)

n – вкупен број на темиња на делницата

Оваа вредност претставува референтна точка за секоја делница и помага во утврдувањето на области подложни на екстремни температури. Делниците со повисока средна надморска висина имаат поголема изложеност на ниски температури и ризик од подмрзнување во зимски услови, додека пак делниците на пониски висини можат да бидат подложни на повисоки температури и загревање на коловозот во летен период.

- **Бочен потисок (Lat_Force)**

Бочниот потисок претставува странична сила што делува врз возилата при нивно движење низ хоризонтални кривини. Овој потисок е резултат на центрифугалната сила што се јавува како последица на закривеноста на патот и брзината на движење[93]. Високите вредности на бочниот потисок можат да доведат до намалена стабилност на возилата, зголемена динамичка оптовареност врз тежиштето и зголемен ризик од

пролизгување или превртување, особено кај возила со високо тежиште, како што се товарни возила и автобуси.

Бочниот потисок L_f (N/kg) се пресметува врз основа на претходно определените параметри: просечната ограничена брзина V_g (km/h) и средниот хоризонтален радиус R_a (m), според следната релација:

$$L_f = \frac{V^2}{12.96 \cdot R} \quad 5.7$$

L_f – бочен потисок (N/kg)

V_g – просечна ограничена брзина (km/h)

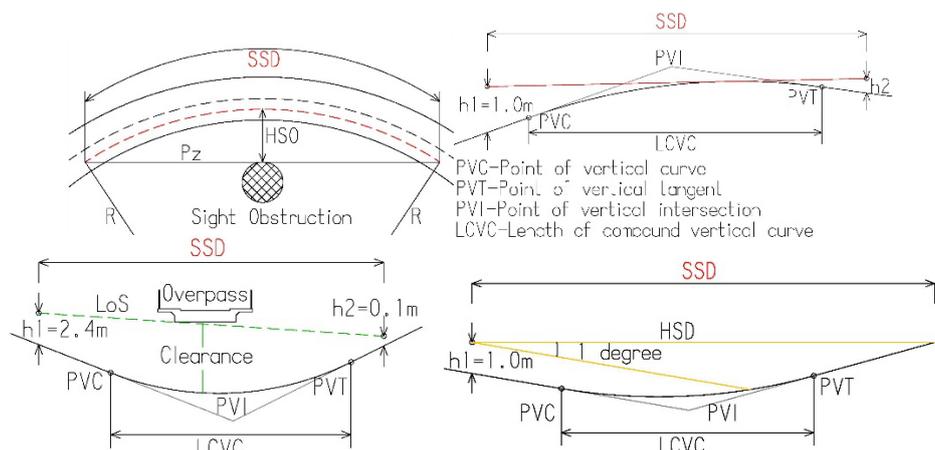
R_a – среден радиус на кривини (m)

Висок бочен потисок укажува на намалена стабилност, особено при повисоки брзини и остри кривини.

- **Прејледност на застанување (SSD)**

Анализата за прејледност на застанување спроведена е во хоризонтален и вертикален смисол, со цел да се оцени дали возачот има доволна прејледност за безбедно запирање при неочекувана пречка на патот. Пресметката на прејледноста е изведена според дефинирана линија на прејледност, утврдена во Правилникот за проектирање на патишта[94].

Прејледноста на застанување анализирана е преку геометриска оцена на линијата на прејледност, кој треба да овозможи забележување на препрека со висина од 10 cm. Доколку во пресметката се утврди дека возачот не може да ја забележи пречката на потребното растојание, тогаш не е задоволена хоризонталната или вертикалната прејледност.



Слика 5.13 Карактеристични случаи на проверка на прејледност на застанување

Во надолжниот профил, должината на застанување овозможена со прејледноста треба да одговара на линијата на прејледност. Во ситуација, должината на застанување може да отстапи поради геометријата и пречките во страничното видно поле. Прејледноста на застанување P_z се определува како збир на должината потребна за реакција на возачот и должината на форсирано кочење:

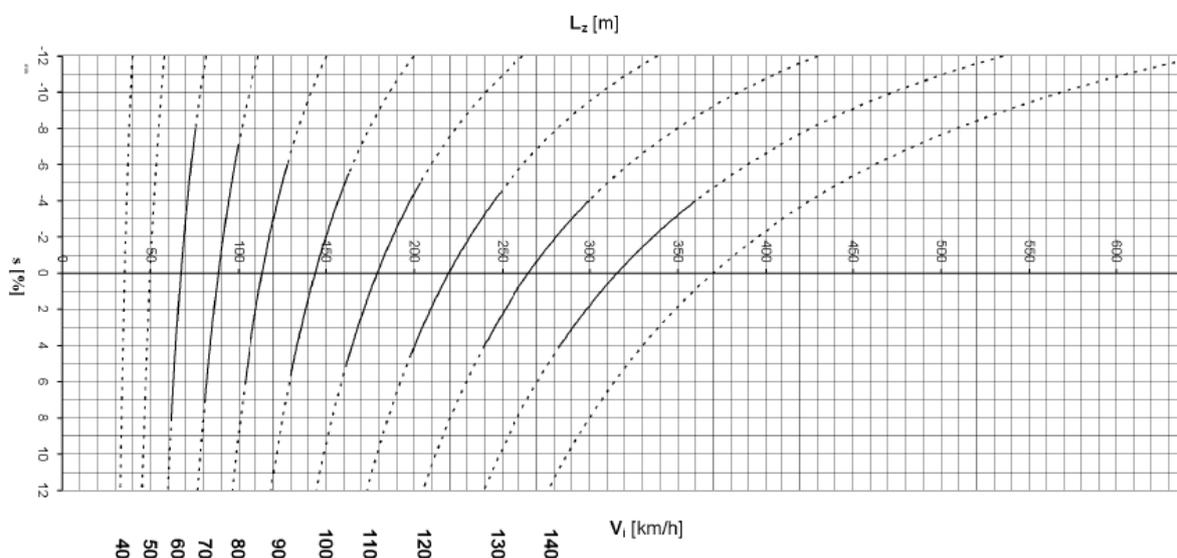
$$P_z = L_r + \Delta L \quad 5.8$$

P_z – потребна должина за застанување (m)

L_z – должина на застанување (m)

ΔL – безбедно растојание до пречката (m)

Должината на застанување L_z се пресметува како функција од брзината, коефициентот на триење и надолжниот наклон на патот.



Слика 5.14 Должина на застанување за најбизитна од техничка група А (МТВ, 2009)

За секоја анализирана делница, направено е броење на непрегледните локации во ситуација и надолжен профил, при што е пресметана средна вредност во однос на вкупната должина на делницата.

Коефициентите на хоризонтална и вертикална прегледност се определуваат според следните изрази:

$$K_h = \frac{\sum n_h}{L} \quad 5.9$$

$$K_v = \frac{\sum n_v}{L} \quad 5.10$$

K_h – коефициент на хоризонтална прегледност (n/km)

K_v – коефициент на вертикална прегледност (n/km)

$\sum n_h$ – сума на сите непрегледни локации во хоризонтална проекција

$\sum n_v$ – сума на сите непрегледни локации во вертикална проекција

L – вкупна должина на делницата (km)

Оваа методологија овозможува квантитативна проценка на прегледноста на патштата, при што високите вредности на коефициентот укажуваат на делници со значително намалена прегледност, додека ниските вредности укажуваат на делници со задоволителна прегледност.

- **Рамноста на коловозот (IRI/RDB_RUT)**

Рамноста на коловозната површина се оценува како разлика меѓу проектниот и реалниот надолжен/напречен профил на коловозната површина на зададена делница [95].

Надолжната нерамност се однесува на вертикалните отстапувања на коловозната површина долж правецот на патот. Таа се мери преку Меѓународниот индекс на рамност (IRI), кој претставува стандардизиран показател за оценка на рамноста на коловозот и се изразува во метри на километар (m/km). IRI се пресметува врз основа на мерења на надолжниот профил на патот и ја одразува кумулативната вертикална дислокација на возилото по единица растојание.

Напречната нерамност, позната како колотрази, се однесува на надолжните вдлабнатини што се формираат на коловозната површина во однос на ширината на патот, обично во траекториите на тркалата на возилата. Колотразите се мерат како длабочина на вдлабнатините и се изразуваат во милиметри (mm). Тие можат да влијаат на управувањето со возилото, одводнувањето, а со тоа и на безбедноста при возење.

Податоците за IRI и колотразите се добиваат преку снимање со Dynatest Multi Function Vehicle (MFV), кое користи ласерски сензори и акцелерометри за детекција на вертикалните и напречните отстапувања на коловозот. Ова возило овозможува мерење со висока прецизност, додека инерцијалните сензори и GPS модулите овозможуваат точна просторна геолокација на секоја снимка [96].



Слика 5.15 Мултифункционално возило Dynatest (Dynatest International, 2009)

Мерењата се изведени на приближно фиксен интервал од 50 m. Бидејќи, годишните промени на IRI и колотразите се мали, пресметани се средни вредности врз основа на десетгодишните податоци за секоја делница, што овозможува подобра репрезентација на долгорочните трендови на состојбата на патната површина.

- **Коефициент на триење (SKIDRES)**

Триењето на возната површина е клучен фактор за безбедноста, бидејќи директно ја дефинира способноста на возилото да одржи контрола при движење, особено во услови на влажна или нечиста коловозна површина. Овој феномен се јавува како резултат на интеракцијата меѓу пневматиците и коловозот, а неговата вредност зависи од карактеристиките на пневматикот, состојбата на коловозната површина и присуството на

вода или други нечистотии. Се изразува преку коефициентот на триење, кој се дефинира како односот меѓу силата на триење и нормалната сила што дејствува врз пневматикот [97].

Вредностите на коефициентот на триење можат значително да варираат зависно од текстурата на абечкиот слој, староста на асфалтот и интензитетот на сообраќајот. Кај нови или подобрени асфалтни површини, тие можат да достигнат до 0.7 – 0.8, што укажува на висок степен на адхезија и безбедност, додека вредности под 0.3 – 0.35 се сметаат за критични и значително го зголемуваат ризикот од пролизгување, особено при влажни услови. Вообичаено, вредностите на триењето кај асфалтните коловози се движат во опсегот од 0.4 до 0.7, при што подобрувањето на текстурата на површината може да го одржи коефициентот на триење во безбедни граници.

Според податоците добиени од ЈПДП, вредностите на триењето за анализираните делници се движат во граници од 0.35 до 0.5. Овој интервал покажува дека на одредени делници триењето е на долната граница на прифатливите вредности, што може да биде знак за потреба од санација или подобрување на површинските карактеристики на возната површина.

- **Оцена на состојба на коловоз (PCI)**

Индексот за состојба на коловозот (PCI) претставува нумеричка оценка за вреднување на состојбата на возната површина врз основа на визуелна инспекција и мерење на оштетувањата. Овој индекс се изразува во поени (од 0 до 100), каде што повисоката вредност означува подобра состојба на коловозот, додека пониските вредности укажуваат на потреба од санација или реконструкција.

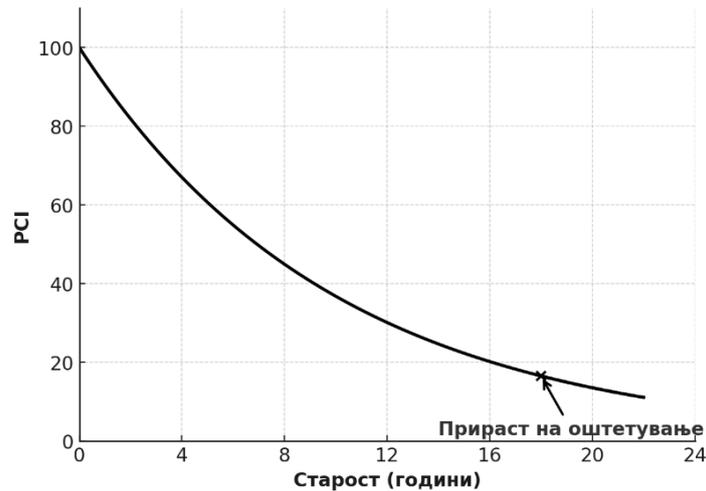
За определување на PCI на анализираните делници, применета е методологија што вклучува неколку клучни чекори. Првично е извршено видео снимање со GPS координати, при што секоја делница е документирана со геопросторни податоци. Ова снимање овозможува точна геолокација на сите оштетувања и нивно понатамошно анализирање. Понатаму, анализираните делници се поделени на пократки делници со должина од 500 до 1000 метри, со што се обезбедува попрецизна проценка [98]. За секоја од овие делници направена е визуелна проценка и приближно мерење на оштетувањата, при што се бележат типот, сериозноста и распространетоста на деформациите, како што се:

- пукнатини (надолжни, напречни, мрежести),
- деформации (колотрази, вдлабнатини, испакнатини),
- површинска деградација (абење на површината-полирање),
- проблеми со одводнувањето (задржување на вода, ерозија).

Секое оштетување се класифицира според намалена вредност (Deduct Value - DV), која се одбива од максималната вредност на PCI (100 поени). PCI се пресметува со следнава формула:

$$PCI=100-\Sigma DV$$

5.11



Слика 5.16 Опаѓање на PCI во зависност од староста на коловозот (C&CPIC 2021)

Анализата на состојбата на коловозот направена е во 2023 година, а врз основа на годината на обнова (рехабилитација, реконструкција или нова градба) проценети се преостанатите вредности на PCI според дијаграмот за приказ на деградација на коловозот по години и кривата на опаѓање, земајќи ја предвид староста на секоја делница и очекуваното намалување на PCI низ годините пресметана е десетгодишна средна вредност на PCI.

Оваа анализа овозможува идентификација на трендовите на промена на состојбата на коловозот и претставува основа за приоритизирање на активностите за санација, како и за подобрување на управувањето со патната инфраструктура.

- **Параметар за густина на патни јазли и наплатни станици (K.Int. P.T)**

Бројот на патни јазли, крстосници и наплатни станици претставува значаен фактор што може да влијае на концентрацијата на возачите и на безбедноста во сообраќајот. Овие точки се карактеризираат со зголемена сообраќајна комплексност, каде што возачите се изложени на дополнителна сообраќајна сигнализација, промени во режимот на движење и потреба од брзо донесување одлуки. Наплатните станици, исто така претставуваат зони на одвлекување на внимание.

Со цел да се квантифицива нивното влијание, дефиниран е параметар за густина на патни јазли, крстосници и наплатни станици, кој го претставува односот меѓу нивниот збир и должината на анализираната делница. Овој параметар се пресметува со формулата:

$$K = \frac{\sum N_p + \sum N_n}{L} \quad 5.12$$

K – параметар за густина на патни јазли, крстосници и наплатни станици (1/km)

$\sum N_p$ – збир на сите патни јазли и крстосници на анализираната делница

$\sum N_n$ – збир на сите наплатни станици на анализираната делница

L – должина на анализираната делница (km)

Пресметаниот параметар за секоја делница овозможува проценка на оптовареноста на возачите со сообраќајни точки кои бараат зголемена концентрација. Поголеми

вредности на параметарот укажуваат на почести прекини во континуитетот на возењето, што може да придонесе за замор на возачот и за зголемен ризик од ненавремени одлуки.

Анализата на добиените вредности овозможува идентификација на делници, каде што честата изложеност на возачите на комплексни сообраќајни ситуации може да влијае на безбедноста, што може да се искористи при идни инфраструктурни подобрувања или при дефинирање на дополнителни мерки за предупредување на учесниците во сообраќајот.

- **Параметар за густина на мостови и вијадукти (*K.Bridges*)**

Мостовите и вијадуктите претставуваат критични точки на патната инфраструктура, особено во зимски услови, бидејќи побрзо се ладат и први замрзнуваат, затоа што од долу и од горе се изложени на ниска температура. За разлика од класичните коловозни површини, мостовите се изложени на студен воздух од сите страни, што предизвикува побрзо ладење и создавање на ледена површина, зголемувајќи го ризикот од пролизгување и губење контрола на возилата.

Со цел да се квантифицира влијанието на овие објекти врз безбедноста во сообраќајот, дефиниран е параметар за густина на мостови и вијадукти, кој го претставува односот меѓу збирот на сите мостови и вијадукти и должината на анализираната делница. Овој параметар, се пресметува со формулата:

$$K = \frac{\sum N_m}{L} \quad 5.13$$

K – параметар за густина на мостови и вијадукти (1/km)

$\sum N_m$ – збир на сите мостови и вијадукти на анализираната делница

L – должина на анализираната делница (km)

Параметарот овозможува проценка на делниците каде што мостовите и вијадуктите можат да претставуваат зголемен безбедносен ризик во одредени временски услови. Поголеми вредности укажуваат на почеста изложеност на возачите на потенцијално замрзнати површини, што може да доведе до зголемен број на сообраќајни незгоди во зимски периоди.

Анализата на добиените вредности, може да помогне во идентификација на делници што бараат дополнителни безбедносни мерки, како што се засилено одржување во зимски услови, подобрена сигнализација или употреба на површински третмани за зголемено триење.

- **Сообраќајна сигнализација (*V sign/H sign*)**

Во анализата на сообраќајната сигнализација опфатени се вертикалната и хоризонталната сигнализација, со фокус на нивната старост и фреквенција на обновување.

Вертикалната сигнализација се внесува како број на години од нејзиното поставување. Важно е да се напомене дека ефективностa на ретрорефлектирачката фолија, која се користи за изработка на сообраќајните знаци, значително се намалува по 7 години, а потоа фолијата почнува да ја губи својата способност за ретрорефлексија. Овој период може да варира во зависност од изложеноста на надворешни влијанија, но

генерално се смета дека по овој временски интервал, потребна е замена на знаците за да се задржи нивната функционалност.

Хоризонталната сигнализација е предмет на редовно обновување, чија фреквенција зависи од категоријата на патот и обемот на сообраќајот. Податоците се добиени од ЈП за одржување и заштита на магистрални и регионални патишта, каде според план и стратегија се дефинирани периодите на обновување. Овие вредности се внесени во табела што покажува колку пати годишно се обновува сигнализацијата за различни категории на патишта, земајќи го предвид обемот на сообраќајот.

- **Просечен годишен дневен сообраќај (PGDS)**

Просечниот годишен дневен сообраќај (ПГДС) претставува мерка за сообраќајната оптовареност на патната мрежа, дефинирана како просечен број на возила што минуваат низ одредена сообраќајна делница во текот на еден ден, земајќи го предвид сообраќајниот интензитет во текот на целата година. Овој параметар е од суштинско значење при планирањето, проектирањето и управувањето со патиштата, овозможувајќи проценка на капацитетот на инфраструктурата, потребата од одржување и воспоставување на безбедносни мерки.

Прибирањето на податоците за ПГДС се врши со метод на непрекинато броење во текот на целата година, со помош на автоматски бројачи што континуирано го регистрираат сообраќајот. Дополнително, за попрецизно утврдување на структурата на сообраќајот, се користат репрезентативни рачни броења, со кои возилата се категоризираат по тип на локацијата. По собирањето, податоците се обработуваат преку одредување на средна вредност на дневните вредности за да се добијат годишни показатели, со што се намалуваат сезонските и дневните варијации, овозможувајќи подетална анализа на оптовареноста на патиштата.

$$PGDS = \frac{\sum_{i=1}^{365} Q_i}{365} \quad 5.14$$

PGDS - просечен годишен дневен сообраќај (возила/ден)

Q_i - број на возила забележани на денот i

365 - број на денови во годината

Во Македонија, официјалните податоци за ПГДС се добиваат преку системот за броење на сообраќајот управуван од ЈПДП, кој опфаќа 188 локации, од кои 58 се на магистралните патишта. Податоците се собираат преку фиксни и мобилни бројачи, овозможувајќи континуиран мониторинг и анализа на сообраќајните текови во период од десет години [99]. Следењето на овие параметри овозможува препознавање на сообраќајните трендови, проценка на растот на интензитетот на сообраќајот и подобрување на стратегиите за управување со патната инфраструктура со цел зголемување на нејзината ефикасност и безбедност.

Добиените податоци се распределуваат соодветно за секоја делница што се анализира, овозможувајќи просторна распределба на сообраќајниот интензитет и идентификација на критичните точки на патната мрежа.

- **Податоци за климатски услови**

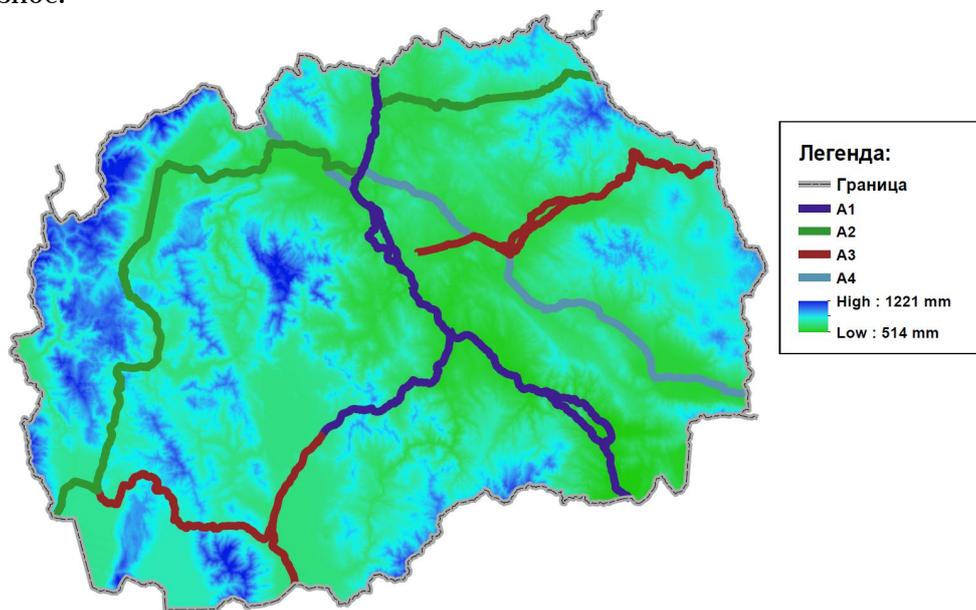
Климатските фактори, како врнежите од дожд и снег, и температурните варијации (средна, максимална и минимална температура), имаат директно влијание врз безбедноста на сообраќајот. Неповолните временски услови можат да предизвикаат намалена видливост, промена на адхезијата меѓу пневматиците и коловозот, зголемување на должината на застанување, задржување на вода и формирање на мраз. Овие фактори се особено критични на делници со поголеми надолжни наклони, остри хоризонтални кривини и зони со чести снежни врнежи.

Анализата опфаќа податоци од главните мерни станици, кои овозможуваат континуирано следење на врнежите и температурните параметри во текот на последните десет години. Пресметана е средна вредност од податоците за да се добијат просечни годишни вредности и се распределени во однос на надморската височина, овозможувајќи прецизна анализа на варијациите, според релјефните карактеристики [100, 101].

Преку ГИС анализа, за секоја точка (теме) од патните осовини се дефинирани параметрите за врнежи од дожд, снежен покривач и температурни вредности. Овие податоци се анализирани за да се добие средната годишна количина на врнежи, просечната, максималната и минималната температура за секоја делница.

Годишни врнежи од дожд

Просторната анализа покажува дека врнежите се менуваат со надморската височина, при што повисоките предели акумулираат повеќе врнежи, а пониските имаат значително помал износ.

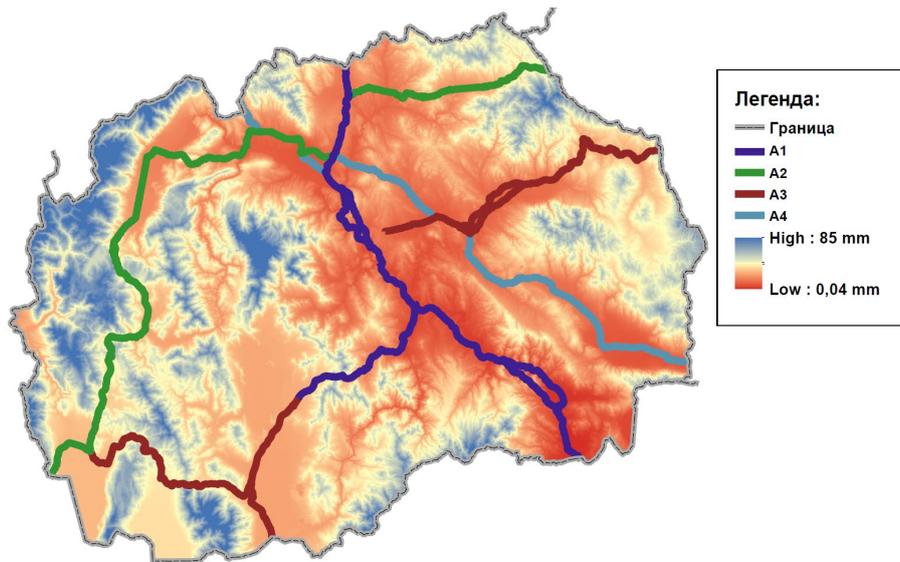


Слика 5.17 Просечни годишни врнежи од дожд за период 2014-2023 (mm)

Од слика 5.17 се гледа дека анализираниите делници претежно минуваат низ зелени зони, односно низ области со пониски годишни врнежи и мала варијабилност.

Годишни врнежи од снег

Според ГИС обработката, снежните врнежи се најизразени во планинските и висорамнинските предели, каде што снежниот покривач останува подолг временски период.



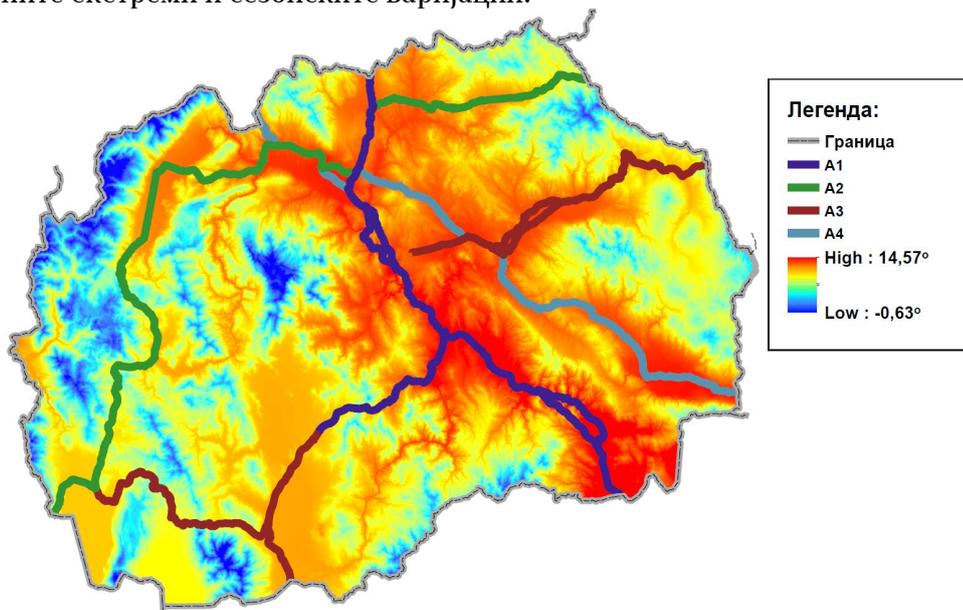
Слика 5.18 Просечни годишни врнежи од снеѓ за период 2014-2023 (mm)

Од прикажаното на слика 5.18 се гледа дека, патиштата најчесто минуваат низ подрачја со низок снежен нанос. Поголеми проблеми се појавуваат претежно на кратки делници кај превоите (Плетвар, Стража, Пресека, Кулата, Ѓавато, Буково и Деве Баир).

Средна, максимална и минимална температуратура

Температурните промени влијаат на стабилноста на коловозот и на условите за возење. Различни температурни параметри се користат за проценка на можните ризици на патиштата. Во продолжение се прикажуваат три температурни карти за десетгодишен период, кои ја следат зависноста од надморската височина.

Средните годишни температури даваат општ приказ на климатските услови во кои се одвива сообраќајот. Тие служат како основа за анализа на долгорочните влијанија врз патната инфраструктура и можат да укажат на можни предизвици поврзани со температурните екстреми и сезонските варијации.

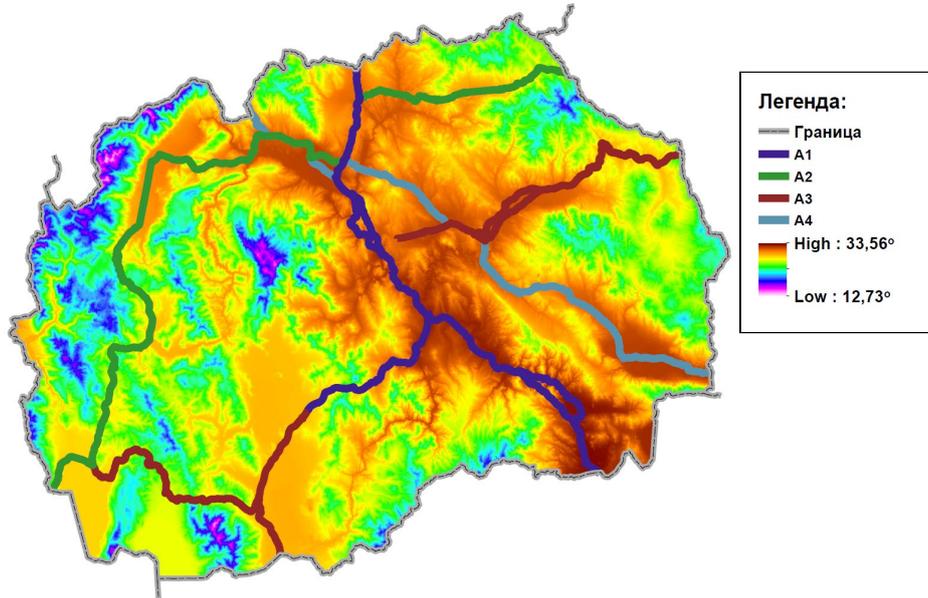


Слика 5.19 Средни годишни температури за период 2014-2023 (°C)

На слика 5.19 прикажана е просторна градација на средните годишни температури. Потопол појас се забележува во низините и котлините, а постудени вредности во

планинските масиви. Приказот е корисен како општа климатска карта за нормални услови на одвивање на сообраќајот и како основа за долгорочни процени.

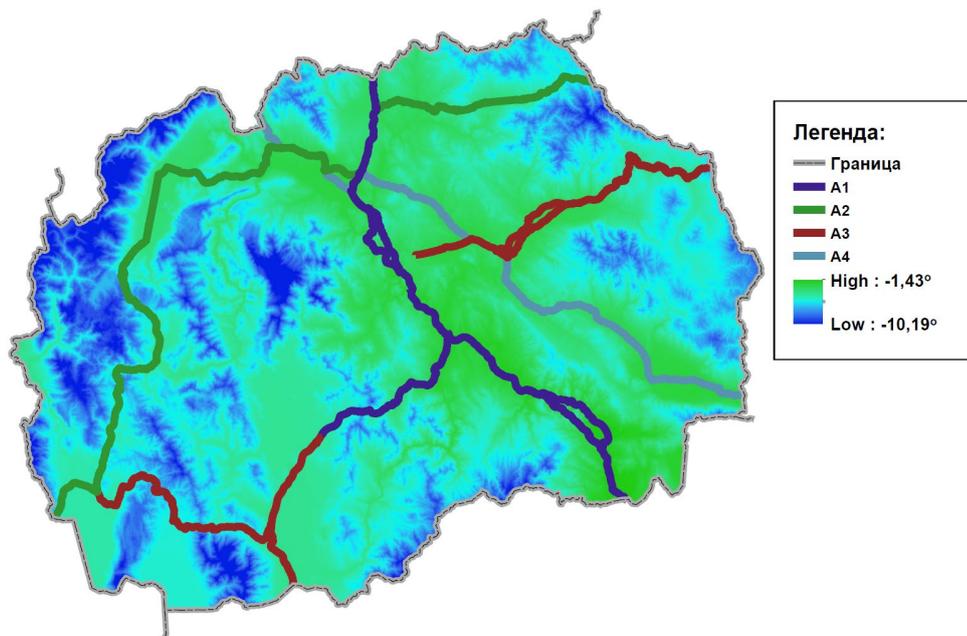
Високиите ѓемјерајтури можат да доведат до размекнување на битуменот во асфалтните слоеви, што резултира со деформации на коловозот и намалена стабилност на возилата.



Слика 5.20 Максимални годишни ѓемјерајтури за ѓериод 2014-2023 (°C)

На слика 5.20 прикажани се зоните со највисоки годишни температури, со изразен максимум по долината на Вардар и во дел од котлините. Овој просторен распоред е корисен за планирање на летни интервенции, избор на асфалтни мешавини со зголемена отпорност на топлина и одредување на приоритети за надзор на критични делници.

Нискиите ѓемјерајтури се најчеста причина за намалување на адхезијата и зголемување на ризикот од незгоди, особено при температури под нула, кога на коловозот настанува замрзнување.



Слика 5.21 Минимални годишни температури за период 2014-2023 (°C)

На слика 5.21 прикажани се минималните годишни температури, со најстудени вредности на високите планински предели и повисоки вредности во котлините и низините. Превоите се издвојуваат како локации со изразено ниски температури, каде во студени периоди се очекува почеста појава на мраз.

Овие податоци овозможуваат проценка на ризиците поврзани со екстремните температурни услови и планирање на соодветни мерки за одржување на безбедноста на патиштата.

- **Податоци за сообраќајни незгоди (Wi)**

Податоците за сообраќајните незгоди се преземени од официјалната база на Министерството за внатрешни работи (МВР) и се однесуваат на регистрираните незгоди во рамките на главната патна мрежа. Овие податоци вклучуваат приближна локација на секоја незгода, времето на нејзиното случување, како и сериозноста на последиците врз учесниците во сообраќајот.

- **локација:** дефинирана според најблискиот километарски показател или координати,
- **време на случување:** вклучувајќи датум и временски интервал,
- **сериозност на последиците:** категоризирана според типот на повредите и бројот на загинати лица.

Податоците се организирани табеларно, а тежината на секоја сообраќајна незгода е пондерирана преку тежински индекс кој ги зема предвид последиците од незгодата [102].

Вредностите на индексот се дефинирани врз основа на следната категоризација:

- незгода со полесни повреди: 1 поен,
- незгода со тешки повреди: 10 поени,

- незгода со жртви: 85 поени.

Врз основа на оваа категоризација, сите регистрирани сообраќајни незгоди се внесени во база на податоци, со цел да се идентификуваат критичните сегменти на патната мрежа. Примената на тежинскиот индекс овозможува да се креира хиерархија на локации со највисок ризик, што е основа за детектирање на таканаречените „црни точки“.

За секоја делница се пресметува вкупниот тежински индекс, според формулата:

$$W_i = \frac{N_l + 10N_t + 85N_z}{L} \cdot 1000 \quad 5.15$$

каде што:

N_l – број на незгоди со полесни повреди

N_t – број на незгоди со тешки повреди

N_z – број на незгоди со жртви

L – должина на делницата (km)

На овој начин, се добива рангирање на делниците според сериозноста на незгодите, што овозможува ефективно утврдување приоритети на патиштата, каде се потребни мерки за подобрување на безбедноста. Со цел да се овозможи споредливост на индексот меѓу различните делници, добиените вредности се поделени со должината на секоја делница. Одредувањето средна вредност ги елиминира варијациите предизвикани од разликите во должината на делниците и овозможува пообјективна анализа на безбедносниот ризик. Овој процес претставува вид на нормализација на податоците, бидејќи овозможува индексираниите вредности да бидат релевантни и директно споредливи независно од должината на делниците. Врз основа на резултатите, податоците служат како основа за предлог мерки и стратегија за интервенции на најкритичните делници.

5.5 Подготовка и почетна анализа на податоците

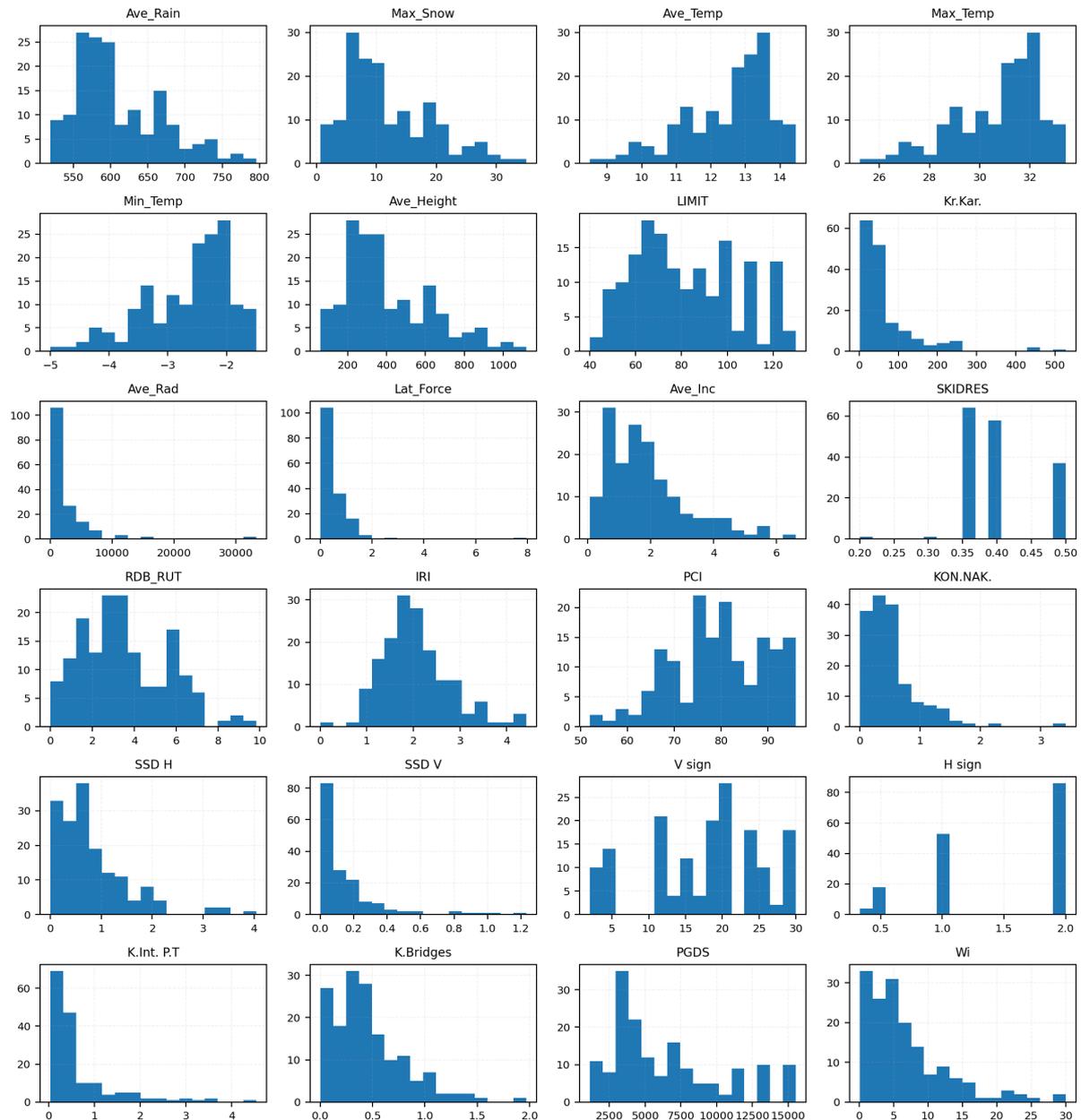
Пред примената на моделите, податоците се средуваат и се утврдуваат заеднички правила за обработка и проверка. Се обезбедуваат чисти и усогласени податоци, со јасно разбрана структура и варијабилност, а однапред се поставуваат правила за да нема протекување на информации (мешање на тренинг и тест-податоци), ниту пристрасни заклучоци.

Подготовката овозможува да се донесат јасни одлуки кои променливи се користат и во која форма, кога е потребна стабилизација или дискретизирање, како се постапува со невообичаени вредности и како се одржува ист протокол во сите методи. Со тоа, се обезбедуваат споредливи резултати, повторливост и разбирливост. Се применува поделба 80/20 (тренинг-тест), со однапред зададен случаен избор (fix-seed) за резултатите да бидат сродни при секое повторување.

5.5.1 Распределби по параметар

На слика 5.22 прикажани се хистограмите за сите нумерички променливи. На дијаграмите се забележуваат неколку форми на распределба на податоците. Првиот тип се десно искосените распределби, каде што најголем број случаи се околу мали вредности,

а мал број се протегаат кон многу големи вредности (PGDS, Ave_Rad и Lat_Force). Ова покажува дека повеќето делници имаат умерени карактеристики, а неколку имаат екстремно високи вредности што се највлијателни во распределбата. Вториот тип се распределби со повеќе врвови, кога податоците не се собрани околу еден центар, туку има повеќе јасни групи (H sign и V sign). Тоа значи дека постојат различни подгрупи со свои типични вредности. Третиот тип се приближно нормални распределби, со еден јасен врв во средината и постепено опаѓање на двете страни (PCI или IRI). Кај нив повеќето делници се околу просечната вредност, а мал број се екстремно ниски или високи.



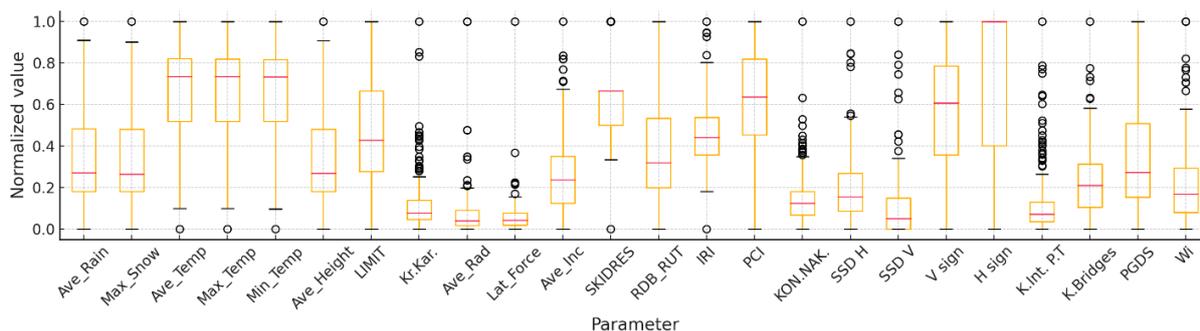
Слика 5.22 Хистограми на параметриите

Овие три типа се најизразени на прикажаните дијаграми и даваат корисна насока за понатамошна обработка. Тие укажуваат кога е соодветна лог-трансформација кај силно искосени променливи, кога е потребно внимателно моделирање при податоци со повеќе групи, и кога приближно нормалните распределби може директно да се користат во анализата.

5.5.2 Интервал и атипични вредности

Атипични вредности (outliers) се оние што значително отстапуваат од вообичаениот опсег и претставуваат ретки или екстремни случаи во податоците. За да се утврдат границите на атипичните вредности, сите параметри се прикажуваат на заедничка (0,1) скала (min-max), што овозможува директна споредба на средниот опсег и екстремните точки меѓу променливи со различни единици.

На слика 5.23 се забележува дека, кај дел од параметрите има повеќе екстремни (K.Int. P.T, K.Bridges и PGDS), што укажува на присуство на ретки, но многу високи вредности. Спротивно на тоа, кај дел од геометриските показатели (SSD H, SSD V и Ave_Inc) распределбите се тесни и стабилни, па доминира типично однесување без големи отстапувања.



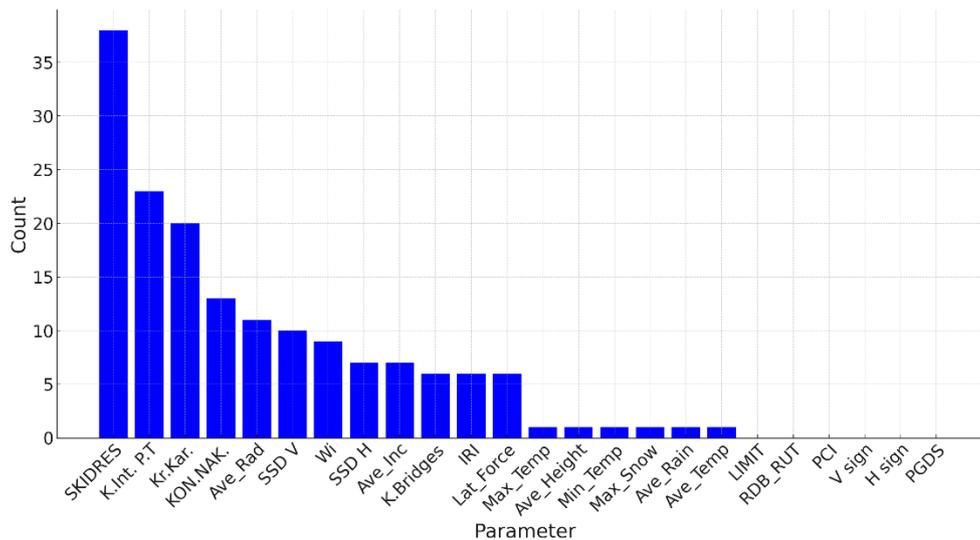
Слика 5.23 Опсег и атипични вредности

Заедничкиот приказ ја покажува различноста меѓу променливите и јасно ги издвојува оние со изразени екстремни, што помага при избор на соодветни техники и правила за обработка. Резултатите упатуваат на употреба на постабилни постапки за стандардизација и ограничување на екстремите, за ретките, но многу високи случаи да не ги нарушат резултатите кај линеарните методи.

5.5.3 Атипични податоци според IQR

Атипичноста се проценува со правилото за интерквартилен опсег – $1.5 \times IQR$ (interquartile range) – врз нормализирани вредности на секој параметар, при што се издвојуваат случаи под долната и над горната очекувана граница. Постапката е стабилна, не зависи од мерни единици и добро ги означува реалните екстремни во распределбите [103].

На слика 5.24 прикажан е бројот на атипични случаи по параметар, што овозможува брзо рангирање според чувствителност на екстремните вредности. Најмногу атипични вредности има кај SKIDRES и K.Int. P.T, потоа кај Kt.Kar. и KON.NAK., што бара повнимателно толкување и проверка на изворните записи кога отстапувањата се групирани.



Слика 5.24 Атипични вредности по параметар според IQR

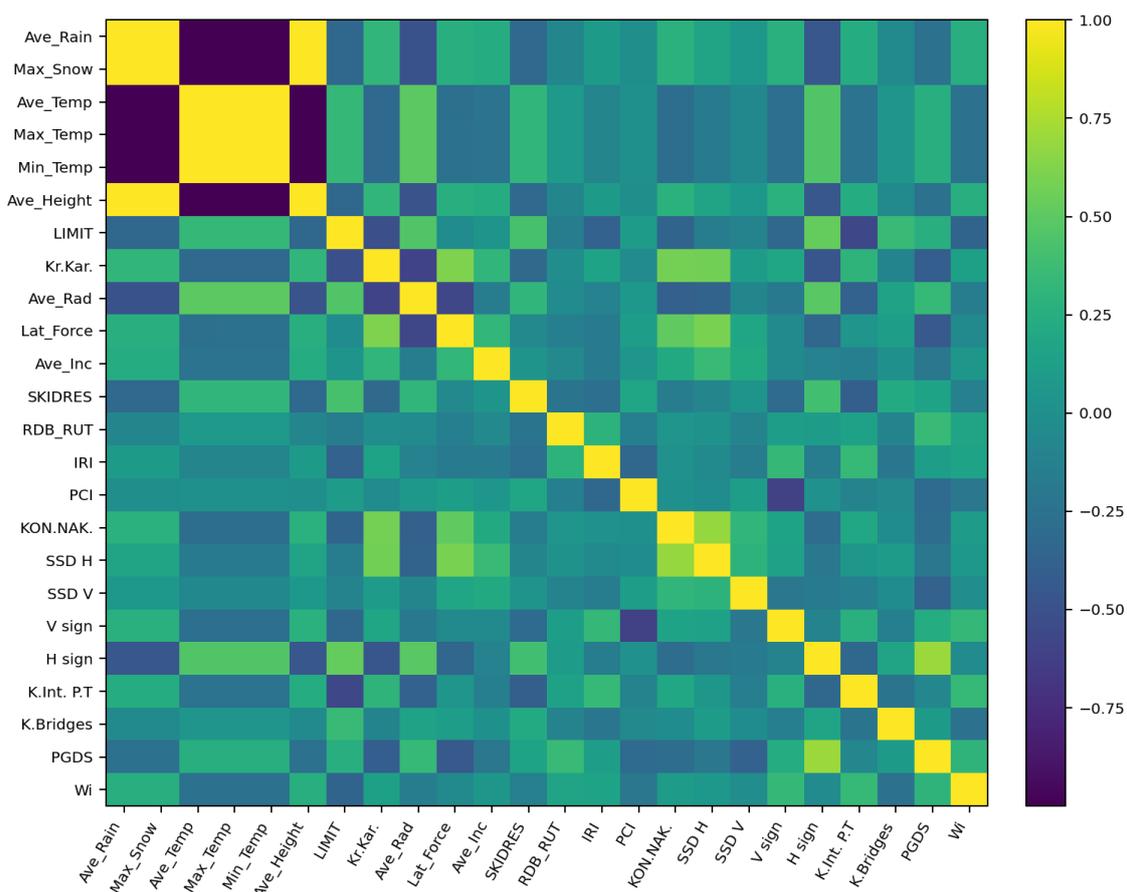
Сумираниот приказ, јасно покажува, каде има многу атипични вредности и каде распоните се умерени, па стабилноста на секој параметар се чита лесно и без дополнителни пресметки.

Овие резултати упатуваат кон примена на постабилни постапки за стандардизација и ограничување/отсечување на екстремите, за ретките, но многу големи случаи да не ги искриват резултатите во линеарни модели.

5.5.4 Зависности меѓу параметриите (Спирманова корелација)

Испитувањето на зависностите меѓу променливите помага да се препознае кога два или повеќе параметри носат иста или блиска информација и навреме да се забележи можната мултиколинеарност (силна меѓусебна поврзаност на предикторите што може да го збунува моделот). Користена е Спирманова корелација (ρ) затоа што работи со рангови, помалку е чувствителна на екстремни вредности и открива еднонасочни врски (и кога не се строго линеарни) [104].

На слика 5.25 прикажана е топлотна мапа со корелациите за сите параметри. Потемни полиња означуваат посилна врска по апсолутна вредност а посветли полиња означуваат послаба врска. Се забележуваат групи, каде поврзаните показатели имаат изразито слично однесување (Ave_Temp–Max_Temp, Ave_Temp–Min_Temp, SSD H–SSD V), додека поголем дел од останатите парови се со умерени до ниски врски (PGDS–H sign, PCI–H sign, Ave_Rain–Wi). Овој приказ овозможува брзо издвојување на групи со слично однесување и сведување на влезовите преку задржување на еден показател по група, со што се поедноставува моделот.



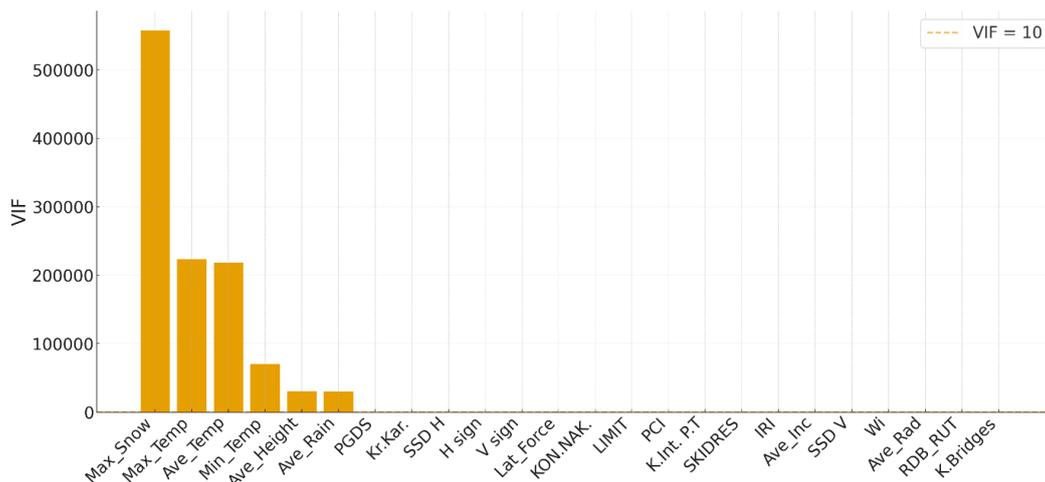
Слика 5.25 Сирманова матрица на корелации (ρ) за сите параметри.

По овој преглед, силно поврзаните парови се проверуваат дополнително за мултиколинearност, а по потреба се намалува бројот на влезови (VIF или избор на еден показател по група). Паровите со слаба врска се третираат како информативно независни.

5.5.5 Мултиколинearносii (VIF)

Целта на ова постапка е да се процени колку предикторите пренесуваат иста или многу слична информација. За секоја променлива пресметан е Variance Inflation Factor (VIF), добиен со регресија на таа променлива/параметар врз сите останати. Вредност VIF = 1 означува отсуство на колinearност, VIF = 5 е предупредувачка зона, а VIF \geq 10 се смета за висока колinearност [105].

На слика 5.26 прикажани се VIF-вредностите по параметар со означена референтна линија на 10. Највисоки VIF се јавуваат кај групата климатски параметри (температури и снег), што укажува на силно повторување на информацијата во таа подгрупа. Најниски VIF се добиваат кај дел од геометриските параметри (SSD H, SSD V, Ave_Inc), што значи дека тие внесуваат нова, самостојна информација за условите на патот и неговата геометрија без суштинско преклопување со останатите. Останатите параметри имаат пониски до умерени вредности, па придонесуваат со различна информација.



Слика 5.26 VIF по параметар со референцна линија на $VIF = 10$

Овој резултат упатува или на избор на претставник од силно поврзаните групи или на користење пристапи помалку чувствителни на колинеарност. Визуелниот приказ овозможува, брзо да се препознаат параметрите кај кои најмногу се зголемува грешката на проценките и полесно да се оправда нивното намалување или преработка.

5.5.6 Клучни резултати и насоки

Прегледот открива три карактеристични распределби и повторувања на информацијата кај дел од параметрите. Прво, се забележуваат десно искосени распределби со многу случаи на мали вредности и мал број екстреми. Потоа, се среќаваат распределби со повеќе врвови, што упатува на подгрупи со различни својства. Се појавуваат и приближно нормални распределби со врв во средината и постепено опаѓање кон краевите. Боксплот дијаграмот на иста скала покажува, каде се јавуваат екстреми и каде опсегот е тесен, што бара постабилни постапки и ограничување на крајните вредности. Спирмановата корелација издвојува групи што се движат заедно и парови со слаба врска, а VIF укажува на преклопување кај климатските променливи и самостоен придонес кај дел од геометриските.

Од овие резултати, произлегуваат неколку практични чекори:

- лог-трансформација за силно искосени променливи,
- третман на променливи со повеќе врвови со раздвојување на групи или соодветно моделирање,
- стандардизација со медијана и IQR,
- избор на еден показател од силно поврзани групи,
- задржување променливи што носат нова информација.

На овој начин резултатите остануваат споредливи и повторливи, полесни за толкување и обезбедуваат стабилна основа за понатамошно моделирање.

6. МЕТОДОЛОГИЈА

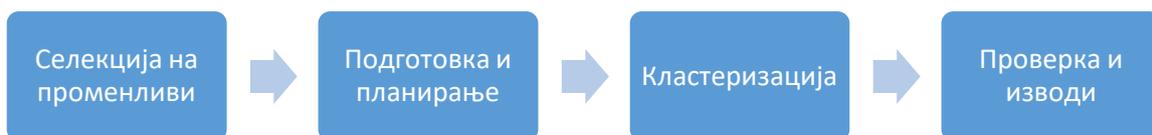
6.1 Општ аналитички пристап

Со методолошкиот пристап најпрво се прочистува и усогласува сетот на податоци. Потоа, се проверува квалитетот, се пополнуваат вредностите што недостасуваат, мерните скали се доведуваат на исто ниво и категориите јасно се означени. Понатаму се проверува, дали делниците природно се групираат со кластер анализа по патни, сообраќајни и карактеристики од околина. Овие групи служат како влез во моделите и како основа за одделна обработка по групи. Во продолжение се издвојуваат влијателните параметри за тежинскиот индекс на незгоди W_i и се користи ист протокол со поделба 80/20, се пресметуваат мерки на важност од повеќе модели, како и одредување на средна вредност на резултатите во една рангирана листа. Следува примена на три пристапи што се надополнуваат меѓусебно. Интегративна формула, модели од машинско учење и Баесови мрежи. За да се задржи контекстот, информацијата од кластерите влегува во сите пристапи по тип на делница. Точноста се проверува со вкрстена валидација и со поделби по простор и време. Сите подготовки се учат само на тренинг делови, а резултатите се прикажуваат на независни валидациони или тест поделби. На овој начин се поставува кратка, стабилна и применлива рамка, што овозможува споредливи и повторливи резултати за практична употреба.

6.2 Кластер анализа

Кластер анализата идентификува природни групи на делници со слични патни, сообраќајни и својства од околина. Групирањето се изведува само врз влезни параметри, а индикаторот W_i (тежински индекс на незгоди) се користи подоцна за проверка на валидноста на групите. Како пристапи се користат хиерархиска кластеризација со поврзување на Ward (метод што минимизира внатрешна варијанса), k-means (метод што ги групира податоците околу средни точки) и Гаусов мешан модел (Gaussian Mixture Model, GMM – пристап што ги претставува групите како комбинации на нормални распределби) [106]. Кога има многу параметри или силни врски меѓу нив, се користи анализа на главни компоненти за да се задржи најважната информација со помал број влезови. По потреба, блиските параметри се спојуваат во еден заеднички показател за полесно да се види нивното заедничко влијание.

На слика 6.1 прикажан е скратен работен тек: Прво се избираат соодветни параметри и се намалува преклопената информација. Потоа следува подготовка, односно усогласување на скали, решавање на пропусти и задавање опсег за број на групи. Потоа се прави самото групирање со избраните методи. На крај се проверува резултатот, се споредуваат решенијата, се тестира стабилноста и одделно се проверува со тежинскиот индекс на незгоди W_i .



Слика 6.1 Методолошка шема на кластер анализа

Овој редослед ја намалува можноста за пристрасни одлуки и дозволува чекорите да се повторат со исти правила. На крај, јасно се добива колку групи постојат, кои делници

припаѓаат во секоја група, кои им се главните одлики и како тие групи ќе се користат во понатамошните модели и во анализата на влијателни параметри.

6.2.1 Податоци и променливи

Во анализата се групираат делниците, така што влезниот сет ги содржи континуираните променливи што го опишуваат патот, сообраќајот и условите од околина. Сите континуирани променливи се доведуваат на иста скала со стандардизирање (се одзема средината и се дели со стандардната девијација) за различните мерни единици да не влијаат на резултатот од групирањето. Кога има многу променливи или силни меѓусебни врски, се применува анализа на главни компоненти. Така се задржува поголемиот дел од информацијата и се добива постабилна основа за групирање. Толкувањето на добиените групи се прави преку оригиналните променливи (непреработени податоци). Индикаторот W_i не влегува во влезниот сет, туку служи само за надворешна проверка, дали групите се оправдани.

6.2.2 Пристапи и избор на број на кластери

За да се издвојат природни групи кај делниците, се користи кластеризирање со неколку пристапи. Групирањето се врши врз основа на сите независни променливи од целиот сет на податоци. Индикаторот W_i подоцна се користи за проверка, дали групите имаат смисла. За кластеризација се користат трите пристапи (Ward, k-means и Гаусов мешан модел, GMM), при што фокусот е ставен на изборот на бројот на кластери и на оцената на стабилноста на решенијата [106].

Податоците се стандардизираат и се користат континуирани показатели за патните, сообраќајните и околините својства. Кај k-means се користи еуклидово растојание (растојание по права линија меѓу две точки) и контролирана случајност за резултатите, да може да се повторат.

Кога променливите се многу или силно зависни, прво се прави анализа на главни компоненти. Така се задржува поголемиот дел од информацијата (90-95%), па кластеризирањето продолжува во тој поедноставен простор. Кај мешаните модели се разгледуваат повеќе форми на варијација (сферична, дијагонална, заедничка, целосна) за да се опфатат групи со различна форма и големина.

Бројот на кластери се разгледува во опсег $k = 2-10$. За Ward и k-means изборот се води од три внатрешни мерки: силуета, Калински-Харабас и Дејвис-Боулдин [110]. За секое k се пресметуваат вредности, мерките се доведуваат на заеднички размер и се зема просек за појасен сигнал. Подобро решение има, висока силуета, висок Калински-Харабас и низок Дејвис-Боулдин. Ако кривите се израмнуваат за неколку соседни k , се избира помалата вредност.

Кај Гаусовите мешавини, изборот е според најниско BIC по k и по форма на варијација (баланс меѓу добро вклопување и едноставност) [106]. За финална одлука се споредуваат сигналите од внатрешните мерки и од BIC и се избира k што најдоследно изгледа најдобро.

Мерка силуета за објект i :

$$s(i) = \frac{b(i) - a(i)}{\max\{a(i), b(i)\}} \quad 6.1$$

Каде $a(i)$ е просечно растојание до точките во сопствената група, а $b(i)$ е најмалото просечно растојание до точките во најблиската друга група. Високи вредности на $s(i)$ значат дека групите се компактни и јасно одделени.

Информационен критериум BIC:

$$BIC = -2l n L + pl n n \quad 6.2$$

Каде lnL е лог-веројатноста на моделот (логаритам од веројатноста да се добијат овие податоци, ако важи моделот), p е бројот на параметри, а n е број на записи. Пониско BIC означува поповолен однос меѓу, описната моќ и сложеноста.

Бројот на групи (k) го одредуваат самите податоци за делниците. Внатрешните мерки покажуваат колку се збиени и колку се одвоени групите. BIC проверува дали моделот е доволно добар без непотребна сложеност. Кога и внатрешните мерки и BIC укажуваат на исто решение, се избира тој k и се добива поделба што е јасна и практична за понатамошна употреба.

6.2.3 *Стабилност, надворешна проверка и изводи*

Стабилноста на кластерите се оценува со повеќекратни повторувања и споредба меѓу пристапите. За k -means се прават повеќе почетни анализи со различни почетни вредности и се следи дали решенијата се повторуваат. Се пресметува прилагоден Рандов индекс (сличност на поделби, 1 блиску, 0 случајно) за споредба на поделбите меѓу различни почетни вредности и меѓу трите пристапи, при што повисоки вредности значат постабилни групи [108]. Дополнително се следи, дали силуетата (-1 до 1, припадност) и другите индикатори остануваат стабилни кога се менува бројот на кластери во тесен опсег околу избраното k .

Надворешната проверка се врши со споредба на кластерите по индикаторот W_i , кој претхоно не е влез во групирањето. Прво се тестира дали постојат разлики по кластери со непараметарски тест на Краскал-Волис, а потоа, ако разликите се значајни, се прават споредби по парови (post-hoc) со Dunn со соодветна корекција [109, 110]. Паралелно се прикажуваат мерки за големина на ефект за да се процени интензитетот и насоката на разликите, а не само нивната статистичка значајност [111]. На тој начин се проверува дали добиените групи се смислени и по однесување во однос на W_i (иако W_i не учествувал во самото групирање).

Изводи од кластеризацијата се: ознаки на кластери за секоја делница, растојанија до центарот за Ward и k -means, како и веројатности за припадност за Гаусови мешани модели. Овие резултати служат за опис на видовите на делници, за толкување на разликите меѓу групите и за понатамошна примена во моделирање и донесување одлуки.

6.2.4 *Класифицирање на параметри (feature clustering)*

Овој чекор ја организира листата на параметри така што, сличните се групираат заедно, а повторувањата се намалуваат. Покрај групирањето на делниците, се групираат и самите параметри според тоа колку меѓусебно се поврзани, за да се добијат компактни групи со сродно однесување. Како мерка на блискост се користи корелационо растојание, дефинирано преку Спирманова рангова корелација, па параметрите што се менуваат на сличен начин имаат мало, а оние што се независни имаат големо растојание. Групирањето се врши со хиерархиско спојување по чекори со просечно поврзување, при што

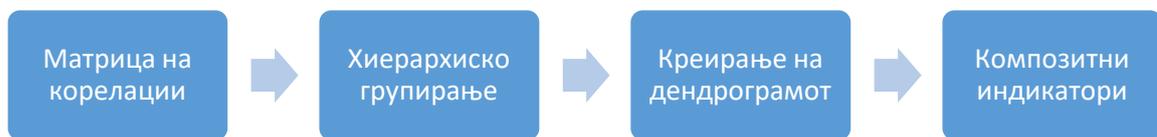
дендрограмот се сече на висина што овозможува добра рамнотежа меѓу сличноста во групите и разликата меѓу нив [107, 112].

Корелационо растојание меѓу два параметра, се одредува со формулата:

$$d(X, Y) = 1 - |\rho(X, Y)| \quad 6.3$$

Каде $\rho_s(x,y)$ е Спирманова рангова корелација. Вредност блиска до 0 значи силна сродност, вредност блиска до 1 значи слаба сродност.

Во примената на податоците за делниците, најнапред се пресметува матрица на $d(X, Y)$ за сите парови параметри, потоа се спојуваат најсличните по редослед, и се избира број на групи што овозможува пократка листа на претставници без суштинска загуба на информација. За секоја група се прави еден заеднички показател (просек или прва главна компонента) и тој се користи како претставник во следните анализи. Слика 6.2 ја прикажува постапката, од пресметување на растојанијата и градење на дендрограм, до избор на групи и создавање на композитни индикатори.



Слика 6.2 Мейџолошка шема за класифицирање на параметри

Овој пристап го намалува бројот на променливи, ја отстранува преклопената информација и го прави сетот на параметри појасен и полесен за употреба во следните анализи.

6.3 Иницијална проверка на техники

Во оваа фаза се прави брза и споредлива почетна проверка на техники за моделирање на W_i . Податоците се делат на тренинг и тест во однос 80/20 со контролирана случајна поделба. Сите подготовки се учат само на тренинг делот, а потоа без промени се применуваат на тест делот, за да нема мешање на информации. Под подготовка спаѓа пополнување на вредности кои недостасуваат, центрирање и скалирање каде што е потребно и неколку мали технички чекори. За сите модели важат исти влезни карактеристики, иста поделба и ист протокол. Стандардизација се применува само кај техники што се чувствителни на размери (не кај дрвовидни методи).

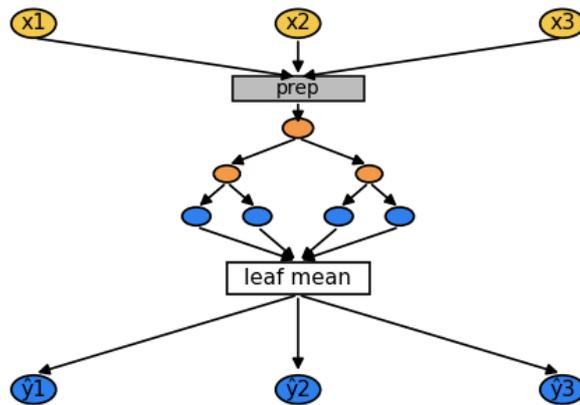
Во проверката учествуваат Decision Tree, Random Forest, Gradient Boosting, XGBoost, CatBoost, LightGBM, AdaBoost, Bagging Regressor и ExtraTrees, а по потреба и HistGradientBoosting и Explainable Boosting Machine, како и споредбени нелинеарни базни MLP и SVR [113]. Поставките на хиперпараметрите се задржани умерени за споредба под исти услови. Главна метрика е R^2 , а MAE и RMSE се прикажуваат како мерки на грешка. Моделите што постигнуваат $R^2 \geq 0.45$ продолжуваат во натамошна обработка на влијателни фактори, при што нивниот придонес во усреднувањето се тежи според R^2 , додека моделите под прагот се прикажани за целосна слика, но не влегуваат во спојувањето. Резултатот од проверката е, јасно рангирање според R^2 (со MAE и RMSE) и кратка листа техники што ќе се користат за извлекување влијателни параметри и понатамошна подготовка на разбирлива формула.

6.3.1 Дрво на одлука (Decision Tree)

Дрвото на одлука ги дели податоците етапно според правило „ако-тогаш“ (if-then). Во секој внатрешен јазол се избира особина и праг што најмногу ја намалува грешката, при што примероците се распоредуваат во две групи. Постапката продолжува додека поделбите носат значајно подобрување или додека не се достигнат однапред поставените ограничувања на сложеноста. Техниката е соодветна за податоците, затоа што природно работи со континуални и категоријални влезови без стандардизирање, опфаќа нелинеарни врски и интеракции и останува разбирлива преку правила по гранки [114].

$$y(x) = (1/n_m) * \sum_{i:x_i \in R_m} y_i \quad 6.4$$

Во формулата: $y(x)$ е проценетата вредност за влез x , R_m е крајната област во која припаѓа x , n_m е бројот на тренинг-примери во R_m , сумата е над сите i за кои важи $x_i \in R_m$, а y_i се нивните вистински вредности.



Слика 6.3 Дрво на одлука (Decision Tree)

На слика 6.3 прикажан е моделот дрво на одлука (Decision Tree), каде може да се увиди принципот на поделби по јазли и крајни групи.

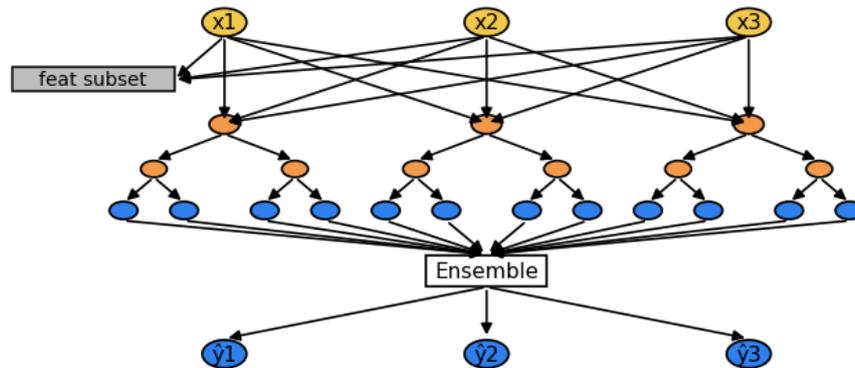
Во анализата применета е конфигурација насочена кон стабилност и јасно читање на условите по гранки: `DecisionTreeRegressor (criterion = squared_error, max_depth = 6, min_samples_split = 10, min_samples_leaf = 5, random_state = 42)`. Овие ограничувања ја држат сложеноста под контрола и спречуваат премали јазли, со што се намалува пренавикнување (меморирање на тренинг податоците) и се задржува прегледна структура на поделбите. За податоците за делници, дрвото само избира граници за мерки како PGDS, SSD_H и SSD_V. Категоријалните влезови ги дели по групи, па така секоја гранка е услов што води до проценка на W_i .

6.3.2 Наизменично генерирани дрва на одлуки (Random Forest)

Random Forest гради многу дрва на одлуки врз различни случајно извлечени подмножества од редови (bootstrap примероци), при што во секој јазол се разгледува само дел од особините, а конечната проценка за тежински индекс на незгоди (W_i) е просек од излезите на сите дрва. На тој начин се намалува варијабилноста и се добива постабилен просек [115]. Со ова, секое ново дрво гледа поинаква слика од податоците, затоа грешките меѓу дрвата се помалку слични и средната проценка е поиздржана.

$$y(x) = (1/T) * \sum_{t=1..T} tree_t(x) \quad 6.5$$

Во формулата T означува број на дрва, $tree_t(x)$ е предвидувањето од t -тото дрво, а $y(x)$ е аритметичка средина од сите предвидувања.



Слика 6.4 Наизменично генерирани дрва на одлуки (Random Forest)

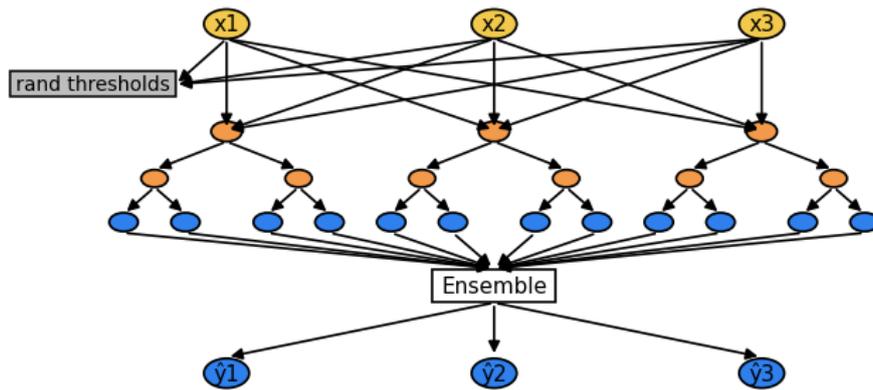
На слика 6.4 е прикажан моделот Random Forest. Во шематскиот приказ се гледа како нов примерок минува низ повеќе дрва, во кои се добиваат поединечни процени и нивната средина е конечниот резултат. Техниката е соодветна за податоците, бидејќи работи со табеларни, претежно континуирани показатели по делници, не бара задолжително стандардизирање на размери и природно опфаќа нелинеарности и интеракции меѓу особините, додека категоријалните влезови се внесуваат со едноставно кодирање и поделби по категории. Во примената, користена е конфигурација која е споредлива со останатите техники: RandomForestRegressor ($n_estimators = 600$, $random_state = 42$, $n_jobs = -1$). Постапката користи многу дрва за стабилен просек. Контролираната случајност дава повторливи резултати. Бидејќи дрвата учат на различни податоци и особини, нивните грешки се разликуваат, а просекот дава поцврста проценка.

6.3.3 Дојолнишелни дрва на одлуки (ExtraTrees)

Extra Trees е групен метод со дрва на одлуки. Во секој јазол се разгледува случаен дел од особините и за секоја се зема праг со случајно избирање, на тој начин дрвата меѓусебно се разликуваат. Конечната проценка е средна вредност од излезите на сите дрва, што го прави резултатот постабилен на табеларни податоци со мешани континуални и категоријални влезови [116].

$$y(x) = (1/N) * \sum_{i=1..N} T_i(x) \quad 6.6$$

Каде $T_i(x)$ е проценка од i -тото дрво, а N е број на дрва.



Слика 6.5 Дојолнишелни дрва на одлуки (ExtraTrees)

На слика 6.5 прикажан е начинот на тестирање во ЕТ моделот. Во секој јазол се тестираат случајни прагови по неколку особини, се избира поделбата што локално ја намалува грешката, а конечниот резултат е просекот од сите дрва.

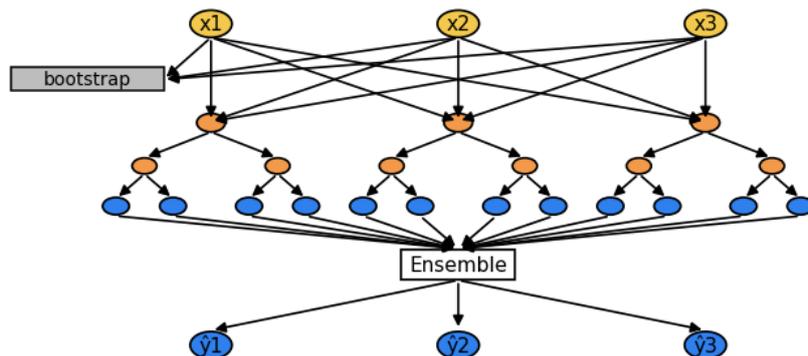
Во анализата применет е ExtraTreesRegressor ($n_estimators = 600$, $random_state = 42$, $n_jobs = -1$), со што се добива доволен број различни дрва и исти правила за случајно избирање за повторлив резултат.

6.3.4 Комбинирани помножесѝва (Bagging Regressor)

Bagging обучува повеќе исти основни модели на различни подмножества со враќање (bootstrap) од оригиналните податоци. Секој модел се обучува на малку поинаква верзија на множеството, со што грешките на моделите се разликуваат. Конечната проценка за W_i се добива така што се зема средна вредност од излезите, што ја намалува променливоста и го прави резултатот порамномерен на табеларни податоци [117].

$$y(x) = (1/T) * \sum_{t=1..T} h_t(x) \quad 6.7$$

Во изразот T е број на модели, а $h_t(x)$ е проценка од t -тиот модел за влез x , формулата покажува дека групното предвидување е средна вредност и поради тоа има помала случајна грешка во однос на поединечен модел.



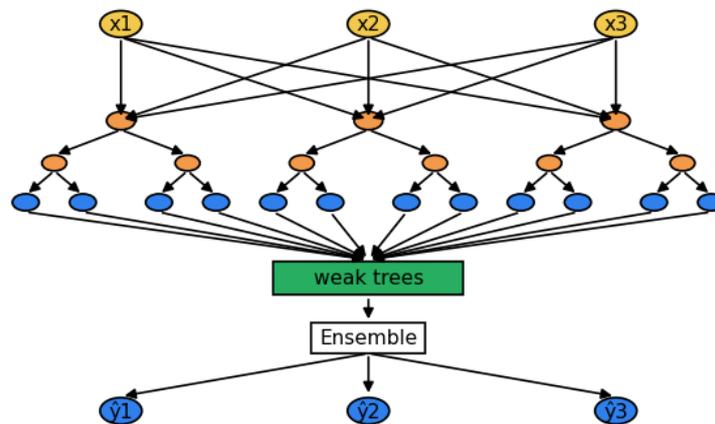
Слика 6.6 Комбинирани помножесѝва (Bagging Regressor)

Сликата 6.6 ја прикажува шемата на Bagging Regressor. Во моделот, од оригиналниот сет се формираат повеќе подмножества со враќање (bootstrapping), на секое подмножество се обучува по еден модел, а новиот запис се оценува со сите модели. Потоа се прави средна вредност на излезните податоци, при што конечниот резултат е постабилен и помалку чувствителен на случајни промени во податоците.

Во анализата применета е поставка насочена кон стабилност: BaggingRegressor (`n_estimators = 300`, `random_state = 42`, `n_jobs = -1`) со основен DecisionTreeRegressor, при што повеќе плитки дрва работат паралелно и излезот е нивна средна вредност. Овој избор обезбедува доволен број независни подмножества и контролирана случајност за повторливи резултати, а структурата останува јасна за инженерска примена.

6.3.5 Адаптивно засилување (AdaBoost)

Адаптивното засилување, гради низа плитки дрва чекор по чекор. Секое наредно дрво дава повеќе тежина на записите што претходно биле погрешно проценети, со што вкупната проценка постепено се подобрува. Овој пристап погоден е за табеларни податоци и дава стабилни резултати со мали дрва и умерен број чекори [118].



Слика 6.7 Адаптивно засилување (AdaBoost)

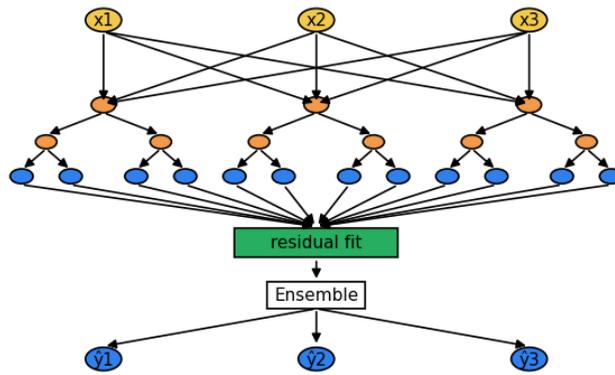
На слика 6.7 прикажана е структурата на AdaBoost. Во овој модел секој нов чекор додава плткo дрво, се зголемува тежината на потешките случаи, и се добива попрецизна проценка. Во иницијалната проверка со показатели по делници и таргет W_i се користи AdaBoostRegressor (`n_estimators=400`, `random_state=42`) со DecisionTreeRegressor (`max_depth=3`), со околу 15 јазли по дрво и доволен број чекори за последователна корекција на преостанатата грешка без преголема сложеност и без потреба од стандардизирање.

$$y(x) = \sum_{t=1..T} \alpha_t * h_t(x) \quad 6.8$$

Каде што T е број на чекори (дрва), $h_t(x)$ е процената од дрвото на чекор t , а α_t е тежината со која чекорот учествува во вкупната проценка. Ваквата постапка добро се вклопува во иницијалната споредба на техники и обезбедува постепено намалување на грешката со едноставни, плитки дрва прилагодени на показателите по делници и W_i , без дополнителни чекори за стандардизирање.

6.3.6 Итеративно засилување (Gradient Boosting)

Итеративното засилување е техника за комбинирање повеќе модели, при што се создаваат и надградуваат слаби модели (плткo дрва на одлука) за да се добие конечен, прецизен модел за предвидување. Основата е постепено подобрување преку итеративен процес во кој секоја наредна итерација ги корегира грешките на претходните модели со минимизирање на функцијата на загуба, што овозможува моделирање на посложени односи меѓу влезните параметри и зависната променлива [119].



Слика 6.8 Итеративно засилување (Gradient Boosting)

Слика 6.8 го прикажува моделот Gradient Boosting. Во секоја итерација се додава плитко дрво и на крај се добива збирна проценка од сите дрва.

Во наједноставна форма, моделот се подобрува во секој чекор според следната формула:

$$F_m(x) = F_{m-1}(x) + \gamma \cdot h_m(x) \quad 6.9$$

Каде што:

$F_m(x)$ – нов модел по m -тата итерација

$F_{m-1}(x)$ – претходен модел

$h_m(x)$ – нов слаб модел што се додава за да ги поправи грешките

γ – коефициент на учење што го контролира влијанието на новиот модел

За да се утврди кои делови од податоците предизвикуваат најголема грешка, се пресметува разликата помеѓу вистинската вредност и предвидената вредност од претходниот модел:

$$e_i = y_i - F_{m-1}(x_i) \quad 6.10$$

Каде што e_i претставува остаток на грешка, односно разлика меѓу вистинската вредност y_i и претходно предвидената вредност $F_{m-1}(x_i)$. Новиот модел се обучува врз основа на овие грешки, со цел постепено да ја намалува неточноста.

Во анализата се применува GB-Regressor ($n_estimators=600$, $learning_rate=0.05$, $max_depth=3$, $random_state=42$). Секое дрво е плитко ($max_depth = 3$, приближно до 8 лисја и околу 15 јазли), додека $\gamma = 0.05$ ја контролира големината на чекорот. Вкупно се додаваат 600 дрва за постепено намалување на грешката со стабилна генерализација на табеларни, претежно нумерички особини. По потреба се применува рано запирање, за да се спречи меморирано учење (пренавикнување) од тренинг-податоците.

Постепеното додавање на дрва овозможува да се фатат и послаби сигнали во податоците, а грешката да опаѓа на секој чекор. Моделот останува читлив и применлив за рангирање на делници според тежински индекс на незгоди (W_i).

6.3.7 Екстремно засилување (XGBoost)

Екстремното засилување е варијанта на итеративното засилување со механизми за ограничување на сложеноста и ефикасни постапки за раст на дрва [120]. Моделот се гради адитивно и во секој чекор се додава плитко дрво што ги подобрува деловите со најголема

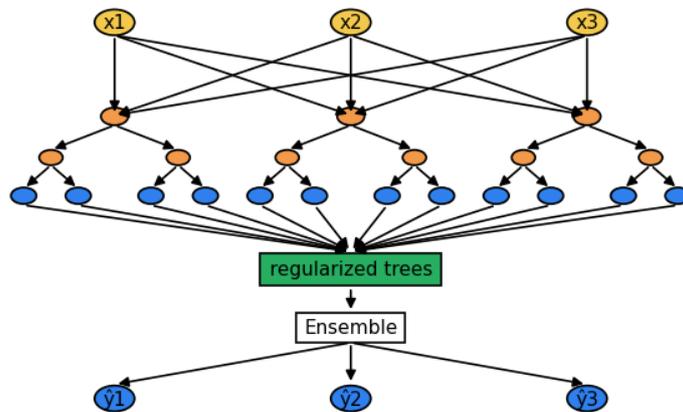
грешка од претходната прогноза, при што во целната функција се вклучуваат термини за контрола на сложеноста за да не се зголеми непотребно големината на дрвата.

$$F_m(x) = F_{m-1}(x) + lr * h_m(x) \quad 6.11$$

Ова значи дека новиот модел F_m се добива со додавање на мало подобрување h_m на претходниот модел F_{m-1} , а lr е мала стапка на учење што го контролира придонесот на тоа подобрување. Со секое ново дрво се намалува преостанатата грешка и моделот се прецизира чекор по чекор.

$$\text{Gain} = 0.5 * \left(G_L^2 / (H_L + \lambda) + G_R^2 / (H_R + \lambda) - G^2 / (H + \lambda) \right) - c \quad 6.12$$

Оваа формула ја проценува користа од кандидат-поделба во јазол. G и H се зборови на прв и втор извод на функцијата на загуба за сите записи во јазолот, G_L и H_L важат за левата страна, G_R и H_R за десната, λ е L2 ограничување на сложеноста, а c е фиксен трошок за секоја поделба. Поделбата се прифаќа ако добивката е поголема од нула, бидејќи тогаш локално се намалува загубата.



Слика 6.9 Екстремно засилување (XGBoost)

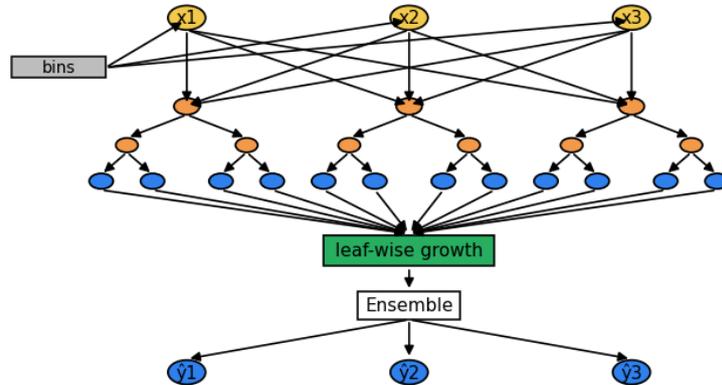
На слика 6.9 прикажан е моделот XGBoost. Прогнозата се надградува чекор по чекор со нови плитки дрва, а функцијата на загуба има додаток што ја ограничува сложеноста, па моделот останува прецизен и на нови податоци.

Во анализата се користи XGBRegressor ($n_estimators=600$, $learning_rate=0.05$, $max_depth=4$, $subsample=0.9$, $colsample_bytree=0.7$, $reg_lambda=1.0$, $reg_alpha=0.0$, $tree_method=hist$, $random_state=42$, $n_jobs=-1$). Секое дрво е плитко ($max_depth=4$; околу 16 лисја и околу 31 јазол), бројот на дрва овозможува постепени надградби, земањето по редови (0.9) и по особини (0.7) ја намалува сличноста меѓу дрвата, а λ и c ја контролираат големината на поделбите. Поставувањето е прилагодено на показателите по делници (PGDS, SSD_H, SSD_V, состојба на коловоз) и таргет W_i . Плитките дрва брзо идентификуваат прагови, а малата стапка на учење ($lr=0.05$) овозможува постепено подобрување без претерано прилагодување на тренинг-податоците.

Овој модел дава прецизни и стабилни проценки на тежинскиот индекс на незгоди (W_i) на табеларни податоци и овозможува јасни правила по гранки со контролирана сложеност, погодни за споредба и инженерска примена.

6.3.8 Лесно агајџивно засилување (LightGBM)

Со оваа техника дрвата се додаваат етапно, со растење по најкорисна гранка (leaf-wise) и се користат хистограмски прагови по особина за побрз избор на поделби [121]. Овој пристап е соодветен за табеларни податоци со показатели по делници и таргет тежински индекс на незгоди (Wi).



Слика 6.10 Лесно агајџивно засилување (LightGBM)

На слика 6.10 прикажана е структурата на LightGBM. Моделот ја продлабочува гранката со најголемо локално подобрување, притоа хистограмите по особина обезбедуваат ограничен сет кандидати за прагови. На овој начин се постигнува брза надградба со умерена сложеност и правилата по гранки се лесни за следење.

Во наједноставна форма, надградбата следи како кај градиентното засилување:

$$F_m(x) = F_{m-1}(x) + lr * h_m(x) \quad 6.13$$

Каде F_m е новиот модел по чекорот m , F_{m-1} претходниот модел, h_m новото плитко дрво, а lr е мала стапка на учење што го контролира чекорот. За избор на поделба s се користи мерка на добивка:

$$G(s) = 0.5 * (G_L(s)^2 / (H_L(s) + \lambda) + G_R(s)^2 / (H_R(s) + \lambda) - G(s)^2 / (H(s) + \lambda)) - c \quad 6.14$$

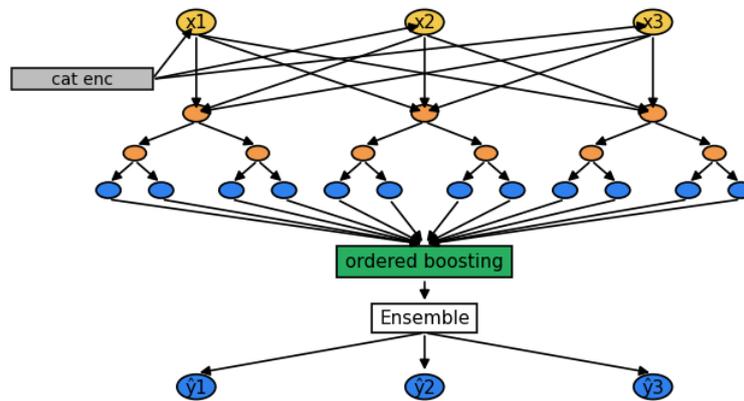
Каде G и H се збир на прв и втор извод на загубата по јазол и страна, λ е L2 ограничување на сложеноста, а c е фиксно намалување на добивката по поделба.

Во анализата се применуваат две поставки за споредливост и стабилност без протекување на информации (80/20, no-leakage): базна LGBMRegressor(random_state=42) и подесена LGBMRegressor(n_estimators=400, learning_rate=0.005011872336272725, num_leaves=7, max_depth=5, min_child_samples=5, subsample=0.9, colsample_bytree=0.6, reg_alpha=0.0, reg_lambda=1.0, random_state=42, n_jobs=-1). Со num_leaves=7 дрвото останува плитко, max_depth=5 и min_child_samples=5 ја одржуваат сложеноста умерена. Делумното земање на редови и особини ја намалува сличноста меѓу дрвата. Поставувањето е усогласено со PGDS, SSD_H, SSD_V и индикаторите за коловоз за да се детектираат граници и постепени промени со мал чекор, при стабилно учење.

6.3.9 Категориско засилување (CatBoost)

CatBoost е метод што прави мали симетрични дрва и ги собира во посилен модел. Кога има категоријални полиња ги кодира по редослед на учење, така што за секој запис користи информации само од претходните и не пренесува информација од тренинг кон

тест. Погоден е за табели со мешавина од броеви и категории не бара стандардизирање и дава стабилни проценки со мали и едноставни правила [122].



Слика 6.11 Категориско засилување (CatBoost)

Сликата 6.11 дава шематски преглед на CatBoost. Моделот прикажува редоследен напредок што ја намалува грешката, при што категоријалните особини се кодираат по ред (ordered encoding) без користење информации од идни редови [125].

Во наједноставна форма, надградбата е како кај градиентното засилување:

$$F_m(x) = F_{m-1}(x) + lr * h_m(x) \quad 6.15$$

Каде F_m е моделот по чекорот m , F_{m-1} претходниот модел, h_m новото плитко дрво што ја подобрува претходната проценка, а lr е мала стапка на учење што го одредува чекорот.

Категоријалните променливи се претвораат во нумерички вредности со редоследна статистика:

$$\text{enc}(\text{cat}_i) = (\sum_{j < i} y_j + a * \text{prior}) / (n_{<i} + a) \quad 6.16$$

Каде $\text{enc}(\text{cat}_i)$ е вредноста за редот i , се користат само претходните редови $j < i$, prior е почетна вредност за стабилизација, а a е коефициент за порамнување. Ваквото редоследно кодирање меѓу тренинг и тест, усогласено е со 80/20 поделба без информационо преклопување.

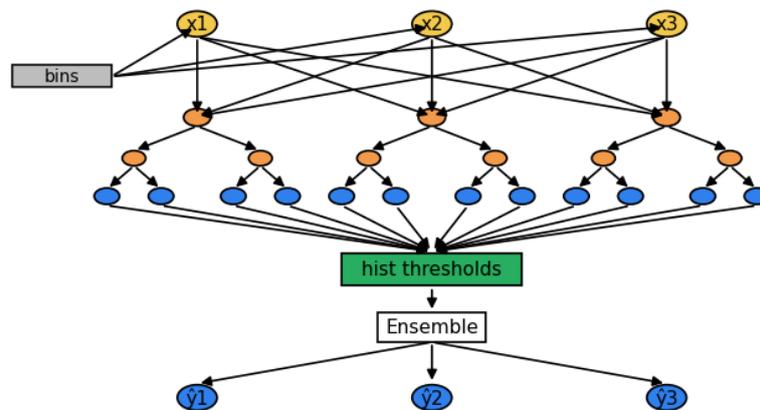
Во иницијалната проверка со показатели по делници и таргет тежински индекс на незгоди (W_i) се применува `CatBoostRegressor(iterations=1000, learning_rate=0.05, depth=6, loss_function=RMSE, random_seed=42, verbose=0)`. Плитките и симетрични дрва, умерениот чекор и доволниот број итерации обезбедуваат постепено подобрување без дополнителни трансформации. Постапувањето е прилагодено на PGDS, SSD_H, SSD_V и индикаторите за состојба на коловоз, при што правилата по гранки се лесни за следење и добро се поврзуваат со инженерската логика.

Со редоследно кодирање и симетрични плитко дрва, CatBoost обезбедува стабилни проценки на тежински индекс на незгоди (W_i) и овозможува сигурна и рамномерна споредба на делници.

6.3.10 Итеративно засилување со хистограмски прагови (HGBosting)

Итеративното засилување со хистограмски прагови работи етапно како градиентното засилување, но границите за поделба ги избира од мал сет вредности добиени од

хистограми по особина. Наместо да пребарува низ сите можни точки, моделот користи бинови, што го забрзува учењето и ја намалува потрошувачката на меморија, со задржана точност на табеларни податоци и мерки по делници како PGDS, SSD_H, SSD_V и состојба на коловоз [123].



Слика 6.12 Итеративно засилување со хистограмски прагови (HistGradientBoosting)

Сликата 6.12 го претставува HistGradientBoosting во шематска форма. Моделот најнапред ги групира вредностите во хистограмски бинови и ги избира праговите од нивните рабови, потоа додава плитко дрво што ја надградува прогнозата, со побрз тренинг и помала потрошувачка на меморија при задржана стабилност на мешани табеларни податоци.

Во најкратка форма, надградбата следи како кај градиентното засилување:

$$F_m(x) = F_{m-1}(x) + lr * h_m(x) \quad 6.17$$

Каде F_m е моделот по чекорот m , F_{m-1} претходниот модел, h_m новото плитко дрво што ја подобрува претходната проценка, а lr е мала стапка на учење што го одредува чекорот. Праговите за поделба се земаат од хистограмските граници:

$$\tau \in t_1, t_2, \dots, t_{K-1}$$

оттука, наместо да се пребаруваат сите можни прагови, изборот се врши од претходно формиран мал сет кандидати.

За споредбата се користи HGBRegressor ($max_iter = 600$, $learning_rate = 0.05$, $max_leaf_nodes = 31$, $max_bins = 255$, $min_samples_leaf = 20$, $random_state = 42$). Ова значи 600 плитки надградби, мал чекор на учење, до 31 краен јазол по дрво и доволно фини кандидати за прагови ($max_bins=255$), додека минималниот број записи по лист спречува премногу ситни листови. Поставувањето е усогласено со показателите по делници за да се детектираат граници и постепени промени без дополнително стандардизирање.

Овој метод овозможува брза обука, умерена сложеност и сигурни проценки на тежински индекс на незгоди (W_i), при што поделбите по гранки се лесно разбирливи и корисни за инженерско рангирање на делници.

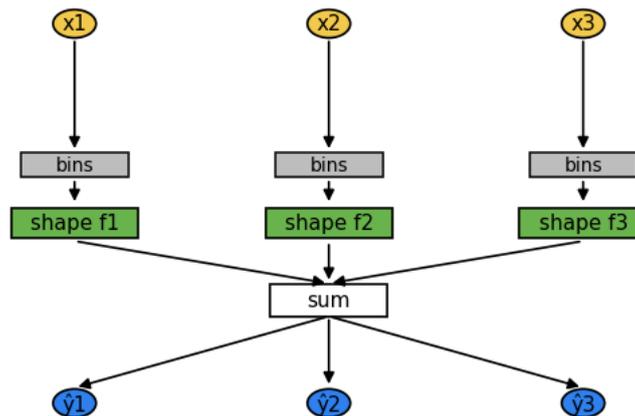
6.3.11 Објаснива машина за засилување (Explainable Boosting Machine, EBM)

Овој пристап гради модел што сумира придонес по секоја особина и по неколку клучни парови. Така се добива прецизна и лесно читлива проценка на W_i , и кога влезовите мешаат броеви и категории [124]. Како формален запис, ова се изразува со сума од

константа, поединечни функции по особина и ограничен број интеракциони функции по парови, прикажано во следниот израз:

$$y(x) = b + \sum_j f_j(x_j) + \sum_{(j,k) \in \text{Pairs}} f_{jk}(x_j, x_k) \quad 6.18$$

Каде b е константа, $f_j(x_j)$ е придонес од поединечна особина j , а $f_{jk}(x_j, x_k)$ е придонес од избрана интеракција помеѓу j и k . Множеството Pairs е мало и контролирано, со цел моделот да остане читлив. На овој начин секој дел од прогнозата може да се види и објасни, а промените во W_i директно се врзуваат со конкретни опсези на PGDS, SSD_H, SSD_V и индикатори за состојба на коловоз.



Слика 6.13 Објаслива машина за засилување (EBM)

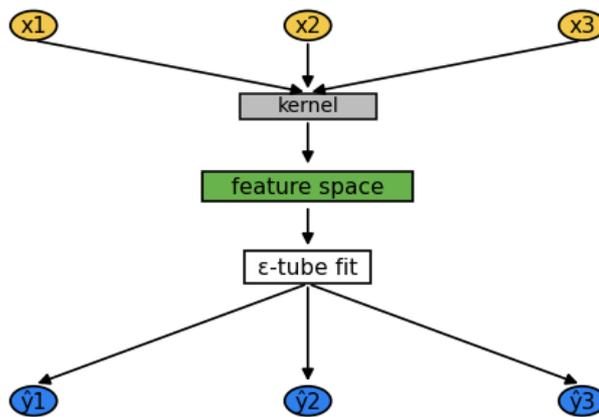
Слика 6.13 го прикажува EBM преку шема на чекори. Секоја особина најнапред се разложува во бинови, од кои се учи нејзиниот облик на влијание (f_1 , f_2 , f_3), а потоа придонесите се собираат во заедничка проценка y .

Во моделот е применет ExplainableBoostingRegressor ($\text{max_bins} = 128$, $\text{max_leaves} = 3$, $\text{learning_rate} = 0.02$, $\text{outer_bags} = 4$, $\text{inner_bags} = 0$, $\text{random_state} = 42$). Овие поставки даваат доволно добри граници по секоја особина (max_bins), мали чекори за сигурно напредување (learning_rate) и дополнително упресечување за помалку шум во кривите (outer_bags). Со таква поставка, влијанијата на особините и нивните врски се јасни, а проценките на W_i остануваат стабилни и корисни без дополнително стандардизирање.

Примената на EBM овозможува да се согледа влијанието на секоја особина по интервали и да се поврзе со инженерската логика. Резултатот е стабилна и транспарентна проценка на тежински индекс на незгоди (W_i), погодна за анализа и споредба на делници.

6.3.12 Појмпорна векторска регресија (Support Vector Regression, SVR)

Методот обезбедува стабилна прогноза, така што ја гради врз мал број клучни точки од податоците. Предвидувањето се смета задоволително ако точките лежат во тесен појас околу линијата, па ситните отстапувања не се третираат како грешки. На границата на тој појас стојат највлијателните точки, познати како вектори за поддршка, и токму тие го одредуваат решението. Кога врската меѓу влезот и излезот е посложена, се применува јадрена функција (kernel) што ги пренесува податоците во повисока димензија и овозможува кривата на предвидување да се обликува пофлексибилно [125].



Слика 6.14 Појторна векторска регресија (SVR)

Шематскиот приказ на SVR покажува, како јадрената функција ги пресликува податоците во нов простор, потоа се поставува линија на предвидување со појас на толеранција ϵ . Точките на границите на појасот се вектори на поддршка и тие ја одредуваат прогнозата.

$$y(x) = \sum_{i=1..n_{sv}} a_i * K(x, x_i) + b \quad 6.18$$

Каде $K(x, x_i)$ е јадрена функција (во анализата RBF), a_i се научени коефициенти, x_i се вектори на поддршка, b е константа, а n_{sv} е нивниот број. Поголемо C значи построго прилагодување кон тренинг-податоците, а поголемо ϵ значи поширок појас и помал број на вектори.

За евалуација се користи SVR (kernel = "rbf", $C = 1.0$, epsilon = 0.1) во pipeline со стандардизирање научено само на тренинг-податоците. Со ваквата поставка моделот директно работи со мерките по делници (PGDS, SSD_H, SSD_V, состојба на коловоз, надморска височина, врнежи и др.) и опишува нелинеарни врски со W_i без дополнителни трансформации.

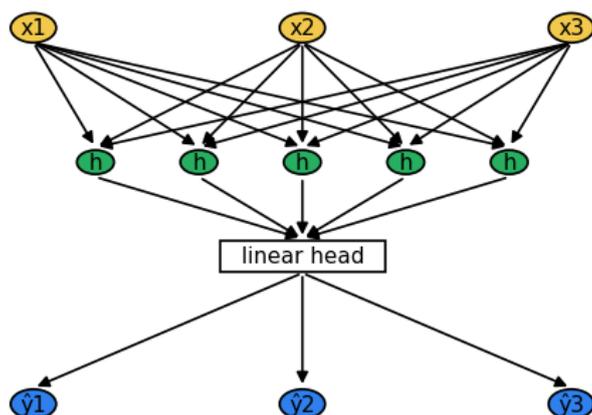
Таквата конфигурација обезбедува стабилна и континуирана прогноза, погодна за споредбена анализа со другите методи. Моделот особено е корисен кога сигналот е умерен, а бројот на влезни параметри поголем, при што влијанието на секоја променлива се толкува преку поддржувачките вектори и избраниот тип на јадрена функција.

6.3.13 Повеќеслоен њерцейџрон (Multilayer Perceptron, MLP)

За стабилна прогноза може да се применат неколку поврзани чекори, при што секој слој ја дообработува претходната слика. Во оваа улога, MLP користи неколку слоеви по ред за да ја изгради врската меѓу влезните показатели и тежинскиот индекс на незгоди (W_i). За сигурна работа влезните променливи прво се доведуваат на сличен размер со стандардизирање, за моделот еднакво да ги следи сите мерки [126].

$$y = W2 * \text{ReLU}(W1 * x + b1) + b2 \quad 6.18$$

Каде x е влезниот вектор, $W1$ и $W2$ се матрици на тежини за првиот и вториот слој, $b1$ и $b2$ се константи, а ReLU пропушта позитивни вредности и ги поставува негативните на нула, со што се додава едноставна нелинеарност.



Слика 6.15 Повеќеслоен перцептрон (MLP)

На слика 6.15 прикажана е шемата на повеќеслојниот перцептрон MLP. Во моделот влезните показатели x_1 , x_2 , x_3 се поврзуваат со скриени јазли h што создаваат покорисни комбинирани сигнали, а потоа линеарниот излез дава проценка на W_i . Врските укажуваат дека тежините се учат од податоците, а скриените јазли постепено извлекуваат поинформативни карактеристики.

Во практичната примена се користи MLPRegressor (`hidden_layer_sizes = (100,)`, `activation = relu`, `alpha = 0.0001`, `max_iter = 2000`, `random_state = 42`) во работен тек со стандардизирање. Со ова поставување моделот директно работи со мерките по делници (PGDS, SSD_H, SSD_V, состојба на коловоз и сл.) и опишува нелинеарни односи со W_i без дополнителни рачни трансформации.

Повеќеслојниот перцептрон дава стабилна прогноза и корисно го дополнува сетот дрвовидни методи, особено кога информацијата е умерена и дел од врските меѓу показателите и W_i не се линеарни.

6.4 Дефинирање на влијателни фактори

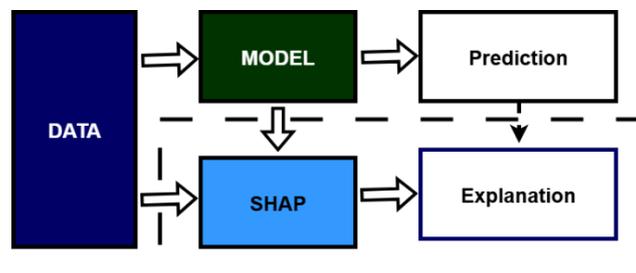
Во овој дел се утврдува кои параметри најмногу влијаат на W_i , под еднакви услови за сите техники и со проверка на стабилност. Анализата се спроведува на иста поделба тренинг-тест 80/20 (no-leakage), со ист протокол за подготовка. Сите трансформации се учат на тренинг-делот и проценките се прават на тест-делот.

Во анализата учествуваат само техники со $R^2 \geq 0.45$. Важностите по модел се добиваат приоритетно со SHAP, а кога тоа не е применливо или е нестабилно, се користи пермутациска важност (permutation importance). За споредливост меѓу моделите, важностите се нормализираат така што нивниот збир е 1. Дополнително се разгледуваат профили на ALE (accumulated local effects), каде сродните променливи по потреба се групираат за заедничко толкување.

Нормализираните важности од сите техники се пресметуваат како пондериран просек, при што тежините растат со R^2 над прагот 0.45. Потоа резултатите повторно се нормализираат да сумираат вредност 1 (по потреба се применува степен за посилен акцент на помоќни модели). Стабилноста се проверува со повторени примероци од тренинг-делот и се бележи зачестеноста на избор по параметар. Потоа се прикажуваат табела и дијаграм со важности по модел и заеднички, како рангирана листа за понатамошна примена.

6.4.1 SHAP анализа

SHAP (објаснувања со Шеплиеве придонеси – *SHapley Additive exPlanations*) се користи за разложување на вредноста што ја враќа моделот на збир од придонеси на поединечните влезни параметри. Методот овозможува читање по примерок и глобална важност пресметна со усреднување на придонесите. Се применува објаснувачки алгоритам за дрвовидни модели (TreeExplainer), бидејќи во анализата се користат дрвовидни модели и техники на засилување (boosting) [127]. Позадинскиот сет за очекувањата се зема само од тренинг-делот, а важностите се пресметуваат на тест-делот од поделбата 80/20 (no-leakage). Поради можна меѓусебна поврзаност меѓу дел од параметрите, паралелно се прегледуваат профили на акумулирани локални ефекти (ALE), и таму каде што е соодветно сродните променливи се прикажуваат како функционална група за заедничко толкување. По пресметката, важностите по модел се нормализираат, така што нивниот збир е вредност 1, со што стануваат директно споредливи меѓу различни техники.



Слика 6.16 Шематски приказ на SHAP алгоритмот

На шематскиот приказ од слика 6.16 е прикажан текот на постапката, каде од податоците се обучува модел и се добива прогноза, при што паралелно обучениот модел со истите влезови влегува во модулот SHAP каде се пресметуваат придонесите на поединечните параметри. Локалните објаснувања потоа се собираат во поле Explanation и може да се усреднат во глобални важности.

Равенката ја претставува излезната вредност, како базно ниво плус збир на поединечни придонеси.

$$y = s_0 + \sum_{j=1}^P s_j(x) \quad 6.19$$

Во равенката y е вредноста на моделот за примерокот x , s_0 е базно ниво добиено како очекување на моделот врз позадински сет од тренинг-делот, а $s_j(x)$ е придонесот на параметарот j за примерокот x .

$$I_j = (1/N_{\text{test}}) * \sum_{i=1}^{N_{\text{test}}} |s_j(x_i)| \quad 6.20$$

Овде I_j е глобална важност на параметарот j , добиена како просек од апсолутните придонеси на тест-делот.

$$\sum_j I_j = 1$$

Ова е нормализација по модел што овозможува директна споредба меѓу различни техники.

SHAP се применува за моделите што ја поминуваат границата $R^2 \geq 0.45$: Random Forest, Gradient Boosting, XGBoost, CatBoost, LightGBM, ExtraTrees и AdaBoost. Во продолжение следат клучните поставки што се користени при анализата.

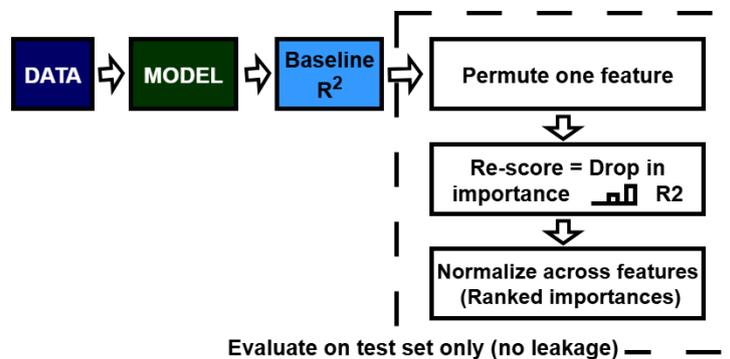
Табела 6.1 Техники и клучни хиперпараметри за SHAP-анализата ($R^2 \geq 0.45$)

Техника	Извор на важности	Клучни хиперпараметри (пример)
Random Forest	SHAP (TreeExplainer)	n_estimators = 500
Gradient Boosting	SHAP (TreeExplainer)	n_estimators = 400, max_depth = 3, learning_rate = 0.05
XGBoost	SHAP (TreeExplainer)	n_estimators = 300, max_depth = 4, learning_rate = 0.07, subsample = 0.9, colsample_bytree = 0.9
CatBoost	SHAP (TreeExplainer)	depth = 6, iterations = 600, learning_rate = 0.07
LightGBM	SHAP (TreeExplainer)	num_leaves = 7, n_estimators = 400, min_child_samples = 5, max_depth = 5, learning_rate = 0.005, subsample = 0.9, colsample_bytree = 0.6
ExtraTrees	SHAP (TreeExplainer)	n_estimators = 500
AdaBoost	SHAP (TreeExplainer)	n_estimators = 400

Овој опис најпрво го поставува методот SHAP и дијаграмот на постапката, потоа следуваат равенките со кратко објаснување на елементите. На крај се наведуваат конкретните конфигурации на моделите за повторлива анализа и споредливи резултати.

6.4.2 Пермутациона анализа

Пермутациона важност (permutation importance) се користи за мерење на промената на точноста кога вредностите на еден параметар се мешаат случајно, додека сите други остануваат непроменети. Постапката не зависи од модел и се изведува на тест-делот од поделбата 80/20 (no-leakage). За стабилност се прават повеќекратни повторувања и каде каде се пресметува средна вредност на резултатите, потоа важностите се нормализираат така што нивниот збир е вредност 1. Како мерка на точност се користи R^2 , со број на повторувања $R = 10$ [128].



Слика 6.17 Шематски приказ на пермутациона анализа (Permutation Importance)

На шематскиот приказ од слика 6.17 прикажан е текот на постапката при која од тест-сетот се добива почетна точност (Baseline R^2), потоа се пермутира еден параметар и се

пресметува падот на R^2 како мерка на важност. Се пресметува средна вредност на резултатите преку повеќе повторувања, при што на крај важностите се нормализираат и се добива рангирана листа.

Со следната формула објаснет е начинот на пресметување на просечната промена на R^2 при пермутирање на еден параметар.

$$I_j^{\text{raw}} = S_{\text{base}} - (1/R) * \sum_{r=1}^R S_j^{(r)} \quad 6.21$$

Каде S_{base} е точноста на моделот на тест-сетот без промени (R^2), $S_j^{(r)}$ е точноста по пермутирање на параметарот j во r -тото повторување, R е бројот на повторувања, а I_j^{raw} е непреработената важност за параметарот j .

$$I_j = I_j^{\text{raw}} / \sum_k I_k^{\text{raw}} \quad 6.22$$

Каде I_j е нормализирана важност на параметарот j , I_j^{raw} е непреработена важност за j , $\sum_k I_k^{\text{raw}}$ е збир на непреработените важности по сите параметри k , а k ги пробројува сите параметри во моделот. Нормализацијата овозможува споредливост меѓу техниките и посочува дека збирот на I_j по сите параметри е 1.

Пермутацијата се применува за техниките што ја поминуваат границата $R^2 \geq 0.45$ и за кои оваа постапка е соодветна за проценка на важностите. Во анализата се користат BaggingRegressor и Explainable Boosting Machine (EBM).

Табела 6.2 Техники и клучни хиперпараметри за пермутациона анализа

Техника	Извор на важности	Клучни хиперпараметри (пример)
Bagging Regressor	Пермутација (R^2 , $n_repeats = 10$)	$n_estimators = 300$
Explainable Boosting Machine (EBM)	Пермутација (R^2 , $n_repeats = 10$)	$max_leaves = 3$, $max_bins = 256$, $learning_rate = 0.01$, $outer_bags = 10$, $interactions = 0$

Со оваа постапка се добива прегледна и стабилна мерка на влијателност заснована на пад на точноста при контролирано пермутирање на влезовите, важностите се проценуваат исклучиво на тест-сет и се нормализираат за директна споредливост меѓу моделите.

6.4.3 Пресметка на средна вредности

Пресметката на средна вредност ги собира нормализираните важности од сите методи проценети на истата поделба тренинг-тест 80/20 (no-leakage). Во просекот влегуваат само модели што ја поминуваат границата на точност $R^2 \geq 0.45$. Овој пристап е вообичаена форма на пондерирано порамнување на модели, сроден на постапките за составни модели и моделирање со тежини според точноста [129, 130].

За секој модел се дефинира тежина што расте со неговата точност на тест-сет, а потоа тежините се нормализираат така што нивниот збир е 1. По избор се користи експонент $p \geq 1$ за да се нагласат помоќните модели:

$$w_m^{\text{raw}} = \max(0, R_m^2 - \tau)^p \quad 6.23$$

Каде w_m^{raw} е непреработена тежина за моделот m , R_m^2 е точноста R^2 на моделот m на тест-сет, τ е прагот за влез ($\tau = 0.45$), а p е експонент што го одредува акцентот.

$$w_m = \frac{w_m^{\text{raw}}}{\sum_k w_k^{\text{raw}}} \quad 6.24$$

Каде w_m е нормализираната тежина, а во именителот е збирот од непреработените тежини на сите модели k , со што важи $\sum_m w_m = 1$.

Збирната важност по параметар се добива како тежинска сума од важностите по модели:

$$I_j = \sum_m w_m \cdot I_{m,j} \quad 6.25$$

Каде $I_{m,j}$ е важноста на параметарот j од моделот m , нормализирана така што $\sum_j I_{m,j} = 1$ за секој модел m . Оттука следи:

$$\sum_j I_j = 1$$

На практичен план, конечната хиерархија се добива со рангирање на параметрите според I_j . Доколку е потребно, се избира најмал број параметри што обезбедуваат целна покриеност (кумулативна важност од 90%) или се користи зависноста помеѓу број на параметри и R^2 за да се одреди разумна точка на запирање, при што сите чекори се изведени на истата поделба тренинг-тест 80/20.

6.5 Интегративна формула - конструкција и калибрација

Во оваа фаза се изведува формула што го пресметува индексот W_i на начин што секој може да го провери и повтори. Влезните променливи се земаат од претходно подготвената листа на најважни фактори. Постапката се изведува на истата поделба 80/20, при што на тренинг-делот се учат сите правила и стандардизација, а на тест-делот се проверува резултатот.

За да се опфатат закривувања и меѓусебни влијанија, се користи полиномна основа од втор степен со поединечни термини, нивни квадрати и сите парни комбинации. Бројот на фактори не се погодува однапред. Наместо тоа, се прават проверки за технички издржани комбинации, се прави модел за секое можно множество од k фактори, се учи на тренинг-делот и се оценува со R^2 на тест-делот. Така се добива крива, која покажува колку фактори, даваат најдобар резултат. Главниот модел е обична линеарна регресија со полиномна основа. За споредба и стабилност се користат уште Ridge и Lasso. Се избира решението со највисок R^2 на тест-делот. Ако повеќе решенија се слични, се зема решение со помал број фактори, за формулата да биде покуса и постабилна.

Кога ќе се изберат факторите, на тренинг-делот се пресметуваат конечните коефициенти и формулата се прикажува со константен член и полиномски членови. За директна примена на нови податоци се објавуваат средина и стандардна девијација за секој влез, за да може правилно да се стандардизира. Потоа се прават кратки проверки на

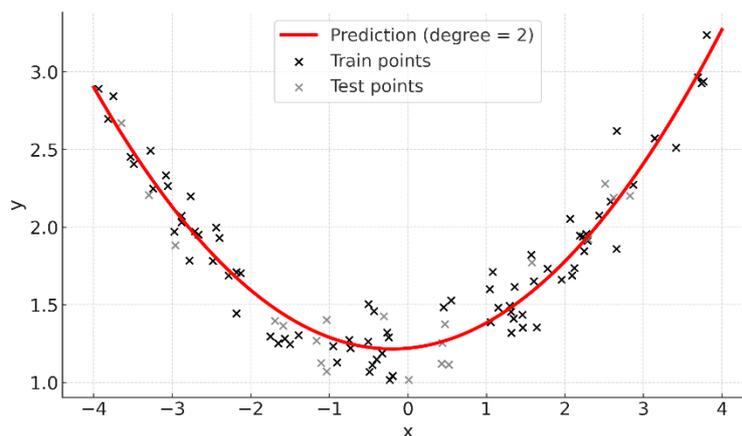
остатоците и влијателните точки, а стабилноста се тестира со неколку малку порамнети поделби 80/20, за да се потврди дека изборот не зависи од една поставка.

Формулата дава континуирана вредност за секоја делница и овозможува рангирање. Ако треба одлука, се задаваат прагови и категории на ризик според распоредот на вредностите на тест-делот. Се прикажуваат ROC график за одвојување, калибрационен график за усогласеност и проценка на нето корист при различни граници. Прагот се избира каде моделот најдобро ги одвојува високите од ниските ризици, а во извештајот се наведуваат погрешните сигнали и пропуштените случаи.

Решението вклучува криви на точност во зависност од бројот на фактори, избраното k со конкретни влезови, целата формула со коефициенти и параметри за стандардизација, како и резултати за дискриминација, калибрација и оперативни граници.

6.5.1 Полиномна регресија

Полиномната регресија е линеарен модел по коефициенти што овозможува нелинеарна врска со тежинскиот индекс на незгоди (W_i), преку проширување на влезовите со полиномски термини. Во оваа примена, се користи втор степен за да се опфатат закривености и парни интеракции меѓу факторите, при што моделот останува разбирлив и со јасна интерпретација на придонесите [131].



Слика 6.18 Полиномна регресија ($deg = 2$) (Hastie, 2009)

Дијаграмот на слика 6.18 прикажува стабилно вклопување на тренинг делот и соодветна генерализација на тест делот. Црвената линија ги прикажува предвидувањата од моделот, црните точки се тренинг-податоци, а сивите точки се тест-податоци според 80/20 поделба.

Формално, моделот се запишува како:

$$W_i = \beta_0 + \sum_{j=1}^m \beta_j \phi_j(z_i) + \varepsilon_i \quad 6.25$$

Каде $\phi_j(z_i)$ се полиномски термини од втор степен формирани од стандардизираните влезови z_i – основни, квадрати и парни интеракции – β_j се коефициенти, а ε_i е резидуалната грешка. Оваа форма обезбедува единствена формула за пресметка на W_i што директно се применува на нови податоци за оценување на делници.

Полиномната основа е од втор степен и ги опфаќа квадратните термини и сите парни интеракции на избраните фактори. Врз оваа основа коефициентите се пресметуваат со метода на најмали квадрати на тренинг делот, а стандардизацијата се учи само на тренинг и непроменета се применува на тест за да нема пренос на информации. Бројот на фактори k се одредува со насочено пребарување по подмножества на истата 80/20 поделба, при што како главна мерка се користи R^2 , а по потреба се следат и MAE и RMSE, со цел да се одржи добар баланс меѓу точност и едноставност на моделот.

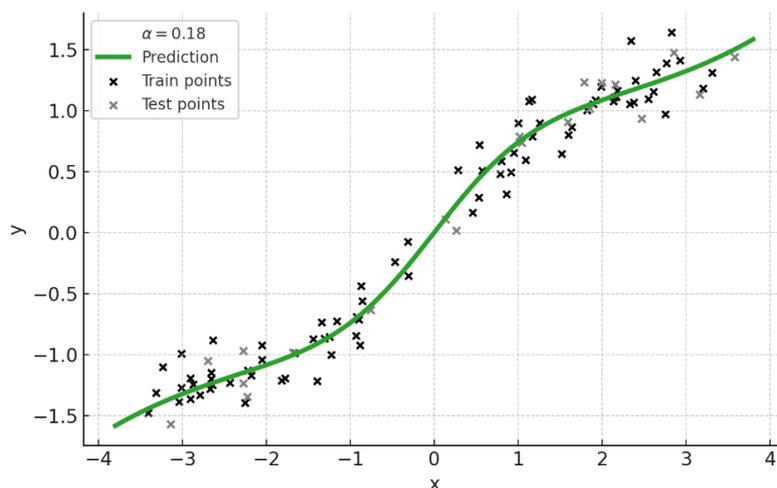
6.5.2 Lasso регресија (L1-регуларизација)

Lasso регресијата е тип на линеарен модел со L1-регуларизација кој, освен што ги намалува преголемите коефициенти, автоматски ги отстранува оние што немаат значаен придонес. Така, формулата станува пократка и појасна, бидејќи остануваат само најважните „термини“ што точно влијаат на резултатот. Овој метод посебно е корисен, кога во моделот има многу можни фактори, а само дел од нив реално придонесуваат за предвидувањето. Во ситуации со ограничен број податоци и меѓусебно поврзани фактори, Lasso ја намалува варијансата и го спречува преобучувањето. На тој начин моделот станува постабилен, поедноставен и подобро ги опфаќа вистинските врски меѓу факторите и индексот W_i [132].

Формулата што ја опишува Lasso регресијата, е:

$$\hat{\beta} = \arg \min_{\beta} \frac{1}{n} \sum_{i=1}^n \left(W_i - \beta_0 - \sum_{j=1}^m \beta_j \phi_j(z_i) \right)^2 + \alpha \sum_{j=1}^m |\beta_j| \quad 6.26$$

Каде $\phi_j(z_i)$ се полиномски термини од втор степен конструирани од стандардизирани влезови, β_0 не се регуларизира, а $\alpha \geq 0$ ја контролира јачината на L1-регуларизацијата и степенот на занемарување на термините.



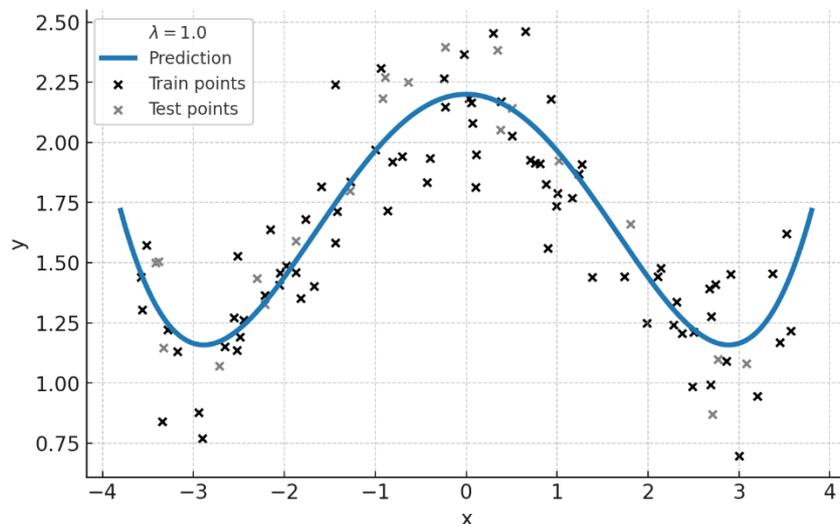
Слика 6.19 Lasso со полиномска основа (deg = 2) (Kitali, 2019)

На слика 6.19 прикажано е предвидувањето на Lasso со $\alpha = 0.18$, тренинг-точките се во црна боја, тест-точките во сива. Кривата е израмнета на краевите од опсегот и тежнее кон попусти облици преку занемарување на дел од термините, што визуелно ја разликува техниката од истата полиномна регресија без регуларизација.

За потребите на оваа анализа, параметрите се адаптирани на полиномска основа од втор степен со сите парни интеракции и 80/20 поделба со фиксно семе `random_state = 42`, стандардизацијата се учи на тренинг-делот и непроменета се применува на тест-делот, α се одбира со LassoCV преку логаритамска мрежа од приближно 10^{-4} до 10^2 со околу 30 точки, а бројот на фактори k се одредува со насочено пребарување по подмножества на истата поделба, при што R^2 на тест-делот е главниот критериум. Ова поставување дава компактна и стабилна формула со јасна и разбирлива структура на избраните термини, при што добиените коефициенти се директно применливи за рангирање на делници според очекуваната тежина на последиците.

6.5.3 Ridge регресија (L2-регуларизација)

Ridge регресијата е тип на линеарен модел кој ја прави формулата постабилна и сигурна. Кога во податоците има многу фактори што се меѓусебно слични или поврзани, обичната регресија може да даде преголеми и нестабилни коефициенти. Ridge ја оптимизира оваа појава, со тоа што, контролирано ги намалува тие големи вредности и на тој начин ја стабилизира врската меѓу променливите. Овој пристап е корисен кога моделот почнува да учи премногу специфични детали од податоците (преобучување), па не може добро да предвиди нови случаи. Со Ridge моделот станува попрост, има помала променливост и подобро ги пренесува врските на нови податоци. Така се добива постабилна и поиздржана проценка, што може сигурно да се користи за анализа и предвидување. [133].



Слика 6.20 Ridge со полиномска основа ($deg = 2$) (Garefalakis, 2022)

На слика 6.20 даден е општ приказ на предикција со Ridge при $\lambda = 1.0$ на примерен сет на податоци. Линијата на моделот е израмнета на краевите, црните точки го прикажуваат тренинг составот, а сивите тест составот, што покажува како L2 регуларизацијата ги намалува големите осцилации и ја стабилизира функцијата. Овој приказ служи за визуелно објаснување на ефектот на L2-регуларизацијата и не носи заклучоци за конкретни делници.

Коефициентите се добиваат во затворена форма:

$$\hat{\beta} = (\Phi^T \Phi + \lambda I)^{-1} \Phi^T W \quad 6.27$$

Каде Φ е дизајн-матрицата со полиномските термини од втор степен конструирани од стандардизирани влезови, β ги содржи коефициентите на кои се применува L2-регуларизација, пресретот I не се регуларизира, а $\lambda \geq 0$ го контролира интензитетот на регуларизацијата.

За потребите на оваа анализа параметрите се поставени, така што се користи полиномска основа од втор степен со сите парни интеракции, со поделба 80/20 и `random_state = 42`. Стандардизацијата се учи на тренинг делот и непроменета се применува на тест делот, λ се избира со RidgeCV преку логаритамска мрежа од 10^{-4} до 10^2 со околу 30 точки, а бројот на фактори k се одредува со насочено пребарување по подмножества на иста поделба со критериум R^2 на тест делот. Овој избор обезбедува контролирана сложеност и постабилна калибрација на крива на предикција.

6.5.4 Тестирање и валидација

Интегративната формула се избира како регресија што дава највисоко R^2 на тест-делот при фиксна 80/20 поделба. Постапката започнува со три највлијателни параметри според однапред добиеното рангирање, се тренираат сите три техники (полиномна OLS, Ridge и Lasso) и се бележи R^2 на тест-делот. Потоа бројот на параметри се зголемува на 4, па 5 и така последователно до 23, при што за секое k се оценуваат трите техники и се запишува тест- R^2 . За секое k се задржува регресијата што дава најголемо R^2 , а конечната формула се дефинира како комбинација од техника и број на параметри k што реализира највисока детерминација на тест-делот, ако резултатите се еднакви се преферира помал k . Рангирањето на факторите и 80/20 протоколот се усогласени со работното подесување за оваа анализа, при што таргет е тежинскиот индекс на незгоди (W_i). Полиномната основа е од втор степен со квадрати и сите парни интеракции, стандардизацијата се учи на тренинг и непроменета се применува на тест, а се користат вградените постапки OLS, RidgeCV и LassoCV.

За финално потврдување на стабилноста се прави уште една валидација само за избраната формула, при што 80/20 поделбата се повторува со неколку различни семиња. Се прикажуваат средната вредност и нормалниот интервал на тест- R^2 , како и дијаграм, предвидено наспроти набљудуваното на тест делот со линија $y = x$ и разлики во однос на предвидувањата, при што не се очекува видлив образец и просекот е блиску до нула. По потреба се додаваат MAE и RMSE за да се потврди калибрацијата и стабилноста на формулата.

За потребите на оваа анализа, параметрите се прилагодуваат на полиномска основа од втор степен со сите парни интеракции и фиксна почетна состојба `random_state`, при што стандардизацијата се учи на тренинг делот и непроменета се применува на тест делот. Потоа λ за Ridge се избира со RidgeCV на логаритамска мрежа од 10^{-4} до 10^2 , α за Lasso со LassoCV со трикратно вкрстување на сличен опсег, а изборот на k се води според тест R^2 на истата поделба. Овој протокол овозможува споредба под исти услови меѓу трите техники и избор на компактна и стабилна формула со соодветен број параметри и највисока детерминација.

6.6 Машинско учење – конфигурација на моделите и избор

Овој дел го опишува пристапот за градење и селекција на модели насочени кон точна проценка на W_i . Параметрите се рангирани по важност, а бидејќи исходот е нумеричка вредност, се користат модели за непрекинати големини. Сите правила за подготовка се учат само на тренинг-делот, а потоа истите правила се применуваат на тест-делот, со што се избегнува протекување на информации.

Прво се одлучува колку параметри ќе влезат во моделот. Се почнува со сите 23, а потоа чекор по чекор се отстранува најмалку важниот. За секој број од 23 до 5 се тренира модел на тренинг-делот и се мери точноста со R^2 на тест-делот. На тој начин се одредува со колку параметри моделот постигнува највисок R^2 на истата поделба 80/20 и таа големина се усвојува. Притоа се подесуваат клучните поставки. Кај моделите со дрва, како Random Forest, Extra Trees и Bagging, се прилагодуваат број на дрва, длабочина, дел од карактеристики по јазол и минимум примери во лист. Кај засилените дрва, како Gradient Boosting, LightGBM, XGBoost и CatBoost, се задава чекор на учење, број повторувања со рано запирање, длабочина или број лисја, делумно земање редови и карактеристики, како и заштитни поставки за да нема преобучување. Кај AdaBoost се земаат во предвид длабочината на базното дрво, бројот на чекори и големината на чекорот на учење. Кај EBM се усогласуваат број на повторувања, чекор на учење, број интервали и лисја, број на надворешни порамнувања и број на интеракции. Се испробуваат неколку практични комбинации околу соодветни почетни вредности, за да задржат прегледност и повторливост.

Главен критериум е R^2 на тест-делот со фиксна почетна состојба, за резултатите да бидат повторливи. Паралелно се прикажуваат MAE и RMSE како мерки на грешка. Кога две поставки даваат сличен R^2 , предност добива пократкиот модел со помал број параметри. За стабилност се проверува изборот на број на параметри со уште неколку блиски поставки, а кај методите со рано запирање се следи како се намалува валидационата грешка низ повторувања.

На крај, се прикажува дијаграм со криви на R^2 во зависност од бројот на параметри за сите модели и табела со избраниот број параметри по модел, R^2 пред и по прилагодување и точниот список параметри во финалната верзија. Овие финални конфигурации потоа се користат за анализа на важности на тест-делот и за објективна споредба меѓу моделите.

6.6.1 Хиперпараметри и тренинг-поставки

Моделите се тренираат на иста поделба со фиксна почетна состојба (random_state), а бројот на влезни променливи се избира одозгора надолу со постепено намалување според рангирањето. Подесувањето на хиперпараметрите се прави преку кратки, практични комбинации околу разумни почетни вредности. Кај методите со засилено учење се користи прерано запирање врз мала валидациона подгрупа од тренинг делот.

Табела 6.3 Хиперпараметри и тренинг-поставки по модел

Модел	Влезни сетирања	Опсег за подесување
AdaBoostRegressor	random_state=42	base tree max_depth \in {2,3,4}, n_estimators \in {400,800,1500}, learning_rate \in {0.03,0.1,0.2}
ExtraTreesRegressor	n_estimators=600, random_state=42, n_jobs=-1	n_estimators \in {1000,2000}, max_features \in {0.5,0.7,1.0},

		max_depth ∈ {None,20,40}, min_samples_leaf=1
BaggingRegressor	n_estimators=300, random_state=42, n_jobs=-1	n_estimators ∈ {300,600,1000}
GradientBoostingRegressor	n_estimators=600, learning_rate=0.05, max_depth=3, random_state=42	loss ∈ {squared_error, huber}, learning_rate ∈ {0.02,0.03,0.05}, n_estimators ∈ {1500,3000,5000}, subsample=0.8, min_samples_leaf=2, validation_fraction=0.2, n_iter_no_change=60
RandomForestRegressor	n_estimators=600, random_state=42, n_jobs=-1	n_estimators ∈ {1000,2000}, max_features ∈ {0.5,0.7,1.0}, max_depth ∈ {None,20,40}, min_samples_leaf=1
LGBMRegressor	n_estimators=400, learning_rate≈0.0050, num_leaves=7, max_depth=5, min_child_samples=5, subsample=0.9, colsample_bytree=0.6, reg_alpha=0.0, reg_lambda=1.0, random_state=42	tighten with n_estimators≈5000 with early stopping, learning_rate ∈ {0.01,0.02,0.05}, num_leaves ∈ {15,31,63}, subsample=0.8, colsample_bytree=0.7
XGBRegressor	n_estimators=800, learning_rate=0.05, max_depth=4, subsample=0.9, colsample_bytree=0.7, reg_lambda=1.0, reg_alpha=0.0, eval_metric="rmse", random_state=42, tree_method="hist"	tighten with max_depth ∈ {3,4,5}, learning_rate ∈ {0.03,0.05,0.08}, n_estimators ∈ {1200,2000,3000}, subsample≈0.85, colsample_bytree=0.7, reg_lambda≈1.5, early stopping ≈120 rounds
CatBoostRegressor	iterations=1200, learning_rate=0.05, depth=6, loss_function="RMSE", random_seed=42	tighten with depth ∈ {6,8,10}, learning_rate ∈ {0.03,0.05}, iterations≈4000, l2_leaf_reg≈3.0, subsample≈0.8, early stopping ≈200
ExplainableBoostingRegressor (EBM)	approx. max_rounds=600, max_leaves=3, max_bins=128, learning_rate=0.02, outer_bags=4, interactions=10	tighten with max_rounds ∈ {800,1200,2000}, max_leaves=4, max_bins=256, learning_rate≈0.03, outer_bags=6, interactions≈15

Фокусот е ставен на јасни комбинации на поставки за секој модел, со ист начин на подготовка на податоците и со контролирана случајност. На тој начин моделите може лесно да се повторуваат и да се проверуваат. Кога треба да се избере меѓу неколку решенија, изборот се прави според тоа, колку се стабилни и како се однесуваат конечните предвидувања. Ако резултатите се блиски, предност има поедноставното решение.

6.6.2 Тестирање и валидација

Оценувањето и проверката на моделите се прават со поделба 80/20, каде што 80% од податоците се користат за учење, а 20% за проверка. Случајноста е фиксирана, а подготовката на податоците се врши без протекување на информации. Главната цел е да

се избере техниката што дава највисок R^2 со соодветен број параметри. Од листата со параметри се земаат различно долги делови и за секој дел се обучуваат сите модели на иста поделба. Потоа, точноста на тест делот се бележи и се избира комбинацијата од модел и број параметри со најдобар резултат. Кога има повеќе, речиси еднакви решенија, се избира пократкиот модел, а таа поставка се смета за финална.

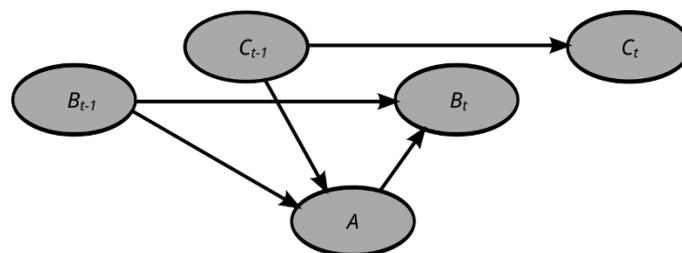
За стабилност се прави кратка проверка со неколку различни почетни состојби, а потоа се прикажува споредба помеѓу предвидените и вистинските вредности. Разликите не треба да покажуваат никаква шема, при што просечната вредност треба да биде блиску до нула, што покажува дека моделот правилно е усогласен.

Покрај тоа, се прави и дополнителна проверка за класификација. Се одредува праг за висок ризик, како најгорните 20% од индексот W_i , и се гледа колку добро моделот ги открива овие делници. Се користат неколку мерки за тоа: колку точно ги разликува повисоките ризици (AUROC и PR-AUC) и колку добро ги рангира првите 10% од најризичните делници (Precision@10%, Recall@10%, Lift@10%), при што резултатите се прикажуваат во вид на график.

Поставките на моделите се одржуваат едноставни и практични. Кај моделите што учат постепено, како засилените дрва, се применува рано запирање врз мала валидациона група, за да не продолжи учењето кога нема подобрување. Се поставува XGBClassifier со околу 1200 дрва, длабочина 3, стапка на учење 0.05 и неколку заштитни параметри за стабилност. Делниците со висок ризик се одредуваат со праг на горните 20% од индексот W_i . Овие поставки се користат за да се пресметаат сите мерки и да се прикажат графици што покажуваат како моделот ги рангира ризичните делници.

6.7 Општи карактеристики на анализата со Баесови мрежи

Баесовите мрежи се насочени ациклични структури (Directed Acyclic Graph) во кои јазлите претставуваат променливи, а насочените врски означуваат условни зависимости. Заедничката распределба се дели според врските меѓу јазлите, па може да се пресметуваат веројатности, да се прават сценарија со дадени податоци и полесно да се види како една променлива влијае на друга [134].



Слика 6.21 Основен пример за Баесова мрежа (Carrodano, 2024)

На слика 6.21 прикажан е едноставен пример каде секоја точка има насочена врска кон друга, а преку тие врски целата веројатност се запишува како производ од помали, условни веројатности. Со додавање нов податок, мрежата ги ажурира сите поврзани вредности според насоката на стрелките, што овозможува да се пресметаат нови веројатности и да се испитаат можни сценарија.

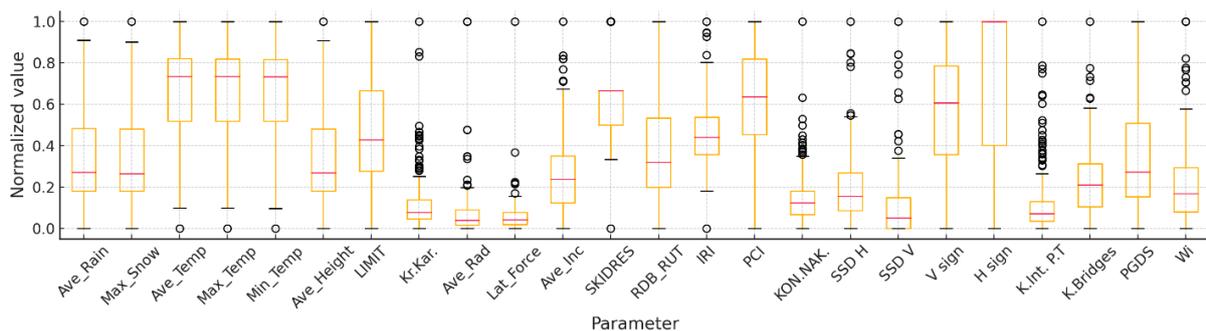
Во анализата најнапред се подготвуваат податоците, се проверува дали недостигаат вредности или има невообичаени отстапувања. Потоа вредностите се претвораат во три

групи со приближно еднаков број на записи и со однапред зададени прагови. Овие групи се користат за учење на мрежата. Структурата се гради со три пристапи, GTT, PC и TAN, а паралелно се проценуваат и условните веројатности за секоја врска. Потоа се проверува колку добро мрежата ги одвојува различните состојби и колку се точни нејзините пресметки. Проверката се спроведува со повеќе поделби на податоците, за да се види дали резултатите се стабилни. Сите чекори се изведени во GeNIe 5, што овозможува јасен приказ на мрежите и споредба на резултатите од различните пристапи.

6.7.1 Дискретизација

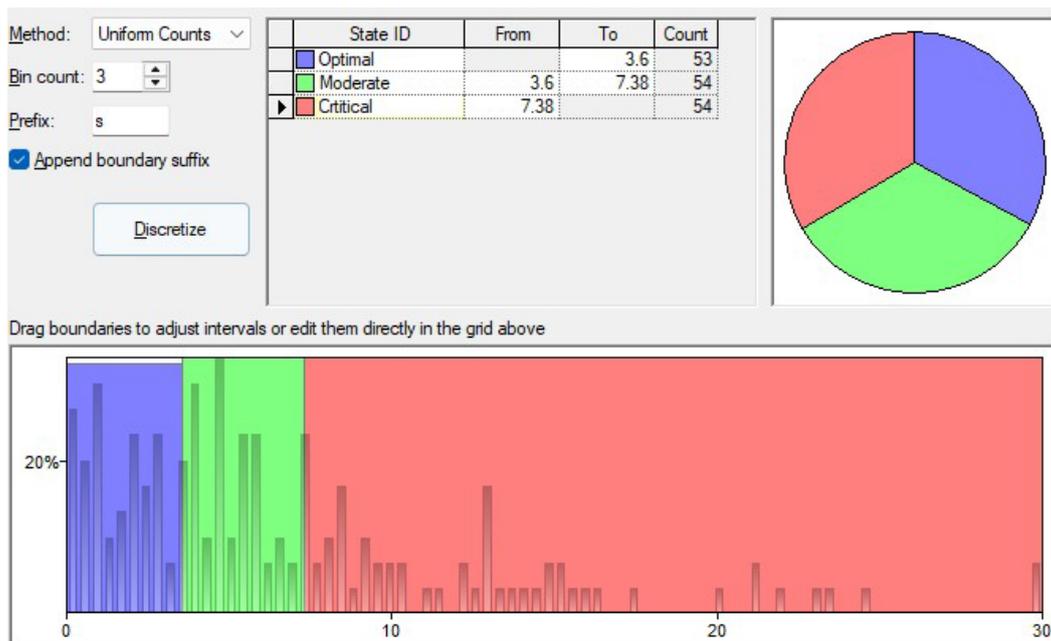
Дискретизацијата е постапка со која континуираните податоци се претвораат во неколку групи за полесна анализа. Кај Баесовите мрежи ова е многу важно, затоа што тие не работат со броеви на цел спектар, туку со групи или категории. Наместо секоја вредност да се третира посебно, податоците се делат во три состојби со приближно еднаков број записи (equal-frequency). На овој начин, екстремните вредности немаат преголемо влијание и пресметките на веројатностите стануваат постабилни [135]

На сликата 6.22 прикажан е боксплот дијаграм, кој ја покажува средината на податоците, нивниот опсег и можните отстапувања. Средните вредности се означени, што овозможува да се види дали податоците се рамномерно распределени или има поместување на една страна. Ова помага при одредување на границите со кои се делат групите.



Слика 6.22 Нормализиран боксплот со средни вредности

На сликата 6.23 прикажано е како се распределуваат записите по трите состојби, откако ќе се направи дискретизацијата. Најчесто секоја состојба содржи приближно еднаков број записи, што значи дека податоците се добро поделени и дека секоја група има доволно примери за учење. Ваквата рамнотежа помага моделот да не фаворизира една состојба, туку да учи стабилно во сите три. Малите отстапувања се појавуваат кога има многу исти вредности или врзани рангови, но влијанието е мало бидејќи распределбата останува близу еднаква. Ако на одредено место има поголема група исти вредности, праговите се одбираат така што, групите и понатаму остануваат практично урамнотежени.



Слика 6.23 Распределба на зависиие во три состојби во дискретизација

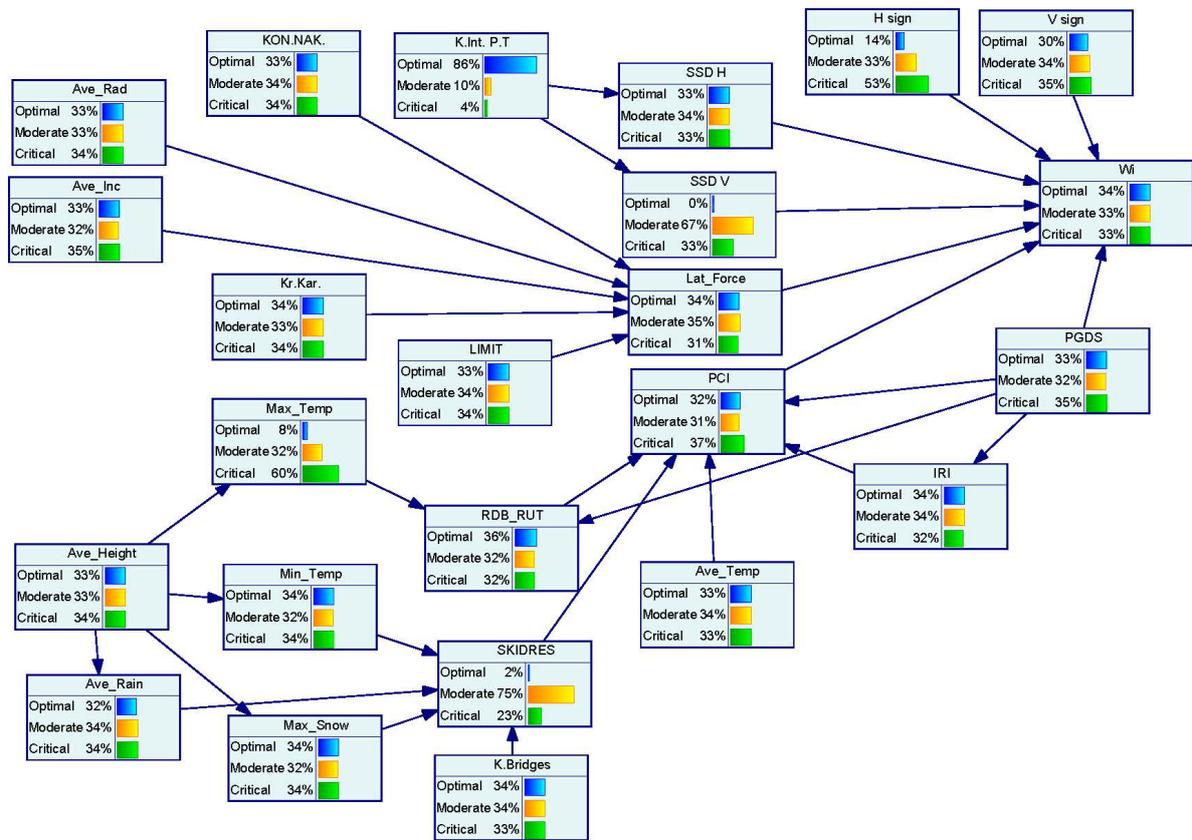
Границите меѓу групите се определуваат однапред според податоците за учење и остануваат исти во целата анализа, за да нема протекување на информации. Веќе дискретизираните податоци остануваат во своите природни категории, а вредностите што недостасуваат се средуваат уште пред дискретизацијата. Истите правила се применуваат за сите влезни фактори и за резултатот W_i , што овозможува мрежата да учи на ист начин за сите променливи. Со ваков пристап, Баесовите мрежи даваат постабилни резултати и полесно се читаат врските меѓу параметрите.

6.7.2 Greedy Thick Thinning (GTT)

GTT е метод за учење на структурата на Баесова мрежа. Прво се подготвуваат податоците и се проверува кои променливи се поврзани, потоа се одредува насоката на зависноста меѓу нив, при што се избегнуваат кругови за мрежата да остане ациклична. Методот работи етапно, додава врски кога има доволно доказ дека постои зависност и отстранува врски кога зависноста е слаба или се објаснува преку други променливи [136]. Потоа се создаваат насочени врски што водат кон целната променлива, односно тежинскиот индекс на незгоди (W_i).

Мрежата е уредена, така што сите влезни параметри имаат пат до W_i , понекогаш директно, понекогаш преку една или повеќе посредни врски. На овој начин лесно се чита кои фактори делуваат непосредно врз W_i и кои влијаат преку други параметри. Кога ќе се додаде нова информација во некој јазол, веројатностите по патеките се ажурираат, па може да се процени како промената се пренесува низ мрежата.

Овој пристап помага да се разложи сложениот систем на разбирливи чекори: кои параметри се клучни, кои се секундарни и каде се создаваат синџири на влијание. Така толкувањето на резултатите е поедноставно, а препораките за мерки може да се насочат прво кон врските што се најблиску до W_i , а потоа кон оние што делуваат посредно.



Слика 6.24 Сиркуиџура на Басова мрежа добиена со алгоритмот GTT

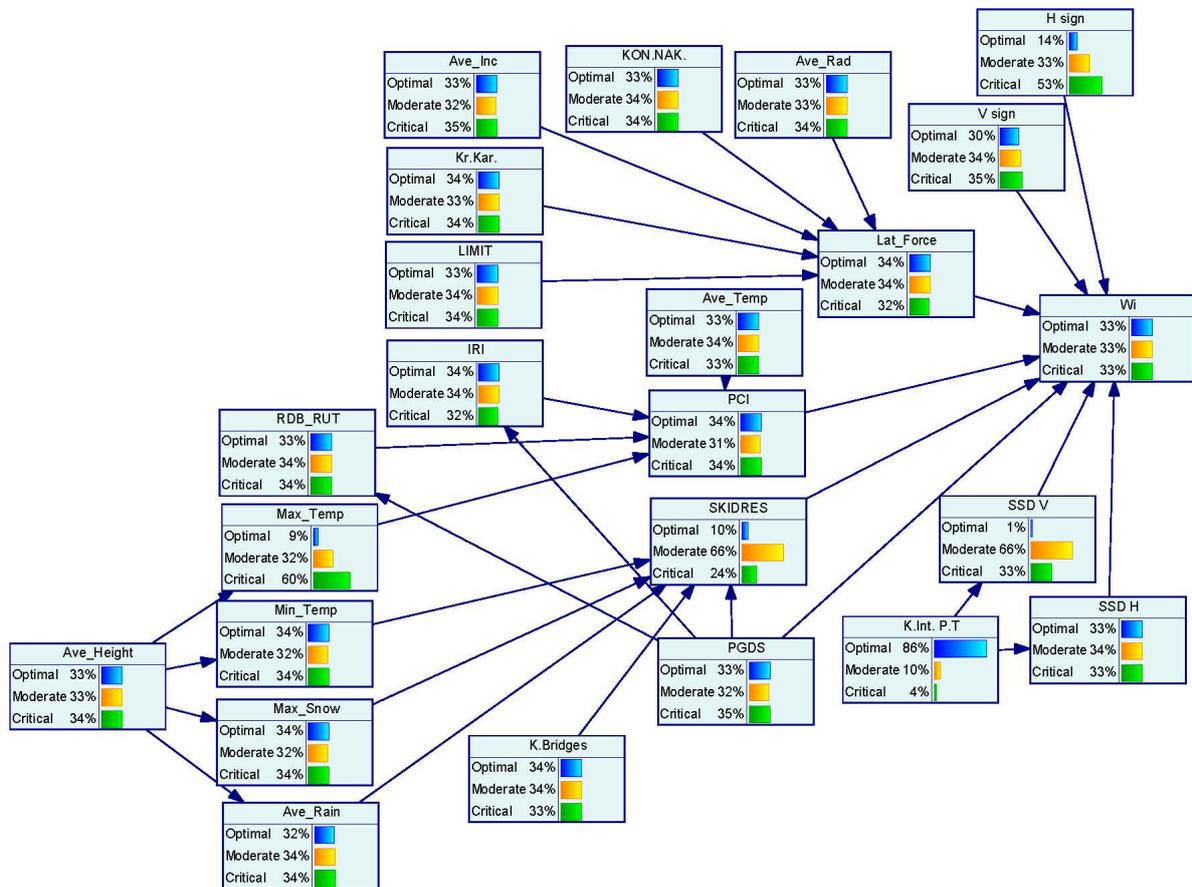
Во оваа поставка Wi е експлицитно целен јазол. Најизразени директни јазли се Lat_Force, PCI, PGDS, SSD_H, SSD_V, H_sign и V_sign, додека индексот на попречна нерамност IRI влегува посредно преку PGDS наместо да се поврзе директно со Wi. На слика 6.24 прикажани се патеките каде факторите од околина како Max_Snow, Min_Temp и Ave_Rain се собираат во SKIDRES, продолжуваат преку RDB_RUT кон PCI и потоа кон Wi, додека геометриските влијанија следат покомпактна траса: Ave_Height, Ave_Rain, SKIDRES и Ave_Rad, Lat_Force, Wi. Препознаени се две доминантни патеки на ризик, патека поврзана со состојба на коловоз и нерамнини и патека поврзана со прегледност и опрема, кои се засилени со влијанијата на силата од бочен потисок. Овие врски ја даваат методолошката подлога за читање на локалните и составените ефекти без внесување на мерки на точност.

Во анализата се користи критериум заснован на оценка со фиксни поставки и фиксна почетна состојба (random_state) за повторливост, без дополнителни ограничувања надвор од наведените правила на домен.

6.7.3 PC (Peter-Clark)

PC пристапот може да се претстави како метод што прво проверува дали две променливи се навистина поврзани или нивната врска се појавува случајно кога се земаат предвид и останатите фактори. Ако се утврди дека не се зависни, таа врска се отстранува. Ако се зависни, врска се задржува и потоа се определува насоката, односно кој фактор влијае на кој.

Овој пристап корисен е кога има многу можни врски меѓу променливите, бидејќи успева да ги издвојува само оние што се повторуваат и постојано се јавуваат. Таквите врски најчесто укажуваат на вистинска причинско-последична зависност, па мрежата што се добива со РС пристапот е појасна и полесна за толкување [137].



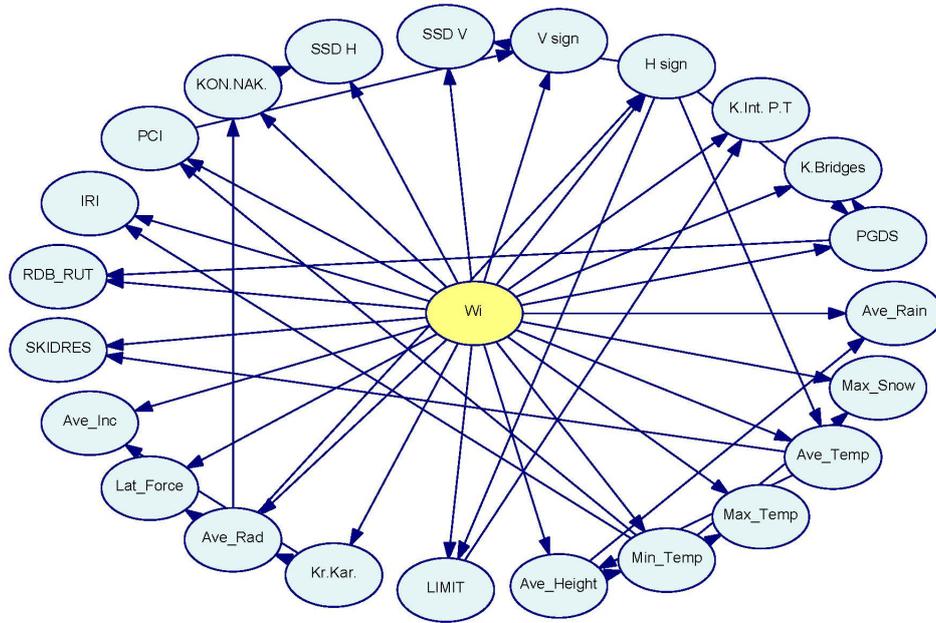
Слика 6.25 Структура на Басова мрежа добиена со алгоритмот РС

Во добиената структура на слика 6.25 се прикажани повеќе променливи кои се поврзуваат не само со Wi, туку и меѓусебно, при што Max_Snow се поврзува со Min_Temp и Ave_Rain, PCI се поврзува со PGDS и SKIDRES, а SSD_H, SSD_V и K.Int.P.T образуваат група што упатува на комбинирани влијанија од прегледност и патна опрема. Директни јазли на Wi се Lat_Force, PCI, PGDS, SKIDRES, SSD_V, H_sign и V_sign, што укажува на поширок сет на непосредни предиктори. Во рамки на поставката се издвојуваат две доминантни траси на ризик: геометрија и коловоз, каде радиусот на кривини, наклонот, силата од бочен потисок и отпорноста на лизгање влијаат на тежинскиот индекс на незгоди, и група околина и прегледност, каде климатските екстрими ги менуваат условите за прегледност за запирање по хоризонтала и вертикала и квалитетот на хоризонталното обележување и вертикалната сигнализација, со посредно зголемување на ризикот. Овој пристап открива поширока и комплексна мрежа на фактори, без потреба од навлегување во мерки на точност.

При РС се користи тест за условна независност со фиксно ниво на важност, а ориентацијата на насоките се спроведува со стандардни правила, со што се обезбедува усогласена структура низ поставките.

6.7.4 TAN (Tree-Augmented Naive Bayes)

Овој пристап ја комбинира разбирливоста на Naive Bayes со едноставно дрво околу целната класа. Секој предиктор, покрај врската со класата, добива уште една врска кон најблискиот сроден предиктор, избрана според тоа колку информации споделуваат. Потоа, околу класниот јазол се оформува плитко дрво со мала сложеност, кое подобро ги опишува зависностите меѓу променливите и остава моделот лесен за читање и толкување [138].



Слика 6.26 Структура на Баесова мрежа добиена со алгоритмот TAN

Слика 6.26 го дава генерираниот распоред, каде класниот јазол W_i е заеднички на сите предиктори, а секој предиктор има уште една врска од друг предиктор, па се оформува плитко дрво што радијално се разгранува од целиот јазол и овозможува едноставни зависности меѓу сродни влезови.

Мрежата генерира три карактеристични патеки, што ги сумираат главните врски:

- прва патека - коловоз: PCI - IRI - RDB_RUT,
- втора патека - прегледност и сигнализација: SSD_H - SSD_V - H_sign - V_sign,
- трета патека - околина и геометрија: Max_Snow - Min_Temp - Ave_Rain - Ave_Rad - Ave_Height - W_i .

Овој распоред претставува практичен баланс меѓу едноставност и можност на моделот, со поголем фокус на тоа како се поврзани променливите отколку на бројките за успех. Дополнителната врска за секој предиктор се избира според мерка на меѓусебна информација, која покажува колку информации споделуваат променливите, па дрвото на врски меѓу влезовите се гради систематски.

6.7.5 Тестирање и валидација

Оценувањето се спроведува за трите модели GTT, PC и TAN, со иста подготовка, иста дискретизација и ист протокол. Се користи петкратна вкрстена валидација, што значи дека податоците се делат на пет дела и секој дел по ред се проверува, додека останатите служат за учење. Пресечните точки за дискретизација се учат само од тренинг-делот,

остануваат фиксни, и потоа непроменети се применуваат на валидациониот дел. Структурата и параметрите се учат исклучиво од тренинг-записи. На валидациониот дел се добиваат предвидувања и за секој од петте делови се пресметуваат мерки на точност, потоа се извлекуваат средна вредност и нормален опсег за да се провери колку се стабилни мрежите и клучните врски. Во оваа фаза не се додаваат нови чекори за подготовка, важат истите правила од подготовката и дискретизацијата.

Мерките и графичките прикази се на три нивоа. За одделување на класите се користи ROC-AUC во поставка една класа наспроти останати, со макро-просек за трите класи [139]. Дополнително се прикажуваат конфузни матрици, точност по класа, прецизност, одзив и макро-F1 [139]. За калибрација се користат reliability криви и Brier скор и понови тестови за калибрација што се темелат на постабилни дијаграми или заменски постапки за Hosmer-Lemeshow [140,141]. Приказите вклучуваат ROC криви, точност по класа и калибрациони криви со иста скала и легенда за трите пристапи, за полесна споредба.

По вкрстената валидација, секој пристап повторно се обучува на целиот тренинг сет со истите фиксни пресечни точки и истата подготовка, при што се оценува на посебен тест сет што не бил користен во учење или избор на поставки. Споредбата меѓу трите мрежи, најпрво се прави според просечниот ROC-AUC од валидацијата и според тоа колку стабилно ги одвојуваат класите. Кога резултатите по дискриминација се блиски, предност има онаа мрежа што е подобро калибрирана и има повисока стабилност на врските низ преклопите. Ако показателите се изедначени, се гледа едноставност на структурата и практичност за инженерска интерпретација. Сите чекори, од учење до проверка на калибрација, се спроведуваат во GeNIe 5, што обезбедува усогласена обработка и споредлива евалуација за GTT, PC и TAN [142]. Сликите и табелите за трите мрежи се подготвени со ист распоред и исти оски, за полесна визуелна споредба.

7. РЕЗУЛТАТИ И ВАЛИДАЦИЈА

Во продолжение, даден е преглед на добиените резултати. Изложувањето е по техника, во ист редослед како методологијата, со проверка на резултатите. Целта е резултатите и влијанието на параметрите да бидат јасни и лесно споредливи. За секој метод накратко се опишува подесувањето, потоа резултатите, а на крај проверката според заеднички критериуми за споредливост меѓу техниките.

7.1 Кластер анализа

Во кластер анализата се прикажуваат резултатите од групирањето на делниците според патни, сообраќајни и својства од околина. Прво се наведува избраниот број групи и нивната големина по техника, потоа се прикажува распоредот на делниците и дали групите се разликуваат според тежинскиот индекс на незгоди (W_i) преку боксплотови и соодветни тестови. Потоа се проверуваат стабилноста и согласноста меѓу техниките, а на крај се дава краток резултат од групирање на параметри за полесно толкување.

7.1.1 Поставки и параметри

Кластер-анализа се спроведува врз финален, стандардизиран (z -скор) сет од 23 влезни параметри. Тежинскиот индекс на незгоди (W_i) не е вклучен во влезниот сет, туку се користи подоцна како надворешна проверка на разликувањето помеѓу кластерите. Разгледан е интервал $k=2-10$ и се применуваат три техники: Ward (хиерархиско групирање), K-Means и Гаусова мешавина (GMM-Gaussian Mixture). Изборот на кластери се води од стандардни индикатори (силуетен коефициент и преглед на дендрограм), а кај GMM од BIC, со конкретни вредности прикажани во следните подточки.

Табела 7.1 Клучни поставки за кластер-анализа

Техника	Растојание/Модел	Иницијализација/линк	Опсег на k	Критериуми за избор
Ward (хиерархиско)	Евклидово на z -скорирани влезови	Ward linkage	2-10	Силуетен коефициент и преглед на дендрограм
K-Means	Евклидово на z -скорирани влезови	k-means++ со повеќекратни иницијализации	2-10	Силуетен коефициент и стабилност на решението
GMM	Коваријантни матрици: full, tied, diag, spherical	Expectation Maximization (EM)	2-10	BIC и стабилност на членства

Со овие поставки резултатите остануваат споредливи меѓу техниките, изборот на k е објективно потврден, а добиените кластери може директно да се користат во понатамошните анализи и проверката со W_i .

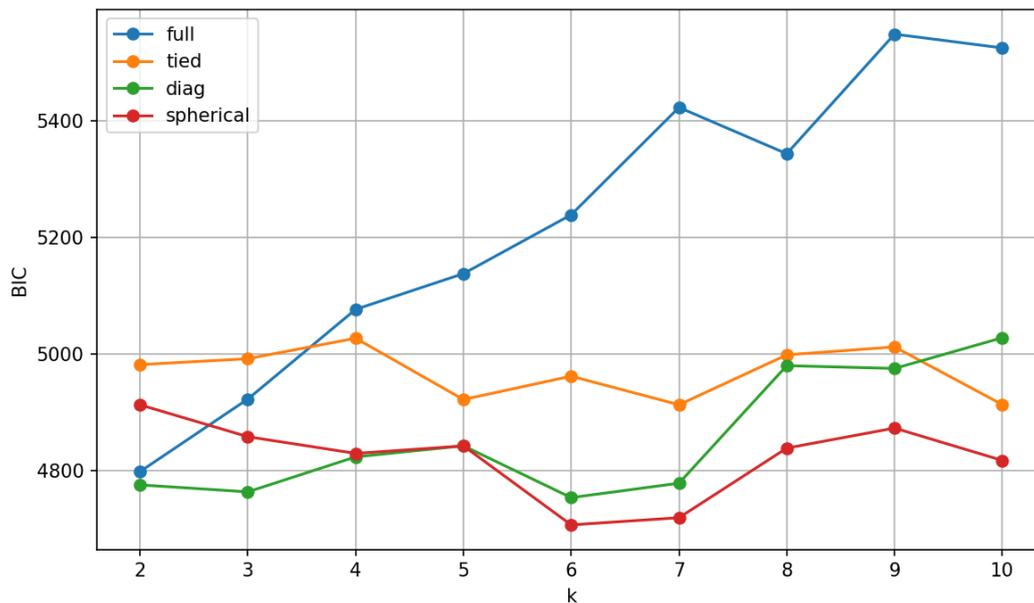
7.1.2 Избор на број на кластери

Во опсегот на број на групи k од 2 до 10, резултатите покажуваат различно однесување кај трите техники. Кај Ward, највисока вредност на силуетата се добива кога има две групи, што значи дека поделбата е појасна, но и поедноставна, бидејќи дели на само две главни целини. Кај K-Means, вредноста на силуетата постепено расте и достигнува врв кај

девет групи, што покажува дека овој метод повеќе ја дели структурата на податоците и создава повеќе ситни групи. Кај GMM, најниска вредност на критериумот BIC се добива кога има шест групи и сферична форма на распределбата, што претставува најдобра рамнотежа меѓу тоа, колку добро моделот ги опишува податоците и колку е сложен.

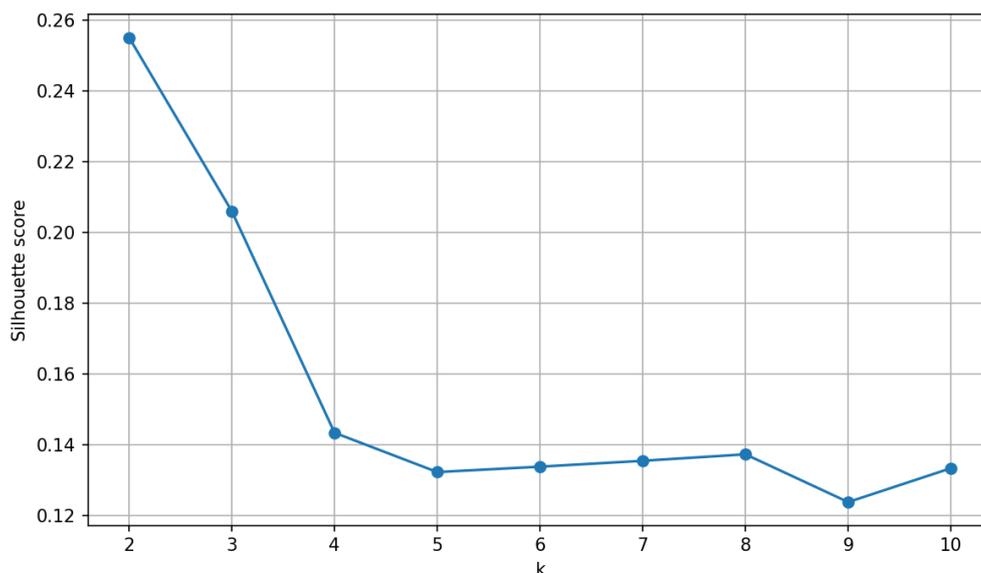
На сликата 7.1 прикажано е како се менува вредноста на BIC со бројот на групи (k) и со различни форми на распределба во GMM. Овде лесно се препознава најсоодветниот модел, односно моделот со најниска вредност на BIC, бидејќи најдобро го опишува односот меѓу податоците и сложеноста.

Од графикот се гледа дека, моделот со сферична форма на распределба, црвената линија, има најниска вредност кај $k = 6$, што упатува на најдобра рамнотежа меѓу точност и едноставност. Кај останатите форми, дијагонална, tied и full, вредностите на BIC се повисоки, што укажува на поголема сложеност или послабо вклопување во податоците.



Слика 7.1 BIC по k и коваријансни форми (GMM)

Од графикот на слика 7.2 се гледа дека кривата на силуетата за Ward има највисока вредност кога бројот на групи е два. Тоа значи дека Ward најдобро ги дели податоците кога се формираат две јасно одвоени, но поголеми групи. Ова укажува дека методот создава поедноставна поделба, каде разликите меѓу групите се изразени и лесно се забележуваат, додека внатре во групите податоците се послични. Со зголемување на бројот на групи силуетата се намалува, што покажува дека поделбата станува посложена и помалку јасна.



Слика 7.2 Просечна силуета по k (Ward)

Со оглед на добиените резултати, за понатамошна анализа се избира решението GMM со шест групи ($k = 6$). Оваа поставка овозможува доволно јасно раздвојување на податоците и притоа задржува разбирлив број на групи, што ја прави погодна за толкување и визуелна анализа. Шесте групи обезбедуваат рамнотежа меѓу точност и едноставност на моделот, односно доволно детална, но и стабилна поделба која може да ги прикаже главните разлики меѓу кластерите и нивната поврзаност со тежинскиот индекс на незгоди (W_i).

7.1.3 Резултатите по метод и проверка со W_i

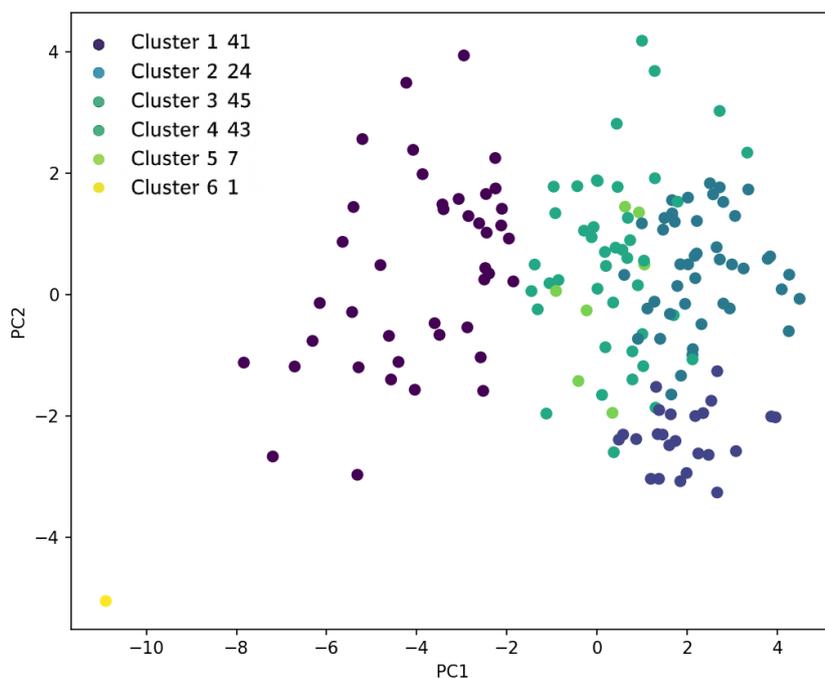
Во согласност со решението со Gaussian Mixture (GMM) при $k = 6$, овде се прикажуваат големините на кластерите, распоредот на делниците во приказ на главни компоненти и распределбата на W_i по кластери.

Табела 7.2 покажува дека најголем удел имаат кластер 3 (28.0%) и кластер 4 (26.7%), по што следи кластер 1 (25.5%), додека кластер 5 е мал (4.3%), а кластер 6 е маргинален (0.6%). Овој распоред упатува дека, повеќето делници припаѓаат на неколку доминантни групи, додека малите кластери најчесто содржат специфични услови.

Табела 7.2 Големини на кластери – GMM ($k = 6$)

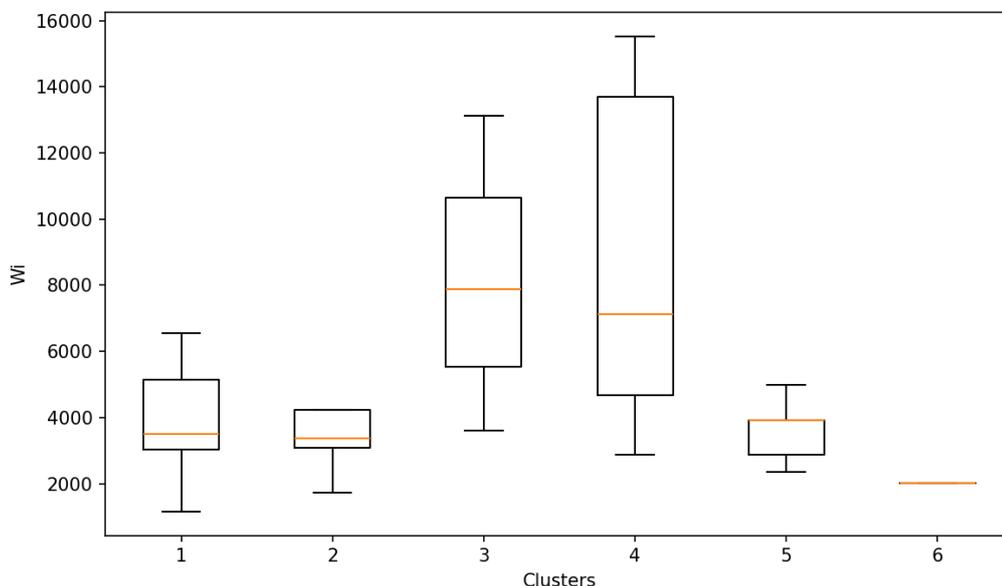
Кластер	Број (N)	Удел (%)
1	41	25.5
2	24	14.9
3	45	28.0
4	43	26.7
5	7	4.3
6	1	0.6

Како што се гледа на слика 7.3, проекцијата на две главни компоненти издвојува компактни облаци со јасни граници меѓу повеќето групи, што овозможува визуелна проверка на раздвојувањето и препознавање на гранични случаи.



Слика 7.3 PCA проекција на делнициите – GMM ($k = 6$)

На слика 7.4 прикажана е распределбата на тежинскиот индекс на незгоди (W_i) по кластери, со видливи разлики во централните вредности и распоните, што укажува дека групите се разликуваат и според последиците, а не само според влезните својства. За формална проверка се применува Kruskal-Wallis, со последователни парни проверки по потреба.



Слика 7.4 W_i по кластери – GMM ($k = 6$)

Сумирано, решението GMM со $k = 6$ дава разбирлива структура со визуелно јасно и статистички поткрепено разликување на W_i , што е корисно за профилирање на кластерите и за планирање на следните чекори.

7.1.4 Профили на класиери

Во продолжение се сумирани профилите на шесте кластери според централните вредности на избрани параметри. Во табела 7.3 се прикажани вредностите по параметар (во редови) и по кластер (во колони), со што споредбата е прегледна и директна.

Табела 7.3 Профили на класиери – медијани на избрани параметри (GMM, $k = 6$)

Параметар	C1	C2	C3	C4	C5	C6
N	41	24	45	43	7	1
Wi	3497	3367	7893	7136	3910	2009
Ave_Rain	679.00	587.50	564.00	600.00	595.00	796.00
Max_Snow	20.21	9.02	6.18	10.67	9.98	34.98
Ave_Temp	11.07	13.01	13.50	12.72	12.84	8.50
Max_Temp	28.74	31.41	32.09	31.02	31.18	25.22
Min_Temp	-3.49	-2.35	-2.06	-2.52	-2.45	-5.00
Ave_Height	671.00	319.50	231.00	369.00	349.00	1118.00
LIMIT	61.43	110.00	90.00	77.14	60.00	100.00
Kr.Kar.	115.00	33.93	33.25	32.58	46.53	450.01
Ave_Rad	526	1900	3264	1293	623	97
Lat_Force	0.61	0.50	0.18	0.26	0.49	7.96
Ave_Inc	2.30	1.67	1.57	1.47	2.02	5.43
SKIDRES	0.35	0.50	0.40	0.40	0.35	0.40
RDB_RUT	3.12	1.50	3.35	5.73	2.66	0.45
IRI	2.28	1.14	2.05	2.00	1.86	2.22
PCI	81	92	74	78	80	68
KON.NAK.	0.58	0.34	0.34	0.38	0.61	1.27
SSD H	1.03	0.54	0.54	0.55	0.92	3.16
SSD V	0.10	0.15	0.05	0.00	0.92	0.00
V sign	21	4	19	20	11	23
H sign	1	2	2	2	1	1
K.Int. P.T	0.64	0.18	0.30	0.46	0.98	0.63
K.Bridges	0.28	0.49	0.46	0.36	0.07	0.63

Табелата овозможува брзо да се споредат клучните карактеристики по кластери и да се поврзе нивниот профил со Wi. Во продолжение е наведен краток опис на кластерите по нивни специфични карактеристики и ориентационен ризик:

- C1 (N=41, Wi=3497). Повисока Ave_Height и пониски LIMIT вредности, со помал Ave_Rad и поизразени Kr.Kar. Коловозните мерки се умерени (PCI=81, RDB_RUT=3.12), а снежните услови се повисоки (Max_Snow~20). Профилот упатува на ридско/планински делници со поостри кривини и пониски брзини, што е во согласност со средно низок Wi.

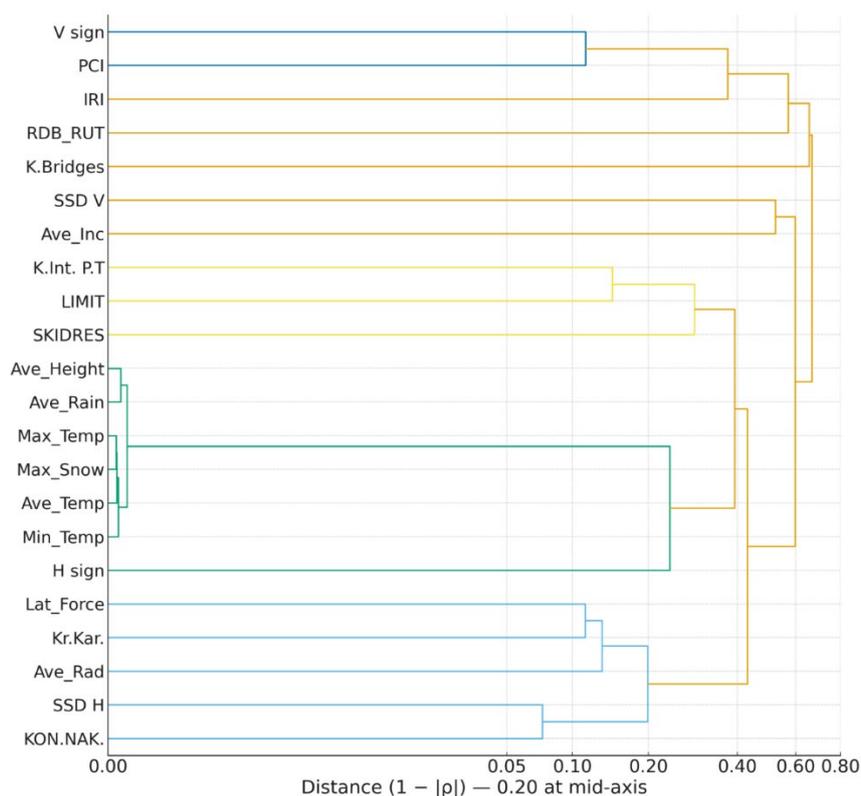
- C2 (N=24, Wi=3367). Понишка Ave_Height, потопли услови и многу големи радиуси (Ave_Rad≈1900), со највисок PCI и најниска RDB_RUT. LIMIT е највисок (110), што е очекувано за испружени делници со поволна геометрија. Оваа комбинација е поврзана со понизок ризик во однос на повеќето други групи.
- C3 (N=45, Wi=7893). Многу големи радиуси и највисока Ave_Temp, но понизок PCI и умерен RDB_RUT. И покрај геометриски поволни услови, Wi е највисок, што упатува дека други фактори во сетот или надворешни услови придонесуваат за ризик. Ова е кластер со највисок ризик и служи како референтна споредба со C4.
- C4 (N=43, Wi=7136). Средни радиуси и LIMIT, но највисока RDB_RUT и понизок PCI во споредба со кластерите со низок ризик. Профилот е во согласност со изразени коловозни оштетувања и повисок оперативен ризик, што се гледа и во повисокиот Wi.
- C5 (N=7, Wi=3910). Помал Ave_Rad и најниско LIMIT, со повисока вредност на K.Int. P.T, што упатува на зголемена интеракција и сложеност во движењето. Ризикот е умерен, а профилот е блиску до C1, но со повеќе управувачки точки.
- C6 (N=1, Wi=2009). Највисока Ave_Height, најниска Ave_Temp, најмал Ave_Rad и највисок Max_Snow, а сепак има најнизок Wi и покажува највисоки безбедносни перформанси. Иако N=1 не дозволува генерализација, овој резултат е важен показател дека комбинацијата на ниски температури, голема надморска висина и мали радиуси не мора да значи висок ризик, ако другите услови и мерки се поволни.

Овие профили овозможуваат насочено толкување и ќе послужат за избор на релевантни влезови во следните анализи.

7.1.5 Класифицирање на параметри (*feature clustering*)

Овде се прикажува резултатот од групирањето на параметрите за полесно препознавање на сродни мерки и за избегнување на непотребно преклопување во понатамошните анализи.

Како референтен приказ, дендрограмот е изграден со метрика $distance = 1 - |\text{Spearman } \rho|$. Кратките гранки означуваат силна поврзаност, без разлика на насоката на корелацијата. Како што е прикажано на слика 7.5, климатските параметри Ave_Temp, Max_Temp, Min_Temp, Ave_Rain и Max_Snow се групираат многу блиску, а Ave_Height се поврзува со нив на мало растојание, што упатува на силна, главно негативна врска со температурите. Одделно се оформуваат компактни блокови кај геометријата на трасата (Ave_Rad, Kг.Kar., Ave_Inc, KON.NAK., Lat_Force), кај мерките за коловоз (PCI, IRI, RDB_RUT), како и кај управувачките и инфраструктурни индикатори (LIMIT, K.Int. P.T, K.Bridges, H sign, V sign), додека SSD_H, SSD_V и SKIDRES се поизолирани јазли со појасно дефинирани врски во рамките на своите подгрупи.



Слика 7.5 Дендрограм на параметри

Овој увид служи како водилка за избор на репрезентативни променливи, за толкување на профилите на кластерите и за формирање составни индикатори каде што е оправдано, со што, следните анализи остануваат пократки и појасни.

7.1.6 Работен избор за понатамошни анализи

За натамошните чекори се утврдува јасен работен план, во кој се наведува како ќе се именуваат кластерите и како ќе се користат. Табела 7.4 накратко прикажува кој метод за групирање е избран, колку групи има и кои податоци се пренесуваат понатаму, со кратко објаснување, зошто и за што ќе се користат.

Табела 7.4 Работен избор и употреба

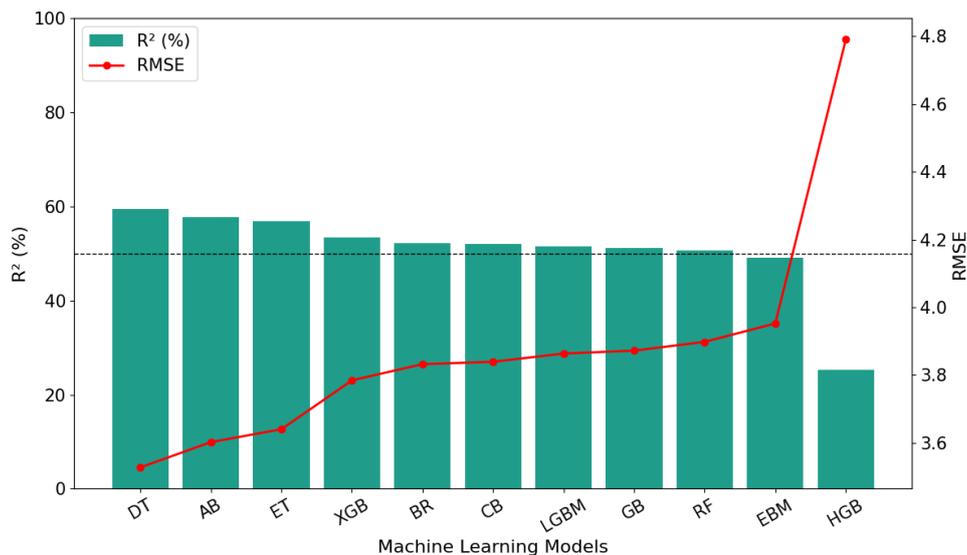
Елемент	Избор
Техника	GMM
Број на кластери (k)	6
Елем. за понатамошна употреба	Категоријална етикета Cluster_ID (C1-C6) и веројатности $p(C1) \dots p(C6)$
Критериуми за избор	Разликување на W_i , поволен BIC, стабилна и читлива структура
Понатамошна употреба	Cluster_ID за групирање и интерпретација, $p(C1) \dots p(C6)$

Ознаките од кластерите ќе се користат како имиња на групи за полесно објаснување. Веројатноста за припадност ќе се користи кога е потребна подетална слика за гранични делници и за проверки на стабилност.

7.2 Иницијална проверка на техники

Во продолжение дадена е почетна слика за перформансите на моделите. Фокусот е на општата споредливост и ориентацијата кои пристапи да се задржат за понатамошна работа. Ова значи дека се дава голема слика, за да се види кој модел во просек работи подобро и кој е постабилен во различни услови. Резултатите се оценуваат со едноставни мерки како точност на предвидувањата и просечната грешка. Намерата не е да се навлегува во технички подробности, туку брзо да се препознае, кои модели вреди да продолжат понатаму, а кои може да се изостават. При читање, корисно е да се следи дали постои најдобро рангиран модел, колкави се разликите меѓу моделите и дали има стабилни трендови.

На слика 7.6 се дадени резултати на иста 80/20 поделба, при што R^2 (%) е прикажан како столбови, а RMSE како линија, за да се види способноста на моделот да ја објасни варијацијата и грешката на предвидување. Од приказот се гледа јасно рангирање: DT е најдобар ($R^2 = 59.5\%$), по него AB ($R^2 = 57.8\%$) и ET ($R^2 = 56.9\%$). Групата модели на итеративно засилување и модели со повеќе дрва (XGB, BR, CB, LGBM, GB, RF) дава компактни вредности околу 51–53% со умерен RMSE. EBM е веднаш под референтната линија од 50%, додека HGB е забележливо послаб. Водилка за ориентација служи испрекинатата линија на 50% R^2 .



Слика 7.6 R^2 (%) и RMSE по модел на 80/20 поделба

Како основен сет за понатамошните анализи се задржуваат моделите што имаат вредност на $R^2 \geq 0.50$ (DT, AB, ET, XGB, BR, CB, LGBM, GB, RF). Моделот EBM се користи како дополнителен показан модел, затоа што добро ги објаснува врските меѓу податоците, иако е под зададениот праг, додека HGB не се смета за приоритетен. Моделите MLP и SVR покажаа многу пониско R^2 (под 50%) и затоа не се користат понатаму во анализите.

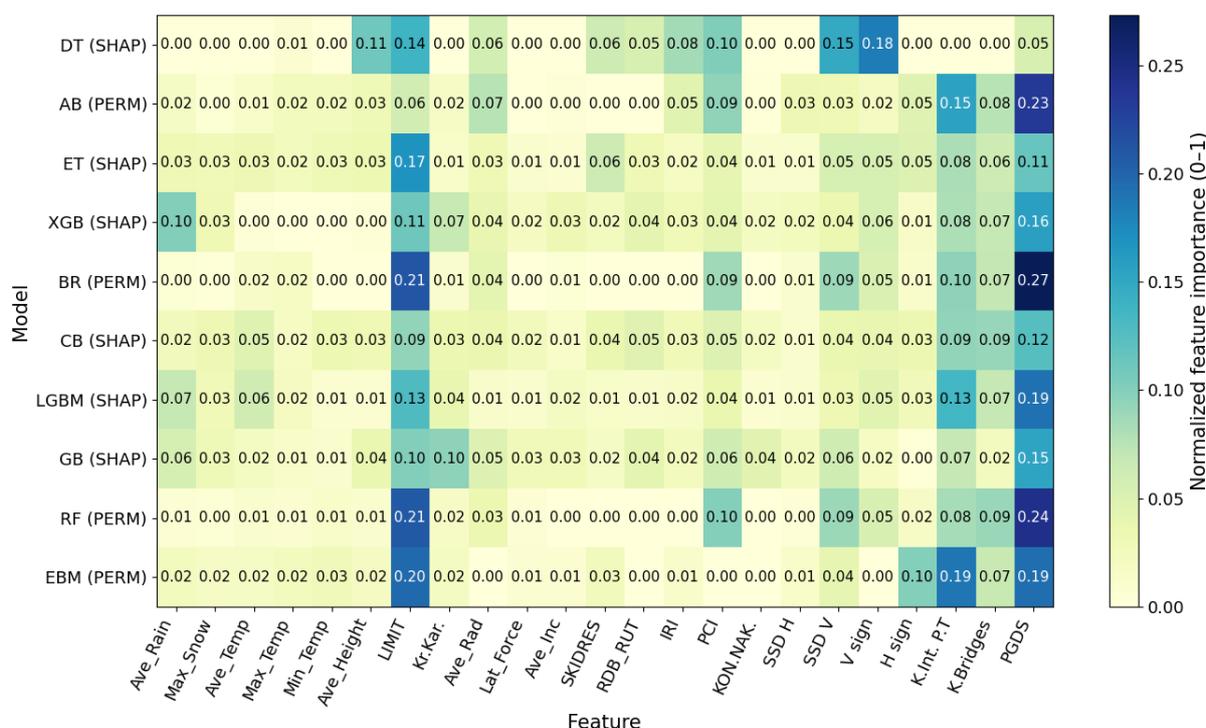
7.3 Влијателни параметри

Следниот чекор ги покажува параметрите што најмногу влијаат на тежинскиот индекс на сообраќајни незгоди (W_i), користејќи модели со $R^2 \geq 0.45$ на истата поделба 80/20.

Влијанието на параметрите се пресметува на два начини: со SHAP (TreeExplainer) за дрвовидни модели и модели на итеративно засилување (DT, ET, XGB, GB, LGBM, RF, CB), и со пермутациона важност кога SHAP не е достапен или е помалку стабилен (AB, BR, EBM). За секој модел влијанијата се доведуваат на иста скала, потоа се прави просек со тежини според Test R², за да се добие една заедничка и споредлива ранг-листа на параметри. Моделите со пониска точност (HGB, MLP, SVR) не се користат во проценката. При толкувањето се зема во предвид сродноста меѓу параметрите, за да не се дуплира или погрешно припише влијанието.

7.3.1 Поединечни резултати по модел

Резултатите по модел се прикажуваат за да се добие јасен увид во влијанијата на параметрите. Вредностите се добиени со SHAP кај моделите со дрва и засилувачки модели (DT, ET, XGB, CB, LGBM, GB, RF), а со пермутациона важност кај AB, BR и EBM. На слика 7.7 дадена е топлотна мапа со модели по редови и параметри по колони, при што влијанијата се нормализирани, така што нивниот збир е 1 во рамките на секој модел, а потемната боја означува поголемо влијание.



Слика 7.7 Влијание на параметри по модел (80/20)

Од сликата се издвојуваат PGDS и LIMIT како најсилни низ повеќето техники. Често повисоки вредности имаат и K.Int. P.T, PCI, SSD V и V sign, додека K.Bridges покажува стабилно средно влијание. Параметрите поврзани со климата, најчесто се со ниски вредности, со умерен показател на Ave_Rain во дел од засилувачките модели. IRI, Ave_Rad и Ave_Height даваат средни влијанија зависно од техниката, а SKIDRES е умерен до низок.

Краток осврт по модел:

DT (SHAP) – највлијателни V sign, SSD V и LIMIT, со видно учество на Ave_Height, PCI и PGDS.

AB (PERM) – доминираат PGDS и K.Int. P.T, потоа PCI, K.Bridges, Ave_Rad и LIMIT.

ET (SHAP) – изразени LIMIT и PGDS, следат K.Int. P.T, SKIDRES, K.Bridges и SSD V.

XGB (SHAP) – повисоки PGDS, LIMIT и K.Int. P.T, со умерено Ave_Rain, K.Bridges и Kr.Kar.

BR (PERM) – силни PGDS и LIMIT, потоа SSD V, PCI и V sign.

CB (SHAP) – нагласени PGDS, LIMIT и K.Int. P.T, со умерени V sign, K.Bridges и SSD V.

LGBM (SHAP) – водечки PGDS, LIMIT и K.Int. P.T, со средни Ave_Rain и Ave_Temp.

GB (SHAP) – повисоки PGDS, LIMIT, PCI и SSD V, со видлив Ave_Rad.

RF (PERM) – изразени PGDS, LIMIT, PCI, SSD V и K.Bridges.

EVM (PERM) – најмногу тежина на PGDS, LIMIT и K.Int. P.T, со умерено H sign и V sign.

Овие резултати прегледно покажуваат кои параметри се најважни и овозможуваат објективна споредба на моделите.

7.3.2 Средна вредност на влијанијата

За споредлив ранг на параметрите, влијанијата најпрво се нормализираат во рамките на секој модел, со што сите модели се на иста скала и споредбата е директна. Потоа се пресметува тежински просек, каде учествуваат само моделите што го исполнуваат дефинираниот критериум. Тежината на секој модел е пропорционална на неговиот Test R² и потоа се нормализира, па моделите со повисок R² добиваат поголема тежина.

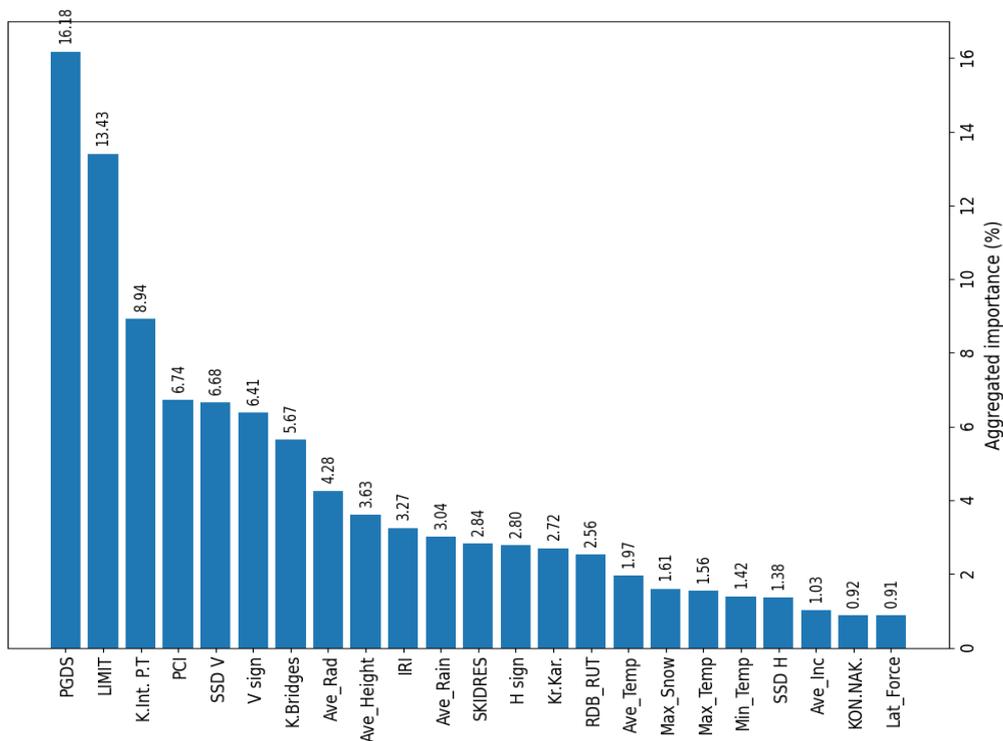
Во табела 7.5 прикажани се Test R² и нормализираните тежини по модел, за полесно да се види учеството на секоја техника во конечниот резултат. Нормализирана тежина значи дека збирот на сите тежини е 1, а повисока тежина укажува на поголемо влијание во просекот. Test R² ја покажува точноста на моделот на тест множеството, па затоа прво се чита оваа вредност. Практичниот редослед за читање е едноставен, прво се гледа R², потоа тежината.

Табела 7.5 Тежини по модел според точноста на 80/20 поделба

Модел	R ²	Нормализирана тежина	Сурова тежина
DT	0.595	0.172	0.145
AB	0.578	0.152	0.128
ET	0.569	0.141	0.119
XGB	0.534	0.100	0.084
BR	0.522	0.086	0.072
CB	0.520	0.084	0.070
LGBM	0.514	0.076	0.064
GB	0.512	0.074	0.062
RF	0.506	0.066	0.056
EVM	0.492	0.050	0.042

Слика 7.8 прикажува обединет ранг на параметрите добиен со овие тежини. Повисок ранг значи посилено влијание врз Wi. Највисоко се позиционирани PGDS и LIMIT, потоа следат K.Int, P.T, PCI, SSD_V и V_sign. Меѓу индикаторите од патот се издвојуваат K.Bridges и Ave_Rad со средно влијание, додека од мерките за коловоз IRI и RDB_RUT имаат умерен придонес. Климатските променливи се со понизок вкупен ефект, иако се корисни при

проверка на одделни случаи и локални модели. Суштинскиот наод е дека обемот на сообраќај и брзинскиот режим најчесто го определуваат ризикот, а геометријата, состојбата на коловоз и прегледноста додаваат дополнително влијание.



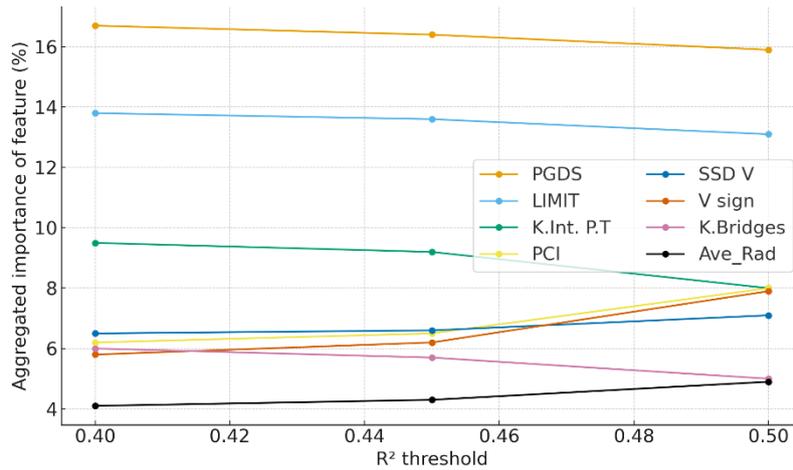
Слика 7.8 Усреднеј ранг на параметрите според усреднето влијание

Оваа комбинација на влијанијата овозможува, постабилен и полесен избор на најважните параметри. На овој начин се создава прегледна слика кои фактори најмногу влијаат, а резултатите може лесно да се споредуваат и користат во следните анализи.

7.3.3 Чувствителност на обединувањето на влијанијата

Дополнителен услов што овозможува постабилна проценка на влијанието на параметрите е отпорноста на резултатите на промени во прагот за тежините и на исклучување на поединечни модели. Овој чекор проверува дали заклучоците остануваат исти и кога правилата за избор на модели малку се менуваат. Најпрво обединувањето се повторува со три различни прага на R^2 : 0.40, 0.45 и 0.50.

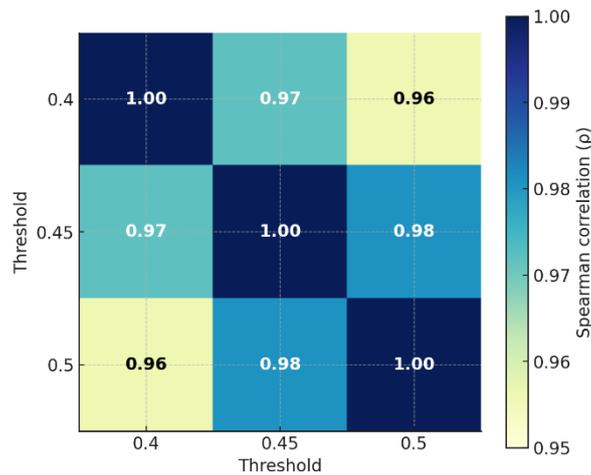
Слика 7.9 прикажува како се менуваат процентите за осумте највлијателни параметри кога се менува прагот. За читање, најпрво се споредуваат линиите за PGDS и LIMIT, потоа се гледа дали редоследот на параметрите останува приближно ист и дали разликите меѓу праговите се мали. PGDS и LIMIT остануваат на врвот со мало намалување, K.Int. P.T има умерено намалување, а SSD V и V sign покажуваат благ пораст. Кога линиите за еден параметар се блиску една до друга за различните прагови, тоа значи стабилно влијание и сигурен избор на параметри. Овие разлики се мали и не доведуваат значајно менување на редоследот на најважните параметри, што покажува дека резултатите се стабилни.



Слика 7.9 Чувствителност на усреднети важности по прагови

За да се провери стабилноста на рангирањето меѓу праговите, се пресметува Спирманова корелација за трите граници.

На слика 7.10 е прикажана Спирмановата матрица на корелации, со вредности во интервал 0.96–1.00, што покажува речиси ист редослед на параметрите без оглед на избраниот праг.



Слика 7.10 Спирманова корелација на рангирањата меѓу прагови на R²Ок

Понатаму се прави проверка со исклучување еден по еден модел (leave-one-out проверка), каде секој модел се исклучува поединечно и се споредува добиеното рангирање со базното обединување, корелациите со базата се движат помеѓу 0.95 и 1.00. Најниска корелација се добива при исклучување на DT (0.95), додека за повеќето модели корелацијата е 0.98–1.00, вклучувајќи случаи без промена на редоследот (RF, BR, CB и EBM). Ова упатува дека ниту еден модел поединечно не го диктира резултатот.

Рангот на параметрите останува стабилен и при ограничени варијации на границите и при различен избор на модели, што потврдува дека највлијателните параметри и нивниот редослед не зависат од конкретните поставки.

7.4 Интегративна формула

Во овој дел се развива и проверува интегративна формула за W_i врз основа на влезните параметри достапни за секоја делница. Целта е да се добие краток и точен модел

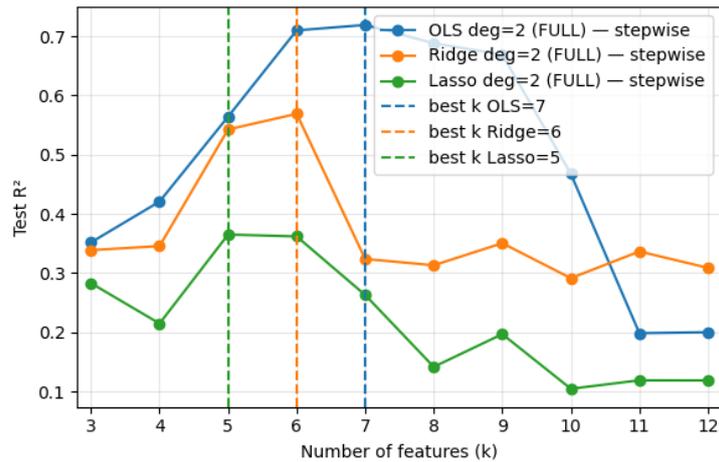
кој ја објаснува варијабилноста на W_i на независните тест-податоци. Интегративната формула ги спојува патните, сообраќајните и параметрите од околина во еден критериум за разграничување на безбедни и небезбедни делници. Оценувањето се прави со Test R^2 на фиксна поделба тренинг-тест 80/20 со зададен random state = 42.

Примарната формула е калибрирана со метод на најмали квадрати (OLS – Ordinary Least Squares). OLS ги одредува коефициентите, така што ја минимизира сумата на квадратите на остаточните грешки меѓу набљудуваните вредности на W_i и излезот од моделот. Примената е врз полиномна основа од втор степен со целосен опфат, со главни ефекти, квадрати и сите парни интеракции. Главните влезови најпрво се стандардизираат со параметри добиени само од тренинг-податоците, потоа се конструира полиномната основа. За проверка на стабилност и како споредбени линии се користат и варијанти со член за ограничување Ridge и Lasso, при што на полиномната матрица се применува дополнителна стандардизација.

Покрај анализата со строг префикс (редоследно проширување според ранг на влијателност), се спроведува и постепено додавање со локално најдобар избор на параметри. Постапката започнува од три највлијателни параметри PGDS, LIMIT, K.Int. P.T и во секој следен чекор се додава уште еден параметар што во тој чекор обезбедува најголемо зголемување на Test R^2 . На овој начин се следи развојот за k од 3 до 23 по сите три техники, а крајниот избор се темели на највисок постигнат Test R^2 со фокус на стабилност и компактност на решението. За идна употреба, пресметаната вредност на W_i се споредува со оперативните прагови, за да се означат небезбедни делници.

7.4.1 Евалуација на модели

Анализата најпрво се изведува со строг префикс, при што добиените вредности на Test R^2 се умерени и не достигнуваат задоволително ниво. Поради тоа се применува пристап со почетен сет од три највлијателни параметри и постепено додавање (seed=top-3 + stepwise), кој започнува од три највлијателни параметри и во секој следен чекор додава по еден нов параметар што во тој чекор обезбедува најголемо зголемување на Test R^2 . Пресметките се вршат на полиномна основа од втор степен со целосен опфат, при поделба на податоците на тренинг и тест 80/20 (со зададен random state=42). Влезните променливи се стандардизираат на тренинг делот, а кај Ridge и Lasso, стандардизирањето се применува на сите карактеристики што влегуваат во моделот.



Слика 7.11 Test R^2 во зависност од бројот на параметри k (OLS, Ridge, Lasso)

Дијаграмот на слика 7.11 ги прикажува кривите на Test R^2 како функција од бројот на параметри k за OLS, Ridge и Lasso, со ограничување на приказот до $k \leq 12$ за појасен фокус околу максимумите.

Највисока детерминација се добива со OLS при $k=7$, со Test $R^2 = 0.72$. Соодветното множество влезови е (PGDS, LIMIT, K.Int. P.T, PCI, V sign, Ave_Inc, KON.NAK.). Кај Ridge и Lasso се добиваат пониски максимуми, што ги позиционира овие техники како споредбени линии за проверка на стабилност, а изборот на финален модел се темели на највисок постигнат Test R^2 со фокус на компактност и стабилност.

7.4.2 Избран модел и предиктивна формула за W_i

Избраниот модел е OLS на полиномна основа од втор степен со $k = 7$ влезови. Во продолжение се дава формулата во поедноставена текстуална форма, каде се прикажани влијанијата, интеракциите и квадратите. Сите параметри се стандардирани со влезни параметри (z -score вредности), пресметани врз основа на тренинг-податоците.

Овој пристап овозможува попрецизна и споредлива проценка на влијанието на секој поединечен параметар врз W_i , при што се задржува стабилноста на моделот и се избегнува преголемо влијание на поединечни влезови.

Формула за пресметка на W_i (OLS, deg=2 FULL, $k=7$; Test $R^2 = 0.719$):

$$W_i = 8.3924 + 3.1366 \cdot \text{K.Int. P.T.} + 2.4905 \cdot \text{PGDS} - 2.4730 \cdot (\text{PCI} \cdot \text{KON_NAK}) + 2.4313 \cdot (\text{LIMIT} \cdot \text{K.Int. P.T.}) + \dots + 0.1384 \cdot (\text{LIMIT} \cdot \text{PCI}) + 0.1014 \cdot (\text{PGDS} \cdot \text{K.Int. P.T.}) - 0.0705 \cdot (\text{Ave_Inc}^2) - 0.0409 \cdot (\text{PGDS} \cdot \text{LIMIT})$$

*Целоснаџа формула е дадена во џрилози

Каде:

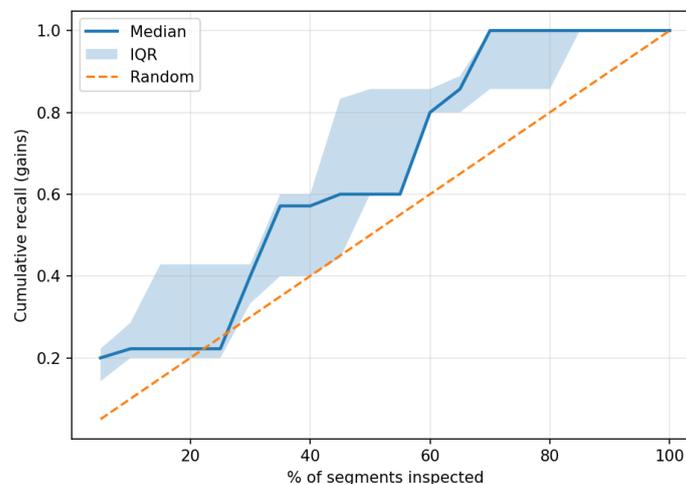
- W_i - тежински индекс на сообраќајни незгоди [поени]
- PGDS - просечен годишен дневен сообраќај [возила/ден]
- LIMIT - просечна ограничена брзина [km/h]
- K.Int. P.T. - густина на јазли/крстосници/наплатни станици [1/km]
- PCI - стандардизиран индекс за состојба на коловоз [поени]

- V_sign – вертикална сигнализација [години]
- Ave_Inc – просечен надолжен наклон [%]
- KON_NAK – индикатор за контранаклони [1/km]

Пред да се пресмета формулата, сите седум влезни параметри се стандардизираат според тренинг-податоците, односно се користат нивните просечни вредности (μ) и стандардни девијации (σ). Од стандардизираните вредности се создаваат нови променливи (квадратни термини и интеракции меѓу параметрите), кои потоа се внесуваат во формулата со соодветните коефициенти. Овој чекор овозможува сите влезови да бидат на иста скала, така што ниеден параметар не доминира поради различни единици или интервал на вредности, што ја прави пресметката попрецизна и резултатите полесни за толкување.

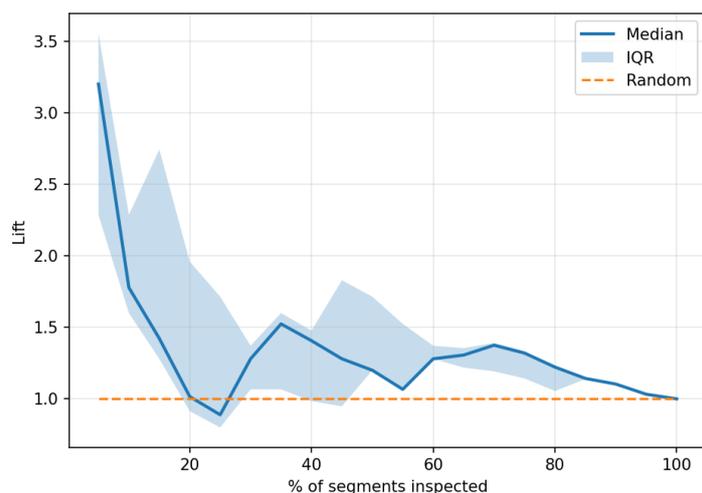
7.4.3 Валидација на предиктивна формула

За да се провери колку добро функционира моделот и дали може да се користи во пракса, тој е тестиран со петкратна вкрстена валидација. Во секој круг на валидацијата податоците се делат на дел за учење и дел за проверка, моделот учи од првиот дел, а потоа се проверува на податоци што претходно не се користени. Поставката е иста во сите кругови, се применува OLS на полиномна основа од втор степен со седум влезови. Делниците со највисоки 20 проценти вредности на W_i се сметаат за високоризични, останатите се со низок ризик. Резултатите од сите кругови се обединуваат и се прикажуваат со просек и со типичен распон на вредности, за да се покаже вообичаеното однесување на моделот. На овој начин се добива објективна слика за точноста, бидејќи проверката се прави на податоци што не се користени за учење.



Слика 7.12 Кумулативна крива на добивки во зависност од покриеноста на сегментите

Дијаграмот на слика 7.12 покажува како расте уделот на пронајдени високоризични сегменти кога се разгледуваат првите x проценти од листата според моделот. Кривата што ја претставува средната вредност по повторувања се наоѓа над линијата на случаен избор, што значи дека моделот ги собира ризичните сегменти на почетокот од листата. Кога покриеноста (recall) е ниска до умерена, во опсег од 10 до 30 проценти, моделот опфаќа поголем дел од високоризичните сегменти отколку што би се очекувало при случаен избор.



Слика 7.13 Крива на зголемување во зависноста од покриеноста

Слика 7.13 покажува како се менува зголемувањето (lift), односно колку е прецизноста на моделот поголема од онаа што би се добила при случаен избор. Кога се гледаат горните делови од рангирањето, вредностите на зголемувањето се поголеми од еден, што значи дека моделот подобро ги препознава делниците со повисок ризик. Како што се зголемува покриеноста, зголемувањето постепено се приближува кон еден, што покажува дека точноста се израмнува со очекуваната вредност. Овој резултат значи дека моделот прво ги издвојува делниците со најголем ризик и овозможува практичен избор на граница за работа, во зависност од достапните ресурси и целите на анализата.

Консолидирани метрики (5-fold, централна вредност \pm среден 50 % опсер): Accuracy = 0.732 ± 0.119 , Balanced accuracy = 0.585 ± 0.138 , F1 = 0.343 ± 0.202 , AUC = 0.673 ± 0.160 , Top-20% capture = 0.314 ± 0.235 . Резултатите покажуваат колку моделот е точен и колку добро ги препознава делниците со висок ризик. Вкупната точност е околу 73 проценти, што значи дека во просек три од четири предвидувања се точни. Избалансираната точност е над 0.5 и е подобра од случаен избор, што укажува дека моделот релативно добро ги разликува ризичните од безбедните делници. Вредноста F1 е пониска бидејќи високоризичните делници се околу 20 проценти од вкупните, па се потешки за препознавање. AUC околу 0.67 значи дека моделот во повеќето случаи им доделува повисок резултат на ризичните отколку на неризичните делници. Показателот Top 20% capture е 0.314, додека при случаен избор би бил околу 0.200, што е околу 57 проценти повеќе и директно ја покажува добивката (gain) кога вниманието се насочува кон врвот на листата.

Табела 7.6 Прецизност, покриеност и лифт на 10, 20 и 30 проценти покриеност

Покриеност на сегменти (%)	Прецизност	Покриеност (recall)	Зголемување (lift)
10	0.500 [0.250–0.500]	0.222 [0.200–0.286]	1.778 [1.600–2.286]
20	0.286 [0.143–0.429]	0.222 [0.200–0.429]	1.016 [0.914–1.959]
30	0.300 [0.200–0.300]	0.400 [0.333–0.429]	1.280 [1.067–1.371]

Табела 7.6 ги прикажува прецизност, покриеност и зголемување за покриеност од 10, 20 и 30 проценти. Прецизност е уделот на реално високоризични меѓу избраните делници. Покриеност е уделот на идентификувани „високоризични“ од нивниот вкупен број. Зголемување (lift) покажува колку пати прецизноста е поголема од базната стапка. Овие мерки помагаат да се одреди каде е најпрактично да се постави границата за акција, односно колку делници најнапред да се разгледаат.

Во целина, резултатите од кривите за покриеност и зголемување потврдуваат дека моделот ги поставува најризичните делници на почетокот од листата. Тоа значи дека формулата е практично применлива за дефинирање на приоритети, а тестирањето потврдува дека моделот функционира сигурно во рамките на зададените услови. Колку делници ќе се разгледаат, најпрво зависи од расположливите ресурси и од длабочината на планираната проверка.

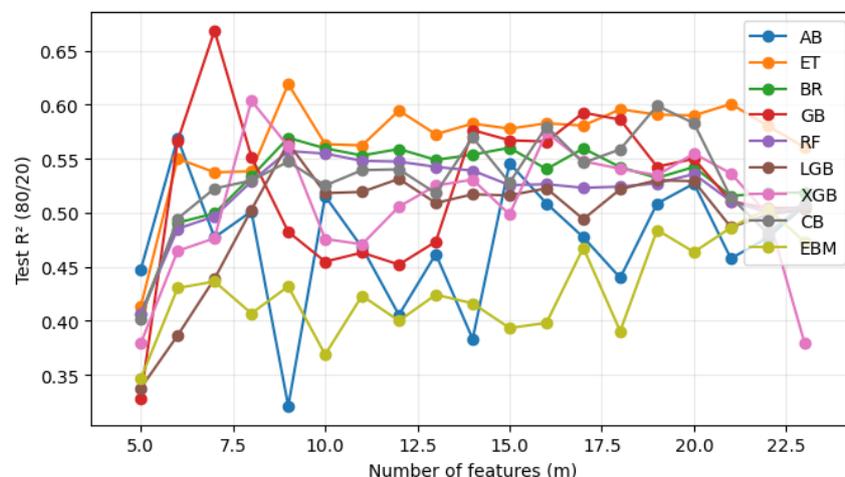
7.5 Избор на оптимален модел од техники на машинско учење

Во продолжение наведен е процесот за избор на најсоодветниот модел од различни техники на машинско учење. Процесот се спроведува врз основа на поделба 80/20 и подготовка без протекување на информации. На почеток се прави споредба меѓу повеќе модели со различен број фактори и со мало прилагодување на параметрите, при што моделот со највисока вредност на Test R^2 и со поедноставна структура се избира како подобар. Потоа следи вториот дел, каде моделот се проверува и потврдува низ повеќе тестови. Се прави калибрација за да се усогласи излезот, потоа се анализираат грешките во предвидувањата и се применуваат метрики што покажуваат колку добро моделот ги распознава делниците со висок ризик. Целта на овој процес е, да се обезбеди стабилен и применлив модел кој сигурно може да ги препознае делниците со висок тежински индекс на незгоди (W_i).

7.5.1 Резултатна анализа и избор на модел според шест и перформанси

Овој дел ја опишува основната постапка за изработка и избор на модели насочени кон W_i . Параметрите се подредени по важност, а се користат модели за непрекинати вредности. Постапката започнува со 23 параметри и чекор по чекор се отстранува најмалку влијателниот сè до $m=5$. За секој m се тренира на тренинг и се мери R^2 на тест, со што се избира m со највисоко R^2 на поделба 80/20. Потоа се подесуваат клучните поставки по тип на модел и се прави споредба со јасни критериуми.

Резултатите од скенирањето по број на параметри (m) и насоченото подобрување служат за споредбена проценка на техниките и за издвојување стабилни конфигурации. Притоа се избира оптимален број параметри по техника и се селектира носечки модел за пресметка на W_i и рангирање на делници, со фиксна 80/20 поделба.



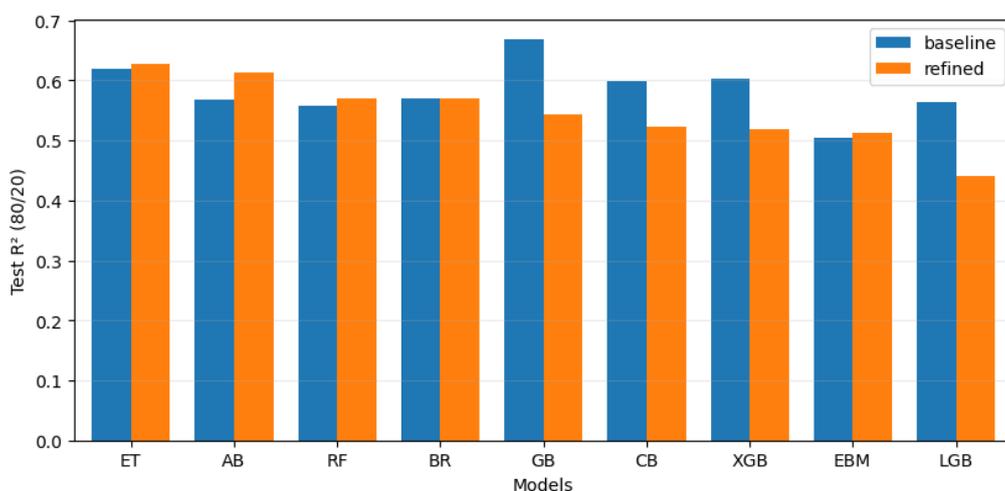
Слика 7.14 R^2 според број на параметри

На слика 7.14 се гледа како различните методи реагираат кога постепено се менува бројот на параметри m . Моделот GB го достигнува највисокото ниво на точност кога $m=7$, ET се стабилизира со високи вредности околу $m=9$, а XGB се однесува рамномерно при $m=8$. Кај CB за да се постигне добар резултат потребно е да се вклучат повеќе параметри, додека EBM, и покрај својата сложеност, дава пониски резултати. Овој распоред покажува дека најуспешните техники не бараат голем број параметри за да го достигнат својот максимум и дека нивните криви во средината се стабилни и без брзи промени. Тоа значи дека малите промени во бројот на параметри не влијаат значително на резултатот, што е добро за практична примена, бидејќи изборот на m во овој интервал е сигурен и лесен за примена при оценување на безбедноста на делниците.

Табела 7.7 Максимално R^2 (базно) и m^* по техника

Техника	m^*	R^2 (базно)
GB	7	0.67
ET	9	0.62
XGB	8	0.60
CB	19	0.60
BR	9	0.57
AB	6	0.57
RF	9	0.56
LGB	9	0.56
EBM	22	0.50

Во табела 7.7 прикажано е кој модел го постигнува најдобриот резултат и со кој m , при фиксна 80/20 поделба и иста подготовка. Во делот со пониски вредности моделите почесто се менуваат кога m се поместува за еден чекор и повеќе зависат од ситни промени во подготовката, па сигурноста на изборот е помала. Кога кривите имаат стабилен дел низ повеќе m , одлуката е посигурна, бидејќи не се заснова на една тесна точка туку на поширок интервал со стабилни резултати.



Слика 7.15 Насочено ипрецизирање

За подобрување на резултатите, се прави дополнително подесување на хиперпараметрите на моделите, а потоа за секој модел се зема повисоката вредност од

двете анализи, базната и подобрената. На слика 7.15 се прикажува како ова локално подобрување делува околу бројот на параметрите. Кај ET се добива мало зголемување околу $m = 9$, додека за AB се забележува пораст околу $m = 8$. Во RF подобрувањето е умерено, а за BR резултатот останува ист. Кај техниките на засилување како GB и XGB има намалување во тесен опсег на поставки. Ова укажува дека кога кривата веќе има стабилен дел со високи вредности, дополнителните промени ретко носат значаен напредок, па базните поставки често се посигурен избор

Табела 7.8 Корективен преглед

Техника	m^* (базно)	R^2 (базно)	m (подобро)	R^2 (подобро)	m (финално)	R^2 (финално)
GB	7	0.67	5	0.54	7	0.67
ET	9	0.62	9	0.63	9	0.63
AB	6	0.57	8	0.61	8	0.62
XGB	8	0.60	8	0.52	8	0.60
CB	19	0.60	18	0.52	19	0.60
RF	9	0.56	9	0.57	9	0.57
BR	9	0.57	9	0.57	9	0.57
LGB	9	0.57	9	0.44	9	0.56
EBM	22	0.50	20	0.51	20	0.51

Во табела 7.8 применето е правило на контролиран избор (guarded selection), при кое за секој модел се споредуваат резултатите пред и по дополнителното подесување. Исто така, се задржува повисоката вредност на R^2 и соодветниот број параметри m . На тој начин се добива јасен и стабилен избор кој не се менува при мали промени во поставките. Овој пристап обезбедува постојаност и сигурност при понатамошната анализа на моделите.

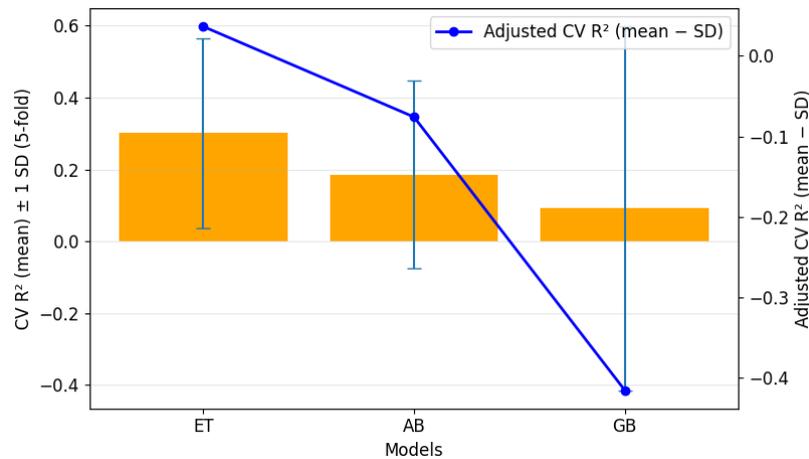
За понатамошна работа се задржуваат три модели со највисоки резултати и стабилност: GB со $m=7$, ET со $m=9$ и AB со $m=8$. Во споредба со XGB, AB покажува нешто повисок резултат и добивка по дополнителното подесување, додека XGB е блиску, но почувствителен е на мали промени. Во следната фаза на валидација ќе се утврди кој од овие три модели ќе се избере како носечки за определување на небезбедни делници со соодветен број параметри.

7.5.2 Избор на стабилен носечки модел

Резултатите од вкрстената валидација со 5 преклопи се користат за споредба на трите финални техники и за избор на стабилен носечки модел при фиксна 80/20 поделба и унифицирана подготовка без протекување на информации. Оценувањето се води од CV R^2 (просек) и неговата стандардна девијација, корегираниот показател Adjusted CV $R^2 = \text{просек} - SD$, метриците Top-k capture за $k = 5\%, 10\%, 20\%$ и распределбата на R^2 по преклопи.

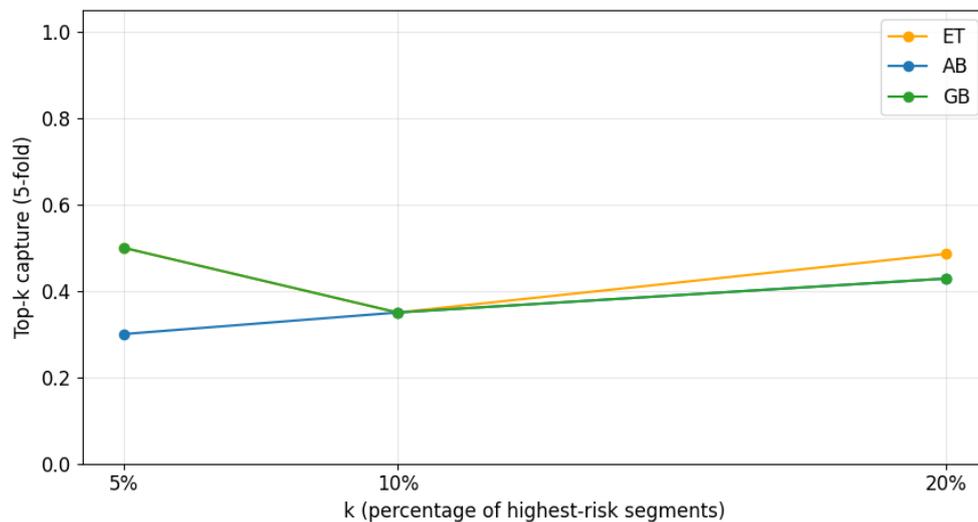
Како што е прикажано на слика 7.16, столбовите го претставуваат CV R^2 (просек) $\pm 1 SD$, а сината линија го дава Adjusted CV R^2 . ET се издвојува со највисока точност и најголем корегирани показател, што покажува дека има мала чувствителност на преклопување и добра способност да ги генерализира резултатите. AB е блиску до него, но има малку понизок корегирани показател и умерена нестабилност. GB има пониска средна точност и

поголема варијабилност, па неговите проценки се помалку сигурни и постојани во однос на другите модели.



Слика 7.16 Точности и стабилности по модел (5-fold)

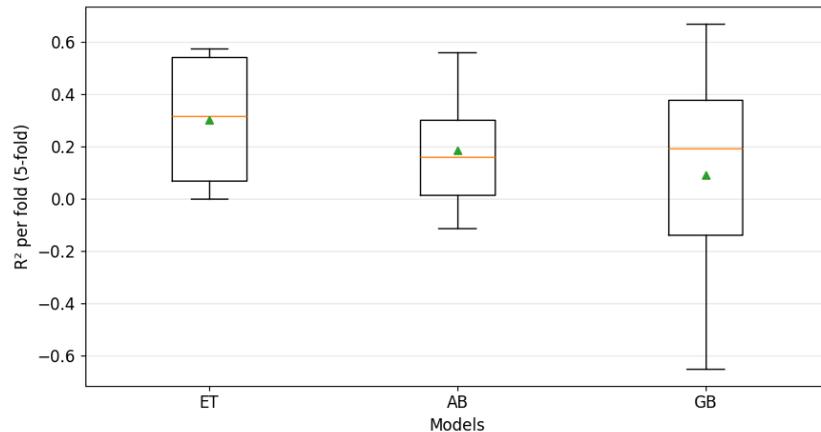
На слика 7.17 се дадени криви што го покажуваат колку добро секој модел ги открива најризичните делници за даден дел од врвот на рангирањето, односно опфат на врвните k делници (top k capture). ET одржува предност за сите вредности на k и препознава поголем дел од најопасните делници во врвот на ранг-листата. AB постепено се подобрува кога k се зголемува, што значи дека добро функционира кога се разгледуваат пошироки групи на делници. GB е почувствителен на изборот на праг, резултатите се намалуваат кај k=10%, а потоа делумно се подобруваат кај k=20%, што укажува на послаба точност во најтесниот врв во споредба со ET и AB. Овој показател е важен, бидејќи помага при избор на ограничен број делници што први ќе се проверуваат на терен, каде повисок опфат на врвните k делници значи поголема ефикасност во приоритетното таргетирање.



Слика 7.17 Точности и стабилности по модел (5-fold)

На слика 7.18 прикажана е распределбата на R² по преклопи, која покажува колку се стабилни резултатите при повеќекратна проверка на моделите. Кај ET кутијата е потесна, а крајните интервали се покуси, што значи дека резултатите се поблиску еден до друг и дека моделот дава усогласени вредности. Кај AB ширината е умерена, што укажува на контролирана променливост и стабилни перформанси. Кај GB распонот е најголем, со појава на неколку пониски вредности, што укажува на поголема нестабилност и помала

сигурност во проценките. Овие разлики покажуваат дека ЕТ има најконзистентни резултати, АВ умерена стабилност, а GB најголема варијабилност при повеќекратно тестирање.

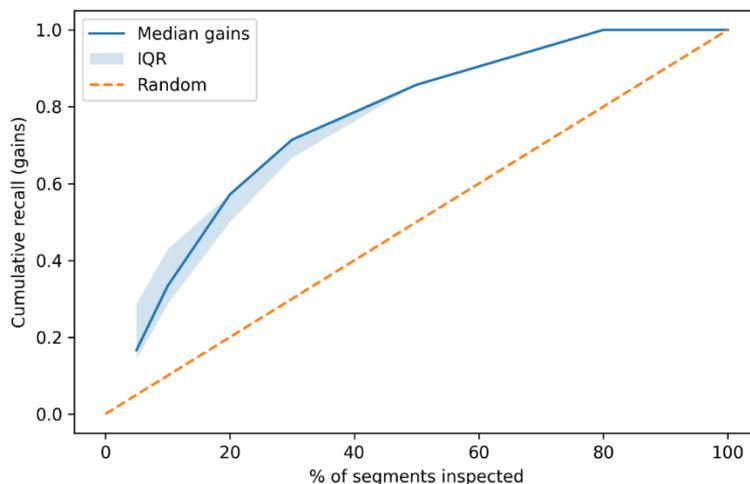


Слика 7.18 Распределба на R^2 по преклопи (5-fold)

Според двата критериуми за споредба, збир на рангови по четири ставки (CV R^2 , Spearman ρ , Top-10% capture, SD) и коригиран показател од вкрстена валидација (Adjusted CV), ЕТ има најдобар пласман, АВ е втор, а GB заостанува поради пониска точност и поголема променливост. Овој распоред ја потврдува стабилноста на ЕТ во однос на останатите избори. Во оваа фаза, ЕТ се усвојува како носечки модел, а АВ и GB се задржуваат како споредбени техники за следните чекори, калибрација, дијагностика на остаточни грешки и проверка со метрики за рангирање на ризик, со цел практична идентификација на делници со висок W_i .

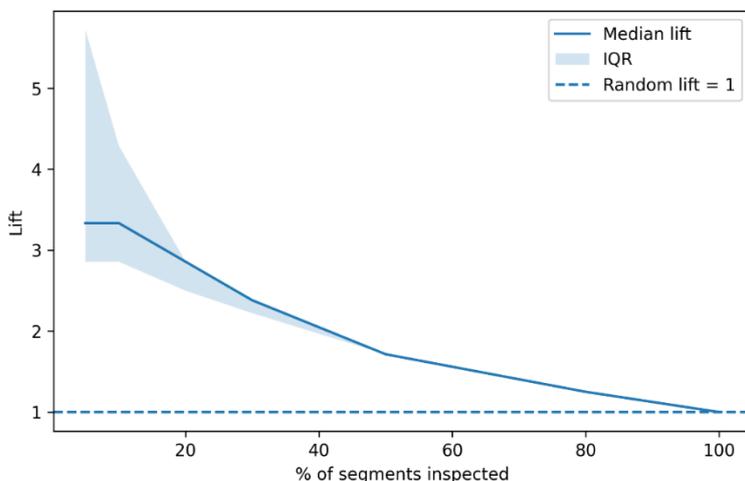
7.5.3 Тестирање и валидација

Во овој дел се оценуваат перформансите преку тестирање на усвоената техника ЕТ. Подготовката на податоците е унифицирана и направена без протекување на информации а валидацијата се спроведува како вкрстена проверка со 5 преклопи при што се задржува истиот однос на класите. Континуалната цел W_i се претвора во бинарна цел, при што високоризичната класа ги опфаќа горните $k=20\%$ делници со највисоки вредности на W_i . Моделот се оценува со повеќе показатели: AUROC, PR-AUC, опфат на врвните делници (Top-k capture) за $k=5\%$, 10% и 20% , како и со прецизност, опфат и F1 при праг што опфаќа околу 20% од делниците. Дополнително се прикажуваат добивка и зголемување за полесно толкување на рангирањето и за процена на практичната применливост на моделот.



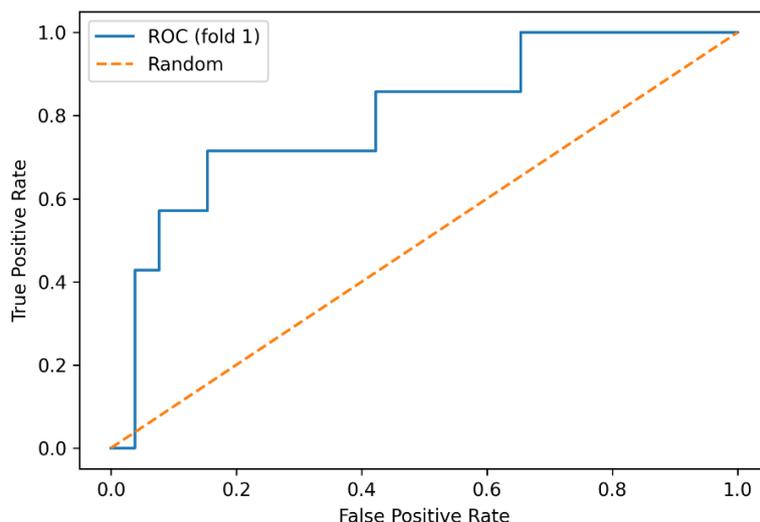
Слика 7.19 Крива на добивка (Gains крива) за ET

Дијаграмот на слика 7.19 прикажува како расте бројот на откриени високоризични делници кога постепено се земаат првите x проценти од рангираната листа според моделот. Кривата со средна вредност по преклопи е над линијата која претставува случаен избор, што значи дека моделот успева да ги концентрира најризичните делници на врвот на листата. При ниска до умерена покриеност, меѓу 10% и 30% од делниците, моделот препознава значително поголем дел од високоризичните случаи во однос на тоа што би се очекувало при случаен избор. Ова укажува дека рангирањето е ефективно и овозможува приоритетна идентификација на делници со зголемен ризик.



Слика 7.20 Крива на зголемување (Lift крива) за ET

На слика 7.20 прикажана е кривата на зголемување (Lift). На почеток вредностите се јасно над 1, а со зголемување на покриеноста постепено опаѓаат, што е очекуван распоред за модел со добра способност за разликување. Вредности над 1 значат предност во изборот на приоритетни делници, особено при мали k .



Слика 7.21 ROC крива со референџна дијагонала

Приказот на слика 7.21 покажува како се менуваат стапката на точно откриени ризични делници и стапката на лажно означени делници кога се менува границата на одлука (пр. од ризична кон безбедна). Кривата е кон горниот лев агол и е усогласена со резултатите од вкрстената валидација, што ја потврдува способноста на моделот да разликува високо ризични, од останати делници.

Главните вредности од вкрстената валидација со 5 преклопи (централна вредност \pm среден 50% опсер) се: AUROC = 0.82 ± 0.08 , PR-AUC = 0.59 ± 0.15 при удел на високоризични околу 20%. За Top-k capture се постигнати приближно 0.21, 0.37 и 0.55 за k = 5%, 10%, 20%. При праг што селектира околу 20% делници, прецизност = 0.51, опфат = 0.55 и F1 = 0.53. Овие резултати се усогласени со добивката и со кривата на зголемување и покажуваат добар баланс меѓу точност и стабилност на моделот, што потврдува дека тој успешно ги издвојува делниците со зголемен ризик и обезбедува доследни резултати при различни поставки.

Валидираниот ET модел јасно разликува делници со повисок ризик и е практичен за нивно приоритизирање, па претставува соодветен избор за целите на оваа анализа.

7.6 Баесови мрежи

Класификацијата на Wi се поставува во три нивоа со однапред утврдени граници 3.60 и 7.38 (квантилна дискретизација), така што класите се приближно подеднакво застапени. Овој распоред овозможува споредливост меѓу модели и полесна интерпретација на матриците на забуну и на мерките по класа. Проценката вклучува дискриминација (ROC-AUC), вкупна точност, прецизност, опфат и F1 по класа, како и калибрација на веројатностите (reliability-криви и Brier-скор).

7.6.1 GTT (Greedy Thick Thinning)

Во оваа мрежа, главниот јазол Wi има најсилни директни јазли Lat_Force, PCI, PGDS, SSD_H, SSD_V, H_sign, V_sign. Влијанијата од околината и коловозните нерамнини главно следат SKIDRES - RDB_RUT - PCI - Wi и Max_Snow - Min_Temp - Ave_Rain - SKIDRES - PCI - Wi, додека геометриското влијание најчесто следи Ave_Rad - Lat_Force - Wi. Се издвојуваат

две доминантни патеки: од коловоз и нерамнини кон Wi и од прегледност и опрема кон Wi.

Табела 7.9 Матрица на забуну - GTT

Предвидено	Оптимална	Умерена	Критична	Вкупно
Оптимална	47	5	1	53
Умерена	6	43	5	54
Критична	2	8	44	54
Вкупно	55	56	50	161

Како што се гледа од табела 7.9, GTT постигнува вкупна точност околу 0.83 и AUROC околу 0.98, при што промашувањата се претежно меѓу соседни нивоа.

Со оглед на високата точност и AUROC, мрежата е практична за класификација на ризик и за насочување на мерки кон делници со послаб коловоз, недоволна сообраќајна опременост и неповолни геометриски услови.

7.6.2 PC (Peter-Clark)

Во оваа мрежа, главниот јазол Wi прима поширок сет непосредни влијанија. Покрај Lat_Force, PCI, PGDS, SKIDRES, SSD_V, H_sign, V_sign, се активираат и зависности меѓу предикторите. Се издвојуваат две пошироки патеки: Ave_Rad - Ave_Inc - SKIDRES - Lat_Force - Wi и Ave_Rain и Max_Snow - SSD_H - SSD_V - H_sign - V_sign - Wi, што овозможува комплетно читање на комбинирани ефекти.

Табела 7.10 Матрица на забуну - PC

Предвидено	Оптимална	Умерена	Критична	Вкупно
Оптимална	48	4	1	53
Умерена	5	45	4	54
Критична	1	9	44	54
Вкупно	54	58	49	161

Од табела 7.10 се чита вкупна точност = 0.85 и AUROC = 0.97, со карта на зависности што овозможува јасно толкување на заедничките влијанија, а грешките се јавуваат претежно меѓу соседни нивоа.

Комбинирано, PC задржува висока точност, покажува добар резултат на AUROC и овозможува читање на заедничките ефекти од геометријата, состојбата на коловоз и прегледноста преку поврзаните јазли. Овој резултат има практична вредност за утврдување приоритетни делници, бидејќи покрај рангирањето на ризик помага да се утврди каде истовремено дејствуваат повеќе фактори и каде мерките треба да се комбинираат, како подобрување на состојбата на коловоз со засилена сигнализација и контрола на брзина.

7.6.3 TAN (Tree-Augmented Naive Bayes)

Во оваа мрежа, главниот јазол Wi е во улога на заеднички главен јазол, а секој предиктор има уште една зависност од друг предиктор, што создава плитко дрво. Се препознаваат три јасни патеки: PCI - IRI - RDB_RUT - Wi (коловоз и нерамнини), SSD_H - SSD_V - H_sign - V_sign - Wi (прегледност и сигнализација) и Max_Snow - Min_Temp -

Ave_Rain - Ave_Rad/Ave_Height - Wi (клима и геометрија). Оваа конфигурација задржува едноставност и транспарентност на толкувањето.

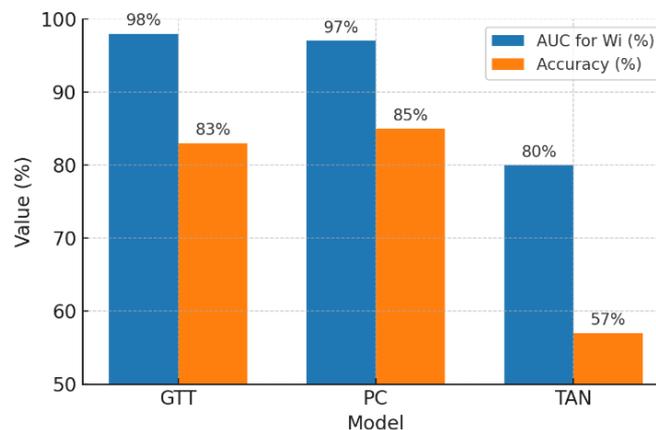
Табела 7.11 Матрица на забуну - TAN

Предвидено	Оптимална	Умерена	Критична	Вкупно
Оптимална	31	11	11	53
Умерена	11	28	15	54
Критична	3	18	33	54
Вкупно	45	57	59	161

Како што се гледа од табела 7.11, TAN има вкупна точност околу 0.57 и AUROC околу 0.80, при што ја задржува најдобрата калибрација на веројатностите. Иако способноста за разликување е пониска од GTT и PC, калибрацијата е добра, па моделот е корисен за поставување граници на ризик и за оперативна распределба на проверки според очекуваната добивка.

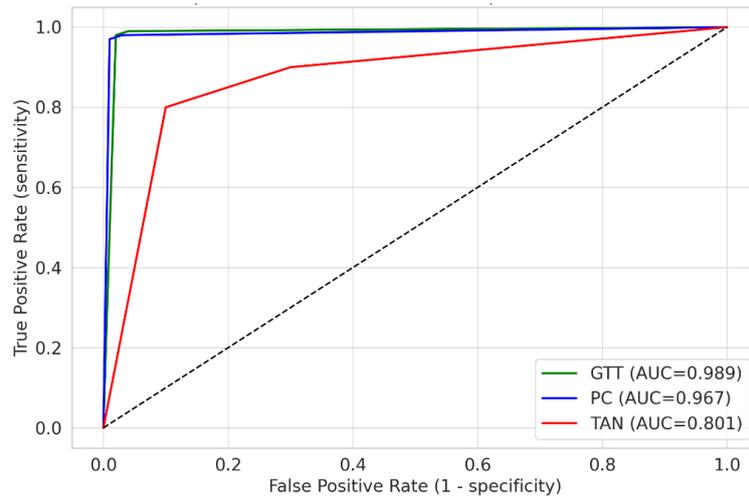
7.6.4 Споредбени дијаграми и калибрација

Овој споредбен преглед ги обединува три показатели: способноста за разликување AUC, вкупната точност и усогласеноста на веројатностите со набљудуваните вредности. Показателите се прикажани со исти оски и легенда за трите пристапи, што овозможува нивна директна споредба.



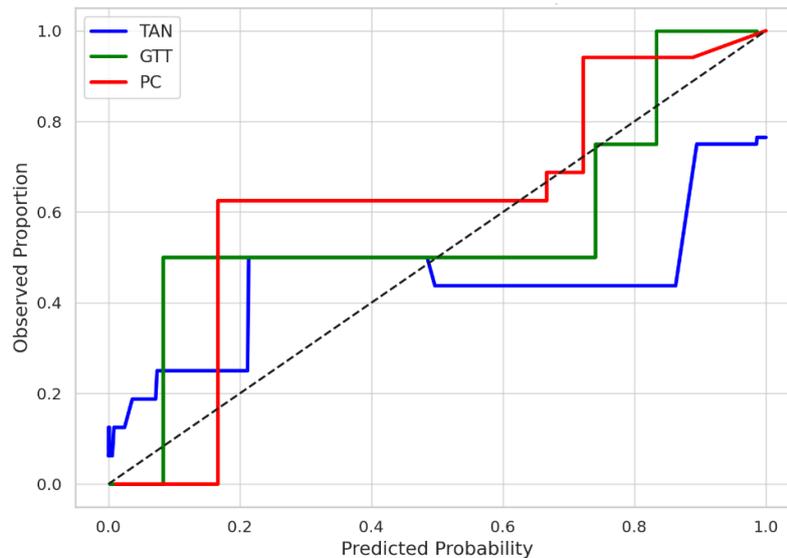
Слика 7.22 Споредба на AUC и вкупна точност

На слика 7.22 прикажана е споредбата меѓу трите баесови мрежи. GTT и PC имаат речиси идентична способност за разликување, при што GTT покажува мала предност според AUC, а PC има повисока вкупна точност. TAN заостанува и по AUC и по точност, што го потврдува послабото совпаѓање со реалните вредности. Овие резултати се усогласени со сумарните метрики и дополнително ја потврдуваат стабилноста и точноста на првите два пристапи.



Слика 7.23 ROC-криви за класа $W_i = Optimal$

На слика 7.23 се гледа дека кривата на GTT е најблиску до идеалниот агол, што значи висока чувствителност при многу ниски погрешно позитивни стапки и многу висок AUC. PC е веднаш зад него со речиси иста форма на кривата, што укажува на слична способност за разликување. TAN има порамнета траекторија со поголема оддалеченост од идеалниот агол, што значи пониска чувствителност кога се бара мала стапка на погрешно позитивни случаи и понизок AUC во однос на GTT и PC.



Слика 7.24 Калибрациски (reliability) криви и HL-тест

На слика 7.24 TAN има најниска HL-статистика и линија најблиска до идеалната дијагонала, што значи најдобро совпаѓање меѓу пресметаните веројатности и набљудуваните фреквенции. GTT и PC задржуваат повисока способност за разликување, со мало отстапување во усогласувањето.

Трите дијаграми покажуваат размена меѓу највисока способност за разликување кај GTT и PC и најдобра калибрација кај TAN. Ако приоритет е откривање на најризични делници, се избира GTT или PC, а ако е важна усогласеноста на веројатностите за прагови и оперативна распределба, предност има TAN.

7.6.5 Избор на водечки модел

Од практична гледна точка, водечки модели се GTT и PC. GTT погоден е за брз преглед на мрежата со висока способност за разликување, па брзо ги издвојува делниците со највисок ризик. PC обезбедува подетален приказ на релации и е посоодветен кога се бара подлабока анализа и објаснување на заедничките влијанија. TAN останува корисен кога приоритет се калибрацијата и јасниот графички приказ, бидејќи веројатностите се подобро усогласени со исходите. Независно од изборот, сите модели даваат резултати во три нивоа и веројатности по делница, што е применливо за приоритизација на теренски прегледи и планирање на мерки.

7.7 Категоризација на безбедноста на делниците

Категоризацијата овозможува брзо и јасно да се види кои делници имаат зголемен ризик, полесно да се споредат резултати од различни техники и да се одредат приоритети за мерки во планирање и управување со мрежата. Континуалните вредности на W_i се претвораат во три нивоа на ризик, се разгледуваат четири начини на поделба и се усвојуваат практични граници што се лесни за примена.

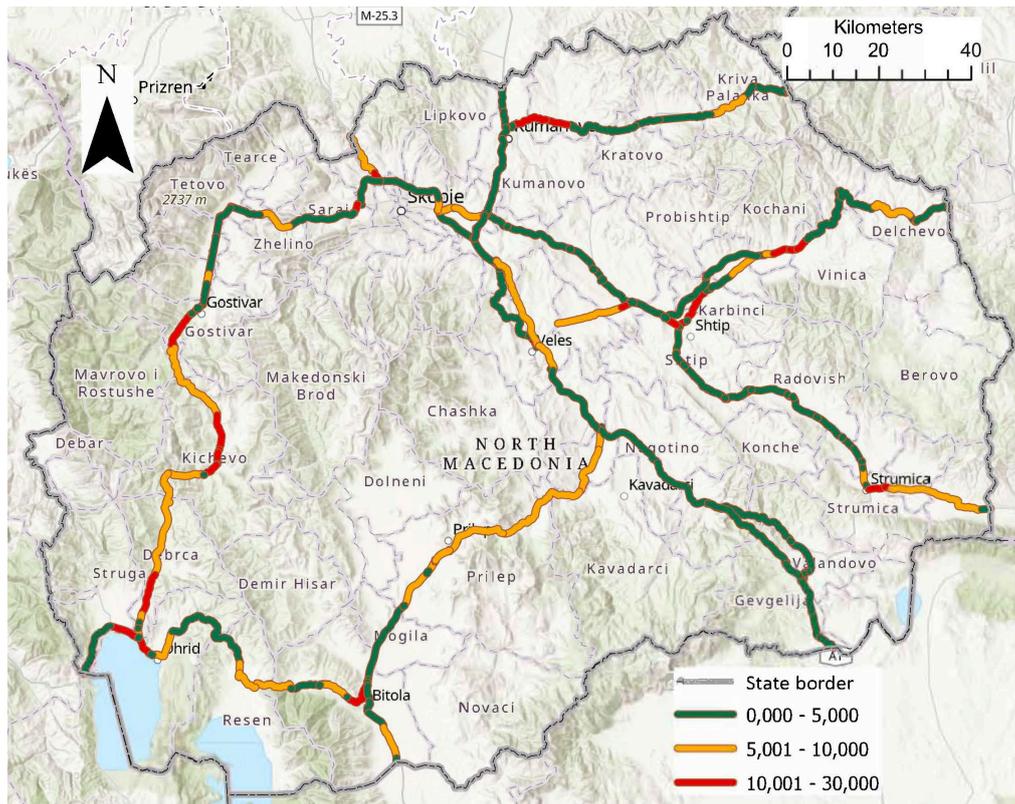
Методи за категоризација:

- **Квантилна распределба:** поделба според процентни граници p_{33} и p_{66} , со приближно еднаков број делници по класа, погодна кога распределбата е изразено нерамномерна.
- **Еднакви интервали:** поделба на опсегот 0–30 во три еднакви сегменти, логична и лесна за примена.
- **Логаритамска распределба:** лог-трансформација за намалување на влијанието на високи вредности и појасно одделување при изразено издолжени краеве на распределбата.
- **Z-score распределба:** класификација според отстапување од средината, изразено во стандардни девијации, корисна за издвојување значајни отстапувања.

Табела 7.12 Категоризација во три оперативни граници

Категорија	Опис (ризик)	Квантилна	Еднакви интервали	Логаритамска	Z-score	Оперативни граници
Безбедна	Низок	0.09 – 4.09	0 – 10	0.10 – 0.98	0.09 – 6.72	0 – 5
Умерена	Среден	4.09 – 6.10	10 – 20	0.98 – 3.06	6.72 – 12.70	5 – 10
Небезбедна	Висок	6.10 – 30.00	20 – 30	3.06 – 30.00	≥ 12.70 –30.00	≥ 10 – 30

Како што е прикажано во табела 7.12, четирите пристапи даваат слична рамка за одделување на делниците. Усоените оперативни граници 0–5, 5–10 и ≥ 10 –30 претставуваат практично решение што ја поедноставува интерпретацијата, овозможува единствена основа за споредливост на резултатите и јасно укажува на границата за приоритетни интервенции.



Слика 7.25 Карта на безбедноста на делниците според оперативните граници (W_i)

На слика 7.25 прикажана е просторната распределба на ризикот по делници со обоено обележување на трите категории: безбедна (0-5), умерена (5-10) и небезбедна ($\geq 10-30$). Ваквиот приказ овозможува едноставна визуелна оцена на критичните делници и подобро разбирање на географската распределба на ризикот, што директно помага во изборот на мерки на терен и комуникацијата со носителите на одлуки.

Овој пристап ги поврзува резултатите од анализата со конкретната локација на мапата и дава јасни визуелни знаци за приоритет. Така се создава рамка што е постојана и корисна за следните чекори во процената и управувањето со безбедноста на патната мрежа.

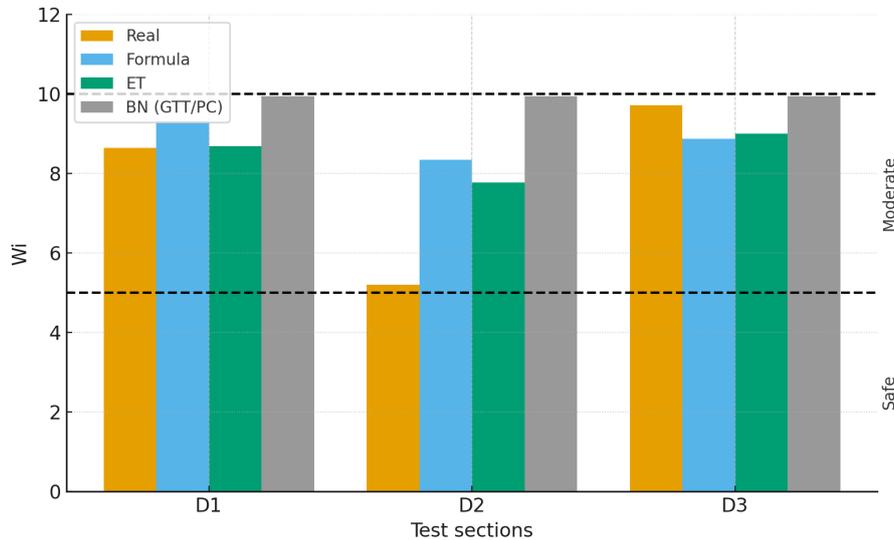
7.8 Надворешна валидација

Надворешната проверка покажува како моделите се однесуваат на други регионални делници што не биле користени за учење. Целта е да се види, дали резултатите се усогласени и корисни кога треба да се одредат приоритети за мерки. За секоја делница се споредува реалната вредност на W_i со вредностите добиени со формула, со моделот ET и со баесовите мрежи при поделба во три нивоа 0-5, 5-10 и $\geq 10-30$.

Табела 7.13 Надворешна валидација за три регионални делници

Делница	Реален W_i	Кл. (real)	Формула: W_i	Кл. (Φ)	$ \Delta W_i $ (Φ)	МЛ (ET): W_i	Кл. (ET)	$ \Delta W_i $ (ET)	Баес (GTT)	Баес (PC)	Баес (TAN)
D1	8.64	Умер.	9.31	Умер.	0.67	8.69	Умер.	0.05	Умер.	Умер.	Умер.
D2	5.20	Умер.	8.35	Умер.	3.15	7.77	Умер.	2.57	Умер.	Умер.	Умер.
D3	9.71	Умер.	8.87	Умер.	0.84	9.00	Умер.	0.71	Умер.	Умер.	Небезбедна

Како што е прикажано во табела 7.13, споредбата на реалните и пресметаните вредности по делници и апсолутните разлики $|\Delta W_i|$ дава јасен увид во отстапувањата на нумеричка скала. Формулата и ЕТ најчесто доделуваат умерена класа за трите делници, со најголемо отстапување кај D2 бидејќи W_i е во граничниот појас 5–10. Баесовите мрежи GTT и PC исто така упатуваат на умерена класа, додека TAN кај D3 означува небезбедна класа, што сигнализира можен граничен ризик околу прагот 10 и потреба од дополнителна претпазливост.



Слика 7.26 Реално и пресметано W_i по делници

На слика 7.26 се гледа спореден приказ на резултатите за трите делници и начинот на кој методите го определуваат ризикот. Може да се констатира дека сите техники генерално ја препознаваат разликата меѓу побезбедни и поризични делови, но со мали разлики во нивните проценки. Формулата и моделот ЕТ покажуваат усогласени вредности со реалниот W_i и добро ги опфаќаат делниците со среден ризик. Баесовите мрежи најчесто се блиску до границите меѓу категориите, со што сигурно ги откриваат потенцијално ризичните случаи. Дијаграмот потврдува дека методите сигурно и усогласено го детерминираат нивото на ризик по делница.

Комбинирањето на табеларниот приказ, кој ги прикажува разликите во податоци, со графичкиот приказ по метод, кој ја покажува позицијата на делниците во однос на границите, обезбедува сеопфатен увид во однесувањето на техниките на независни регионални делници. Ова го олеснува препознавањето на гранични случаи околу 5 и 10, поддржува образложено рангирање на ризик и создава основа за планирање интервенции на најрелевантните делници.

Иако во нумеричкиот приказ се забележуваат разлики во вредностите на W_i меѓу методите, при примената на оперативните граници 0–5, 5–10 и ≥ 10 овие разлики ретко влијаат врз категоризацијата, бидејќи добиените класи се усогласени со реалните. Ова покажува дека трите пристапи, односно формулата, ЕТ и Баесовите мрежи, ја вршат својата функција за процена на безбедноста и можат да се користат како доверливи и меѓусебно комплементарни алатки во практиката.

8. НАЧИН НА ИМПЛЕМЕНТАЦИЈА

Овој сумарен преглед ја прикажува постапката и начинот на разгледување на резултатите за утврдување на небезбедни делници и редоследот на приоритети за подобрување. Целта е да се прикаже процесот со кој, од податоците се доаѓа до конечна одлука за имплементација на терен. Во прегледот, редоследно се опишуваат обработка на податоците, кластер анализа, анализа на влијателност, главната анализа, валидација, категории и карти за работа на терен. Акцентот е ставен на карактеристиките на патот и сообраќајот, бидејќи преку нив најмногу се намалува ризикот.

Обработка на податоци

Податоците се собрани од повеќе извори, при што поголемиот дел потекнува од институционални бази на податоци. За неколку параметри потребна е теренска посета за дополнително снимање и утврдување на конкретни вредности (PCI, SSD и KON.NAK.). Мерењата се изведуваат по едноставна постапка на лице место, според оперативни упатства. Упатствата опфаќаат насоки за евидентирање на коловозни оштетувања (PCI), попис на непрегледни локации, според прегледност на застанување (SSD) и броење на контранаклони (KON.NAK.). Теренското снимање може да го спроведе и лице без посебна стручна квалификација. Доволно е да се следат упатствата за евиденција и внесување на податоците.

Секоја вредност е геореференцирана за соодветна делница, а мерните единици се изедначени, односно сите мерења се претворени на исти основи за споредба (брзини во km/h, должини и бројни показатели по km). Дополнително, сите поединечни податоци се нормализирани по должина на делница и се изразени по километар. Постапката вклучува отстранување на дуплирани податоци и очигледни грешки, пополнување на вредности кои недостасуваат преку споредба на повеќе извори и ограничување на екстремните вредности на реални граници за да не се наруши распределбата на податоците. На овој начин се добива чиста база на податоци во која патните, сообраќајните и карактеристиките на околината можат да се споредуваат. Тежинскиот индекс на незгоди се пресметува по делница и се прилагодува според изложеност по километар. Патните параметри опфаќаат прегледност, радиуси на хоризонтални кривини, надолжни наклони, состојба на коловоз, како и хоризонтална и вертикална сигнализација. Сообраќајните податоци ги опфаќаат ограничувањето на брзините и ПГДС, а карактеристиките на околината опфаќаат врнежи од дожд и снег, средни температури и надморска височина. Вака подготвената база овозможува основана проценка на влијателност и стабилни споредби меѓу делници.

Кластер анализа

Делниците се групирани според патни, сообраќајни и карактеристики на околина, користејќи финален стандарден сет од 23 влезни параметри, при што тежинскиот индекс на незгоди не е вклучен во анализата. Поделбата е направена во природни групи на делници со слични карактеристики. Кластерите се користат како второ ниво за одредување приоритети за подобрување на делниците поради безбедносниот ризик. По почетното рангирање по W_i , секој кластер се разгледува одделно и предност добиваат делниците што најмногу отстапуваат над типичната вредност за својот кластер. Овој пристап избегнува споредба на делници кои не се сродни, ги насочува мерките според типот на делницата и овозможува побрза селекција на интервенции во рамки на хомогени групи.

Класиџер 1: Умерен ризик. Делници на поголема надморска висина со пониско ограничување на брзина (60 km/h) и помали радиуси на хоризонтални кривини (530 m). Коловозните показатели се во умерени граници, а тежинскиот индекс на сообраќајни незгоди е понизок во однос на групите со изразени колотрази. Профилот упатува на траси со повеќе хоризонтални кривини, каде брзините се минимални и одржувањето на коловозот е на прифатливо ниво.

Класиџер 2: Умерен ризик. Делници на пониска надморска висина со повисоки средни температури и поголеми радиуси на хоризонтална кривина (1900 m). Во овој профил се дефинирани повисоки брзини (110 km/h), висок индекс на квалитет на коловоз и мали колотрази. Ваквата комбинација поврзана е со понизок ризик од сообраќајни незгоди и упатува на траси со поголеми радиуси на хоризонтална кривина и добра состојба на коловоз.

Класиџер 3: Висок ризик. Овие делници имаат големи радиуси на хоризонтална кривина (3250 m) и повисоки средни температури, а на коловозот се забележува понизок индекс на квалитет и појава на колотрази. Овде тежинскиот индекс на сообраќајни незгоди е највисок, што покажува дека понизок индекс на квалитет на коловоз и присутни колотрази го зголемуваат ризикот иако геометријата е поволна. Во овој профил најпрво треба да се рехабилитира коловозната површина и да се промени ограничувањето на брзина на критичните места.

Класиџер 4: Висок ризик. Делници со минимални ограничувања во хоризонталната геометрија (радиуси од 1300 m) и просечни брзини на движење (околу 80 km/h), со изразени колотрази и понизок индекс на квалитет на коловоз. Овој профил е поврзан со повисоки вредности на тежинскиот индекс на незгоди и укажува дека колотразите и изразената површинска оштетеност се главен извор на ризик, дури и кога геометријата е во умерени граници.

Класиџер 5: Умерен ризик. Делници со помали радиуси на хоризонтални кривини (630 m) и пониско ограничување на брзина (60 km/h). Тежинскиот индекс на сообраќајни незгоди е со средни вредности, а профилот укажува на потреба од локални корекции на геометријата и на хоризонталната сигнализација. Во вакви услови, мали корекции на хоризонталното решение и почеста обнова на хоризонталната сигнализација можат брзо да дадат ефект во намалување на тежинскиот индекс на сообраќајни незгоди.

Класиџер 6: Низок ризик. Ова е најмал профил и опфаќа делница на голема надморска висина (1120 m) со ниска средна (годишна) температура, мали радиуси на хоризонтални кривини (97 m) и поголема дебелина на снежен нанос. Поради ниските вредности на тежинскиот индекс на сообраќајни незгоди, кластерот во разгледуваниот примерок се смета за најбезбеден. Застапеноста е многу мала, затоа резултатот се третира како описен и не се користи за пошироки заклучоци.

Сумирано, профилите најмногу се разликуваат по комбинација од радиуси на хоризонтални кривини, состојба на коловоз, ограничување на брзина, надморска висина и температури. Овие разлики помагаат делниците да се читаат по профил и да се насочат мерки што одговараат на конкретниот тип.

Анализа на влијателности

Овој дел има за цел практично да утврди кои параметри најмногу го менуваат тежинскиот индекс на сообраќајни незгоди.

Прво се посочуваат сообраќајните карактеристики. Сообраќајното оптоварување и ограничувањето на брзина најчесто се меѓу водечките фактори. Кога ПГДС е висок и ограничувањето на брзина е повисоко, растат изложеноста и енергијата при судир, па ризикот е поголем. Потоа се разгледуваат патните карактеристики. Најмногу влијаат прегледноста, радиусите на хоризонталните кривини, надолжните наклони и состојбата на коловозот. Од коловозните показатели важни се индексот на квалитет на коловоз (PCI), рамноста на коловозот (IRI) и колотразите. Своја улога имаат и прегледноста на застанување и сообраќајната сигнализација.

Оваа подреденост на влијателните параметри може директно да се примени во пракса како унифицирана стратегија за иднина, без потреба од дополнителна анализа. Најпрво, се фокусира на управување со сообраќајното оптоварување и ограничување на брзината, што овозможува брзи и ефикасни промени во тековните услови. Потоа, се преминува кон рехабилитација на коловозната површина и обновување на хоризонталната и вертикалната сигнализација, за да се подобри видливоста и стабилноста при возење. Доколку по овие мерки ризикот сè уште останува висок, се пристапува кон посложени интервенции, како реконструкција или корекции на геометријата на патот, со цел долгорочно зголемување на безбедноста.

На крај, рангирањето по влијателност служи и како влез во останатите анализи и постапки како избор на параметри за пресметки, поставување граници и категории, подготовка на карти за работа и насочување на мерки по тип на делница.

Интегративна формула

Формулата служи за определување на тежинскиот индекс на сообраќајни незгоди (W_i) по делница, со што се добива единствена бројна вредност за споредба и дефинирање на приоритети за интервенции. Резултатот може да се следи во текот на времето за да се види тренд на подобрување или влошување на безбедносниот ризик.

Во формулата се применети 7 параметри. Заедно со константата, нивните квадрати и парни интеракции, се добива полином од втор степен со 36 членови. Параметрите се однесуваат на карактеристиките на патот и сообраќајот: ПГДС, ограничувањето на брзина, густина на патни јазли/крстосници/наплатни станици, индекс на квалитет на коловоз (PCI), вертикална сигнализација, просечен надолжен наклон и индикатор за контранаклони.

Податоците директно се внесуваат во датотека (Excel/CSV), во која е поставена готова пресметка според формулата. Образецот содржи преддефинирани полиња и проверки на внесување за да се избегнат грешки. По внесувањето на податоците, пресметката автоматски го дава тежинскиот индекс W_i за анализираната делница и за него може веднаш да се дефинира категоријата на безбедност според утврдени граници.

Формулата е погодна за брзи проверки на безбедносниот ризик и за избор на следна интервенција по делница. На овој начин се добива оперативен увид за планирање на мерки и за следење на ефектот во текот на времето. Резултатите можат веднаш да се прикажат на карта или табеларно (како листа на приоритети).

Машинско учење (ET метод)

ET е техника од машинско учење со дрвовидна структура. Моделот гради повеќе дрва на одлука, при што секое дрво учи правила на поделба преку прагови на вредности на поединечни параметри и при секоја поделба разгледува случајно избрано подмножество

од параметрите, при што конечната проценка на W_i е просек од сите дрва. Обуката се изведува на тренинг-податоци со подготовка по исти правила за влезните податоци, а тестирањето е одвоено за да нема пренос на информации.

Анализата се спроведува со 9 параметри, собрани и нормализирани по километар во еден влезен фајл (Excel/CSV). Овој состав на податоци опфаќа сообраќајно оптоварување, геометрија и состојба на коловоз. Користените параметри се: ПГДС, ограничување на брзина, густина на патни јазли/крстосници/наплатни станици, индекс на квалитет на коловоз (PCI), прегледност на застанување (SSD_V), вертикална сигнализација (V_sign), индикатор за мостови, просечен радиус на хоризонтална кривина и просечна надморска висина.

Подготвениот фајл се внесува во платформа за извршување „скрипти“ (Google Colab, Jupyter Notebook или Kaggle Notebooks), каде скриптата го обучува или повикува ET-моделот и враќа излез во табели и дијаграми.

За моделот да функционира на ист начин, потребно е да се чува историјата од претходните обучувања, вклучувајќи зачувани модели, поставки и статистики од подготовката. Новите податоци се проценуваат со истите правила или се прави повторно тренирање врз целата достапна историја, за резултатите да останат споредливи во текот на времето.

Баесови мрежи

Во анализата со Баесови мрежи тестирањето потврди дека трите структури GTT, PC и TAN се соодветни за следни проверки на безбедноста. Иако сите три пристапи се валидирани и точни, за понатамошна примена се препорачува да се усвои еден носечки модел како стандарден избор, заради усогласеност, одржливост и полесно следење на резултатите. Во оваа поставеност практичен избор е PC поради балансирана точност и стабилност, додека GTT и TAN може да се задржат како споредбени пристапи при повремени проверки. Моделот учи од тренинг-податоци, со иста подготовка за сите влезни вредности, а проверката е одвоена за да нема пренос на информации. На излез се добива очекувана вредност на W_i и веројатност дека W_i е над зададен праг, што овозможува и рангирање на делниците по ризик.

Влезните податоци го опфаќаат целиот сет со 23 параметри, собрани и нормализирани по километар во еден влезен фајл (Excel или CSV). Сите 23 податоци се дискретизираат во неколку нивоа со однапред зададени граници, научени од тренинг-делот, а истите граници потоа се применуваат на сите делници, за резултатите да бидат споредливи во текот на времето.

Подготвениот фајл со дискретизирани податоци се внесува во алатка за Баесови мрежи (GeNIe 5 или Hugin), или во работна околина со библиотеки како bnlearn или pgmpy. Се избира носечкиот модел (PC), каде се учат условните табели на веројатности и се пресметува W_i по делница. На излез се добиваат табели и дијаграми: категории на W_i , веројатности за ризик според зададен праг, како и графици за чувствителност што покажуваат кои параметри најмногу го менуваат W_i .

За моделот да се однесува на ист начин во текот на времето, се чуваат обучената структура, табелите со веројатности и правилата за дискретизација. Новите фајлови треба да имаат исти имиња на колони и истите граници на нивоа при дискретизација. Делниците може директно да се проценуваат со постоечката мрежа или периодично да се прави

повторно тренирање со дополнетата историска база, за резултатите да останат споредливи.

Оваа методологија е соодветна за национални програми за безбедност на патиштата, каде е неопходно да се воспостави механизам за комбинирање на сите релевантни податоци од институциите во единствена база што континуирано ќе се ажурира. Потребни се усогласени формати и мерни единици, договорени правила за дискретизација, јасни процедури за проверка на квалитет, следење на промени на податоците и редовни извештаи. Со ваква поставеност, Баесовите мрежи трајно обезбедуваат механизам за дефинирање на приоритети по делници, планирање на мерки и следење на ефектите во текот на времето.

Валидација

Методите се тестирани и ги исполнуваат критериумите за практична примена, сепак во идни анализи се препорачува редовна проверка за следење на стабилноста на резултатите и навремено воочување на можни отстапувања. За регресија, се препорачува да се користат метриците R^2 , MAE, RMSE, а за класификација Accuracy, F1, ROC-AUC и PR-AUC. Тестирањето треба да се спроведува како дополнителна контрола, со цел да се открие евентуално потфрлање во методологијата, при нормализирањето или дискретизацијата.

Категории и карти за работи

Границите за тежинскиот индекс на сообраќајни незгоди (W_i) дефинирани во методологијата треба да се усвојат како официјални на државно ниво. Тие треба да се применуваат на ист начин во сите анализи, без оглед дали W_i е пресметан со Интегративната формула, со ET моделот, со Баесови мрежи или е добиен при тековен увид во настанати незгоди. Истите граници треба да важат за сите извештаи и циклуси на ажурирање, да обезбедуваат споредливост во текот на времето и да воспоставуваат единствен систем за проверка на безбедносниот ризик.

Овие официјални граници можат да се користат за споредување меѓу методите. За секоја делница да се спроведе анализа по трите пристапи, да се пресмета W_i , да се мапира на истите граници и да се споредат категориите. Крајниот ризик да се утврди врз основа на усогласените категории и ако се поклопуваат да се смета дека проценката е сигурна. Ако не се поклопуваат, делницата да се упати на дополнителен увид, најпрво да се проверат податоците, а потоа да се изврши теренска потврда. На овој начин границите да се применуваат и како заедничка мерка за сите анализи и како алатка за усогласено споредување на резултатите.

За практична примена картите за работа треба да се подготват во ГИС. Делниците да се геореференцираат, а ризикот да се прикажува просторно според официјалните граници со јасна разграничена градација на бои. Да се изработат одделни слоеви за печатење и за веб-приказ, со легенда, размер и датум на верзија. При секое ажурирање картите да се обновуваат од единствената база на податоци, за приказот да остане усогласен во текот на времето.

Како завршна проверка, на сите делници, извршена е споредба по трите методологии (Интегративна формула, ET и Баесови мрежи). Резултатите се усогласени, односно за W_i е добиена иста категоризација според официјалните граници. Картата за работа ги прикажува истите состојби, во согласност со категоризација прикажана на слика 7.25.

9. ДИСКУСИЈА

Во ова поглавје се сумираат главните резултати од анализите и се објаснува нивното значење во контекст на мрежата од државни патишта во Македонија. Целта е да се покаже како добиените модели помагаат при препознавање делници со зголемен ризик и како рангирањето може реално да се искористи при избор на приоритети, при што толкувањето се темели на параметрите што најчесто излегуваат како највлијателни и на условите под кои нивниот ефект е најизразен.

Рамката на пристапот го прикажува начинот на кој се спроведува целата анализа. Најпрво се прави истражувачка кластер анализа, што ги групира делниците во препознатливи профили според три области: геометријата на трасата, состојбата на коловозот и обемот на сообраќај. На овој начин резултатите полесно се гледаат по групи и полесно се применуваат, наместо секоја делница да се разгледува одделно, што е важно бидејќи ризикот не се појавува на ист начин во различни средини. Паралелно се избираат само техники што имаат доволна точност и стабилност, за споредбата да дава стабилни листи, јасен редослед и разумно толкување на влијанијата. Потоа, кај моделот се користат три пристапи. Полиномниот модел дава фино рангирање и покажува како факторите делуваат заедно. Моделите од машинско учење препознаваат посложени и нелинеарни врски и дозволуваат да се види придонесот на секој влез. Баесовите мрежи додаваат јасна структура и калибрирани веројатности, корисни за одлуки на три нивоа и за едноставна комуникација на ризик. Рангирањата главно се поклопуваат, што покажува стабилен избор на најприоритетни делници, според ризик без фаворизирање на еден метод.

Клучните параметри што најмногу го објаснуваат ризикот, се обемот на сообраќај и брзината, потоа геометријата на трасата и прегледноста, како и густината на крстосници, надолжните наклони, состојбата на коловозот и одводнувањето, заедно со елементи на сигнализација. Ризикот расте токму кога овие фактори се спојуваат, па контролата се намалува и веројатноста за незгода се зголемува.

Во практична употреба фокусот на делување се одредува со криви на покриеност и добивка, кои покажуваат колку од ризичните делници се опфатени на врвот на листата, потоа се избира зона за проверка во горниот дел, најчесто првите 10–30%. На овие места има најголема шанса да се откријат реални слабости во безбедноста. За полесно управување, W_i се дели на неколку нивоа за резултатите полесно да се пренесат во реална употреба. Просторниот приказ во ГИС ги поврзува податоците со точни локации, и овозможува на надлежните институции да увидат како се креира ризикот и како се прикажува во простор.

Комбинацијата на табеларниот приказ и графичкиот дел по метод дава целосна слика за однесувањето на методите на независни делници, овозможува основано рангирање на ризик и дава јасна основа за планирање мерки.

Иако постојат разлики меѓу предвидените (детерминирани) и реалните вредности, во најголем дел тие спаѓаат во иста категорија, што покажува дека методите се издржани и меѓусебно се надополнуваат во практика.

9.1 Синтеза на главните резултати

Овој дел ги обединува најважните резултати од анализите и ги претвора во преглед што може директно да се користи за одредување приоритети.

Кластер-анализа со GMM издвојува шест профили на делници со јасни разлики по W_i , при што C3 и C4 имаат централна вредност $W_i=7\ 893/ W_i=7\ 136$ со најголем ризик, додека C6 со $W_i=2\ 009$ служи како граничен показател со најмал ризик. Овие профили овозможуваат резултатите да се читаат и по групи делници, не само поединечно, што помага да се предлагаат мерки со слична логика на ризик.

Интегративна формула со седум влезови обезбедува континуирано рангирање со Test $R^2 = 0.72$. Влезните параметри се PGDS, LIMIT, K.Int.P.T, PCI, V_sign, Ave_Inc, KON_NAK, а нивното заедничко дејство објаснува зошто ризикот расте, кога неповолни состојби се преклопуваат. Во врвот на листата прецизноста е 0.50 на 10% опфат, 0.29 на 20% и 0.30 на 30%, што е подобро од случајно избирање и овозможува реално планирање на опфатот според ресурсите.

Во делот со машинско учење, прво беа тестирани девет алгоритми на фиксна 80/20 поделба, а потоа трите најдобри техники беа проверени со вкрстена валидација со пет преклопи за споредба и стабилност. По минимално прилагодување на клучни параметри, Gradient Boosting постигна Test $R^2 = 0.67$, Extra Trees = 0.63, а AdaBoost = 0.61. Иако Gradient Boosting има највисока детерминација, Extra Trees се оценува како подобар избор за скрининг поради постабилни листи при промена на семиња и поделби. Дополнително, моделот задржува висока сличност на горните 10–30% со базното рангирање и дава конзистентни важности на факторите. Во класификациска рамка за Extra Trees со висока класа околу 20% се добива AUROC = 0.82, PR-AUC = 0.59, Top-k capture = 0.21 на 5%, 0.37 на 10% и 0.55 на 20%. При праг со околу 20% избрани делници, прецизноста е 0.51, опфатот 0.55 и F1 0.53, што ја потврдува оперативната корист на моделот.

Баесовите мрежи обезбедуваат класификација во три нивоа со граници $W_i = 3.60$ и $W_i = 7.38$. Во оваа рамка GTT има точност = 0.83 и ROC-AUC = 0.98, PC точност = 0.85 и ROC-AUC = 0.97, а TAN точност = 0.57 и ROC-AUC = 0.80, при што TAN има најдобра калибрација, а GTT и PC поголема дискриминација. Овој излез погоден е кога треба едноставно соопштување на ризик и јасни граници за дејствување.

Влијателните параметри се стабилни низ техниките. PGDS и LIMIT најчесто се на врвот, по нив следат K.Int.P.T, PCI, SSD_V и V_sign, а тоа е потврдено со SHAP и со пермутациска важност.

Надворешната проверка на три регионални делници покажува добро совпаѓање меѓу анализа и терен. За D1: реален $W_i = 8.64$, формулата $W_i = 9.31$, ET $W_i = 8.69$, и сите класи остануваат умерени. За D2: реален $W_i = 5.20$, формулата $W_i = 8.35$, ET $W_i = 7.77$, со умерена класа што е очекувана за зона блиску до границата 5–10. За D3: реален $W_i = 9.71$, формулата $W_i = 8.87$, ET $W_i = 9.00$, а TAN сигнализира повисоко ниво во близина на $W_i = 10$, што бара поголема оперативна претпазност и увид на лице место.

Методите даваат стабилни и точни резултати, со усогласени рангирања меѓу интегративната формула ($R^2 = 0.72$), техниките од машинско учење ($R^2 = 0.61$ – 0.67) и баесовите мрежи (ассигасу = 0.57–0.85). Врвот на листата има јасна предност, што се гледа преку покриеност и прецизност на првите 10–30 проценти од делниците. Оваа конзистентност овозможува рангирањето по W_i , сигурно да се претвори во конкретни чекори за увид на лице место, планирање мерки и следење на ефектите од интервенциите.

9.2 Толкување на механизмите

Ризикот на една делница произлегува од неколку поврзани дејства што делуваат истовремено. Кога низ една делница поминуваат голем број возила, бројот на потенцијални конфликти расте, а секоја грешка има поголема шанса да заврши со незгода. Ако притоа, ограничувањето на брзина е поголемо од она што го дозволуваат условите на трасата, времето за реакција се скратува, патот на сопирање се издолжува, а маневрите стануваат покритични. Места каде брзината не е усогласена со трасата и околината создаваат варијации на брзините, што води кон судири во задниот дел и повеќе конфликти при престојување и претстигнување.

Геометријата на трасата го обликува просторот за управување со возилото. Помали радиуси на хоризонтални кривини и кратки правци меѓу кривини, бараат почести корекции на насоката и прецизно држење на патеката. Последователни кривини со кратки меѓуправци го зголемуваат заморот и го намалуваат квалитетот на проценката. Надолжните наклони влијаат на можноста за сопирање и забрзување, а напречните наклони ја намалуваат стабилноста при дожд. Овие услови директно се поврзуваат со прегледност на хоризонтални кривини и превои со растојанијата за запирање, па таму каде што линијата на прегледност е кратка, секоја грешка побрзо прераснува во проблем.

Прегледноста и сообраќајната сигнализација ја насочуваат перцепцијата на возачите. Ограничена прегледност на хоризонтални кривини и превои го намалува времето за одлука, а истрошена или недоволна сигнализација ја отежнува проценката на опасноста, особено навечер и при врнежливо време. Кога сигнализацијата и ограничувањето на брзина не се во согласност со геометријата и приклучувањата, возачите доцна ја намалуваат брзината, што ја зголемува веројатноста за погрешна проценка и конфликт.

Приклучувањата и јазлите ја зголемуваат сложеноста на движењето. Поголема густина на влезови и излези создава повеќе точки на пресек и мешање на текови. Во тие зони има почести промени на брзина и често престојување, што носи зголемен ризик за судири во задниот дел и странични судири. Овој ефект е поизразен кога приклучувањата се блиску до кривини или во зони со ограничена прегледност.

Состојбата на коловозот и одводнувањето влијаат на контактот меѓу пневматикот и асфалтот и на стабилноста на возилото. Задржана вода во вдлабнатини и на ниски точки, деформации на коловоз и намалена рамност го забрзуваат деградирањето на коловозот и ја влошуваат контролата на возилото, па патот на сопирање се продолжува. Овој ефект е изразен на делници со повисоки брзини и во надолнини, каде малку посилено кочење или ненадејно избегнување, лесно водат кон пролизгување.

Климатските услови и надморската височина најчесто делуваат посредно. Пониски температури и снег го забрзуваат деградирањето на коловозот а дождот и маглата ја намалуваат прегледноста и триењето меѓу пневматикот и асфалтот. Иако овие фактори генерално имаат понизок директен показател во моделите, во одредени комбинации можат да бидат пресудни, како при истрошен коловоз и поголем надолжен наклон.

Објектите на патот често имаат споеви и поинаков коловоз од околната траса. На премините кај дилатации, во сенки и на места кај реки и долови коловозот е постуден, а влагата се задржува подолго, што ја намалува стабилноста, особено при дожд. Овие точки бараат понагласено сообраќајно означување, подобро одводнување и редовно одржување, иако поединечно ретко се најсилниот фактор.

Ризикот најмногу расте кога неколку неповолни состојби се појавуваат заедно. Типични примери се делници со големо сообраќајно оптоварување и погрешно ограничување на брзина и при тоа со повеќе приклучувања, со кривини и ограничена прегледност, или делници со поголеми наклони и послаб контакт со коловозот поради вода и оштетувања. Во такви услови и мала промена на брзината, растојанието меѓу возилата или вниманието може значително да го промени исходот.

Од практичен аспект, овој збир на механизми укажува на јасни точки за теренска проверка од стручен тим. Онаму каде брзината не одговара на трасата, потребна е ревизија на ограничувањата и означувањето. Онаму каде има многу приклучувања, корисни се мерки за управување со пристапите и подобрување на прегледноста. Таму каде има вода и оштетувања, приоритет се одводнување и обновување на коловозната површина. На делници со последователни кривини и кратки правци, треба да се разгледаат мали геометриски корекции и решенија за предупредување што го задржуваат вниманието на возачите.

Оваа констатација помага, редоследот во рангирањето да биде заснован на причините, а не само на податоците и ја насочува практичната селекција на делници со највисок приоритет.

9.3 Усогласеност со светската литература

Резултатите се во согласност со светските истражувања. Обемот на сообраќај и режимот на брзините се главни поттикнувачи на ризик. Геометријата и прегледноста одредуваат колку простор и време има возачот за реакција, а густината на приклучувања, наклоните и состојбата на коловозот го засилуваат ефектот според локалниот контекст. Времето и надморската височина најчесто делуваат посредно, преку контактот пневматик-асфалт и преку прегледноста, што е во согласност со трудовите каде климата има улога на посредник, а не на главен фактор.

Пристапот со три меѓусебни техники на надополнување е во согласност со современата практика. Интегративната формула обезбедува фино рангирање и читливи врски меѓу фактори, моделите од машинско учење детектираат нелинеарности и меѓузависности што често се занемаруваат во поедноставени форми, а Баесовите мрежи додаваат структурна слика и толкување на веројатност што е корисно кога одлуките се на три нивоа. Овој спој на транспарентност, процена и калибрација, се смета за добра пракса и ги прави резултатите употребливи и за анализа и за избор на мерки.

Постапките за валидација и споредба се усогласени со светските насоки. Користењето поделба на податоци без протекување на информации, вкрстена валидација и споредба на метрики на тест, заедно со процена на важности по влезови со SHAP или пермутации, претставува стандард во апликации од овој тип. Проверка со криви на покриеност и добивка е вообичаен начин да се процени дали врвот на листата навистина концентрира поголем дел од ризичните делници од случајниот избор, а тоа директно го поддржува скринингот и планирањето на стручен тим. Просторниот приказ во ГИС и врзувањето со елементи контролирани на терен ја позиционира работата рамноправно со студии што инсистираат на интеграција меѓу аналитика и теренска верификација.

Категоризацијата на ризик во три нивоа е во согласност со практиките за брза комуникација и оперативни одлуки. Употребата на јасни граници, заедно со континуирано рангирање за фин избор во рамките на приоритетните делници е честа комбинација во

мрежни процени, каде ресурсите се ограничени, а потребна е споредливост меѓу техники и периоди.

Карактеристично е спојувањето на кластер-профили со моделските резултати и употребата на стабилни правила за избор на техники, така што рангирањата се поклопуваат и кога се менува методот. Ова е во согласност со светските препораки за потврда од повеќе извори и дава поголема доверба при селекција на делници со највисок приоритет. Дополнително, акцентот на теренска проверка и локален контекст е во согласност со досегашната пракса, дека добрите показатели од анализата треба да се потврдат на лице место и да се препорачаат во мерки што одговараат на конкретните траси и услови во Македонија.

9.4 Стабилност на резултатите и ограничувања

Проверките на стабилност покажуваат дека, изборот на делници со највисок приоритет останува сличен и кога се менуваат подесувањата, границите за толкување и избраните техники на моделирање. Најголемите разлики се појавуваат кај делници со блиски вредности (гранични позиции). Во горниот дел од листата рангирањата од интегративната формула, од избраниот модел на машинско учење и од Баесовите мрежи во голема мерка се поклопуваат. Кривите на покриеност и добивка потврдуваат дека врвот концентрира поголем дел од реално ризичните делници отколку случајниот избор.

Стабилноста се обезбедува така што, податоците се подготвуваат на ист начин, без мешање меѓу тренинг и тест, со фиксна поделба и со нормализација научена само на тренинг, а резултатите се споредуваат со исти тест-метрики. Дополнително, приоритетите по влезови, се сметаат на ист начин и се проверуваат со исклучување на поединечни модели. Во такви услови, малите технички промени не го менуваат изборот на делници што први влегуваат во приоритетната група. Методот, соопшто може да се пренесува на државни патишта со слично движење и одржување, а за урбани делници или средини со поинаква клима и работа потребна е кратка надворешна проверка и мало прилагодување.

Постојат ограничувања што треба да се имаат предвид. Податоците за незгоди може да имаат неусогласености, а соседните делници често се однесуваат слично затоа што имаат исти услови во простор. Тежинскиот индекс на незгоди W_i собира повеќе аспекти во една мерка, па некои локални разлики може да останат во позадина. Кога W_i се дели на неколку нивоа за полесна примена, се губи дел од деталноста во рамките на групите. Важни се и достапноста и квалитетот на податоците за сигнализација, одводнување и објектите на патот, особено ако недостигаат нови мерења.

Практично, резултатите служат како појдовна основа, а не како замена за увид на лице место. За поединечни локации потребна е теренска проверка од стручен тим, со дополнителна процена и споредба со локалните услови. За делници блиску до границите, корисна е кратка проверка на чувствителност, разгледување на алтернативни поставки и потврда на состојбата на терен. На тој начин се одржува рамнотежа меѓу аналитиката и инженерскиот пристап, а рангирањето по W_i се претвора во навремена и соодветна мерка за анализираната делница.

10. НАУЧЕН ПРИДОНЕС И ПРИМЕНА НА РЕЗУЛТАТИТЕ ОД ИСТРАЖУВАЊЕТО

Ова истражување развива методологија што ги поврзува податоците за сообраќај, геометрија, состојба на коловоз и услови од околина со исходите од незгоди, со цел да се препознаат делници со зголемен ризик и да се поддржи избор на приоритети. Пристапот комбинира статистички модели, техники од машинско учење и просторна обработка во ГИС, така што сложените врски се претвораат во читливо рангирање и јасни насоки за теренска проверка и мерки.

10.1 Научен придонес

Во продолжение се прикажани придонесите на трудот во научната обработка на безбедноста на патиштата, со фокус на постапките, правилата и алатките што овозможуваат практична примена и проверени резултати:

- Методологијата за проверка и приоритизација на делници претставува целосен пристап што ги опфаќа сите фази од обработката на податоците до донесување заклучоци. Со овој методолошки модел се опишува постапка која започнува со подготовка и избор на соодветни податоци, продолжува со моделирање и анализа, а завршува со изработка на приоритетна листа на делници, според нивото на ризик. Притоа, се врши проверка на точноста, стабилноста и валидноста на резултатите, без базирање само на еден модел, со што се обезбедува поголема сигурност и објективност во изборот на приоритети.
- Примената на SHAP и пермутациона важност се користи за мерење на влијателноста. Приоритетите се пресметуваат поединечно за секој модел, се доведуваат на иста скала, а потоа се усреднуваат со тежини според R^2 на тест-сетот. На тој начин се добива стабилен и споредлив ранг на фактори, корисен и за избор на влезови и за објаснивост на резултатите.
- Интегративната формула се користи како референтен модел со мал број влезови и јасно дефинирани интеракции, научен на тренинг и проверен на тест. Овој пристап помага да се постават претпоставки и полесно да се види што се случува кога се додава или се отстранува одреден параметар.
- Краткиот протокол за избор и проверка на модели од машинско учење се води според R^2 на тест-сет и вкрстена валидација, со дополнителни мерки за покриеност, добивка и стапка на препознавање на ризичните случаи (Top-k capture), како и проверки на стабилност на резултатите. Постапката е пренослива и може да се примени на нови податоци без измени.
- Профилирањето на делниците со кластер-анализа создава јасни групи според геометрија, состојба на коловоз и сообраќајно оптоварување. За секоја делница се доделуваат кластер-етикети и веројатности што може да се користат како објаснувачки променливи во моделите и како основа за насочени мерки по типови делници.
- Категоризацијата на ризикот го претвора континуираниот индекс W_i во три нивоа со граници 0-5, 5-10 и 10-30, што овозможува брзо читање, полесна споредба меѓу техники и директно поврзување со оперативни избори.

- Интеграцијата на ГИС со аналитиката ги поврзува просторните прикази со резултатите од моделите, овозможувајќи да се видат групирањата на ризик и да се лоцираат точки за увид на лице место.
- Правилата за подготовка на податоците и проверките на чувствителност опфаќаат стандардизација и групирање на сродни показатели за да се избегне преклопување и двојно влијание на истите фактори. Во рамките на проверките се менуваат границите и се исклучуваат поединечни модели за да се испита стабилноста на резултатите и да се потврди доследноста на избраните параметри.
- Основите за изработка на упатства и стандарди се темелат на јасно дефинирани чекори на одлучување, кои започнуваат со проверка и завршуваат со увид на лице место. Секој чекор поврзан е со конкретни индикатори и граници, со што се создава стабилна основа за усогласување на минимален сет мерки и процедури што може да се применуваат на национално ниво. Ваквиот пристап овозможува стандардизирана и транспарентна постапка при оценување и подобрување на безбедноста на патиштата.

Овие придонеси заедно обезбедуваат јасен и повторлив начин на работа што ги поврзува аналитичките резултати со практичниот избор на приоритети и со проверка на терен.

10.2 Практична примена

Практичната примена на алатките и резултатите се реализира преку поврзани цели и активности, насочени кон идната теренска и планска работа, во рамки на редовни постапки за одржување и унапредување на безбедноста.

- **Планирање и проектирање на патишта.** Со пресметаниот тежински индекс на незгоди (W_i), профилите на делници и категориите на ризик се предвидуваат слабости уште во фаза на проектирање. Се проверува усогласеноста на ограничувањата на брзината со геометријата, се проценуваат потребите за одводнување и сигнализација, како и влијанието на приклучувањата и наклоните. На терен, категориите на ризик се препознаваат преку контролни листи и ГИС-прикази, за да се обезбеди усогласена проценка по делници.
- **Инспекција и ревизија на постојната инфраструктура.** Рангирањето по W_i служи како листа за увид на лице место и за редовна контрола. Проверките се насочуваат кон делници со повисок приоритет, се бележат мерки за коловоз, прегледност, сигнализација и управување со приклучувањата, а ефектот се следи со повторна пресметка по интервенции и споредба по циклуси.
- **Одржување и мерки на терен.** За делници со оштетен коловоз или со задржување вода, приоритет се поправки и подобро одводнување. На места со ограничена прегледност и тесни кривини, акцентот се става на дополнителна сигнализација и прилагодување на режимот на брзина. Делниците со густы приклучувања бараат управување со пристапите и организирање безбедни влезно-излезни движења.
- **Поддршка на институциите за сообраќајна безбедност.** Одлуките за приоритети се темелат на единствен индекс и јасно дефинирани категории на ризик, а мапираните делници овозможуваат поефикасна подготовка на програми и распределба на буџет. На овој начин се поедноставува координацијата меѓу служби

и региони, се поставуваат контролни табли со клучни показатели што се обновуваат периодично и се подготвуваат кратки образложенија за избор на мерки.

- **Превенција и едукација на јавноста.** Категориите на ризик и просторните прикази носат јасни пораки за внимателно возење на делници со ограничена прегледност, надолжни наклони и присуство на вода на коловоз. Материјалите се користат за информирање на возачите, за кампањи и за внатрешни обуки на стручниот тим.

Како целина, овој пристап создава едноставен циклус што ги поврзува сите клучни чекори: процена, избор на делници за увид, спроведување мерки и следење на промените низ време. Само на таков начин, анализата се претвора во практична алатка за постепено намалување на ризикот.

11. ИДНИ ИСТРАЖУВАЊА

Овие насоки за идни истражувања се базираат на направените анализи и алатки и имаат цел подобро и појасно да се разбере ризикот по делница, да се подобри начинот на процена, за да биде едноставен и контролиран, како и да се зајакне врската меѓу аналитиката, теренскиот увид и планирањето во заедничкиот процес, од податок до мерка.

- **Просторно-временска динамика.** Истражувањата во оваа област треба да ги опфатат сезонските услови и временските појави што влијаат на ризикот од сообраќајни незгоди во различни региони и години. Потребно е да се моделира поврзаноста меѓу врнежите, снегот и температурата со тежинскиот индекс на незгоди (W_i), како и со фактори како надолжниот наклон на патот, прегледноста и ограничувањата на брзината. Треба да се изработи просторна и временска анализа со карти по месеци и да се споредат зимските и летните периоди на различни надморски височини. Ваквиот пристап би овозможил идентификација на периодите со зголемен ризик и изработка на препораки за превентивни мерки и прилагодување на одржувањето, според временските услови.
- **Напредни модели за предвидување.** Моделите за предвидување треба да користат архитектури на длабоко учење и други современи техники, како надградба на постојните методи засновани на интегративната формула, ExtraTrees и Баесовите мрежи. Потребно е да се изврши калибрација на излезните вредности и да се испита нивната разбирливост со пристапи, слични на SHAP-анализата. Исто така, треба да се направи споредба со истиот сет на метрики за да се види, дали новите модели подобро ги препознаваат посложените врски и меѓусебните влијанија меѓу факторите што го одредуваат ризикот.
- **Следење во реално време.** Идниот пристап треба да овозможи систем за следење на состојбите во реално време преку поврзување на моделите со ГИС-платформи и со оперативни извори на податоци како временски станици, сензори за проток и извештаи од терен. Потребно е да се постават јасни правила за предупредување при појава на ризични комбинации како силен дожд, зголемен проток и познато тесно грло на патната мрежа. Ваквиот систем би овозможил навремено известување на стручен тим и преземање привремени мерки за зголемување на безбедноста.
- **Сценарија и симулации.** Симулациите и сценаријата треба да послужат за процена на ефектите од различни мерки пред нивната примена. Потребно е да се испита како промените во намалување на ограничувањето на брзината, поставување знаци за предупредување, подобрување на одводнувањето или измена на режимот на приклучувања, влијаат врз ризикот од сообраќајни незгоди. Со ваков пристап може да се пресмета очекуваното намалување на ризикот и да се избере мерка со најповолен однос меѓу корист и трошоци.
- **Ефект на мерки пред и потоа.** Анализите на ефектот од спроведените мерки треба да се базираат на споредба на состојбите пред и по нивната примена. Потребно е да се спроведуваат систематски анализи за делници каде се реализирани нови интервенции, со однапред утврден протокол за избор на контролни делови, времески периоди и метрики. Треба да се следи промената на тежинскиот индекс

на незгоди (Wi) и на категоријата на ризик, со изработка на извештаи што ќе покажат кои типови мерки даваат најбрз и најстабилен ефект.

- **Пократки делници.** Анализите во иднина треба да преминат кон пократки делници за да се опфатат и малите, но значајни промени на теренот. Таквиот пристап овозможува подобро забележување на места каде што геометријата, прегледноста, одводнувањето и приклучувањата се менуваат на кратки растојанија. За таа цел, треба да се користат ГПС-податоци, видеоснимања и локални мерења со кои би се откривале мали, но критични точки што можат да имаат значајно влијание врз ризикот од сообраќајни незгоди.
- **Регионални и локални патишта.** Примената на методологијата треба да се прошири и на регионални и локални патишта, каде условите и карактеристиките често се разликуваат од оние на главната мрежа. Потребно е да се провери квалитетот на достапните податоци, да се прилагодат границите за оценување и да се процени дали истиот сет на параметри може да го објасни дел од ризикот или пак се потребни дополнителни влезни показатели специфични за овие патишта. На тој начин ќе се утврди применливоста и флексибилноста на пристапот во различни категории на патна инфраструктура.
- **Градски сообраќајници.** Методологијата треба да се примени и во градски услови, каде сообраќајот е посложен и зависи од голем број дополнителни фактори. Во урбаните коридори потребно е да се вклучат елементи како раскрсници со семафори, пешачки премини, автобуски стојалишта, паркирање на коловоз и реални брзини на движење. Ваквиот пристап би овозможил создавање индикатори што го опишуваат степенот на конфликтност во градската средина и би помогнал во препознавање на критични точки, каде се преклопуваат различни типови на сообраќај.
- **Преносливост и надворешна валидација.** Потребно е да се изврши проверка на преносливоста и надворешната валидација на методите во различни услови и средини. Треба да се испита нивното однесување во области со различна клима и режими на движење, за да се утврди дали резултатите се стабилни и надвор од првичниот опфат. Кога распределбата на тежинскиот индекс на незгоди (Wi) се разликува, неопходно е да се прилагодат границите и да се изврши повторно тренирање на моделите. Сите промени и услови под кои е постигната преносливоста треба јасно да се документираат, за да се обезбеди контрола и споредливост на резултатите.
- **Попрецизна категоризација на ризик.** Категоризацијата на ризикот треба да стане попрецизна со премин од три на пет нивоа, за да се овозможи попрецизно планирање и подобра споредливост со системот iRAP, кој користи пет категории. За оваа цел, потребно е да се соберат повеќе податоци и да се утврдат јасни граници меѓу нивото на ризик. Ваквиот пристап бара и проверка дали новата поделба е доволно разбирлива и практична за користење во реални услови при процена и планирање на мерки за безбедност.
- **Квалитет и збогатување на податоци.** Податоците што се користат за анализите треба да се унапредат и збогатат со подобар квалитет и поголема точност. Потребно е да се подобри геолокацијата на сообраќајните незгоди и да се усогласат записите што доаѓаат од различни извори. Треба да се воспостави поврзување меѓу

базите за одржување, временски услови и контроли на брзина, со што би се добила покомплетна слика за условите на патот. Дополнително, треба да се вклучат нови влезни податоци како реална брзина добиена од сензори, точни радиуси на хоризонтални кривини, реална прегледност и мерења за присуство на вода на коловозот, со цел да се зголеми точноста и практичната вредност на моделите.

- **Поддршка на одлучување.** Потребно е да се развие практичен систем за поддршка на одлучувањето што ќе го поврзува рангирањето на делниците според ризик, проценетата цена на мерките и очекуваниот ефект од нивната примена. Таквиот пристап треба да овозможи, прикажување на листа со делници каде за секоја ќе бидат предложени соодветни мерки и пресметано очекувано намалување на ризикот за соодветен буџет и временска рамка. Со тоа, процесот на избор ќе стане транспарентен и контролиран, овозможувајќи приоритетите да се одредуваат врз основа на објективни податоци и мерливи критериуми.

Овие насоки го прошируваат методолошкиот и применливиот опсег на истражувањето, овозможуваат поефикасно планирање на интервенции и придонесуваат кон континуирано намалување на ризикот на најважните делници во државната патна мрежа.

12. ЗАКЛУЧОЦИ

Заклучоците ги обединуваат најважните резултати од целата анализа и објаснуваат како податоците, моделите и ГИС помагаат да се изберат приоритетите на едноставен и контролиран начин. Прво се подготвуваат податоците, потоа се применуваат моделите и се прави поделба на нивоа, а на крај резултатите се ставаат на карта за полесно читање и користење. Со ваков пристап, резултатите директно се користат и во планирањето и на терен. На овој начин се воспоставува врска меѓу анализата, увидот на терен и изборот на мерки, па затоа статистиката и инженерската практика делуваат усогласено за подобра безбедност.

Моменталната состојба на сообраќајните незгоди во Македонија покажува дека е потребна систематска процена по поединечни делници. Незгодите не се распределени рамномерно, туку се концентрираат на одредени места каде последиците често се потешки. На ризикот силно влијаат локалните услови, како геометријата на патот, прегледноста, одводнувањето и начинот на движење. Податоците покажуваат дека Македонија има повисоки вредности од европскиот просек. По население има приближно 70 жртви на милион жители, додека во ЕУ се околу 46. По должина на мрежата има околу 9 жртви на 1 000 километри патишта, во ЕУ се околу 5. По број на возила разликата е најголема, кај нас околу 197 жртви на милион регистрирани возила, во ЕУ околу 52. Затоа, важен е пристап што истовремено ги разгледува и целата мрежа и критичните локации, за резултатите брзо да се претворат во приоритети за увид и конкретни мерки на терен.

Претходните истражувања во регионот и пошироко најчесто ги разгледуваат факторите што влијаат на сообраќајните незгоди одделно, па најчесто се истакнуваат обемот на сообраќај, ограничувањата на брзина, геометријата на патот, прегледноста и состојбата на коловозот. Во поглед на методите, најчесто се користат регресији, дрвовидни модели и Баесови пристапи, кои често се поврзуваат со ГИС за просторна анализа. Овој труд ја следи таа насока, но оди чекор понатаму така што ги обединува сите техники во заеднички процес со исти метрики и јасни правила за споредба. На овој начин резултатите се претставени преку усогласени рангирања и разбирливи ГИС-прикази. Наместо одвоени анализи, факторите се разгледуваат заедно во постапка што овозможува проверка на точноста, објаснување на влијанието на секој параметар и просторен увид. Со тоа се добива практична поддршка за одлучување, која е истовремено споредлива и лесна за користење.

Податоците што се користат се од различни извори и со различен квалитет. Записите за незгоди се земени од МВР, податоците за геометријата на трасата и состојбата на коловозот од ЈПДП, а климатските серии од УХМР. Локалните карактеристики како прегледност, сигнализација и напречни наклони се снимени со GPS-видео и мобилно мапирање. Сите овие извори се спојуваат, така што секоја делница добива една табела со свои параметри, а дополнително се додаваат временските услови и надморската височина за поширока слика. Во анализата се користат 23 влезни параметри и една целна променлива, тежинскиот индекс на незгоди (Wi). Бидејќи има празнини и разлики во квалитетот, прво се прави уредување: чистење и усогласување на записите, проверка на локациите, ставање на сите мерки на иста скала и поделба на податоците на тренинг и тест без мешање помеѓу нив. За да се види колку резултатите се стабилни, се прават проверки на чувствителност и споредби меѓу различни модели. Целта е да се намали пристрасноста и да се добијат резултати што се споредливи и веродостојни.

Експлоративната кластер-анализа издвојува шест препознатливи профили на делници, добиени со Гаусов мешан модел (GMM, Gaussian Mixture Model) со шест кластери и избор воден од ВИС. Профилите се разликуваат според спој на својства: геометријата на трасата како радиуси и наклони и густина на приклучувања, прегледноста и сигнализацијата, сообраќајното оптоварување и состојбата на коловозот. Според тежинскиот индекс на незгоди W_i , постојат јасни разлики меѓу групите, а тие разлики се повторуваат и кога се применуваат различни техники и поставки, што укажува на стабилна и смислена поделба. Ваквиот распоред дава прегледна слика за типовите делници и нивниот ризик, го олеснува толкувањето на резултатите во контекст, и дозволува објективна споредба внатре во група и меѓу групи, без да се сведува оценката само на поединечни случаи. Со тоа се добива практична основа за насочување мерки таму каде што ризикот е најзабележлив.

Влијанието на параметрите оценето е со SHAP како прв избор, бидејќи овозможува јасен придонес на секој влез. Каде што не е можно да се примени SHAP, се користи пермутациона важност како алтернативен пристап. Резултатите од двете мерења се усогласуваат и како највлијателни се дефинирани ПГДС (обем на сообраќај), LIMIT (ограничување на брзина) и K.Int.P.T (густина на приклучувања по тип), додека Ave_Rain, IRI и RDB_RUT имаат најмало влијание. Ова усогласување дава сигурен редослед на факторите што може директно да се користи за избор на клучни влезови, поедноставување на модели и планирање на мерки, при што параметрите со мало влијание може да се изостават без значајна загуба на точност.

Интеграативната формула ја заокружува анализата како јасна и контролирана постапка за процена на ризикот: влезовите најнапред се стандардизираат, се гради полиномна основа од втор степен со квадрати и парни интеракции, а изборот на променливи тргнува од трите највлијателни (PGDS, LIMIT, K.Int.P.T) и се надградува со вклучување на секој следен параметар што ја подобрува точноста на независниот тест. При 80/20 поделба, најдобар резултат дава OLS полином со седум влезови и $\text{Test } R^2 = 0.72$, при што Ridge и Lasso служат како споредливи техники за стабилност. Формулата обезбедува континуирано рангирање и разбирливо толкување на интеракциите, а сигурноста дополнително е потврдена со вкрстена валидација со 5 преклопи во класификациона рамка со три нивоа на ризик. Со тоа се добива компактна и прецизна референтна основа за споредба со моделите од машинско учење и Баесовите мрежи во заклучните оценки.

Сеопфатната проверка и вкрстено тестирање на девет модели од машинско учење потврдува дека пристапот е точен и стабилен за практична селекција на делници со зголемен ризик. Како најефективни три методи се истакнуваат Gradient Boosting со $R^2 = 0.67$, Extra Trees со $R^2 = 0.63$ и AdaBoost со $R^2 = 0.61$ на тест-сет. По дополнителна вкрстена валидација со 5 преклопи и исти метрики, Extra Trees избран е како носечки модел поради помала варијабилност меѓу преклопи и усогласени врвови во рангирањата. Способноста за разликување потврдена е со AUROC = 0.82, PR-AUC = 0.59, Top-20% capture = 0.55, Top-30% = 0.69 и F1 = 0.53 при граница со околу 20% избрани делници, а редоследот на важните влезови се совпаѓа со претходните резултати, што ја олеснува интерпретацијата и планирањето на мерки. По сите тестирања, Extra Trees прикажува најголема стабилност и се предлага како носечки модел, кој во иднина ќе се користи за детерминација на безбедноста по делници. Со ова се обезбедува сигурна и применлива основа за насочување на увид на лице место и планирање интервенции.

Во Баесовите мрежи се анализираат три структури: GTT, PC и TAN, при што дискретизацијата конкретно е подесена во три категории, а распределбите се израмнети за да се зголеми точноста без губење на читливост. По перформанси, GTT постигнува вкупна точност 0.83 и AUROC 0.98, PC има вкупна точност 0.85 и AUROC 0.97, а TAN вкупна точност 0.57 и AUROC 0.80, но со најдобра калибрација на веројатностите. Практично, GTT и PC се најсоодветни кога е важно сигурно издвојување на делниците со повисок ризик, додека TAN корисен е кога се бараат добро усогласени веројатности за јасна комуникација и планирање на мерки.

Категоризацијата на Wi во три нивоа со граници 0-5, 5-10 и $\geq 10-30$ го прави преносот на информации поедноставен, а споредбата меѓу пристапите појасна. Резултатите се усогласени кога се користат сите три модели за детерминација, па насоките за проверка на терен се јасни и применливи за стручен тим и носители на одлуки. Просторните прикази во ГИС ги истакнуваат зоните со зголемен ризик и упатуваат на конкретни точки за увид и мерки. Со овие оперативни граници се создава стабилна основа и за поставување приоритети и за споредба меѓу различни модели.

Поврзувањето со ГИС им дава на анализите јасен просторен контекст и на карта покажува каде точно се проблемите. Рангирањата и категориите се прикажуваат по делници, теренските проверки се организираат по координати, а на истите точки се означуваат забелешки за прегледност, состојба на коловоз, одводнување и приклучувања. Во тој случај брзо се препознаваат критичните сегменти и нивните групирања, полесно се избираат делници за увид и се планираат мерки по региони со проценка на трошок и очекуван ефект.

Главните делници што излегуваат како приоритет за подобрување остануваат речиси исти, независно дали се применува интегративната формула, Extra Trees или Баесови мрежи, и кога се менуваат влезните параметри и границите. Мали разлики се појавуваат само кај делници со слични параметарски карактеристики и многу блиски вредности на Wi. Податоците за незгоди понекогаш се нецелосни или неусогласени, а соседните делници може меѓусебно да влијаат. Дополнително, Wi претставува единечен индекс што спојува повеќе различни аспекти на ризик. Поради тоа, резултатите треба да се користат како насока при избор, додека конечната одлука треба да се потврди со увид на лице место и стручно-инженерска проценка, особено кога делницата е блиску до границите меѓу категориите.

Надворешната проверка на избрани регионални делници покажува усогласување меѓу проценетиот ризик и затекнатите услови. По формулата, Extra Trees и Баесовите мрежи од тип GTT/PC, категориите се целосно усогласени за сите три делници. Разликите во континуалниот Wi се мали и не ги поминуваат границите, па затоа класификацијата останува иста. Наведеното потврдува дека пристапите усогласено го детерминираат индексот во соодветна категорија и се применливи за приоритизација и планирање на проверки на лице место.

Во целина, трудот спојува ГИС, статистички анализи и повеќе пристапни модели во единствен процес што води од податоци до проверка и приоритизација на делници. Резултатите обезбедуваат стабилно рангирање, доверлив избор на приоритети и јасен пат до мерки на терен. Со ова се гради солидна основа за систематско намалување на ризикот на најважните делници во Македонија и за подобрување на податоците, методите и начинот на носење одлуки.

13. СПИСОК НА ЛИТЕРАТУРА

- [1] World Health Organization. (2024). Road traffic injuries. <https://www.who.int/news-room/factsheets/detail/road-traffic-injuries>
- [2] European Commission, European Road Safety Observatory. (2023). Road Safety Statistics 2023. https://road-safety.transport.ec.europa.eu/european-road-safety-observatory_en
- [3] Државен завод за статистика. (2025). Макстат – Статистичка база на податоци. <https://makstat.stat.gov.mk/PXWeb/pxweb/mk/>
- [4] European Commission, DG MOVE (ERSO). (2023). Key performance indicators (KPIs). https://road-safety.transport.ec.europa.eu/european-road-safety-observatory/data-and-analysis/key-performance-indicators-kpis_en
- [5] OECD/International Transport Forum. (2023). Road Safety Annual Report 2023. <https://www.itf-oecd.org/road-safety-annual-report-2023>
- [6] Државен завод за ревизија. (2025). Безбедност во сообраќајот на патиштата во РСМ – Конечен извештај. https://dzt.mk/sites/default/files/2025-05/62_RU_Bezbednost_vo_soobrakajot_na_patistata_vo_Makedonija_2024.pdf
- [7] World Health Organization. (2023). Global status report on road safety 2023. <https://www.who.int/publications/i/item/9789240086517>
- [8] International Transport Forum, OECD. (2024). Road Safety Annual Report 2024. <https://www.itf-oecd.org/sites/default/files/docs/irtad-road-safety-annual-report-2024.pdf>
- [9] European Court of Auditors. (2024). Special report 04/2024: Reaching EU road safety objectives. <https://www.eca.europa.eu/en/publications?ref=sr-2024-04>
- [10] Belin M A, Tillgren P, Vedung E. (2012). Vision Zero – a road safety policy innovation. *International Journal of Injury Control and Safety Promotion*, 19(2), 171-179. <https://doi.org/10.1080/17457300.2011.635213>
- [11] Ji A, Levinson D. (2020). An energy loss-based vehicular injury severity model. *Accident Analysis & Prevention*, 146, 105730. <https://doi.org/10.1016/j.aap.2020.105730>
- [12] Van den Broek B, Aarts L, Silverans P. (2023). Baseline report on the KPI Speeding. Vias institute. https://road-safety.transport.ec.europa.eu/system/files/2023-03/Baseline_KPI_Speeding.pdf
- [13] World Health Organization. (2004). World report on road traffic injury prevention. <https://www.who.int/publications/i/item/world-report-on-road-traffic-injury-prevention>
- [14] Runyan C W. (1998). Using the Haddon Matrix: introducing the third dimension. *Injury Prevention*, 4(4), 302-307. <https://injuryprevention.bmj.com/content/4/4/302>
- [15] Sleet D A., Dahlberg L L., Basavaraju S V., Mercy J A., McGuire L C., Greenspan A. (2011). Injury Prevention, Violence Prevention, and Trauma Care: Building the Scientific Base. *MMWR Supplements*, 60(04), 78-85. <https://www.cdc.gov/mmwr/preview/mmwrhtml/su6004a13.htm>
- [16] Postransky, T. & Vovk, Y. (2020). Changing of the bus driver's functional state in city conditions. *Transport Technologies*, 1(1), 12-21. <https://doi.org/10.23939/tt2020.01.012>
- [17] International Transport Forum, OECD. (2018). Speed and Crash Risk. <https://www.itf-oecd.org/speed-crash-risk>
- [18] PIARC (World Road Association). (2025). Road Safety Manual. <https://roadsafety.piarc.org/en>
- [19] European Transport Safety Council. (ETSC). (2024). 18th Road Safety Performance Index (PIN) Report. <https://etsc.eu/pin18>

- [20] WHO & UN Road Safety Collaboration. (2018–2019). Voluntary Global Performance Targets for Road Safety Risk Factors and Service Delivery Mechanisms and corresponding indicators. <https://cdn.who.int/media/docs/default-source/documents/un-road-safety-collaboration/targets-and-indicators-visual-clean.pdf>
- [21] Elvik R. (2009). The Power Model of the relationship between speed and road safety. TØI Report 1034/2009. <https://www.toi.no/getfile.php?mmfileid=13206>
- [22] Eurostat, ITF, UNECE. (2019). Glossary for Transport Statistics (5th ed.). [https://unece.org/sites/default/files/2021-12/Glossary for Transport Statistics EN FINAL WEB2 1.pdf](https://unece.org/sites/default/files/2021-12/Glossary%20for%20Transport%20Statistics%20EN%20FINAL%20WEB2%201.pdf)
- [23] European Union. (2019). Directive (EU) 2019/1936 of the European Parliament and of the Council of 23 October 2019 amending Directive 2008/96/EC on road infrastructure safety management. <https://eur-lex.europa.eu/eli/dir/2019/1936/oj>
- [24] AASHTO. (2023). Highway Safety Manual (2nd ed.). <https://store.transportation.org/Item/PublicationDetail?ID=4719>
- [25] iRAP. (2020–ongoing). iRAP Methodology and Star Ratings (incl. SR4D). <https://irap.org/methodology/>
- [26] WHO, GRSP; IFRC. (2023). Speed management: a road safety manual (2nd ed.). <https://www.who.int/publications/m/item/speed-management--a-road-safety-manual-for-decision-makers-and-practitioners.-2nd-edition>
- [27] World Bank, GRSF. (2024). Guide for Safe Speeds. <https://www.globalroadsafetyfacility.org/sites/default/files/2024-05/Guide%20for%20Safe%20Speeds%20-%20Managing%20Traffic%20Speeds%20to%20Save%20Lives.pdf>
- [28] ETSC. (2022). PIN Flash 42: How traffic law enforcement can contribute to safer roads. [https://etsc.eu/wp-content/uploads/ETSC PINFLASH42 v2TH JC FINAL corrected-060522.pdf](https://etsc.eu/wp-content/uploads/ETSC_PINFLASH42_v2TH_JC_FINAL_corrected-060522.pdf)
- [29] Dingus T. A., et al. (2016). Driver crash risk factors and prevalence evaluation using naturalistic driving data. PNAS, 113(10), 2636–2641. <https://doi.org/10.1073/pnas.1513271113>
- [30] Theeuwes J. (2021). Self-explaining roads: What does visual cognition tell us about designing safer roads? Cognitive Research: Principles and Implications, 6, 29. <https://doi.org/10.1186/s41235-021-00281-6>
- [31] Klauer S G, et al. (2014). Distracted Driving and Risk of Road Crashes among Novice and Experienced Drivers. New England Journal of Medicine, 370(1), 54–59. <https://doi.org/10.1056/NEJMSa1204142>
- [32] Farmer C M. (2004). Effect of Electronic Stability Control on Automobile Crash Risk. Traffic Injury Prevention, 4(5), 317–325. <https://doi.org/10.1080/15389580490896951>
- [33] Cicchino J. B. (2017). Effectiveness of forward collision warning (FCW) and autonomous emergency braking (AEB) in reducing police-reported crash rates. Accident Analysis & Prevention, 99(A), 142–152. <https://doi.org/10.1016/j.aap.2016.11.009>
- [34] Fildes B., et al. (2015). Effectiveness of low speed AEB in real-world rear-end crashes. Accident Analysis & Prevention, 81, 24–29. <https://doi.org/10.1016/j.aap.2015.03.029>
- [35] Kahane C J. (2015). Lives Saved by Vehicle Safety Technologies and Associated Federal Motor Vehicle Safety Standards, 1960–2012. NHTSA DOT HS 812 069. <https://crashstats.nhtsa.dot.gov/Api/Public/ViewPublication/812069>
- [36] Andrey, J. (2010). Long-term trends in weather-related crash risks. Journal of Transport Geography, 18(2), 247–258. <https://doi.org/10.1016/j.jtrangeo.2009.05.002>

- [37] Dell'Acqua G., Russo F. (2011). Safety Performance Functions for Low-Volume Roads. *The Baltic Journal of Road and Bridge Engineering*, 6(4), 225–234. <https://doi.org/10.3846/bjrbe.2011.29>
- [38] Pathivada B K., Maddox B., Yamashita E. (2024). Impact of real-time weather conditions on crash injury severity in U.S. *Accident Analysis & Prevention*. <https://doi.org/10.1016/j.aap.2023.107453>
- [39] Edwards J B. (2010). Weather-related road accidents in England and Wales: a spatial analysis. *Journal of Transport Geography*, 3(4), 201-212. [https://doi.org/10.1016/0966-6923\(96\)00006-3](https://doi.org/10.1016/0966-6923(96)00006-3)
- [40] Amoros E., Martin J-L., Laumon B. (2006). Under-reporting of road crash casualties in France. *Accident Analysis & Prevention*, 38(4), 627-635. <https://doi.org/10.1016/j.aap.2005.11.006>
- [41] Soltani, S., et al. (2022). What is counted counts: An innovative linkage of police, hospital, and spatial data for transportation injury prevention. *Journal of Safety Research*, 82, 275–287. <https://doi.org/10.1016/j.jsr.2022.08.002>
- [42] Eurostat, International Transport Forum (ITF). (2019). Glossary for Transport Statistics. 5th ed. Luxembourg, ISBN 978-92-76-00206-2. [https://unece.org/sites/default/files/2021-12/Glossary for Transport Statistics EN FINAL WEB2 1.pdf](https://unece.org/sites/default/files/2021-12/Glossary%20for%20Transport%20Statistics%20EN%20FINAL%20WEB2%201.pdf)
- [43] Perez K., et al. (2019). Implications of estimating road traffic serious injuries from hospital data. *Accident Analysis & Prevention*, 130, 125–135. <https://doi.org/10.1016/j.aap.2018.04.005>
- [44] OECD/ITF. (2016). Zero Road Deaths and Serious Injuries. <http://dx.doi.org/10.1787/9789282108055-en>
- [45] ISO. (2025). ISO 39001 Startup Guide: Road traffic safety (RTS) management systems. https://www.iso.org/files/live/sites/isoorg/files/developing_standards/docs/en/ISO_39001_Startup_Guide_2025.pdf
- [46] Костадинова Љ. Влијание на елементите на патот врз бројот на сообраќајните незгоди. М-груд, Градежен факултет, УКИМ, Скопје.
- [47] Milenković, M., Glavić, D., Kočić, A., & Petković, M. (2017). Impact of Road and Traffic Characteristics to Traffic Accidents. *Put i saobraćaj*, 63(1), 5–12. <https://www.putisaobracaj.rs/index.php/PiS/article/view/54>
- [48] Othman, S., Thomson, R., & Lanner, G. (2009). Identifying Critical Road Geometry Parameters Affecting Crash Rate and Crash Type. *Annals of Advances in Automotive Medicine*, 53, 155–165. https://www.researchgate.net/publication/41577546_Identifying_Critical_Road_Geometry_Parameters_Affecting_Crash_Rate_and_Crash_Type#fullTextFileContent
- [49] Tsubota T., Fernando C., Yoshii T., Shirayanagi H. (2018). Effect of Road Pavement Types and Ages on Traffic Accident Risks. *Transportation Research Procedia*, 2018(34), 211–218. <https://doi.org/10.1016/j.trpro.2018.11.034>
- [50] Jima, A. B., & Sipos, T. (2022). The Impact of Road Geometric Formation on Traffic Accidents. *Sustainability*, 14, 8475. <https://doi.org/10.3390/su14148475>
- [51] Dumitrascu, D. I. (2024). Influence of Road Infrastructure Design on Traffic Safety. *Infrastructures*, 9, 154. <https://doi.org/10.3390/infrastructures9090154>
- [52] Charlton, S. G. (2007). The role of attention in horizontal curves: A comparison of advance warning, delineation, and road marking treatments. *Accident Analysis & Prevention*, 39(5), 873–885. <https://doi.org/10.1016/j.aap.2006.12.007>
- [53] Јолакоски, Стевчо. (2014). Значењето на средствата за смирување на сообраќајот во насока на подобрување на безбедноста на секундарната патна мрежа. Докторска дисертација, УКЛО – Битола. <https://eprints.uklo.edu.mk/id/eprint/9013/1/Stevco%20Jolakoski%20DT.pdf>

- [54] Gataric D., Ruskic N., Aleksic B., Djuric T., Pezo L., Loncar B., Pezo M. (2023). Predicting Road Traffic Accidents–Artificial Neural Network Approach. *Algorithms*, 16(5):257. <https://doi.org/10.3390/a16050257>
- [55] Aarts L., van Schagen I. The relation between speed and crashes. SWOV Fact sheet. <https://swov.nl/publication/relation-between-speed-and-crashes>
- [56] Retallack A E., Ostendorf B. (2020). Relationship Between Traffic Volume and Accident Frequency at Intersections. *IJERPH*, 17(4), 1393. <https://doi.org/10.3390/ijerph17041393>
- [57] Sosolovski L., Jovanovski Z., Cikarski Z. (2020). Analysis of traffic accidents in North Macedonia 2010–2019. *Journal of Criminology and Criminal Law*, 58(3), 97–110.
- [58] Deretić, N., Stanimirović, D., Awadh, M. A., Vujanović, N., & Djukić, A. (2022). SARIMA Modelling Approach for Forecasting of Traffic Accidents. *Sustainability*, 14(8), 4403. <https://doi.org/10.3390/su14084403>
- [59] Qiu L., Nixon W A. (2008). Effects of Adverse Weather on Traffic Crashes: Systematic Review and Meta-Analysis. *Journal of Transportation Engineering*, 134(10). <https://doi.org/10.3141/2055-16>
- [60] Theofilatos A., Yannis G. (2014). A review of the effect of road weather conditions on road safety. *Accident Analysis & Prevention*, 72, 244–256. <https://doi.org/10.1016/j.aap.2014.06.017>
- [61] Keay K, Simmonds I. Road accidents and rainfall in a large Australian city. *Accident Analysis & Prevention*, 2006; 38(3):445–454. <https://doi.org/10.1016/j.aap.2005.06.025>
- [62] Malin F., Norros I., Innamaa S. (2019). Accident risk of road and weather conditions on different road types. *Accident Analysis & Prevention*, 122, 181–188. <https://doi.org/10.1016/j.aap.2018.10.014>
- [63] Malin F. Norros I., Innamaa S. (2017). Identification of accident risk on the Finnish main road network in different weather conditions. *Finnish Transport Agency Reports 41/2017*. https://www.doria.fi/bitstream/handle/10024/143844/lts_2017-41_978-952-317-440-5.pdf?sequence=2
- [64] Babanoski, K., Ilijevski, I., & Dimovski, Z. (2016). Analysis of Road Traffic Safety through Direct Relative Indicators for Traffic Accidents Fatality: Case of Republic of Macedonia. *Promet – Traffic & Transportation*, 28(6), 661–669. <https://traffic.fpz.hr/index.php/PROMTT/article/view/2137>
- [65] Montella A. (2010). A comparative analysis of hotspot identification methods. *Accident Analysis & Prevention* 42 (2), 571-581 . <https://doi.org/10.1016/j.aap.2009.09.025>
- [66] Amorim B. S. P., et al. (2023). A Machine Learning Approach for Classifying Road Accident Hotspots. *ISPRS IJGI*, 12(6):227. <https://doi.org/10.3390/ijgi12060227>
- [67] iRAP. iRAP Methodology and Star Ratings (incl. SR4D). <https://irap.org/methodology/>
- [68] Caliendo C., Guida M., Parisi A. (2007). A crash-prediction model for multilane roads. *Accident Analysis & Prevention*, 39(4), 657–670. <https://doi.org/10.1016/j.aap.2006.10.012>
- [69] Yan X., et al. (2023). Relationship between Highway Geometric Characteristics and Accident Risk: MLP Approach. *Sustainability*, 15(3):1893. <https://doi.org/10.3390/su15031893>
- [70] Tsubota, T., Fernando, C., Yoshii, T., & Shirayanagi, H. (2018). Effect of Road Pavement Types and Ages on Traffic Accident Risks. *Transportation Research Procedia*, 34, 211–218. https://www.cee.ehime-u.ac.jp/~keikaku/ists18/pdf/ISTS_IWTDCS_2018_paper_67.pdf
- [71] Kim D., Yoon S., Kim B. (2021). Risk Prediction for Winter Road Accidents on Expressways. *Applied Sciences*, 11(20):9534. <https://doi.org/10.3390/app11209534>

- [72] Lord, D., & Mannering, F. (2010). The statistical analysis of crash-frequency data: A review and assessment of methodological alternatives. *Transportation Research Part A: Policy and Practice*, 44(5), 291–305. <https://doi.org/10.1016/j.tra.2010.02.001>
- [73] Miaou, S.-P. (1994). The relationship between truck accidents and geometric design of road sections: Poisson versus negative binomial regressions. *Accident Analysis & Prevention*, 26, 471–482. [https://doi.org/10.1016/0001-4575\(94\)90038-8](https://doi.org/10.1016/0001-4575(94)90038-8)
- [74] Zou, Y., Zhang, Y., & Lord, D. (2013). Application of finite mixture of negative binomial regression models with varying weight parameters for vehicle crash data analysis. *Accident Analysis & Prevention*, 50, 1042–1051. <https://doi.org/10.1016/j.aap.2012.08.004>
- [75] Dong, C., Clarke, D. B., Nambisan, S. S., Huang, B., & Yang, C. (2014). Multivariate random-parameters zero-inflated negative binomial regression model: An application to estimate crash frequencies at intersections. *Accident Analysis & Prevention*, 70, 320–329. <https://doi.org/10.1016/j.aap.2014.04.018>
- [76] Ma, J., Kockelman, K. M., & Damien, P. (2008). A multivariate Poisson–lognormal regression model for prediction of crash counts by severity, using Bayesian methods. *Accident Analysis & Prevention*, 40(3), 964–975. <https://doi.org/10.1016/j.aap.2007.11.002>
- [77] Zhang, S., Khattak, A., Matara, C. M., Hussain, A., & Farooq, A. (2022). Hybrid feature selection-based machine learning classification system for the prediction of injury severity in single and multiple-vehicle accidents. *PLOS ONE*, 17(2), e0262941. <https://doi.org/10.1371/journal.pone.0262941>
- [78] Yu R., Abdel-Aty M. (2013). Using SVM in real-time crash risk evaluation. *Accident Analysis & Prevention*, 51, 252–259. <https://doi.org/10.1016/j.aap.2012.11.027>
- [79] Li P., Abdel-Aty M., Yuan J. (2020). Real-time crash risk prediction on arterials based on LSTM-CNN. *Accident Analysis & Prevention*, 135, 105371. <https://doi.org/10.1016/j.aap.2019.105371>
- [80] Mannering, F., & Bhat, C. R. (2016). Analytic methods in accident research: Methodological frontier and future directions. *Analytic Methods in Accident Research* <https://doi.org/10.1016/j.amar.2013.09.001>
- [81] Ye Z., Zhang Y., Lord D. (2013). Goodness-of-fit testing for accident models with low means. *Accident Analysis & Prevention*, 61, 78–86. <https://doi.org/10.1016/j.aap.2012.11.007>
- [82] Bergmeir C., Benítez J. M. (2012). On the use of cross-validation for time series predictor evaluation. *Information Sciences*, 191, 192–213. <https://doi.org/10.1016/j.ins.2011.12.028>
- [83] Saito T., Rehmsmeier M. (2015). The Precision-Recall Plot Is More Informative than the ROC Plot When Evaluating Binary Classifiers on Imbalanced Datasets. *PLoS ONE* 10(3) <https://doi.org/10.1371/journal.pone.0118432>
- [84] Davis J., Goadrich M. (2006). The relationship between PR and ROC curves. *ICML 06*, 233–240. <https://doi.org/10.1145/1143844.1143874>
- [85] Cheng W., Washington S. (2005). Experimental evaluation of hotspot identification methods. *Accident Analysis & Prevention*, 37, 870–881. <https://doi.org/10.1016/j.aap.2005.04.015>
- [86] Elvik R. (2013). International transferability of accident modification functions for horizontal curves. *Accident Analysis & Prevention*, 59, 487–496. <https://doi.org/10.1016/j.aap.2013.07.010>
- [87] Hauer E., Harwood D. W., Council F. M., Griffith M. S. (2002). Estimating safety by the Empirical Bayes method: A tutorial. *Transportation Research Record*, 1784(1), 126–131. <https://conferences.ict.illinois.edu/spfsummit09/SPF%20Basic%20Tutorial%202001%20by%20Ezra%20Hauer.pdf>

- [88] Chen, S., Saeed, T. U., Alinizzi, M., Lavrenz, S., & Labi, S. (2019). Safety sensitivity to roadway characteristics: A comparison across highway classes. *Accident Analysis & Prevention*, 123, 39–50. <https://doi.org/10.1016/j.aap.2018.10.020>
- [89] Berhanu, Y., Schroder, D., Wodajo, B. T., & Alemayehu, E. (2024). Machine learning for predictions of road traffic accidents and spatial network analysis for safe routing on accident and congestion-prone road networks. *Results in Engineering*, 23, 102737. <https://doi.org/10.1016/j.rineng.2024.102737>
- [90] ЈП за државни патишта. „Веб-ГИС платформа за просторни анализи и визуализации“. <http://62.77.137.99/pesr/webgis/#/map>
- [91] Министерство за локална самоуправа. (2021). „Програма за развој на плански региони 2021–2026“.
- [92] Дончева Р., Огњеновиќ С. (2024). Проектирање патишта. ГФ, УКИМ. ISBN 978-608-4510-60-4.
- [93] Wright P H., Dixon K., Meyer M. (2004). *Highway Engineering* (7th ed.). Pearson.
- [94] Министерство за транспорт и врски. (2009). Правилник за техничките елементи за изградба и реконструкција на јавните патишта и на објектите на патот. Службен весник на Република Македонија бр. 110/2009 (3.9.2009). <https://dki.gov.mk/wp-content/uploads/2015/03/110-2009.pdf>
- [95] Огњеновиќ С. (2014). „Калибрирање на моделите на оштетување на коловозни конструкции“. PhD, ГФ УКИМ.
- [96] Dynatest International A/S. (2013). *Dynatest MFV RSP III/IV – Owner’s Manual* (Version 2.6.12). København: Dynatest.
- [97] Tobias P., de Leon Izeppi E., Flintsch G., Katicha S., McCarthy R. (2023). *Pavement Friction for Road Safety: Primer on Friction Measurement and Management Methods*. FHWA, FHWA-SA-23-007. <https://highways.dot.gov/sites/fhwa.dot.gov/files/2023-06/FHWA%20Pavement%20Friction%20for%20Road%20Safety%20Primer.pdf>
- [98] City & County Pavement Improvement Center. (2021). *Pavement Condition Index (PCI): Best Practices for Pavement*. UCPRC-BP-2021-01. <https://www.ucprc.ucdavis.edu/ccpic/pdf/PCI%204-Pager%20final%20v2.pdf>
- [99] ЈПДП. Traffic Data Processing System (TDPS) – Платформа за сообраќајни податоци (ПГДС). <https://tdps.roads.org.mk/Home/Index>
- [100] Ѓешовска В., Тасески Г., Илиоски Б. (2024). „Интензивни врнежи во РСМ“. УКИМ – Скопје.
- [101] Студија за просторен план – Природни и техничко-технолошки непогоди. (4.12.2023).
- [102] Government of the Republic of North Macedonia, MoT, PIU. (2024). *Handbook on Black Spot Management (BSM)*.
- [103] Wilcox R R. (2001). *Fundamentals of Modern Statistical Methods*. Springer. <https://doi.org/10.1007/978-1-4757-3522-2>
- [104] Hunter M., Shaik S., Ksaibati K. (2021). A Proactive Approach to Evaluating Intersection Safety Using Hard-Braking Data. *Journal of Infrastructure Systems*, 3, 81-94. <https://doi.org/10.1007/s42421-021-00039-y>
- [105] Li P., Chen S., Chen F., Xie K. (2024). Analysis of pedestrian accident severity by spatiotemporal effects. *Heliyon*, 10(4), e26438. <https://doi.org/10.1016/j.heliyon.2024.e26438>
- [106] Fraley, C., & Raftery, A. E. (2002). Model-based clustering, discriminant analysis, and density estimation. *Journal of the American Statistical Association*, 97(458), 611–631. <https://doi.org/10.1198/016214502760047131>

- [107] Arbelaitz O., Gurrutxaga I., Muguerza J., Perez J M., Perona I. (2013). Comparative study of cluster validity indices. *Pattern Recognition*, 46(1), 243–256. <https://doi.org/10.1016/j.patcog.2012.07.021>
- [108] Steinley D. (2004). Properties of the Hubert–Arabie adjusted Rand index. *Psychological Methods*, 9(3), 386–396. <https://doi.org/10.1037/1082-989X.9.3.386>
- [109] McKnight P E., Najab J. (2010). Kruskal–Wallis Test. *Corsini Encyclopedia of Psychology*. <https://doi.org/10.1002/9780470479216.corpsy0491>
- [110] Dinno, A. (2015). Nonparametric pairwise multiple comparisons in independent groups using Dunn’s test. *The Stata Journal*, 15(1), 292–300. <https://doi.org/10.1177/1536867X1501500117>
- [111] Tomczak M., Tomczak E. (2014). The need to report effect size estimates revisited. *Trends in Sport Sciences*, 21(1), 19–25. https://www.wbc.poznan.pl/Content/325867/5_Trends_Vol21_2014_%20no1_20.pdf
- [112] Chavent, M., Kuentz-Simonet, V., Liquet, B., & Saracco, J. (2012). ClustOfVar: An R package for the clustering of variables. *Journal of Statistical Software*, 50(13), 1–16.. <https://doi.org/10.18637/jss.v050.i13>
- [113] Roberts D. R., et al. (2017). Cross-validation strategies for data with temporal, spatial, hierarchical, or phylogenetic structure. *Ecography*, 40(8), 913–929. <https://doi.org/10.1111/ecog.02881>
- [114] Loh W-Y. (2011). Classification and regression trees. *WIREs: DMKD*, 1(1), 14–23. <https://doi.org/10.1002/widm.8>
- [115] Breiman L. (2001). Random Forests. *Machine Learning*, 45(1), 5–32. <https://doi.org/10.1023/A:1010933404324>
- [116] Geurts P., Ernst D., Wehenkel L. (2006). Extremely randomized trees. *Machine Learning*, 63(1), 3–42. <https://doi.org/10.1007/s10994-006-6226-1>
- [117] Sagi O., Rokach L. (2018). Ensemble learning: A survey. *WIREs: DMKD*, 8(4), e1249. <https://doi.org/10.1002/widm.1249>
- [118] Schapire R E., Freund Y. (2012). *Boosting: Foundations and Algorithms*. MIT Press. <https://doi.org/10.7551/mitpress/8291.001.0001>
- [119] Bentejac C., Corgo A., Martinez-Munoz G. (2021). Comparative analysis of gradient boosting algorithms. *AI Review*, 54, 1937–1967. <https://doi.org/10.1007/s10462-020-09896-5>
- [120] Chen T., Guestrin C. (2016). XGBoost: A Scalable Tree Boosting System. *KDD 2016*. <https://doi.org/10.1145/2939672.2939785>
- [121] Ke G., et al. (2017). LightGBM: A Highly Efficient Gradient Boosting Decision Tree *NeurIPS 2017*. https://proceedings.neurips.cc/paper_files/paper/2017/file/6449f44a102fde848669bdd9eb6b76fa-Paper.pdf
- [122] Prokhorenkova L., Gusev G., Vorobev A., Dorogush A V., Gulin A. (2018). CatBoost: Unbiased boosting with categorical features. *NeurIPS 2018*. <https://doi.org/10.48550/arXiv.1706.09516>
- [123] Pedregosa F., Varoquaux G., Gramfort A., et al. (2011). scikit-learn: Machine Learning in Python. *JMLR*, 12, 2825–2830. <https://jmlr.org/papers/volume12/pedregosa11a/pedregosa11a.pdf>
- [124] Lou, Y., Caruana, R., Gehrke, J., & Hooker, G. (2013). Accurate intelligible models with pairwise interactions. *Proceedings of the 19th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD ’13)*, 623–631. <https://doi.org/10.1145/2487575.2487579>
- [125] Yang, Z.-Q., Sun, Y., & Song, L. (2012). Highway traffic accident prediction based on support vector regression (SVR). *Advanced Materials Research*, 433–440, <https://doi.org/10.4028/www.scientific.net/AMR.433-440.5886>

- [126] Assi K. (2020). Traffic Crash Severity Prediction—A Synergy by Hybrid Principal Component Analysis and Machine Learning Models. *IJERPH*, 17(20), 7598. <https://doi.org/10.3390/ijerph17207598>
- [127] Hasan A S., et al. (2024). Application of machine learning models and SHAP to examine crashes involving young drivers in New Jersey. *International Journal of Transportation Science and Technology*, 14, 156-170. <https://doi.org/10.1016/j.ijtst.2023.04.005>
- [128] Zhu S., Wang K., Li C. (2021). Crash Injury Severity Prediction Using an Ordinal Classification ML Approach. *IJERPH*, 18(21), 11564. <https://doi.org/10.3390/ijerph182111564>
- [129] Dietterich T G. (2000). Ensemble methods in machine learning. *Multiple Classifier Systems, LNCS 1857*, 1-15. https://doi.org/10.1007/3-540-45014-9_1
- [130] Hoeting J A., Madigan D., Raftery A E., Volinsky C T. (1999). Bayesian model averaging: A tutorial. *Statistical Science*, 14(4), 382-417. <https://doi.org/10.1214/ss/1009212519>
- [131] Hastie T., Tibshirani R., Friedman J. (2009). *The Elements of Statistical Learning* (2nd ed.). Springer. <https://doi.org/10.1007/978-0-387-84858-7>
- [132] Kitali A E., et al. (2019). Identification of Secondary Crash Risk Factors Using Penalized Logistic Regression. *Transportation Research Record*, 2673(11), 901-914. <https://doi.org/10.1177/0361198119849053>
- [133] Garefalakis T., Katrakazas C., Yannis G. (2022). Data-Driven Estimation of a Driving Safety Tolerance Zone Using Imbalanced Machine Learning. *Sensors*, 22(14), 5309. <https://doi.org/10.3390/s22145309>
- [134] Carrodano C. (2024). Data-driven risk analysis of nonlinear factor interactions in road safety using Bayesian networks. *Scientific Reports*, 14, 18948. <https://doi.org/10.1038/s41598-024-69740-6>
- [135] Nojavan F., Qian S S., Stow C A. (2017). Comparative analysis of discretization methods in Bayesian networks. *Environmental Modelling & Software*, 87, 64-71. <https://doi.org/10.1016/j.envsoft.2016.10.007>
- [136] BayesFusion. (2024). Greedy Thick Thinning (GTT) – GeNIe Modeler Documentation. https://support.bayesfusion.com/docs/GeNIe/structurelearning_gtt.html
- [137] Colombo D., Maathuis M H. (2014). Order-Independent Constraint-Based Causal Structure Learning. *JMLR*, 15, 3921-3962. <https://doi.org/10.48550/arXiv.1211.3295>
- [138] Friedman N., Geiger D., Goldszmidt M. (1997). Bayesian Network Classifiers. *Machine Learning*, 29(2-3), 131-163. <https://doi.org/10.1023/A:1007465528199>
- [139] Huang Y., Li W., Macheret F., Gabriel R. A., Ohno-Machado L. (2020). A tutorial on calibration measurements and calibration models for clinical prediction models. *Journal of the American Medical Informatics Association*, 27(4), 621-633. <https://doi.org/10.1093/jamia/ocz228>
- [140] Dimitriadis T., Gneiting T., Jordan A I. (2021). Stable reliability diagrams for probabilistic classifiers. *PNAS*, 118(8), e2016191118. <https://doi.org/10.1073/pnas.2016191118>
- [141] Henzi A., Puke M., Dimitriadis T., Ziegel J. (2024). A Safe Hosmer-Lemeshow Test. *NEJSDS*, 2(2), 175-189. <https://doi.org/10.51387/23-NEJSDS56>
- [142] Wang Y., Zhang J., Wu G. (2023). Bayesian-Based Traffic Safety Evaluation Study for Driverless Infiltration. *Applied Sciences*, 13(22), 12291. <https://doi.org/10.3390/app132212291>

14. ПРИЛОЗИ

14.1 Опис на параметри

Р.Б.	Симбол	Опис	Единица	Формула	Објаснување на формулата
1	Ave_Rain	Просечни годишни врнежи	mm/m ²	/	/
2	Max_Snow	Максимален годишен снег	mm/m ²	/	/
3	Ave_Temp	Просечна годишна температура	°C	/	/
4	Max_Temp	Максимална годишна температура	°C	/	/
5	Min_Temp	Минимална годишна температура	°C	/	/
6	Ave_Height	Просечна надморска височина	m	/	/
7	LIMIT	Ограничување на брзина	km/h	/	/
8	Kr.Kar.	Коефициент на закривеност	1/km	$\sum\alpha/L*1000$	α - збир од прекршни агли; L - должина на делницата
9	Ave_Rad	Просечен хоризонтален радиус	m	$\sum R/n$	R - хоризонтален радиус; n - број на кривини
10	Lat_Force	Бочен потисок	N/kg	$V^2/(12.96*R)$	V - просечна ограничена брзина; R - просечен хоризонтален радиус
11	Ave_Inc	Просечен надолжен наклон	%	$\sum i_n/n*100$	i_n - секој наклон; n - број на наклони
12	SKIDRES	Коефициентот на отпорот од триење	/	/	/
13	RDB_RUT	Колотрази	mm	/	/
14	IRI	Меѓународен индекс на рамност (IRI)	m/km	/	/
15	PCI	Индекс на состојба на коловоз (PCI)	points	/	/
16	KON.NAK.	Коефициент на попречен наклон	1/km	$\sum n_{kk}/L*1000$	n_{kk} - збир од сите контра наклони; L - должина на делницата
17	SSD H	Коефициент на хоризонтална прегледност	1/km	$\sum n_{hk}/L*1000$	n_{hk} - број на хоризонтално затскриени точки; L - должина
18	SSD V	Коефициент на вертикална прегледност	1/km	$\sum n_{vk}/L*1000$	n_{vk} - број на вертикално затскриени точки; L - должина
19	V sign	Старост на вертикална сигнализација	years	/	/
20	H sign	Стапка на обнова на хоризонтална сигнализација	times/year	/	/
21	K.Int. P.T	Коефициент на крстосници и наплатни станици	1/km	$(\sum n_i + \sum n_{pt})/L*1000$	n_i - број на клучки; n_{pt} - број на патарини; L - должина
22	K.Bridges	Коефициент на мостови	1/km	$\sum n_b/L*1000$	n_b - број на мостови; L - должина на делницата
23	PGDS	Просечен годишен дневен сообраќај (ПГДС)	veh/day	/	/
24	Wi	Тежински индекс на незгоди	1/km	$(F*85 + I*10 + C*1)/L*1000$	F - незгоди со смртни последици, I - незгоди со повредени, C - незгоди без повреди; L - должина на делницата

14.2 Збирна табела на сите податоци

РБ	Пат	БП	Делница	Дол. [m]	A													K.P.	K.B	PGDS	Wl							
					1	2	3	4	5	6	7	8	9	10	11	12	13					14	15	16	17	18	19	20
1	A1		Државна граница РС/МК (Табановце) - Речица (автопат)	5664	602	11	13	31	-3	376	94	35,33	1554	0,44	0,66	0,50	0,94	1,59	92	0,53	0,88	0,18	12,0	2,0	0,18	1,06	7136	2,21
2	A1		Речица (автопат) - Д. Коњаре 1 (автопат)	3489	596	10	13	31	-2	353	130	7,67	1900	0,69	1,38	0,50	1,97	1,76	91	0,29	1,15	0,00	12,0	2,0	0,57	0,86	7136	1,15
3	A1		Д. Коњаре 1 (автопат) - Куманово 1 (автопат)	2775	593	10	13	31	-2	340	130	43,27	1060	1,23	1,47	0,35	6,68	1,16	79	0,72	0,72	0,36	12,0	2,0	0,36	1,44	7136	9,37
4	A1		Куманово 1 (автопат) - Миладиновци 1 (автопат)	17855	598	10	13	31	-2	361	96	30,85	1920	0,37	2,18	0,35	5,82	1,32	76	0,62	0,73	0,11	12,0	2,0	0,17	0,78	8112	2,46
5	A1		Миладиновци 1 (автопат) - Петровец 1 (автопат)	5827	569	7	13	32	-2	251	120	24,54	3264	0,34	1,19	0,35	1,42	2,51	71	0,17	0,51	0,00	21,0	2,0	0,51	0,86	7893	3,35
6	A1		Петровец 1 (автопат) - Катланово 1 (автопат)	7008	564	6	14	32	-2	232	120	12,28	6005	0,18	0,98	0,40	1,30	2,12	68	0,14	0,00	0,00	24,0	2,0	0,14	0,29	11232	10,13
7	A1		Катланово 1 (автопат) - Башино село (автопат)	23794	566	6	13	32	-2	239	79	102,53	688	0,70	1,60	0,50	1,35	1,64	71	0,59	1,85	0,08	30,0	2,0	0,08	0,76	11232	3,19
8	A1		Башино село (автопат) - Велес 1 (автопат)	3286	559	6	14	32	-2	213	79	65,69	6600	0,07	3,23	0,50	0,56	1,58	74	0,61	1,22	0,00	26,0	2,0	0,30	0,91	11232	12,93
9	A1		Велес 1 (автопат) - Д. Каласлари (автопат)	6673	566	7	13	32	-2	238	90	32,82	1458	0,43	3,97	0,40	7,25	2,32	84	0,15	0,75	0,15	20,0	2,0	0,30	0,45	8001	3,75
10	A1		Д. Каласлари (автопат) - Градско 1 (автопат)	18459	549	4	14	33	-2	174	96	30,84	1340	0,53	1,81	0,40	5,76	2,22	86	0,27	0,43	0,05	11,0	2,0	0,16	0,70	8001	4,09
11	A1		Градско 1 (автопат) - Неготино 1 (автопат)	15783	544	4	14	33	-2	154	120	27,14	4165	0,27	2,37	0,40	3,04	2,84	74	0,19	0,57	0,19	19,0	2,0	0,13	0,13	5085	0,25
12	A1		Неготино 1 (автопат) - Демир Капија 1 (автопат)	16110	540	3	14	33	-2	136	110	35,58	5168	0,18	0,90	0,40	2,77	2,20	76	0,37	0,81	0,06	19,0	2,0	0,12	0,31	5526	2,86
13	A1		Демир Капија 1 (автопат) - Миравци 2 (автопат)	22739	570	7	13	32	-2	250	107	41,05	1252	0,71	1,95	0,40	3,53	1,01	93	0,44	0,66	0,04	5,0	2,0	0,04	0,48	3094	1,1
14	A1		Миравци 2 (автопат) - Смоквица (автопат)	5333	538	3	14	33	-2	132	110	15,34	3209	0,29	2,33	0,50	2,75	0,99	94	0,19	0,38	0,19	5,0	2,0	0,19	0,19	3094	0,94
15	A1		Смоквица (автопат) - Прдејци (автопат)	7188	524	1	14	33	-2	77	90	26,84	5043	0,12	0,83	0,50	3,15	1,67	95	0,42	0,56	0,00	13,0	2,0	0,28	0,42	6618	0,56
16	A1		Прдејци (автопат) - Гевгелија 1 (автопат)	6242	519	1	14	33	-1	59	100	21,64	4508	0,17	0,98	0,50	5,65	1,82	86	0,48	0,16	0,00	12,0	2,0	0,16	0,32	6618	5,69
17	A1		Гевгелија 1 (автопат) - Државна граница МК/ГР (Богородица)	3864	520	1	14	33	-1	61	73	46,52	4593	0,09	1,93	0,50	4,39	2,10	81	0,26	0,52	0,00	11,0	2,0	0,26	0,52	7626	1,29
18	A2		Миладиновци (автопат) - Илинден (автопат)	7340	571	7	13	32	-2	258	74	45,06	2613	0,16	1,57	0,40	3,35	2,79	59	0,41	0,41	0,14	18,0	2,0	0,41	0,82	11152	5,93
19	A2		Илинден (автопат) - Инчиково (автопат)	3997	564	6	14	32	-2	231	100	104,82	2062	0,37	1,96	0,40	2,17	3,70	61	0,50	0,25	0,00	18,0	2,0	0,50	0,25	11141	8,01
20	A2		Инчиково (автопат) - Ченто 2 (автопат)	2094	564	6	14	32	-2	231	120	12,85	15900	0,07	1,62	0,40	1,13	4,18	72	0,00	0,96	0,00	18,0	2,0	0,48	0,00	9775	5,73

РБ	Пат	БП	Делница	Дол. [m]	A																PGDS	WI						
					1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16			17	18	19	20	21	22
21	A2	54	Ченто 2 (автопат) - Клучка Стенковец	14363	579	8	13	32	-2	289	72	21,48	3010	0,13	0,88	0,50	1,05	1,71	66	0,28	0,49	0,14	18,0	2,0	0,35	0,63	9775	7,38
22	A2	55	Клучка Стенковец - Сарај 1 (автопат)	9210	588	9	13	31	-2	318	90	33,25	4165	0,15	4,31	0,50	1,53	1,58	71	0,11	0,54	0,11	18,0	2,0	0,33	1,09	10641	5,92
23	A2	56	Сарај 1 (автопат) - Д. Греч (автопат)	2024	577	8	13	32	-2	282	90	26,95	1956	0,32	2,16	0,50	0,98	1,74	70	0,00	0,49	0,00	18,0	2,0	0,49	0,99	13118	15,32
24	A2	57	Д. Греч (автопат) - Грулчин (автопат)	15289	600	11	13	31	-3	367	96	34,71	6338	0,11	1,92	0,40	0,90	1,48	74	0,20	0,46	0,07	24,0	2,0	0,20	0,39	13118	6,61
25	A2	58	Грулчин (автопат) - Желино 1 (автопат)	7917	626	14	12	30	-3	468	97	38,21	1293	0,56	1,84	0,50	0,64	1,61	76	0,38	0,76	0,13	24,0	2,0	0,25	0,88	13706	9,98
26	A2	59	Желино 1 (автопат) - Тетово 1 (автопат)	6857	616	13	12	31	-3	428	110	27,31	3342	0,28	2,06	0,50	1,50	1,62	75	0,29	0,73	0,00	24,0	2,0	0,15	0,73	13706	16,19
27	A2	60	Тетово 1 (автопат) - Блатиште (Камењане) (автопат)	7502	622	13	12	30	-3	451	92	17,95	11630	0,06	0,44	0,50	2,61	2,13	75	0,27	0,40	0,00	24,0	2,0	0,53	0,67	14911	14,66
28	A4	151	Инчиново (автопат) - Петровец 1 (автопат)	10644	564	6	14	32	-2	230	82	53,70	33355	0,02	0,61	0,20	5,76	2,13	71	0,28	0,38	0,19	16,0	2,0	0,28	0,28	15536	10,33
29	A4	152	Миладиновци 1 (автопат) - Средно Коњари (автопат)	6384	590	9	13	31	-2	329	120	72,29	1250	0,89	4,48	0,50	1,50	1,74	92	0,31	0,94	0,16	4,0	2,0	0,16	0,16	3367	0,94
30	A4	153	Средно Коњари (автопат) - Превод (автопат)	15210	613	12	12	31	-3	417	120	32,77	2104	0,53	3,03	0,50	1,50	1,45	93	0,33	0,33	0,13	4,0	2,0	0,13	0,46	3367	0,62
31	A4	154	Превод (автопат) - Свети Николе С (автопат)	6879	588	9	13	31	-2	320	120	18,80	2263	0,49	1,70	0,50	1,50	1,42	92	0,00	0,15	0,15	4,0	2,0	0,15	0,44	3367	5,89
32	A4	155	Свети Николе С (автопат) - Свети Николе Ј (автопат)	3219	576	8	13	32	-2	278	120	56,59	920	1,21	0,80	0,50	1,50	1,16	94	0,00	0,62	0,00	4,0	2,0	0,31	1,24	3367	2,49
33	A4	156	Свети Николе Ј (автопат) - Кадрифаково (автопат)	8340	570	7	13	32	-2	253	120	29,77	2618	0,42	0,63	0,50	1,50	0,89	94	0,36	0,24	0,12	4,0	2,0	0,24	0,60	3367	1,5
34	A3	158	Кадрифаково (автопат) - Штип (Три Чешми - автопат)	8567	580	8	13	32	-2	291	120	1,65	7086	0,16	0,74	0,50	1,50	1,09	96	0,23	0,23	0,12	23,0	2,0	0,23	0,12	6798	4,84
35	A1	301	Демир Капија 3 (Корешница) - Удово	20357	531	2	14	33	-2	105	78	70,51	686	0,68	0,45	0,40	4,94	2,55	66	0,25	0,44	0,00	23,0	0,5	0,05	0,49	3610	1,92
36	A1	302	Удово - Марвинци 1	8899	521	1	14	33	-2	68	83	23,15	2689	0,20	0,20	0,40	3,05	1,89	67	0,11	0,22	0,00	23,0	0,5	0,11	0,22	3610	2,7
37	A1	303	Марвинци 1 - Смоквица	2717	522	1	14	33	-2	71	79	94,26	658	0,72	1,55	0,40	3,07	1,95	64	0,74	1,10	0,00	23,0	0,5	1,10	0,74	3610	0,37
38	A2	311	Државна граница МК/БГ (Деве Баир) - Узем	1580	796	35	9	25	-5	1118	100	450,01	97	7,96	5,43	0,40	0,45	2,22	68	1,27	3,16	0,00	23,0	1,0	0,63	0,63	2009	5,38
39	A2	312	Узем - Киселица (Крља)	8992	726	26	10	27	-4	845	87	229,91	425	1,36	4,47	0,35	2,74	3,01	52	1,22	1,89	0,11	29,0	1,0	0,44	0,56	2009	3,67
40	A2	313	Киселица (Крља) - Крива Паланка	3705	681	21	11	29	-4	674	80	261,94	169	2,92	2,21	0,35	3,07	1,83	88	2,16	4,05	0,81	11,0	1,0	0,54	0,54	2009	4,59
41	A2	314	Крива Паланка - Мождивњак	9398	662	18	11	29	-3	605	70	209,86	377	1,00	2,73	0,35	3,86	2,03	80	1,70	1,92	0,21	16,0	1,0	0,64	0,21	3118	7,45
42	A2	315	Мождивњак - Ранковце	11796	637	15	12	30	-3	510	64	86,55	1230	0,26	1,05	0,35	2,56	1,95	76	0,85	1,36	0,25	21,0	1,0	0,59	1,02	3118	3,43

РБ	Пат	БП	Делница	Дол. [m]	[m]																PGDS	WI													
					A	R	M	S	A	T	Мк	T	Mn	T	A	H	LIM	Кр.Кар.	A	Rad			L	F	A	Inc	SKID	R	R	IRI	PCI	K.N.	S.H	S.V	V.s
43	A2		316 Ранковце - Страцин (Крилатица)	5792	640		16	12	30	-3	520	63	57,95	1244	0,25	5,11	0,40	2,99	2,26	68	0,52	0,69	0,35	28,0	1,0	0,35	0,35	3118	4,83						
44	A2		317 Страцин (Крилатица) - Војник	17714	641		15	12	30	-3	522	70	77,46	601	0,63	2,96	0,40	2,70	1,33	93	0,62	0,85	0,34	2,0	1,0	0,17	0,28	2883	4,74						
45	A2		318 Војник - Младо Нагоричане	4833	600		11	13	31	-3	369	80	18,02	7875	0,06	3,79	0,40	2,80	1,51	94	0,21	0,21	0,21	2,0	1,0	0,41	0,21	2883	21,21						
46	A2		319 Младо Нагоричане - Д. Коњаре 2 (Карпош)	8635	613		12	12	31	-3	420	70	86,28	623	0,61	2,28	0,35	2,01	1,30	93	0,46	1,27	0,46	2,0	1,0	0,35	0,12	2883	12,97						
47	A2		320 Д. Коњаре 2 (Карпош) - Д. Коњаре 1 (автопат)	1299	595		10	13	31	-2	349	80	40,05	6930	0,07	5,54	0,40	0,46	1,87	93	0,77	0,77	0,77	2,0	1,0	1,54	0,77	2883	3,85						
48	A2		321 Блатиште (Каменање) (автопат) - Ременница (Неготино Полошко) (автопат)	8233	630		14	12	30	-3	484	70	1,88	0	0,00	0,66	0,40	2,40	1,86	74	0,00	0,00	0,00	19,0	2,0	0,24	0,36	14916	15,91						
49	A2		322 Ременница (Неготино Полошко) (автопат) - Тумчевиште (автопат)	2165	636		15	12	30	-3	506	100	1,64	0	0,00	1,72	0,35	3,52	2,05	81	0,00	0,00	0,00	19,0	2,0	0,46	0,46	14916	8,31						
50	A2		323 Тумчевиште (автопат) - Гостивар 1 (автопат)	3936	637		15	12	30	-3	510	76	3,96	0	0,00	0,87	0,40	4,85	1,90	77	0,00	0,00	0,00	24,0	2,0	0,51	0,25	14916	9,15						
51	A2		324 Гостивар 1 (автопат) - Гостивар 2 (Млака)	1887	641		16	12	30	-3	524	68	36,93	2048	0,17	1,94	0,35	8,72	1,44	79	0,53	0,53	0,00	20,0	2,0	1,06	0,00	13665	5,3						
52	A2		325 Гостивар 2 (Млака) - Гостивар 3 (Здунье)	1810	645		16	12	30	-3	542	60	20,11	2781	0,10	0,97	0,35	6,07	1,55	79	0,55	0,00	0,00	20,0	2,0	1,66	0,00	13665	11,33						
53	A2		326 Гостивар 3 (Здунье) - Ново Село (Церово)	10519	703		22	11	28	-4	762	57	239,36	208	1,19	3,55	0,40	2,15	1,80	81	1,81	3,42	0,00	30,0	2,0	0,57	0,19	6556	15,35						
54	A2		327 Ново Село (Церово) - Жајас 1 (Д. Строгомиште)	22467	771		30	9	26	-5	1019	58	246,49	212	1,23	3,48	0,35	3,58	1,58	76	1,47	3,25	0,13	30,0	1,0	0,71	0,36	6083	9,79						
55	A2		328 Жајас 1 (Д. Строгомиште) - Кичево 1 (Осломеј)	7834	679		20	11	29	-3	669	70	64,01	590	0,64	2,08	0,40	2,89	1,27	96	0,51	0,89	0,00	15,0	1,0	0,89	0,13	6083	23,36						
56	A2		329 Кичево 1 (Осломеј) - Кичево 2	634	671		19	11	29	-3	640	60	40,73	978	0,28	4,72	0,40	0,72	3,33	96	0,00	0,00	0,00	15,0	1,0	1,58	0,00	6083	30						
57	A2		330 Кичево 2 - Кичево 3 (Бигор Доленци)	2592	668		19	11	29	-3	625	50	67,15	1403	0,14	1,48	0,35	2,06	3,39	87	0,00	0,00	0,00	20,0	1,0	0,77	0,77	5946	17,36						
58	A2		331 Кичево 3 (Бигор Доленци) - Кичево 4	293	668		19	11	29	-3	629	50	526,93	771	0,25	0,07	0,35	2,15	2,71	85	3,41	0,00	0,00	20,0	1,0	3,41	0,00	5946	30						
59	A2		332 Кичево 4 - Другово	3511	669		19	11	29	-3	630	63	53,86	1076	0,29	0,83	0,40	3,60	1,99	86	0,57	0,85	0,00	25,0	1,0	1,42	0,00	5147	14,95						
60	A2		333 Другово - Извор	8471	684		21	11	29	-4	685	65	161,49	290	1,12	3,03	0,40	4,26	2,35	81	1,53	3,42	0,12	25,0	1,0	0,35	0,12	5147	9,27						
61	A2		334 Извор - Ботун	26400	736		27	10	27	-4	886	68	118,22	526	0,67	3,12	0,40	4,29	3,01	77	0,72	1,93	0,11	30,0	1,0	0,11	0,61	4903	7,14						
62	A2		335 Ботун - Требеништа	9535	698		22	11	28	-4	741	61	64,53	690	0,42	0,96	0,35	6,23	3,34	56	0,52	1,26	0,42	30,0	1,0	0,21	0,52	4887	13,95						
63	A2		336 Требеништа - Подмолје	5419	689		22	11	28	-4	709	70	124,01	1403	0,27	2,38	0,35	4,80	2,77	52	0,37	1,29	0,18	30,0	2,0	0,55	0,18	8454	6,92						
64	A2		337 Подмолје - Струга	6717	686		21	11	29	-4	696	59	23,04	2082	0,13	0,61	0,40	8,82	2,81	84	0,45	0,30	0,00	15,0	2,0	0,89	0,30	8454	21,96						
65	A2		338 Струга - Државна граница МК/АЛ (Кафасан)	12900	713		24	10	28	-4	799	60	92,35	597	0,00	2,71	0,35	3,70	2,84	91	1,32	1,32	0,39	30,0	1,0	0,47	0,08	3994	4,03						

РБ	Пат	БП	Делница	Дол. [m]																								
				1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	
66	А3	351	Подмоље - Охрид 1 (Св. Евамо)	5171	686	21	11	29	-4	698	60	48,10	1309	0,21	2,50	0,35	4,05	2,15	64	0,58	0,19	0,00	30,0	2,0	0,39	0,19	9956	15,57
67	А3	352	Охрид 1 (Св. Евамо) - Косел	8488	693	22	11	28	-4	721	49	21,06	1588	0,12	1,45	0,40	4,12	3,33	81	0,47	0,59	0,24	20,0	0,5	1,53	0,47	3497	7,25
68	А3	353	Макази 1 (Кајжани) - Кајжани	9946	768	32	9	26	-5	1010	78	117,95	425	1,09	6,62	0,35	3,79	1,54	81	1,11	1,81	0,20	20,0	0,5	1,10	0,60	3497	8,09
69	А3	354	Кајжани - Превалец 1 (Ротински)	6365	741	28	10	27	-4	906	90	42,53	1234	0,51	2,06	0,35	3,12	1,50	84	0,31	0,79	0,00	15,0	0,5	0,16	0,47	3603	4,63
70	А3	355	Превалец 1 (Ротински) - Снегово (Вилбил Камен)	7090	716	25	10	28	-4	812	90	55,24	1023	0,61	3,35	0,35	3,10	1,41	84	0,28	1,27	0,00	15,0	0,5	0,14	0,28	3603	5,43
71	А3	356	Снегово (Вилбил Камен) - Крклино	7819	679	20	11	29	-3	671	66	58,46	1056	0,32	2,58	0,50	3,10	1,15	81	0,51	1,02	0,38	15,0	0,5	0,26	0,26	3603	13,81
72	А3	357	Крклино - Црна Земја (Тополчани)	18891	659	18	11	29	-3	594	90	19,06	4542	0,14	0,81	0,40	2,61	2,58	80	0,21	0,16	0,00	15,0	1,0	0,05	0,26	4619	4,71
73	А3	358	Црна Земја (Тополчани) - Ново Лагово 2 (Прилеп)	12255	668	19	11	29	-3	627	100	17,75	5691	0,14	0,48	0,35	2,71	2,33	81	0,16	0,08	0,08	20,0	1,0	0,16	0,16	4619	8,89
74	А3	359	Каратманово - Овче Поле	14261	577	8	13	32	-2	279	67	46,53	1704	0,20	1,63	0,35	2,66	1,87	80	0,28	0,63	0,56	11,0	1,0	0,49	0,07	3910	8,38
75	А3	360	Овче Поле - Кадрифаково 1 (Овче Поле)	967	569	7	13	32	-2	250	55	43,95	645	0,36	1,66	0,35	5,55	1,86	84	0,00	1,03	1,03	11,0	1,0	2,07	0,00	3910	21,2
76	А3	362	Штип (Три Чешми - автопат) - Штип 1	2423	585	9	13	31	-2	310	49	8,84	11190	0,02	1,40	0,35	5,82	2,08	94	0,00	0,00	0,00	20,0	2,0	0,41	0,41	6876	12,79
77	А3	364	Штип 2 (Клучка) - Крупиште	11531	578	8	13	32	-2	286	62	69,60	643	0,46	0,88	0,50	2,62	2,07	65	1,04	1,65	0,52	30,0	2,0	0,35	0,17	4864	13,18
78	А3	365	Делчево 1 (Драмче) - Делчево 2 (Голак, Габрово)	855	660	18	11	29	-3	596	50	148,27	0	0,00	1,33	0,35	3,08	4,42	71	0,00	0,00	0,00	25,0	0,3	2,34	0,00	1299	2,34
79	А3	366	Делчево 2 (Голак, Габрово) - Делчево 3 (Тработивиште)	1034	661	18	11	29	-3	600	50	30,91	1797	0,00	2,30	0,35	2,48	2,56	71	0,00	0,00	0,00	25,0	0,3	1,93	0,97	1299	1,93
80	А3	367	Делчево 3 (Тработивиште) - Делчево 4 (Цветор)	339	664	19	11	29	-3	614	45	438,44	262	0,60	2,90	0,35	3,66	1,83	72	0,00	0,00	0,00	25,0	0,3	2,95	0,00	1299	5,9
81	А3	368	Делчево 4 (Цветор) - Државна граница МК/БГ (Рамна Нива)	9709	737	27	10	27	-4	890	59	184,66	296	0,90	5,44	0,35	1,74	2,28	74	1,13	1,03	0,10	20,0	0,3	0,21	0,00	1153	2,37
82	А3	369	Овче Поле - Кадрифаково 1 (Овче Поле)	1025	569	7	13	32	-2	250	55	44,19	472	0,50	1,50	0,40	5,05	3,05	74	0,98	0,98	0,98	11,0	1,0	0,98	0,00	3910	20
83	А4	402	Штип 1 (Три Чешми) - Софилари 1 (Ложиште А4)	6115	576	8	13	32	-2	277	68	18,20	3189	0,11	1,91	0,35	6,05	1,12	94	0,33	0,33	0,16	2,0	1,0	0,49	0,49	4503	1,96
84	А4	403	Софилари 1 (Ложиште) - Софилари 2 (Завој)	1013	571	7	13	32	-2	257	80	175,91	277	1,78	4,75	0,35	9,84	1,80	68	0,00	1,97	0,00	30,0	1,0	1,97	0,00	4503	24,68
85	А4	404	Ново Село (Шопур) - Рудник Бучим	6711	604	11	13	31	-3	383	56	118,68	253	0,95	1,66	0,40	3,22	1,18	68	0,75	1,64	0,00	30,0	1,0	0,30	0,60	4503	12,81
86	А4	405	Радовиш 1 (Инево) - Радовиш 2 (Воиславци)	1312	602	11	13	31	-3	376	100	74,23	709	1,09	2,73	0,40	7,24	1,06	95	0,76	1,52	0,00	2,0	1,0	0,76	1,52	4561	6,1

РБ	Пат	БП	Делница	Дол. А.Р. М.С.А.Т. Мх.Т. Мн.Т. А.Н. ЛИМ. Кс.Кар. А.Рад. L.F.A.Inc SKID R.R. IRI PCI K.N. SH SV V.S H.s K.P. K.B PGDS WI																								
				1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	
				[m]																								
87	A4	406	Радвиш 2 (Вославци) - Владивци	17957	583	9	13	32	-2	302	75	30,02	1896	0,23	0,80	0,40	3,87	2,05	89	0,56	0,56	0,17	11,0	1,0	0,22	0,61	4697	4,61
88	A4	407	Владивци - Градошорци	4781	570	7	13	32	-2	254	68	43,00	629	0,56	0,94	0,40	7,33	2,17	89	0,84	1,46	0,42	11,0	1,0	0,63	0,21	4697	1,67
89	A4	408	Градошорци - Струмица 1 (кружен тек)	3915	564	6	14	32	-2	229	48	85,03	4243	0,04	0,87	0,40	4,22	2,38	86	0,26	0,51	0,00	15,0	1,0	0,51	0,51	4697	8,56
90	A4	409	Струмица 1 (кружен тек) - Струмица 2 (Силоси)	1140	563	6	14	32	-2	225	40	107,96	0	0,00	0,38	0,35	5,26	3,34	86	0,00	0,00	0,00	15,0	1,0	3,51	0,00	4697	2,63
91	A3	410	Штип (Три Чешми - автопат) - Крупиште (автопат)	14119	589	9	13	31	-2	325	100	48,71	1349	0,57	2,84	0,50	0,00	1,12	94	0,35	0,35	0,21	3,0	2,0	0,14	0,50	1992	0,81
92	A4	411	Софилари 1 (Ложиште А4) - Ново Село (Шопур А4)	15173	582	8	13	32	-2	298	70	34,09	2378	0,00	1,78	0,50	0,00	1,12	94	0,33	0,46	0,26	2,0	2,0	0,13	0,99	1724	0,53
93	A4	412	Ново Село (Шопур А4) - Радвиш 1 (Ињево)	14854	614	13	12	31	-3	420	73	45,27	809	0,50	1,63	0,40	0,00	1,45	94	0,47	0,81	0,34	2,0	2,0	0,13	0,88	1724	0,74
94	A3	413	Крупиште (автопат) - Трчање (автопат)	12696	586	9	13	31	-2	315	100	28,53	2628	0,29	1,04	0,50	0,00	0,00	92	0,39	0,39	0,32	3,0	1,0	0,24	0,39	1190	0,91
95	A1	1002	Градско 1 (автопат) - Росоман	4272	545	4	14	33	-2	156	65	36,90	1739	0,19	1,54	0,35	3,60	1,83	74	0,47	0,47	0,23	20,0	1,0	0,47	0,00	5097	7,96
96	A1	1003	Росоман - Дреново (Возарци)	11045	548	4	14	33	-2	171	69	57,98	708	0,52	1,49	0,35	3,66	2,40	74	0,81	1,27	0,09	20,0	1,0	0,54	0,45	5097	9,46
97	A1	1004	Дреново (Возарци) - Беловодица	23468	600	11	13	31	-3	366	63	174,31	333	0,92	2,58	0,35	4,85	2,00	84	1,32	2,26	0,04	15,0	1,0	0,21	0,34	4677	7,63
98	A1	1005	Беловодица - Прилеп 1 (Клучка Могила)	14641	715	24	10	28	-4	805	79	98,10	427	1,12	3,70	0,35	5,45	2,12	84	0,68	1,16	0,34	15,0	1,0	0,34	0,34	4677	5,12
99	A1	1006	Прилеп 1 (Клучка Могила) - Ново Лагово 2 (Прилеп)	3848	676	20	11	29	-3	659	75	17,82	4466	0,10	0,88	0,35	3,58	2,34	75	0,26	0,00	0,00	20,0	1,0	0,52	0,52	4672	6,89
100	A3	1021	Косел - Јанковец	26082	750	29	9	27	-4	937	61	157,20	250	1,16	3,83	0,35	2,80	2,08	82	1,19	2,19	0,19	25,0	1,0	0,35	0,27	3468	4,18
101	A3	1022	Јанковец - Ресен	2158	739	28	10	27	-4	896	55	117,12	228	1,02	0,76	0,35	4,82	2,29	81	0,46	0,46	0,00	25,0	1,0	1,85	0,00	3468	8,34
102	A3	1023	Ресен - Макази 1 (Кажани)	5825	735	27	10	27	-4	883	62	115,00	172	1,71	1,04	0,40	2,76	2,82	80	0,86	0,69	0,17	20,0	1,0	2,92	0,17	3603	5,84
103	A3	1024	Крупиште - Чифлик	5478	581	8	13	32	-2	295	53	43,18	548	0,40	0,47	0,35	2,16	1,81	59	0,91	1,83	0,00	30,0	1,0	0,73	0,37	4741	2,74
104	A3	1025	Чифлик - Трчање (автопат)	6347	585	9	13	32	-2	310	70	24,81	973	0,39	1,51	0,35	2,09	1,60	59	0,32	0,47	0,16	30,0	1,0	1,10	0,32	4741	9,53
105	A3	1026	Трчање (автопат) - Кочани 1 (Бели)	2431	590	9	13	31	-2	331	73	32,58	1551	0,27	1,38	0,35	2,00	2,39	68	0,82	0,82	0,00	30,0	1,0	1,23	0,41	4741	4,53
106	A3	1027	Кочани 1 (Бели) - Кочани 2 (Грдовци)	1782	590	9	13	31	-2	332	80	22,57	1315	0,38	2,62	0,50	5,86	2,89	78	1,12	1,12	0,56	20,0	1,0	1,12	1,12	4741	7,86
107	A3	1028	Кочани 2 (Грдовци) - Кочани 3 (Гранит)	3734	594	10	13	31	-2	344	70	21,49	1134	0,33	1,82	0,50	6,35	2,79	78	0,80	0,54	0,27	20,0	1,0	0,27	0,27	4611	12,72
108	A3	1029	Кочани 3 (Гранит) - Превалец 2 (Кочанско)	1625	601	11	13	31	-3	373	60	63,27	375	0,74	2,02	0,50	1,74	1,82	78	0,62	0,62	1,23	20,0	1,0	1,23	0,00	4985	11,08
109	A3	1030	Превалец 2 (Кочанско) - Истибања	3265	601	11	13	31	-3	368	50	149,61	247	0,78	2,20	0,35	2,78	1,72	79	0,61	0,92	0,92	20,0	0,5	0,61	0,61	2352	12,25
110	A3	1031	Истибања - Македонска Каменица	17857	624	14	12	30	-3	456	52	236,07	185	1,11	1,65	0,35	3,12	1,95	79	1,29	1,51	0,11	25,0	0,5	0,45	0,62	2097	2,77

РБ	Пат	БП	Делница	Дол.	[m]																								WI											
					A	R	M	S	A	T	Мх	T	Мн	T	A	H	LIM	Кс.Кар.	A	Rad	L	F	A	Inc	SKID	R	R	IRI		PCI	K.N.	S	H	S	V	V	S	H	S	K
111	A3	1032	Македонска Каменица - Драмче	12200	647		16	12	30	-3	547	67	203,46	195	1,76	4,13	0,35	3,90	1,71	80	1,07	2,05	0,16	25,0	0,5	0,90	0,49	2286	2,75											
112	A3	1033	Драмче - Делчево 1 (Драмче)	11058	653		17	12	29	-3	568	60	154,07	327	0,85	2,56	0,35	1,74	1,76	80	0,99	1,54	0,09	15,0	0,5	0,54	0,36	2286	5,29											
113	A3	1041	Крчилино - Горно Оризари	712	660		18	11	29	-3	597	55	246,86	298	0,78	0,43	0,35	0,10	2,99	81	0,00	0,00	0,00	14,0	0,5	1,40	1,40	3019	0,1											
114	A3	1042	Горно Оризари - Битола 1 (црква Св. Недела)	3761	661		18	11	29	-3	599	65	92,55	627	0,52	1,27	0,35	3,96	2,55	81	0,53	0,00	0,00	14,0	0,5	2,39	0,27	3019	0,53											
115	A3	1043	Битола 1 (црква Св. Недела) - Битола 2 (Хераклеја Линкестис)	3081	661		18	11	29	-3	601	47	226,30	148	1,14	1,11	0,35	4,22	4,09	81	0,97	1,30	0,00	21,0	0,5	4,54	0,32	3019	0,97											
116	A3	1044	Битола 2 (Хераклеја Линкестис) - Крвари	4742	659		18	11	29	-3	592	63	20,63	831	0,36	0,35	0,35	1,64	2,52	66	0,42	0,00	0,21	27,0	0,5	1,69	0,21	3019	4,64											
117	A3	1045	Крвари - Мешитлија	7484	656		18	12	29	-3	583	70	58,07	834	0,45	0,17	0,35	0,85	2,05	65	0,94	1,34	0,13	30,0	0,5	0,94	0,40	3019	6,41											
118	A3	1046	Мешитлија - Државна граница МК/ГР (Мешитлија)	918	658		18	12	29	-3	588	55	95,08	382	0,61	1,26	0,35	1,63	2,88	81	1,09	1,09	0,00	16,0	0,5	2,18	0,00	3019	2,18											
119	A4	1101	Државна граница МК/КОС (Блаце) - Стенковец (Орман)	11313	589		9	13	31	-2	326	55	133,81	337	0,69	1,61	0,40	8,19	3,55	89	1,41	2,21	0,27	11,0	2,0	0,62	0,18	6406	7,25											
120	A4	1102	Софилари 2 (Завој) - Ново Село (Шопур)	14285	583		8	13	32	-2	303	70	37,70	2378	0,16	0,93	0,35	3,22	1,19	66	0,70	0,70	0,21	30,0	2,0	0,63	0,49	3611	7,46											
121	A4	1103	Струмица 2 (Силоси) - Дебиље	3628	561		6	14	32	-2	221	58	64,77	753	0,34	0,61	0,35	6,53	2,74	83	0,28	0,55	0,00	21,0	2,0	3,58	0,00	9139	12,27											
122	A4	1104	Дебиље - Робово	1452	561		6	14	32	-2	219	65	57,70	1372	0,24	0,74	0,40	3,97	1,40	79	1,38	0,69	0,00	21,0	2,0	2,75	0,00	9139	14,46											
123	A4	1105	Робово - Ново Коњарево	21832	559		6	14	32	-2	212	66	50,82	525	0,64	0,72	0,35	3,59	1,79	88	0,82	1,10	0,09	21,0	2,0	1,74	0,46	9139	6,73											
124	A4	1106	Ново Коњарево - Државна граница МК/БГ (Ново Село)	1946	555		5	14	32	-2	197	65	61,52	1396	0,23	2,64	0,35	1,86	2,65	89	0,51	0,51	0,00	21,0	2,0	1,03	0,00	7049	3,6											
125	A4	1107	Стенковец (Орман) - Клучка	2130	583		9	13	32	-2	304	70	56,38	1659	0,23	0,60	0,30	3,56	4,19	83	0,47	1,41	0,00	21,0	2,0	1,88	0,47	6406	23											
126	A1	9001	Државна граница РС/МК (Табановце) - Речица (автопат)	5678	602		11	13	31	-3	375	100	42,66	1554	0,50	0,66	0,50	3,16	2,40	88	0,53	0,70	0,18	12,0	2,0	0,18	1,06	7136	2,82											
127	A1	9002	Речица (автопат) - Д. Коњаре 1 (автопат)	3493	596		10	13	31	-2	353	115	26,13	1900	0,54	1,38	0,50	3,32	1,79	89	0,29	2,00	0,00	12,0	2,0	0,57	1,15	7136	3,72											
128	A1	9003	Д. Коњаре 1 (автопат) - Куманово 1 (автопат)	2795	592		10	13	31	-2	339	130	43,25	1060	1,23	1,47	0,35	2,20	1,22	78	0,36	0,72	0,00	12,0	2,0	0,36	1,07	7136	3,76											
129	A1	9004	Куманово 1 (автопат) - Миладиновци 1 (автопат)	17865	599		10	13	31	-2	363	97	35,67	1920	0,38	2,18	0,35	2,86	1,57	76	0,50	0,84	0,06	12,0	2,0	0,17	0,56	8112	2,18											
130	A1	9005	Миладиновци 1 (автопат) - Миладиновци 1 (автопат) - Петровец 1 (автопат)	5853	569		7	13	32	-2	251	120	24,95	3264	0,34	1,19	0,35	6,46	1,97	68	0,34	0,51	0,00	21,0	2,0	0,51	0,85	7121	1,03											
131	A1	9006	Петровец 1 (автопат) - Петровец 1 (автопат) - Катланово 2 (автопат)	7288	565		6	13	32	-2	236	112	20,49	6005	0,16	0,98	0,35	4,66	2,38	94	0,00	0,14	0,00	24,0	2,0	0,14	0,27	11192	4,12											
132	A1	9007	Катланово 2 (автопат) - Велес 1 (автопат)	21760	596		10	13	31	-2	350	90	41,18	3967	0,16	3,59	0,35	4,87	1,71	72	0,51	0,87	0,09	30,0	2,0	0,09	0,60	11192	5,03											

РБ	Пат	БП	Делница																WI									
			Дол.	A_R	M_S	A_T	Мх_Т	Мп_Т	A_H	LIM	Кр.Кар.	A_Rad	L_F	A_Inc	SKID	R_R	IRI	PCI		K.N.	SH	S_V	V_s	H_s	K.P.	K.B	PGDS	23
			[m]	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	
133	A1	9009	6641	566	7	13	32	-2	240	88	30,84	1458	0,41	3,97	0,40	3,28	1,72	71	0,45	0,75	0,15	20,0	2,0	0,30	0,45	7478	8,58	
134	A1	9010	18486	549	4	14	33	-2	174	95	30,13	1340	0,52	1,81	0,40	2,89	2,25	70	0,22	0,65	0,05	11,0	2,0	0,16	0,70	8001	1,68	
135	A1	9011	15789	544	4	14	33	-2	153	87	27,86	4165	0,14	2,37	0,40	7,19	3,04	70	0,38	0,57	0,06	19,0	2,0	0,13	0,19	5085	1,9	
136	A1	9012	16110	539	3	14	33	-2	136	100	35,71	5168	0,15	0,90	0,40	3,35	3,24	74	0,37	0,81	0,06	19,0	2,0	0,12	0,31	5526	3,51	
137	A1	9013	22692	570	7	13	32	-2	253	100	40,79	1252	0,62	1,95	0,40	2,24	1,05	93	0,57	0,84	0,09	5,0	2,0	0,04	0,44	3094	0,09	
138	A1	9014	5338	538	3	14	33	-2	132	110	15,34	3209	0,29	2,33	0,50	1,25	1,19	93	0,37	0,37	0,00	5,0	2,0	0,19	0,19	3094	0,37	
139	A1	9016	7191	523	1	14	33	-2	76	80	27,27	5043	0,10	0,83	0,40	5,57	1,85	84	0,42	0,70	0,28	13,0	2,0	0,28	0,42	6618	0,14	
140	A1	9017	6230	519	1	14	33	-2	59	105	21,13	4508	0,19	0,98	0,40	5,39	2,05	84	0,32	0,48	0,00	12,0	2,0	0,16	0,48	6618	0,8	
141	A1	9018	3907	520	1	14	33	-2	63	65	55,18	4593	0,07	1,93	0,50	5,59	2,19	79	0,51	0,77	0,00	11,0	2,0	0,26	0,26	7626	4,22	
142	A2	9051	7966	571	7	13	32	-2	260	78	44,74	2613	0,18	1,57	0,40	5,43	1,92	76	0,13	0,00	0,00	18,0	2,0	0,38	0,75	11152	5,21	
143	A2	9052	4002	564	6	13	32	-2	232	92	101,14	2062	0,32	1,96	0,40	6,22	2,26	89	0,50	0,75	0,00	18,0	2,0	0,50	0,25	11152	7,25	
144	A2	9053	2540	564	6	14	32	-2	232	90	12,87	15900	0,04	1,62	0,40	5,73	2,14	76	0,00	0,39	0,00	18,0	2,0	0,39	0,00	9775	5,51	
145	A2	9054	14328	579	8	13	32	-2	289	80	20,57	3010	0,16	0,88	0,50	5,62	2,07	64	0,14	0,63	0,21	18,0	2,0	0,35	0,63	9775	4,19	
146	A2	9055	9199	587	9	13	31	-2	317	120	33,28	4165	0,27	4,31	0,50	5,75	1,69	69	0,22	0,43	0,00	18,0	2,0	0,33	0,98	10641	2,61	
147	A2	9056	2033	578	8	13	32	-2	283	110	54,72	1956	0,48	2,16	0,50	6,85	2,50	74	0,49	0,49	0,00	18,0	2,0	0,49	1,97	13118	13,03	
148	A2	9057	15274	601	11	13	31	-3	369	92	35,65	6338	0,10	1,92	0,35	5,85	2,14	74	0,46	0,92	0,07	24,0	2,0	0,20	0,46	13118	3,83	
149	A2	9058	7945	626	14	12	30	-3	468	97	37,20	1293	0,56	1,84	0,50	5,43	1,77	79	0,38	0,63	0,00	24,0	2,0	0,25	0,76	13706	5,48	
150	A2	9059	6839	616	13	12	31	-3	428	108	29,22	3342	0,27	2,06	0,50	6,20	1,56	79	0,44	0,73	0,00	24,0	2,0	0,15	0,73	13706	3,95	
151	A2	9060	7500	622	13	12	30	-3	451	97	17,94	11630	0,06	0,44	0,35	7,30	2,29	80	0,13	0,53	0,00	24,0	2,0	0,53	0,80	14911	1,33	
152	A4	9151	11502	563	6	14	32	-2	229	77	32,42	33355	0,01	0,61	0,40	6,33	2,17	66	0,09	0,26	0,09	16,0	2,0	0,26	0,35	15536	4,09	

РБ	Пат	БП	Делница	Дол.	А	Р	М	С	А	Т	Мх	Т	Мп	Т	А	Н	ЛИМ	Кр.Кар.	А	Rad	L	F	A	Inc	SKID	R	R	IRI	PCI	K.N.	S	H	S	V	V	s	H	s	K	P.	K	B	PGDS	WI								
				[m]	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24																								
153	A4	9152	Милединовци 1 (автопат) - Средно Коњари (автопат)	6440	590	9	13	31	-2	328	110	70,41	1250	0,75	4,48	0,40	1,00	1,41	88	0,31	0,62	0,16	4,0	1,0	0,16	0,31	3367	6,13																								
154	A4	9153	Средно Коњари (автопат) - Преод (автопат)	15209	613	12	12	31	-3	417	110	33,77	2104	0,44	3,03	0,40	1,22	1,20	89	0,26	0,33	0,20	4,0	1,0	0,13	0,46	3367	0,39																								
155	A4	9154	Преод (автопат) - Свети Николе С (автопат)	6883	587	9	13	31	-2	319	110	17,44	2263	0,41	1,70	0,40	1,05	0,94	89	0,44	0,29	0,15	4,0	1,0	0,15	0,44	3367	0,15																								
156	A4	9155	Свети Николе С (автопат) - Свети Николе Ј (автопат)	3233	576	8	13	32	-2	278	110	66,35	920	1,01	0,80	0,40	1,75	1,02	88	0,62	1,24	0,00	4,0	1,0	0,31	1,24	3367	1,86																								
157	A4	9156	Свети Николе Ј (автопат) - Кадрифаково (автопат)	8662	570	7	13	32	-2	254	110	19,51	2618	0,36	0,63	0,40	1,23	0,80	89	0,00	0,35	0,23	4,0	1,0	0,23	0,58	3367	0,12																								
158	A3	9158	Кадрифаково (автопат) - Штип (Три Чешми – автопат)	8243	581	8	13	32	-2	294	110	6,21	7086	0,13	0,74	0,40	1,00	1,06	90	0,24	0,12	0,12	23,0	2,0	0,24	0,12	6876	0,85																								
159	A2	9321	Блатиште (Камењане) (автопат) - Ременница (Неготино Полошко) (автопат)	8233	630	14	12	30	-3	483	100	2,07	0	0,00	0,66	0,40	5,73	1,79	61	0,00	0,00	0,00	19,0	2,0	0,24	0,36	14916	3,7																								
160	A2	9322	Ременница (Неготино Полошко) (автопат) - Тумчевиште (автопат)	2167	636	15	12	30	-3	507	107	2,53	0	0,00	1,72	0,35	5,85	2,67	64	0,00	0,00	0,00	19,0	2,0	0,46	0,46	14916	5,54																								
161	A2	9323	Тумчевиште (автопат) - Гостивар 1 (автопат)	3936	637	15	12	30	-3	511	120	7,63	0	0,00	0,87	0,40	6,22	2,07	72	0,00	0,00	0,00	24,0	2,0	0,51	0,25	14916	4,83																								

14.3 Дескриптивна статистика на сите податоци

Параметар	Просек	Варијанса	Стандардна дев.	Мин.	Макс.	Број
Ave_Rain	610,60	3460,44	58,83	519,00	796,00	161,00
Max_Snow	11,88	51,92	7,20	0,61	34,98	161,00
Ave_Temp	12,51	1,56	1,25	8,50	14,46	161,00
Max_Temp	30,73	2,95	1,72	25,22	33,42	161,00
Min_Temp	-2,64	0,54	0,74	-5,00	-1,49	161,00
Ave_Height	408,29	50290,01	224,25	59,00	1118,00	161,00
LIMIT	82,20	501,84	22,40	40,00	130,00	161,00
Kr.Kar.	69,06	6354,82	79,72	1,64	526,93	161,00
Ave_Rad	2646,91	19142448,51	4375,21	0,00	33355,00	161,00
Lat_Force	0,50	0,53	0,73	0,00	7,96	161,00
Ave_Inc	1,89	1,59	1,26	0,07	6,62	161,00
SKIDRES	0,40	0,00	0,06	0,20	0,50	161,00
RDB_RUT	3,52	4,24	2,06	0,00	9,84	161,00
IRI	2,03	0,52	0,72	0,00	4,42	161,00
PCI	79,35	102,49	10,12	52,00	96,00	161,00
KON.NAK.	0,51	0,22	0,47	0,00	3,41	161,00
SSD H	0,80	0,53	0,73	0,00	4,05	161,00
SSD V	0,13	0,04	0,21	0,00	1,23	161,00
V sign	17,57	63,82	7,99	2,00	30,00	161,00
H sign	1,46	0,36	0,60	0,33	2,00	161,00
K.Int. P.T	0,66	0,61	0,78	0,04	4,54	161,00
K.Bridges	0,46	0,13	0,36	0,00	1,97	161,00
PGDS	6459,94	15050406,20	3879,48	1153,00	15536,00	161,00
Wi	6,72	35,73	5,98	0,09	30,00	161,00

14.4 Дескриптивна статистика на тренинг податоци (80%)

Параметар	Мин.	Макс.	Просек	Стандардна дев.
Ave_Rain	519,00	796,00	610,28	57,96
Max_Snow	0,63	34,98	11,84	7,10
Ave_Temp	8,50	14,46	12,52	1,23
Max_Temp	25,22	33,42	30,74	1,69
Min_Temp	-5,00	-1,50	-2,64	0,73
Ave_Height	59,00	1118,00	407,14	220,98
LIMIT	40,00	130,00	82,17	22,07
Kr.Kar.	1,64	526,93	73,40	86,71
Ave_Rad	0,00	33355,00	2595,03	4497,67
Lat_Force	0,00	7,96	0,51	0,79
Ave_Inc	0,07	6,62	1,86	1,32
SKIDRES	0,20	0,50	0,40	0,06
RDB_RUT	0,00	9,84	3,55	2,13
IRI	0,00	4,42	2,00	0,70
PCI	52,00	96,00	78,88	10,40
KON.NAK.	0,00	3,41	0,54	0,50
SSD H	0,00	4,05	0,86	0,77
SSD V	0,00	1,23	0,14	0,21
V sign	2,00	30,00	17,71	8,07
H sign	0,33	2,00	1,45	0,61
K.Int. P.T	0,04	4,54	0,66	0,80
K.Bridges	0,00	1,97	0,46	0,36
PGDS	1153,00	15536,00	6509,23	3994,18
Wi	0,09	30,00	6,58	6,08

14.5 Дескриптивна статистика на тест податоци (20%)

Параметар	Мин.	Макс.	Просек	Стандардна дев.
Ave_Rain	519,00	739,00	611,82	62,99
Max_Snow	0,61	27,70	12,03	7,70
Ave_Temp	9,77	14,46	12,48	1,34
Max_Temp	26,96	33,42	30,70	1,84
Min_Temp	-4,26	-1,49	-2,66	0,79
Ave_Height	59,00	896,00	412,76	240,06
LIMIT	48,75	120,00	82,32	24,01
Kr.Kar.	8,84	149,61	52,21	39,55
Ave_Rad	172,00	15900,00	2848,12	3921,18
Lat_Force	0,00	1,71	0,46	0,40
Ave_Inc	0,76	4,72	2,04	1,01
SKIDRES	0,35	0,50	0,41	0,06
RDB_RUT	0,64	6,22	3,43	1,80
IRI	1,01	4,18	2,16	0,81
PCI	61,00	96,00	81,15	8,91
KON.NAK.	0,00	1,12	0,39	0,29
SSD H	0,00	1,93	0,58	0,46
SSD V	0,00	0,92	0,10	0,19
V sign	4,00	30,00	17,03	7,75
H sign	0,33	2,00	1,49	0,58
K.Int. P.T	0,04	2,92	0,66	0,68
K.Bridges	0,00	1,24	0,45	0,33
PGDS	1299,00	13706,00	6268,76	3448,07
Wi	0,37	30,00	7,27	5,63

14.6 Кластер профили по податоци (GMM m-6)

GMM	Гол.	A_R	M_S	A_T	Mx_T	Mn_T	A_H	LIM	Kr.Kar.	A_Rad	L_F	A_Inc	SKID	R_R	IRI	PCI	K.N.	SH	SV	Vs	Hs	K.P.	K.B	PGDS	Wi
0	41	679,00	20,21	11,07	28,74	-3,49	671,00	61,43	115,00	526,00	0,61	2,30	0,35	3,12	2,28	81,00	0,58	1,03	0,10	21,00	1,00	0,54	0,28	4034	3497,00
1	24	587,50	9,02	13,00	31,41	-2,35	319,50	110,00	33,93	1900,00	0,50	1,66	0,50	1,50	1,14	92,00	0,34	0,54	0,15	4,00	2,00	0,18	0,49	3964	3367,00
2	45	564,00	6,18	13,50	32,09	-2,06	231,00	90,00	33,25	3264,00	0,18	1,57	0,40	3,35	2,05	74,00	0,34	0,54	0,05	19,00	2,00	0,30	0,46	8043	7893,00
3	43	600,00	10,67	12,72	31,02	-2,52	369,00	77,14	32,58	1293,00	0,26	1,47	0,40	5,73	2,00	78,00	0,38	0,55	0,00	20,00	2,00	0,46	0,36	9085	7136,00
4	7	595,00	9,98	12,84	31,18	-2,45	349,00	60,00	46,53	623,00	0,50	2,02	0,35	2,66	1,86	80,00	0,61	0,92	0,92	11,00	1,00	0,98	0,07	3910	3910,00
5	1	796,00	34,98	8,50	25,22	-5,00	1118,00	100,00	450,01	97,00	7,96	5,43	0,40	0,45	2,22	68,00	1,27	3,16	0,00	23,00	1,00	0,63	0,63	2009	2009,00

14.7 Кластер профили по параметри (m-8)

Пар. Кластер	Големина	Сред.Внат.Анс.Кор.	Сред.Општ.Анс.Кор.	Сред.Расп. Wi
0	6	0,549	0,222	0,342
1	7	0,845	0,220	0,251
2	3	0,459	0,247	0,158
3	1	1	0,175	0,211
4	1	1	0,127	0,374
5	2	0,616	0,140	0,268
6	1	1	0,176	0,118
7	1	1	0,105	0,357
8	1	1	0,108	0,092

РБ	Параметри	Кластер
1	Ave_Rain	1
2	Max_Snow	1
3	Ave_Temp	1
4	Max_Temp	1
5	Min_Temp	1
6	Ave_Height	1
7	LIMIT	2
8	Kr.Kar.	0
9	Ave_Rad	0
10	Lat_Force	0
11	Ave_Inc	3
12	SKIDRES	2
13	RDB_RUT	7
14	IRI	6
15	PCI	5
16	KON.NAK.	0
17	SSD H	0
18	SSD V	4
19	V sign	5
20	H sign	1
21	K.Int. P.T	2
22	K.Bridges	8
23	PGDS	0

14.8 Целосна интегративна формула (OLS k=7)

$$\begin{aligned}
 W_i = & 8.3925 + 3.1367 \cdot K.Int. P.T + 2.4906 \cdot PGDS - 2.4730 \cdot (PCI \cdot KON.NAK.) + 2.4313 \cdot (LIMIT \cdot K.Int. P.T) + \\
 & 2.1251 \cdot (K.Int. P.T \cdot PCI) + 1.9008 \cdot (K.Int. P.T \cdot V \text{ sign}) + 1.6295 \cdot V \text{ sign} - 1.4302 \cdot (PGDS \cdot PCI) + 1.2119 \cdot KON.NAK.^2 - \\
 & 1.1400 \cdot LIMIT - 1.0901 \cdot PGDS^2 - 0.9821 \cdot (Ave_Inc \cdot KON.NAK.) + 0.8787 \cdot (LIMIT \cdot V \text{ sign}) - 0.8171 \cdot PCI^2 + \\
 & 0.7685 \cdot (LIMIT \cdot KON.NAK.) + 0.7025 \cdot (PGDS \cdot KON.NAK.) + 0.6868 \cdot Ave_Inc - 0.6337 \cdot (V \text{ sign} \cdot KON.NAK.) + \\
 & 0.5746 \cdot PCI - 0.5728 \cdot (PGDS \cdot Ave_Inc) + 0.5618 \cdot (PCI \cdot Ave_Inc) - 0.5453 \cdot (PCI \cdot V \text{ sign}) + 0.5131 \cdot LIMIT^2 - \\
 & 0.4793 \cdot KON.NAK. + 0.4239 \cdot (K.Int. P.T \cdot Ave_Inc) - 0.3422 \cdot (PGDS \cdot V \text{ sign}) - 0.3346 \cdot K.Int. P.T^2 - \\
 & 0.3272 \cdot (LIMIT \cdot Ave_Inc) - 0.2919 \cdot (K.Int. P.T \cdot KON.NAK.) + 0.2098 \cdot V \text{ sign}^2 + 0.1897 \cdot (V \text{ sign} \cdot Ave_Inc) + \\
 & 0.1385 \cdot (LIMIT \cdot PCI) + 0.1015 \cdot (PGDS \cdot K.Int. P.T) - 0.0705 \cdot Ave_Inc^2 - 0.0409 \cdot (PGDS \cdot LIMIT)
 \end{aligned}$$

Број	Термини	Коефициенти
1	Intercept	8,39
2	K.Int. P.T	3,14
3	PGDS	2,49
4	(PCI·KON.NAK.)	-2,47
5	(LIMIT·K.Int. P.T)	2,43
6	(K.Int. P.T·PCI)	2,13
7	(K.Int. P.T·V sign)	1,90
8	V sign	1,63
9	(PGDS·PCI)	-1,43
10	KON.NAK. ²	1,21
11	LIMIT	-1,14
12	PGDS ²	-1,09
13	(Ave_Inc·KON.NAK.)	-0,98
14	(LIMIT·V sign)	0,88
15	PCI ²	-0,82
16	(LIMIT·KON.NAK.)	0,77
17	(PGDS·KON.NAK.)	0,70
18	Ave_Inc	0,69
19	(V sign·KON.NAK.)	-0,63
20	PCI	0,57
21	(PGDS·Ave_Inc)	-0,57
22	(PCI·Ave_Inc)	0,56
23	(PCI·V sign)	-0,55
24	LIMIT ²	0,51
25	KON.NAK.	-0,48
26	(K.Int. P.T·Ave_Inc)	0,42
27	(PGDS·V sign)	-0,34
28	K.Int. P.T ²	-0,33
29	(LIMIT·Ave_Inc)	-0,33
30	(K.Int. P.T·KON.NAK.)	-0,29
31	V sign ²	0,21
32	(V sign·Ave_Inc)	0,19
33	(LIMIT·PCI)	0,14
34	(PGDS·K.Int. P.T)	0,10
35	Ave_Inc ²	-0,07
36	(PGDS·LIMIT)	-0,04

14.9 Оптимална детерминација

m	AB	ET	BR	GB	RF	LGB	XGB	CB	EBM
5	0,45	0,41	0,41	0,33	0,41	0,34	0,38	0,40	0,42
6	0,57	0,55	0,49	0,57	0,49	0,39	0,46	0,49	0,40
7	0,48	0,54	0,50	0,67	0,50	0,44	0,48	0,52	0,44
8	0,50	0,54	0,53	0,55	0,53	0,50	0,60	0,53	0,47
9	0,32	0,62	0,57	0,48	0,56	0,56	0,56	0,55	0,47
10	0,52	0,56	0,56	0,45	0,55	0,52	0,48	0,53	0,39
11	0,47	0,56	0,55	0,46	0,55	0,52	0,47	0,54	0,45
12	0,41	0,59	0,56	0,45	0,55	0,53	0,51	0,54	0,42
13	0,46	0,57	0,55	0,47	0,54	0,51	0,53	0,52	0,47
14	0,38	0,58	0,55	0,58	0,54	0,52	0,53	0,57	0,46
15	0,55	0,58	0,56	0,57	0,53	0,52	0,50	0,53	0,47
16	0,51	0,58	0,54	0,57	0,53	0,52	0,57	0,58	0,48
17	0,48	0,58	0,56	0,59	0,52	0,49	0,55	0,55	0,46
18	0,44	0,60	0,54	0,59	0,52	0,52	0,54	0,56	0,46
19	0,51	0,59	0,53	0,54	0,53	0,53	0,53	0,60	0,51
20	0,53	0,59	0,54	0,55	0,54	0,53	0,56	0,58	0,51
21	0,46	0,60	0,52	0,51	0,51	0,49	0,54	0,51	0,50
22	0,48	0,58	0,52	0,50	0,50	0,50	0,50	0,48	0,46
23	0,51	0,56	0,52	0,50	0,51	0,50	0,38	0,51	0,49

