

Univerzita Palackého v Olomouci

Přírodovědecká fakulta

Katedra geoinformatiky

Renáta SLEZÁKOVÁ

**NELINEÁRNÍ REGRESNÍ ANALÝZA
V PROBLEMATICE MODELOVÁNÍ NAHODILÉ
TĚŽBY DŘEVA**

Bakalářská práce

Vedoucí práce: Mgr. Pavel Tuček, Ph.D.

Olomouc 2014

Čestné prohlášení

Prohlašuji, že jsem bakalářskou práci bakalářského studia oboru Geoinformatika a geografie vypracovala samostatně pod vedením pana Mgr. Pavla Tučka, Ph.D. a konzultantů Ing. Pavla Samce a Doc. RNdr. Pavla Cudlína, Csc. . Všechny použité materiály a zdroje jsou citovány s ohledem na vědeckou etiku, autorská práva a zákony na ochranu duševního vlastnictví. Všechna poskytnutá i vytvořená digitální data nebudu bez souhlasu školy poskytovat.

V Olomouci dne 20. května 2014

Poděkování

Na začátku své bakalářské práce bych ráda poděkovala za podněty, připomínky a konzultace při vypracování práce především vedoucímu bakalářské práce Mgr. Pavlu Tučkovi, Ph.D. z katedry geoinformatiky. Následné velké dík patří mým konzultantům Ing. Pavlu Samcovi z Ústavu pro hospodářskou úpravu lesů Brandýs nad Labem a panu Doc. RNDr. Pavlu Cudlínovi, Csc. z výzkumného centra CzechGlobe v Brně, kterému bych chtěla zároveň poděkovat za poskytnutá data.

Bakalářská práce vznikla rovněž za podpory projektu Vnitřní grantové agentury Univerzity Palackého v Olomouci s názvem Statistické inference nad daty ze senzorů a senzorových sítí - předzpracování, modelování, verifikace, interpretace a vizualizace. Registrační číslo projektu PrF_2013_024.

Vložený originál **zadání** bakalářské/magisterské práce (s podpisy vedoucího katedry, vedoucího práce a razítkem katedry). Ve druhém výtisku práce je vevázána fotokopie zadání.

OBSAH

SEZNAM POUŽITÝCH ZKRATEK	7
ÚVOD	8
1 CÍLE PRÁCE.....	9
2 SOUČASNÝ STAV ŘEŠENÉ PROBLEMATIKY	10
2.1 Změna druhotné skladby lesů.....	10
2.2 Nahodilá těžba (NT).....	10
2.2.1 Abiotičtí činitelé	11
2.2.2 Biotičtí činitelé.....	11
2.2.3 Antropogenní činitelé	11
2.3 Matematicko-statistické metody k modelování nahodilých lesních jevů	12
3 POPIS DAT, POUŽITÉ PROGRAMY A POSTUPY ZPRACOVÁNÍ.....	16
3.1 Popis dat	16
3.1.1 Základní statistická charakteristika dat.....	19
3.1.2 Zájmová lokalita	22
3.2 Použité programy	24
3.3 Postupy zpracování	24
4 STATISTICKÉ METODY ZPRACOVÁNÍ DAT	26
4.1 Regresní analýza	26
4.1.1 Lineární regresní model.....	27
4.1.2 Logistická regrese.....	28
4.2 Softwarové řešení logistické regrese v prostředí RStudio.....	29
4.2.1 Zobecnění lineární model (GLM).....	30
4.3 Shluková analýza	30
4.3.1 Hierarchické shlukování	31
4.3.2 Nehierarchické shlukování	32
4.4 Softwarové řešení shlukové analýzy v prostředí RStudio.....	33
4.4.1 Funkce cluster.Sim	34
5 VÝSLEDKY	36
5.1 Popisná statistika dat	36

5.2	Predikce kumulativní nahodilé těžby	41
5.3	Rozdělení porostních skupin podle podobnosti	42
5.3.1	Charakteristika shluků dle vlastností stanoviště	43
5.3.2	Charakteristika shluků dle vlastností porostu	45
5.3.3	Charakteristika shluků dle vlastností NT	46
5.3.4	Charakteristika shluků dle vlastností poškození porostu	47
5.3.5	Vyhodnocení pomocí heatmap	47
6	DISKUZE	49
7	ZÁVĚR	50
	POUŽITÁ LITERATURA A INFORMAČNÍ ZDROJE	
	SUMMARY	
	PŘÍLOHY	

SEZNAM POUŽITÝCH ZKRATEK

Zkratka	Význam
AIC	Akaike informační kritérium
Gamm	Generalized Mixed Additive Regression Model
GIS	Geografický informační systém
GLM, GLIM	Generalized Linear Model
CHKO	Chráněná krajinná oblast
LT	Lesní typ
LVS	Lesní vegetační stupeň
LHP	Lesní hospodářský plán
MNČ	Metoda nejmenších čtverců
MS	Microsoft
NT	Nahodilá těžba
RMSE	Root Mean Square Error
SHP	Shapefile
SLT	Soubor lesních typů
SW	Software

ÚVOD

Lesní porost je nedílnou součástí naší přírody a plní funkce, které jsou pro společnost rozsáhlé a nezastupitelné. Svými vlastnostmi a bohatstvím se řadí mezi nepřírozanější ekosystémy, které můžeme na Zemi nalézt. Přístup společnosti však nebyl, není a nebude, vůči lesnímu bohatství, vždy pozitivní. Člověk svými zásahy do přírodního lesního ekosystému narušuje jeho stabilitu a přispívá ke zvyšování ztrát porostu způsobených náhodnými nežádoucími činiteli.

Sledování jednotlivých porostních skupin, jejich vlastností, výši poškození aj. je důležitou složkou, která vstupuje do mnoha statistických analýz. Jednotlivé statistické analýzy vlastností porostu mohou sloužit nejen pro účely lesního hospodářství, ale také v problematice územního plánování a ochrany životního prostředí.

Náplní práce je zpracovat data o desetiletém sledování porostních skupin v nejvýchodnější části České republiky, okresu Frýdek-Místek a na základě proměnné zaznamenávající výši kumulativní nahodilé těžby, vypočítat predikci nahodilé těžby do roku 2025. Součástí je i rozdělení porostních skupin do oblastí podle podobných vlastností pro bližší charakteristiku porostních skupin.

1 CÍLE PRÁCE

Cílem bakalářské práce je zpracovat data o desetiletém pozorování lesního porostu a vyhodnotit na základě doporučených matematicko-statistických metod nahodilou těžbu, která je klíčovým ukazatelem pro modelování poškození porostu. Bakalářská práce bude rozšířena o část zabývající se rozdělením jednotlivých porostních skupin do oblastí s podobnými vlastnostmi, podle které bude možné zjistit vzájemný vztah mezi vlastnostmi porostních skupin.

Práce bude obsahovat teoretickou část zaměřenou na přiblížení pojmu nahodilá těžba, činitelé ji vyvolávající a sepsání krátké rešerše na již využitě matematicko-statistické metody modelování nahodilých jevů v lesnictví.

Hlavním cílem práce je, na základě poskytnutých dat zaznamenaných v průběhu deseti let, vypočítat predikci kumulativní nahodilé těžby pro sledované porostní skupiny do roku 2025. Vedlejším výsledkem je rozdělení porostních skupin na shluky podle podobných vlastností, kde byl kladen důraz na zpětnou charakteristiku jednotlivých shluků. Práce se bude také zabývat základní statistikou charakteristikou poskytnutých dat.

Výsledkem práce budou přehledové tabulky a grafy zaznamenávající popis základní statistiky dat, výsledek predikce a charakteristiku vytvořených shluků. Součástí bude poster s mapovým výstupem výsledku predikce a shlukové analýzy. Přiloženo bude i krátké video průběhu nahodilé těžby během sledovaného a predikovaného období.

K bakalářské práci budou zhotoveny i webové stránky.

2 SOUČASNÝ STAV ŘEŠENÉ PROBLEMATIKY

V první části kapitoly současného stavu řešené problematiky je přiblížen význam pojmu nahodilé těžby, její činitelé a změna druhotné skladby lesů, která výši nahodilé těžby rovněž ovlivňuje. Ve druhé části je souhrn již využitých matematicko-statistických metod k modelování úbytku a poškození lesního porostu, jak ze zahraničních, tak i českých publikací a oblastí nejvíce sužovaných zvýšenou nahodilou těžbou způsobenou náhodnými jevy.

2.1 Změna druhotné skladby lesů

Lesy obecně jsou důležitou složkou krajiny a v jejím přirozeném stavu jsou jednou z dominant, které určují její vzhled (Mezera a kol. 1979), chrání ji před zpustnutím a zachovávají úrodnost půdy (Reichholk, 1999). Nejen v dnešní době, ale již po staletí dochází ke značné ztrátě lesních porostů. Tyto fakta vychází z činitelů záporně působící na lesní ekosystém, které se dělí na *abiotické* (vítr, sněhový závěs, námraza, sucho,...), *biotické* (škodlivý hmyz listožravý i podkorní, parazitické houby, býložravá zvěř,...) a *antropogenní činitele* (vznikající činností člověka), jejichž dopady jsou řešeny tzv. nahodilou těžbou. Výše nahodilé těžby vyplývá také ze změny druhotné skladby lesů. Člověk přirozenou skladbu lesů změnil ve prospěch dřevin, ze kterých měl větší užitek, a proto došlo ke změně z původních přirozených lesů smíšených k lesům tvořených jedním stejnověkým druhem stromů – smrku a borovice (Průša, 1990). Dle M. Martiše (1988) jsou lesy takto pozměněny nazvány jako umělé, méně stabilní lesní porosty, na kterých se stále více projevuje působení škodlivých činitelů.

2.2 Nahodilá těžba (NT)

Pod pojmem NT si lze představit těžbu s obtížně předvídatelným nebo nahodilým charakterem vyvolaným působením přírodních či antropogenních činitelů škodící lesním porostům (Poleno, 1995). Těžba nevychází z plánu, který je určen přirozeným dorůstáním lesního porostu až do okamžiku pro skácení stromů, ale je vynucena předčasně kalamitami, popadanými, vyvrácenými stromy, stromy napadenými brouky apod. (Martiš, 1988).

2.2.1 Abiotičtí činitelé

Celou přírodu trvale obklopuje abiotické prostředí, které na ni působí většinou příznivě. Vítr, srážky, teploty mohou někdy překročit hranici svých příznivých vlivů a působí pak na lesy a stromy škodlivě a v některých případech i zhoubně (Křístek a kol. 2002).

Nepříznivé abiotické vlivy působily a stále působí několika násobně větší NT než veškerí biotičtí činitelé dohromady. Proto je pro lesní hospodářství na prvním místě ochrana proti těmto vlivům (Křístek a kol. 2002). Poškození lesních porostů způsobené abiotickými činiteli se nazývá polom. Polomy dělíme podle rozsahu vzniklé škody a podle původce vzniku na *jednotlivé, skupinové, plošné, ledovkové, námrazové, sněhové a větrné*.

2.2.2 Biotičtí činitelé

Podle Š. Křístka a kol. (2002) změna druhotné skladby vedla k vytvoření podmínek pro přemnožení hmyzích škůdců. Ti primárně napadají stromy oslabené, které tvoří příznivé hostitele. Z tohoto důvodu jsou hmyzí škůdci tzv. sekundární škůdci, neboť jen málokdy je jejich rozmnožování nepodmíněno oslabeným dřevinným jedincem.

Polomy způsobené větrem, námrazou nebo sněhem jsou ideálními místy pro přemnožení kůrovce, kteří poté ohrožují i okolní zdravé stromy (Pfeffer, 1989). Proto je nezbytné z těchto polomů odstranit kůru, k zamezení prostředí vhodného pro kůrovce (Reichhok, 1999), včas zpracovat dříví, aby nedošlo k znehodnocení dříví a aby mohla být oblast co nejdříve zalesněna (Poleno, 1995).

2.2.3 Antropogenní činitelé

Hospodářské lesy na našem území jsou ekologicky málo stabilní. Nejsou odolné, jak proti biotickým a abiotickým činitelům, tak i proti člověku, který má také velký podíl na poškozování lesních porostů. Tyto škody působí přímo i nepřímo, úmyslně i neúmyslně, jako uživatel, majitel lesa nebo organizátor hospodářských aktivit. Odhad škod, které způsobí člověk, je tři čtvrtiny z celkových škod působených na lesích (Křístek a kol. 2002). Mezi hlavní poškozování lesů člověkem patří: znečištění ovzduší (vliv imisí na lesní dřeviny), těžba surovin, stavba liniových staveb a provoz komunikací, rekreace a sportovní činnost, nesprávné lesní hospodářství.

2.3 Matematicko-statistické metody k modelování nahodilých lesních jevů

Matematické úvahy v ekologických teoriích a aplikacích, hrají ústřední roli již velmi dlouho. Mnohé modely mají pomoci k řešení problémů, které navzájem spojují matematiky a ekology. Existuje velká spousta modelů, které můžou být aplikovány na data z oblasti poškození lesního porostu, ale ne každý nám dá výsledek, který se očekává. Ke stále lepším výsledkům dopomáhá stálý rozvoj geografických informačních systémů (GIS) a dále vyspělejší analytické metody pro prostorová data k usnadnění prostorově explicitních analýz lesních poruch v rámci velkých prostorových a časových měřítek.

K nejvíce sledovaným, zpracovávaným a predikovaným složkám v poškozování lesních porostů patří především větrné polomy a s nimi spojené rozšíření napadení dřeva kůrovcem. Samozřejmě jsou řešeny i ostatní činitelé způsobující NT, jako sněhové převisy, houby, apod. Tyto negativní jevy mají velký vliv na rozsah nutné NT, ztráty velkého množství dřeva a zásah do pravidelného chodu ekosystému. Oblasti poškozovány větrnými polomy mají celosvětové měřítko, ale k nejvíce náchylným oblastem z výše zmíněným poškozením se řadí především Evropa a západní část Kanady, Britská Kolumbie, kde je také mnoho studií zaměřených na řešení problému s větrnými polomy a rozmnožením kůrovců, spojené se zvýšením požárů a dalších negativních jevů.

V Evropě patří mezi ekonomicky nejvýznamnější evropskou dřevinu Smrk ztepilý, který je obzvláště náchylný k odolávání vichřic, bouří a napadení hmyzem. Mezi hlavní škůdce evropského smrku se řadí Lýkožrout smrkový (*Ips typographus*). Změna klimatu ke stále vyšším teplotám a delším vegetačním obdobím bude mít za následek častější a vážnější propuknutí tohoto škůdce důsledkem více generací za rok (Overbeck a Schmidt, 2012).

Obdobně jako v Evropě je tomu i v Britské Kolumbii v západní části Kanady, kde je mezi dřevinami nejvíce zastoupena Borovice pokroucená s velmi nízkou obranyschopností proti škůdcům a to především proti tzv. Mountain Pine Beetle (*Dendroctonus Ponderosae*), který má za následek mnoho epidemií a katastrof v této oblasti.

Zmíněnými nepříznivými jevy a dopady na lesní ekosystém se zabývalo a zabývá mnoho vědců z různých profesních odvětví. K dosažení, co nejpřesnějšího a nejefektivnějšího výsledku využívají širokou škálu matematicko-statistických metod a přístupů. Tyto studie jsou důležité pro udržení lesního ekosystému, předejití ztrát velkého množství dřeva a predikci rizik. V následující části je přiblížen jen zlomek prací, které se snažili zmenšit nebo předejít negativním vlivům dopadající na lesní porost.

G. Stadelmana a kol. (2013) tvrdí, že existují opatření proti poškozením kůrovcem: předcházením ohniskům bouří poškozeného dřeva NT a sanitací dřeva napadeného kůrovcem. K analýze dynamiky napadení lesního porostu lýkožroutem použili Poissonovi log-normální modely. Práce G. Stadelman a kol. (2013) je zaměřena na datovou sadu údajů za devět let, obsahující 487 porostních oblastí ve Švýcarsku, které byly poškozené především bouří Lothar. Kvantifikovali škody způsobené deštěm a bouřkami, sanaci kácení, NT, množství smrků a klimatické faktory a u těchto proměnných posoudili relativní význam pro odhad napadení kůrovcem. V 1. části byl použit celý datový soubor, kde byly zahrnuty, jak oblasti poškozené bouří, tak i oblasti nepoškozené. Datový soubor byl použit k analýze účinků sanitace kácení na možné poškození kůrovcem. Ve 2. části byla zahrnuta data, která byla ovlivněna bouří. Analýzy byly provedeny pomocí Poissonova rozdělení zaměřené na účinky NT v bouří poškozených oblastech napadených kůrovcem a účinky sanitace. Ke zjištění vhodnosti modelu použili tzv. AIC (Akaike informační kritérium) a Akaike hmotnosti, která odpovídá pravděpodobnosti, že model nejlépe popisuje data. G. Stadelman a kol. (2013) následné modely využili k výpočtu RMSE (Root Mean Square Error) mezi pozorováním a predikcí.

Podle L. M. Schroedera (2010), vědce pracujícího na dynamice rozšíření kůrovce na větrem pokácených stromech, jsou rozhodujícími nástroji pro lesní hospodářství modely, které hodnotí riziko kolonizace kůrovce. Modely posouzení těchto rizik by mohly na základě dostupných proměnných prostředí rozhodnout, které poškozené oblasti větrem by měly být zpracovány jako první. Použil takové modely, které vyhodnocovaly riziko kolonizace kůrovce z větrem vzniklých mezer po bouři v jižním Švédsku v roce 2005, kde bylo hlavním cílem určit, které mezery a krajinné vlastnosti korelují s pravděpodobností kolonizace kůrovce na větrem pokácených smrků. Svoji studii založil na 36 tzv. *mezerách* (oblasti zničené větrem) ve velikosti od 3 – 1.168 počtu padlých stromů. K analýze byly použity zobecněné lineární smíšené modely.

Přítomnost vaječných chodeb v kůře se charakterizovala pomocí pravděpodobnosti pro stromy j v mezeře i , jejichž variace byla modelována pomocí statisticky logistické funkce. Hustota vaječných chodeb byla modelována podobně za předpokladu, že pro daný strom j z mezery i bude následovat negativní binomická distribuce. Konečné regresní modely byly ověřeny porovnáním odhadovaných hodnot přítomnosti vaječných chodeb a hustota těchto chodeb na m^2 kůry pomocí lineárního regresního modelu mezer a krajinných proměnných s logistickými a exponenciálními transformacemi. L. M. Schroeder (2010) zjistil, že začlenění krajinných proměnných v analýze zvýšila vypovídací schopnost modelů. Velký vliv má na množství kolonizace velikost mezer, mezery velké jsou napadené dříve než malé mezery, a proto mají i vyšší pravděpodobnost napadení kůrovci. L. M. Schroederova (2010) studie ukázala, že větrem pokácené stromy zůstávají kolonizovány déle než 3 roky a většina ze stromů bude v tomto období zamořena.

Kamimura a Shirashi (2007) ve své studii rozdělili metody pro zjištění dynamiky a pravděpodobnosti negativních větrných vlivů poškozujících lesní porost na empirické, statistické a mechanické. Celá datová sada, se kterou pracovali, byla z oblasti Japonska, kde jsou velké větrné výkyvy.

Empirické metody jsou založeny na pozorování oblasti, historických záznamech a literárních recenzích pro definici oblastí, kde hrozí nebezpečí poškození větrem. Většinou jsou vyjádřeny jednoduchými regresními analýzami topologických a biologických faktorů. Jednou z empirických analýz je i Booleanova analýza. Tyto metody jsou však statisticky nejisté a je potřeba návaznost na další metody statistické.

Pomocí statistických metod především regresních analýz a modelů lze hodnotit a předpovídat škody za dlouhé časové období. Avšak jejich využití pro predikci je omezeno dostupností a věrohodností používaných dat.

Poslední metody jsou mechanické, které jsou zaměřené jednak na hodnocení odolnosti oblasti vůči větru a dalším činitelům a také pro předpovídání škod podle kritické rychlosti větru a modelu proudění vzduchu pro odhad místního větrného klimatu.

Odlisný přístup ve své práci aplikoval C. Robertson a kol. (2008), kteří se zaměřili na oblast Britské Kolumbie v Kanadě a zavedli systém pro hodnocení krátkodobého rizika úmrtnosti borovic po napadení kůrovce na základě proporcionálních šancí ordinální regrese. Jako proměnné bylo použito přímé krátkovlnné záření, index místa, průměr kmene ve výšce prsou (1,3 m od terénu) a místní časový vývoj populace kůrovce.

Stejnou oblastí se zabýval i Ch. Bone a kol. (2013), kteří ve své studii použili posouzení krátkodobého rizika napadení horské borovice kůrovcem na základě časoprostorového chování šíření v oblasti Britské Kolumbie. Pro tento odhad použili dvourozměrný místní Moran I, který klasifikuje kategorie rizik popisující populační dynamiku kůrovce z jednoho roku na následující, s určením největší pravděpodobnosti výše populace kůrovce - zvýšení, konstanta a pokles. Přesnost predikce modelu kvalitativního rizika byla v prvních letech studia vyšší (91 % v roce 2002) a v následujících letech se snižovala (72 % v roce 2006). Hodnocení rizika může být průběžně aktualizováno s využitím ročních průzkumů. Touto možností se zajistí relativně vysoká predikce rizik v krátkodobém horizontu.

M. Overbeck a M. Schmidt (2012) na základě potřeby identifikovat a kvantifikovat rizika napadení kůrovcem vyvinuli smíšený binární lineární regresní model popisující účinky různých stanovišť (oblastí) a místních proměnných rizika napadení kůrovce na smrkovém porostu. Tato studie proběhla na desetiletých záznamech z přírodní rezervace Dolního Saska v západním pohoří Harz. Proměnné mají kategorii „nenapadené“ a „zamořené“, a proto byl použit zobecněný aditivní smíšený regresivní model (Gamm - generalized mixed additive regression model), za předpokladu, že binomické rozdělení bylo s použitím logistické funkce. Smíšený regresní model se uvádí i s náhodnými efekty, kterými jsou neznámé pěstební opatření a taktika obhospodařování lesů proti škůdcům. Prahová hodnota v práci byla zvolena tak, aby minimalizovala součet špatných výsledků. Bylo zjištěno, že prognostická proměnná věk vykazuje odchylku od linearity.

Odlíšným problémem od ostatních zmíněných, jako je poškození porostu sněhem v Moravskoslezských Beskydech, se zabýval Š. Křístek a kol. (2013), kteří založili svoji práci na použití neuronové sítě, která odhalila podle citlivosti nejvlivnější faktory pro rok a typ poškození. Vztah mezi nimi byl vyjádřen jednoduchou regresní analýzou založenou na polynomu n-tého stupně. Regresní polynomy použili pro vyrovnání hodnot jednotlivých faktorů na stupnici $\langle 0,1 \rangle$, kde 0 odpovídá žádnému poškození a 1 odpovídá 100% poškození. Z vyrovnaných faktorů vypočetli vážený aritmetický průměr, váhou bylo průměrné skóre z analýzy citlivosti (Hlásný a kol. 2011). Výsledná hodnota představuje poškození porostu jako podíl stromů poškozených daným typem poškození v daném roce na celkový počet stromů v porostní skupině (Křístek a kol. 2013).

3 POPIS DAT, POUŽITÉ PROGRAMY A POSTUPY ZPRACOVÁNÍ

3.1 Popis dat

Data, využitá pro vypracování bakalářské práce, byla poskytnuta výzkumným centrem CzechGlobe – Centrum výzkumu globální změny Akademie věd ČR, v. v. i. sídlící v Brně. Formát dat je v Microsoft Excel 97-2003 (xls.) a jejich aktuálnost se vztahuje na desetileté zaznamenávání jednotlivých vlastností a jevů ve vybraných 736 porostních skupinách. Přesněji data popisují situaci v jednotlivých porostních skupinách od roku 1999-2008. Záznamy byly získány i pro rok 2009, avšak ne pro všechny porostní skupiny, které byly pro práci studovány. Z toho důvodu data za rok 2009 nebyla v práci použita pro další zpracování. Jednotlivé porostní skupiny jsou součástí 5 hlavních revírů, které lze rozeznat podle čísel identifikačních kódů porostních skupin. Celá sada dat se skládá z 36 proměnných, které byly sledovány, zaznamenávány nebo dány podle lesního hospodářského plánu (LHP) v jednotlivých letech pro jednotlivé porostní skupiny. Porostními skupinami je míněna statistická jednotka, na které jsou měřeny určité parametry/proměnné, určující její vlastnosti. Typ proměnných je velmi různorodý. Vyskytují se zde, jak proměnné *nominálního* a *ordinálního* typu, souhrnně značená jako data *kvalitativní*, tak i proměnné *spojité* (*metrické*), které patří do *kvantitativních* neboli *numerických dat*. První 3 sloupečky v souboru podávají základní informace o porostních skupinách a zbylých 33 proměnných odpovídá hodnotám sledovaných vlastností porostních skupin. Popis proměnných je blíže uveden v Tabulce 1 i s barevně odlišenými typy jednotlivých proměnných. K datům byl přiložen i shapefile (SHP) polohového rozložení porostních skupin, který byl dále využit pro mapový výstup.

ID OK	ID_OK rozepsan é	rok	Plocha por skupiny	Zastoupe ní dřeviny	Plocha dřeviny	věk 51-80	věk >=81	Zakm.	Lesní typ	SLT	Hosp. soubor	NT kůrovec (5,9)	NT ost.hmyz (6)	NT tracheom yk. (10)	NT živelná (8)	NT OSTAT.	NT Celkový součet
101/A/06	101/A/06	1999	1,08	100	1,08	1	0	9 581	5B		551	0	0	0	0	25,14	25,14
	101/A/06	2000	1,08	100	1,08	1	0	9 581	5B		551	0	0	0	0	0	0
	101/A/06	2001	1,08	100	1,08	1	0	9 581	5B		551	0	0	0	0	0	0
	101/A/06	2002	1,08	100	1,08	1	0	9 581	5B		551	0	0	0	1,87	1,86	3,73
	101/A/06	2003	1,08	100	1,08	1	0	9 581	5B		551	1,09	0	0	0,85	1,27	3,21
	101/A/06	2004	1,08	100	1,08	1	0	9 581	5B		551	0	0	0	3,33	0	3,33
	101/A/06	2005	1,08	100	1,08	1	0	9 581	5B		551	0	0	0	0	0	0
	101/A/06	2006	1,08	100	1,08	1	0	9 581	5B		551	0	0	0	0	10,65	10,65
	101/A/06	2007	1,08	100	1,08	1	0	9 581	5B		551	0	0	0	0	0	0
	101/A/06	2008	1,08	100	1,08	1	0	9 581	5B		551	0	0	41,07	15,52	0	56,59
	101/A/06	2009	1,08	100	1,08	1	0	9 581	5B		551	0	0	0	0	0	0
101/A/07	101/A/07	1999	3,24	88	2,86	1	0	9 581	5B		2511	0,59	0	0	0	0	0,59
	101/A/07	2000	3,24	88	2,86	1	0	9 581	5B		2511	0	0	0	0	68,58	68,58
	101/A/07	2001	3,24	88	2,86	1	0	9 581	5B		2511	0	0	0	0	0	0
	101/A/07	2002	3,24	88	2,86	1	0	9 581	5B		2511	1,1	0	0	0	5,71	6,81
	101/A/07	2003	3,24	88	2,86	1	0	9 581	5B		2511	0	0	0	0	0	0
	101/A/07	2004	3,24	88	2,86	1	0	9 581	5B		2511	0	0	0	1,12	0	1,12
	101/A/07	2005	3,24	88	2,86	1	0	9 581	5B		2511	0	0	0	0	0	0
	101/A/07	2006	3,24	88	2,86	1	0	9 581	5B		2511	0	0	0	0	3,84	3,84

Obr. 1 Ukázka dat (zdroj: vlastní)

Tab. 1 Popis jednotlivých proměnných v datech (1. část)

POPIS JEDNOTLIVÝCH PROMĚNNÝCH V DATECH	
PROMĚNNÁ	POPIS
ID OK	Identifikační kód porostní skupiny, skládající se z číslic a písmen. První číslo stanoví, do jakého revíru porostní skupina spadá (100 => 1. revír, ..., 500 => 5. revír), poslední číslo zastupuje stáří (06 =>60 let, 05 => 50 let, ...), ostatní písmena zastupují jednotlivé dílce porostní skupiny.
ID_OK rozepsané	Rozepsané identifikační kódy pro jednotlivé roky.
Rok	Jednotlivé roky.
Plocha por skupiny (pl_ID)	Plocha porostní oblasti v ha.
Zastoupení dřeviny (zast_dr)	Ukazatel druhotné skladby porostu, poměr redukovaných ploch pokryté jedním druhem dřevin v % (100 % - zastoupení jednoho druhu dřevin, např.: smrk).
Plocha dřeviny (pl_dr)	Parcela v ha, která odpovídá převedené proměnné Zastoupení dřevin na ha.
Věk 51-80	Věk porostní skupiny v rozmezí od 51 let po 80 let. Pokud odpovídá označení 1, neodpovídá označení 0.
Věk >=81	Věk porostní skupiny roven nebo větší než 81 let. Pokud odpovídá označení 1, neodpovídá označení 0.
Zakm.	Zakmenění. Ukazatel hustoty porostu, stanovení zásob dřeva v porostu. Nabývá celočíselných hodnot v rozmezí 0-10, kde 10 odpovídá velmi hustému porostu. V našich datech máme rozmezí rozšířené až po hodnotu 13. Určuje se, buď výpočtem (poměr redukovaných ploch (Zastoupení dřeviny) / poměr celkové plochy (Plocha por skupiny)) nebo pomocí odhadu, jak je tomu v tomto záznamu.
Lesní typ – LT	Proměnná neměnná v čase (mění se, jen když byla v předešlém případě chybně určena), ale v prostoru. Slouží k určení tzv. „úživnosti“ plochy pro dřeviny. Značen trojmístným kódem kombinací čísel a písmen. První číslo určuje lesní vegetační stupeň v rozmezí hodnot 0-9 (viz. níže), písmeno značí edafické kategorie, neboli hodnotu půdy z pohledu trofických, hydrologických vlastností i polohy v terénu (B-bohatá, H – hlinitá, ...) a poslední číslo odpovídá zařazení lesního typu v rámci přírodní lesní oblasti(is.muni.cz, 2010).
SLT	Soubor LT, vyšší typologická jednotka než LT. Stejný kód s vypuštěním posledního údaje. Shromažďuje lesní typy podle příbuznosti hospodářky významných vlastností stanoviště (is.muni.cz, 2010).
Hosp. soubor	Hospodářský soubor. Generalizovaný SLT. Značený číselným kódem, kde první číselná dvojice značí cílové hospodářství a zbylá čísla značí kategorie lesů, porostní typ a další charakteristiky (mezismrky.cz, 2012).
NT kůrovec (NT_kur), NT ost. hmyz, NT tracheomyk (NT_tracheo), NT živelná(NT_ziv), NT OST (NT_ost)	Proměnné zaznamenávající nahodilou těžbu způsobenou danými činiteli (kůrovec, ostatní hmyz, tracheomyk, živelná NT a ostatní) v m ³ . Zjištěné odhadem.
NT Celkový součet (NT_celk_s)	Součet všech nahodilých těžeb v m ³ .
ZÁSOBA daného roku po odečtení NT z předchozích let (zasoba_1)	Zbytek porostu z těžby z předchozího roku v m ³ .

Tab. 1 Popis jednotlivých proměnných v datech (2. část)

POPIS JEDNOTLIVÝCH PROMĚNNÝCH V DATECH	
PROMĚNNÁ	POPIS
ZÁSoba daného roku po odečtení NT z předchozích let + aktuálního roku (zasoba_2)	Zásoba v daném roce.
%NT ze zásoby postupné (odečítám těžby) (%NT_zas_post)	Procentuální vyjádření NT Celkový součet ze ZÁSoby daného roku po odečtení NT z předchozích let.
%NT ze zásoby počáteční (uvedené v LHP) (%NT_zas_poc)	Procentuální hodnota NT z počáteční zásoby, která byla stanovena lesním hospodářským plánem (LHP), vypracovaným obvykle na dobu deseti let, pomáhající vlastníkům lesů ke správnému hospodaření.
Kumulativní %NT ze zásoby počáteční (uvedené v LHP) (%NT_kum)	Načítané hodnoty %NT z počátečních zásob uvedeny v LHP.
ASPECT	Orientace svahu v °.
SLOPE	Sklon svahu v °.
ELEVATION	Nadmořská výška uvedená v m.
Ozáření	Značí vhodnost produkce smrku z hlediska ozáření. Hodnoty 1-4, kde 1 značí nejvhodnější podmínky.
LVS	Lesní vegetační stupeň, který nabývá hodnot 0-9 (0-bory, 1-dubový, 2-bukovo - dubový, 3-dubovo - bukový, 4-bukový, 5-jedlovo - bukový, 6-smrkovo - bukový, 7-bukovo - smrkový, 8-smrkový, 9-klečový)(is.muni.cz, 2010).
TROF.U.	Jednotlivé trofické úrovně představují úrovně produkce biomasy a pohyb energie uvnitř ekosystému. Z nichž můžeme, zjistit potravní možnosti daného ekosystému neboli úživnost. Stupně trofie jsou jako ordinální data v rozsahu 1-21, čím nižší hodnoty, tím lepší, neboť potravní základna pro vyšší úrovně se snižuje vlivem energetických ztrát (Ulbrichtová, 2005).
HYDR.U.	Hydrologická úroveň určující stav hydrologických podmínek v porostní skupině. Udává hodnotu, jak je oblast např.: podmáčená. Škála je v rozmezí 1-9, čím nižší hodnota, tím lepší.
Věk PŘESNÝ (vek_pr)	Přesný věk porostních skupin.
VERT.STR	Vertikální struktura zachycující rozrůzněnost oblasti. Neměnná v čase, ale jen v prostoru. Zjištěná odhadem, nikoliv výpočtem. Rozmezí je 1-5, kde hodnota 5 vyznačuje extrémní svahy. V poskytnutých datech jsou rozšířené kategorie od 0-11.
POCAT.POSK	Poškození na začátku plánu, kde se řadí vlivy abiotické, biotické a antropogenní. Rozmezí 0-8 podle výše počátečního poškození.
IMISE	Stupně poškození porostů imisemi. Dle dynamiky zhoršení imisemi se oblasti dělí do pásem ohrožení A – D, kde A značí nejhůře poškozenou oblast. Ve stávajícím souboru převedeno na ordinální data v rozsahu 0-12 (eagri.cz, 2009).
ZDR.STAV	Určuje zdravotní stav porostu, kde platí, čím nižší číslo, tím horší zdravotní stav. Rozsah je dán od -1-4.
%pošk. stromů	Procentuální záznam poškození porostů, zjištěné odhadem.

Nominální typ proměnné

Značeny také jako *Jmenovité* proměnné jsou tzv. kategoriálními/slovními proměnnými, které není možné vzájemně porovnávat a nelze určit jejich pořadí. Mohou být i číselné, ale bez kvantitativního významu (Neubauer a kol. 2012).

Ordinální typ proměnné

Oproti nominálnímu typu proměnné lze jejich hodnoty seřadit vzestupně/sestupně. Toto pořadí může být nahrazeno číselnými hodnotami. Můžeme určit i rozdíl mezi dvěma hodnotami, který nám ukazuje rozdíl jejich pořadí. Hodnoty bereme jako určitou formu kvantifikace se zohledněním faktu, že stejné hodnoty u dvou statistických jednotek nebudou považovány za totožné (Neubauer a kol. 2012).

Spojitě proměnné

Tyto hodnoty mohou nabývat z teoretického pohledu libovolných hodnot z intervalu reálných čísel (Hendl, 2004).

3.1.1 Základní statistická charakteristika dat

Před samotnou částí statistického zpracování dat pro predikci a další části práce bylo primárně zapotřebí podrobnější seznámení se s jednotlivými proměnnými v datech, se kterými bylo pracováno. Za tímto účelem byla provedena popisná statistická charakteristika dat, která nám vypovídá o různých aspektech dat a při níž bylo použito základních metod s ohledem na různorodost typů proměnných. Popisná statistika se zabývá zejména charakteristikou centrální tendence neboli střední hodnoty a rozptýleností dat. Jak již bylo zmíněno, data obsahují proměnné různého typu, ke kterým bylo třeba odlišného přístupu. Pro tzv. *spojité* proměnné (*metrické*) byl vypočten aritmetický průměr a medián, jako součást charakteristiky střední hodnoty v datech a pro míru rozptýlenosti dat byl vypočten rozptyl a směrodatná odchylka. K charakteristice proměnných *nominálního* a *ordinálního* typu byly vytvořeny podle četností grafy, které vypovídají o četnosti každé z hodnot v proměnné s porovnáním jednotlivých revírů. (Hendl, 2004).

Potřebné bylo i vyhodnocení vzájemných vztahů mezi všemi proměnnými, které bylo zjištěno pomocí korelační analýzy, respektive prostřednictvím hodnoty korelačního koeficientu.

Z důvodů lepšího vyhodnocení výsledků a rozsáhlé datové sady byly porostní skupiny za období deseti let (1999/2008) seskupeny do jednotlivých revírů, pro které byla následně provedena popisná statistika a četnosti. V kapitole níže jsou vysvětleny jednotlivé výpočty aplikované na datech – průměr, medián, rozptyl, směrodatná odchylka apod.

Aritmetický průměr (\bar{x})

Aritmetický průměr podle Hendla (2004) udává typickou velikost hodnot dat v souboru a jeho výpočet se definuje jako součet všech naměřených hodnot x_1, x_2, \dots, x_n vydělených jejich počtem, značené písmenem n . Negativem je ovlivnění extrémními hodnotami.

$$\bar{x} = \frac{\sum_{i=1}^n x_i}{n}$$

Medián (\tilde{x})

Prostřední hodnota, v souboru uspořádaného vzestupně podle velikosti, rozdělující soubor na dvě stejné poloviny. Pro sudý počet n hodnot platí, že \tilde{x} (Hendl, 2004)

$$\tilde{x} = 0,5 \left(x_{\frac{n}{2}} + x_{\frac{n}{2}+1} \right).$$

Pro lichý počet n hodnot platí

$$\tilde{x} = x_{(n+1)/2}.$$

Rozptyl (s_x^2)

Statistická veličina udávající míru rozptylu hodnot v proměnné v okolí její střední hodnoty. Výsledek je vypočítán jako aritmetický průměr čtverců odchylek hodnot proměnné od aritmetického průměru hodnot. Při průměrování odchylky dělíme číslem $(n - 1)$ (Hendl, 2004):

$$s_x^2 = \frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n - 1}$$

Směrodatná odchylka (s_x)

Veličina používaná jako míra rozptýlenosti kolem průměru. Definována odmocninou z rozptylu s^2 . Platí, že $s_x = 0$, když se všechna data rovnají nule, jinak $s_x > 0$. Jako \bar{x} je i s_x ovlivněna extrémními hodnotami (Hendl, 2004).

$$s_x = \sqrt{\frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n - 1}}, \text{ neboli } s_x = \sqrt{s_x^2}.$$

Absolutní četnost

Dle Neubauera a kol. (2012) pro absolutní četnost n_j , která představuje počet výskytů variant x_j v daném souboru, platí

$$\sum_{j=1}^k n_j = n,$$

kde k odpovídá počtu variant. Dle Hendla (2004) je nejznámějším způsobem zobrazení hodnot jedné proměnné *histogram*, kde na ose x jsou hodnoty proměnné a na ose y absolutní četnosti.

Korelační analýza

Ve statistice představuje korelační analýza míru vzájemného vztahu mezi proměnnými, tedy jestli má jedna proměnná tendenci se vyskytovat s hodnotami jiné proměnné. Míra korelace se hodnotí prostřednictvím korelačního koeficientu. Pearsonův korelační koeficient r je jedním z nejdůležitějších ukazatelů velikosti vztahu mezi proměnnými X a Y . Korelační koeficient nabývá hodnot z intervalu $(-1;1)$. Pokud je hodnota $r = 0$, proměnné mezi sebou nekorelují. Čím více se hodnota r blíží k 1, tím více spolu proměnné korelují a naopak, čím více se hodnota r blíží k -1, tím méně mezi sebou proměnné korelují. Koeficient r se počítá prostřednictvím tzv. kovariance s_{xy} , která pomáhá k porozumění, že r má kladnou hodnotu, jestliže vztah mezi proměnnými je pozitivní a věrodatných odchylek s_x a s_y , proměnných X i Y (Hendl, 2004):

$$s_{xy} = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{n - 1}, \quad r_{xy} = \frac{s_{xy}}{s_x s_y}.$$

3.1.2 Zájmová lokalita

Poloha sledovaných porostních skupin je situována v 16 ti obcích v nejvýchodnější části České republiky, v okrese Frýdek-Místek, jenž je součástí Moravskoslezského kraje. Území tvoří v severovýchodní a východní části státní hranice Polské republiky a v části jihovýchodní státní hranice republiky slovenské.



Obr. 2 Zájmové oblast (zdroj: vlastní)

Okres Frýdek-Místek se vyznačuje především lesním bohatstvím, které zaujímá více než polovinu plochy okresu, a proto je označován za nejlesnatější okres na našem území. Díky vysoké produktivitě dřeva a z ní vyplývajících ekonomických důvodů byly přirozené lesní oblasti nahrazovány výhodnějšími smrkovými porosty (www.czso.cz, 2012). Z toho důvodu území spadá do oblasti se střední až velmi silnou odchylkou od přirozené struktury lesů s převažujícími smrkovými porosty a vzhledem k jejich vysoké náchylnosti patří i do mírných až velmi silných zón ohrožení lesa (Samec, Tuček a kol. 2012). Zbylé oblasti s převažujícím přirozeným porostem jsou chráněné. Ze západní části zasahuje do území CHKO Beskydy, která patří mezi největší CHKO v České republice a náleží celku Moravskoslezských Beskyd. Ve zmíněné části CHKO je i národní přírodní rezervace (NPR) Mionší a mnoho přírodních rezervací (PR) např.: PR Uplaz, PR Velký Polom a přírodní památka (PP) Kyčmol. Mimo oblast CHKO, ve východní části zájmové lokality, se rozprostírá PR Bukovec a v severní části PP Rohovec a PP Filipka (www.treking.cz, 2006). Souhrnně se v okrese Frýdek-Místek vyskytují dvě CHKO, dále pět NPR, 18 PR a 19 PP (www.czso.cz, 2012).

Z geologického hlediska je převážná část území tvořena horninou karpatské flyše, v oblasti Třince se vyskytuje těšínský vápenec. Vrcholový pás Beskyd, zasahující do oblasti, se skládá z odolného pískovce a v pásu pohraničním přibývá navíc slepenec. Z Tabulky 2 lze vyčíst začlenění oblasti do jednotlivých geomorfologických jednotek i s přehledovou mapkou zobrazující členění oblasti na celky (Obr. 3).

Tab. 2 Členění do geomorfologického systému (zdroj: www.treking.cz, 2006)

GEMORFOLOGICKÉ ČLENĚNÍ					
SYSTÉM	SUBSYSTÉM	PROVINCIE	SUBPROVINCIE	OBLAST	CELEK
Alpsko - himalájský	Karpaty	Západní Karpaty	Vnější Západní Karpaty	Západobeskydské podhůří	Podbeskydská pahorkatina
				Západní Beskydy	Slezské Beskydy Moravskoslezské Beskydy Jablunkovská brázda Jablunkovské mezihoří



Obr. 3 Ukázka členění oblasti na celky (zdroj: www.risy.cz, 2012)

Povrch oblasti je značně členitý, snížení nastává ve vnitro horské sníženině Jablunkovské brázdy, která leží mezi Moravskoslezskými a Slezskými Beskydami. Oblast náleží k Baltskému úmoří, neboť tudy protéká přítok řeky Odry, Olše, pramenící v Polské republice.

V klimatických podmínkách dochází k přechodu oceánského a kontinentálního klimatu. K velkému ovlivnění přispívá zasahující horský masiv Beskyd, kdy dochází k velkému množství srážek v oblastech návětrné strany masivu. Tyto oblasti jsou považovány za nejdeštivější oblasti v České republice. Průměrné množství srážek odpovídá přibližně 1459,2 mm/rok. Průměrné roční teploty z hlediska nejteplejších hodnot jsou 11,7 °C v červenci a nejchladnějších v lednu -6,1 °C (www.czso.cz, 2012).

3.2 Použité programy

Pro vypracování cílů práce byly použity čtyři hlavní softwary (SW) – Microsoft Excel, RStudio, ArcGIS a Windows Movie Maker, jejichž výběr byl založen na osobní zkušenosti daných programů a možnostech využití jednotlivých SW k potřebám vyřešení cílů bakalářské práce.

Pro základní úpravu, tvorbu tabulek a z části i pro základní statistické výpočty byl použit tabulkový procesor Microsoft Excel 2007 od společnosti Microsoft, ve kterém mi rovněž byly poskytnuty data k bakalářské práci.

K samotným výpočtům, směřujícím k hlavním výsledkům práce, z části základní charakteristice dat a grafické vizualizaci, jsem využila volně dostupné RStudio, které je integrovaným vývojovým prostředím programu R, zaměřené na statistickou analýzu a grafickou vizualizaci nejrůznějších oblastí zájmu.

Mapové výstupy, použité k vizualizaci některých výsledků, byly zpracovány v prostředí ArcGIS 10.1 od firmy ESRI a jejich případné úpravy byly doplněny ve vektorovém editačním prostředí CorelDRAW X6.

Práce obsahuje i krátké video, vyexportovaných mapových výstupů z ArcGIS 10.1 ve formátu jpeg, které byly dále upravovány v programu Windows Movie Maker 2012, zaměřeného na tvorbu, zpracování a sdílení filmů tvořených z videí a fotografií.

3.3 Postupy zpracování

Na začátku práce bylo potřeba seznámení se s literaturou vysvětlující pojem NT, způsoby jejího vzniku a další důležité informace pro správné pochopení dané problematiky. Nevylímáje studie zabývající se modelováním problematiky lesních porostů spojené s činiteli způsobujícími NT, které dopomohlo k obecné představě, jak je problém zvyšující se NT v určitých oblastech světa řešen v současné době. Na základě těchto studií se zvolila vhodná metoda a vizualizace výpočtu predikce kumulativní NT od roku 2009 do roku 2025.

Následoval krok přiblížení si základní charakteristiky a polohy zájmové lokality a pochopení jednotlivých proměnných, hodnot, kterých nabývají, strukturu a způsob jejich zjišťování (odhad/výpočet), zejména konzultací s odborníkem. Na datech byla vytvořena základní statistická charakteristika pro zjištění statistických vlastností a pro možnost jejich srovnání. Proběhla úprava dat, zvolení proměnné kumulativní NT k odhadu predikce, která byla přetransformována v prostředí MS Excelu, následně

importovaná do poznámkového bloku a dále vložena do programu RStudio, kde daný výpočet proběhl. Byla také spojena a upravena data, která vstupovala do základní statistické charakteristiky a vytvoření skupin/oblastí (oblast podle vlastností stanoviště, porostu, poškození a NT), které obsahovaly proměnné charakteristické pro danou oblast a vstupovaly do shlukové analýzy k rozdělení porostních skupin do shluků dle podobnosti.

Po provedených výpočtech byla řešena vhodná vizualizace a statistické zhodnocení daných výsledků. Poslední částí bylo vytvoření vizualizace pomocí mapových, grafických a tabelárních výsledků a finální sepsání práce.

4 STATISTICKÉ METODY ZPRACOVÁNÍ DAT

Statistické analýzy se obecně málokdy zaměřují pouze na jednu proměnnou. Jejich hlavním cílem je především zkoumání vztahů mezi jednotlivými proměnnými, jejich změn v čase apod. Pro zjištění predikce jsou zapotřebí základní metody pro určení vztahu mezi proměnnými, mezi které se řadí metoda regresní analýzy. V kapitole 4 bude tato metoda obecně popsána, zejména její speciální forma - logistická regrese, použitá pro výpočet predikce kumulativní NT, neboť nejlépe vystihovala modelování pravděpodobnosti závislé proměnné - pro mě proměnná Kumulativní NT na nezávislé proměnné - pro mě proměnná Rok a jejíž křivka nejlépe proložila daná data. Za pomoci matematicko-statistických teorií a definic je popsána i metoda shlukové analýzy. Pomocí této metody byly porostní skupiny rozděleny na předem daný počet shluků podle vzájemně podobných vlastností mnou nadefinovaných „oblastí“, obsahující proměnné charakterizující jednotlivé oblasti (oblast podle vlastností stanoviště, porostu, NT a poškození porostu). Toto rozdělení sloužilo k další interpretaci daných vlastností porostních skupin, podle nichž lze porovnávat, jak moc se odlišují jednotlivé shluky porostních skupin daných oblastí mezi sebou. Zda shluky jednotlivých oblastí vykazují podobné zařazení porostních skupin nebo je zařazení porostních skupin jednotlivých oblastí zcela různorodé.

4.1 Regresní analýza

Pojem regrese poprvé použil na konci 19. století učenec z Velké Británie Francis Galtona. K jeho nejznámější studii patří práce specializovaná na závislost výšky potomků na výšce jejich rodičů (otec/syn, matka/dcera). Objevil, že syn otce s vysokou výškou má obvykle menší vzrůst než otec a naopak, že syn malého otce má výšku vyšší (Hendl, 2004). Pro srovnatelnost výšek mezi otcem/synem a matkou/dcerou zvětšil výšku vždy o 8 % (Zvára, 2008). Tuto skutečnost pojmenoval jako tzv. „regresi k průměru“. Od studie F. Galtona došel vývoj ke vzniku regresní analýzy, zabývající se zkoumáním vztahů a závislostí mezi různorodými proměnnými. Regresní analýza buď samostatně, nebo v kombinaci s jinými metodami patří mezi nejlépe využívané metody na světě. Její hlavní cíl je popsat tvar závislosti (vztahu) dvou nebo více proměnných. Pomáhá stanovit vhodný vzorec pro predikci závislé proměnné a zhodnotit predikční chybu na základě znalosti ovlivňujících proměnných. Obecně jde o analýzu závislosti (vztahu) mezi závisle proměnnou (regresand, náhodná veličina, vysvětlovaná proměnná) značenou Y

a nezávisle proměnnou (regresor, ovlivňující, vysvětlující proměnná) značenou X . Vysvětlovaná proměnná Y je spojena s vysvětlující proměnnou X tzv. *regresní funkcí*, která obsahuje několik neznámých parametrů. Pokud je v těchto parametrech funkce lineární, jedná se o tzv. *lineární regresní model*. (Hendl, 2004)

4.1.1 Lineární regresní model

Využití lineárního regresního modelu ke statistickému modelování závislosti sledované proměnné Y na vysvětlující proměnné X vychází z předpokladu, že pro dané údaje platí rovnice

$$y = \alpha + \beta x + e,$$

kde hodnota chyby e představuje rozdíl mezi naměřenou hodnotou a hodnotou predikovanou tzv. *reziduální hodnotu predikce* neboli *chybu predikce*, proměnné Y . Předpokladem u hodnot chyb e je normální rozdělení a jejich vzájemná nezávislost. Mezi proměnnými musí být také splněna podmínka

$$E(Y|X = x) = \alpha + \beta x,$$

kde β je regresním koeficientem regresního modelu, α absolutním členem a $\sigma_{y,x}$ směrodatnou chybou odhadu. Z rovnice je také zřejmé to, že podmíněná střední hodnota závislé proměnné Y je lineární funkcí hodnoty x .

K odhadu parametrů regresní přímky $\alpha + \beta x$ je nejideálnější odhad a a b pomocí metody nejmenších čtverců (MNC). Tato metoda je založena tak, aby minimální součet druhých mocnin chyby predikce e byl

$$s_r^2 = \sum e_i^2 = \sum (y_i - a - bx_i)^2$$

vůči parametrům a a b . Minimalizace sečtených čtverců úseček vyznačuje velikost reziduálních hodnot e pro hodnoty $\{(x_i, y_i)\}$ od proložené přímky ($y = a + bx$) ve směru osy Y . Minimum vede k hodnotám

$$b = r \frac{s_y}{s_x}, a = \bar{y} - b\bar{x},$$

kde r značí korelaci proměnných a s_y, s_x označují směrodatnou odchylku naměřených hodnot obou proměnných X a Y (Hendl, 2004)

4.1.2 Logistická regrese

Typ logistické regrese byl navržen v 60. letech jako náhradní postup k MNČ v případě, kdy vysvětlovaná proměnná nabývá binárních hodnot (Meloun a Militký 2004). Stejným způsobem, jak je tomu u lineární regrese, bude logistická regrese vyjadřovat střední hodnotu závislé (vysvětlované) proměnné jako funkci nezávisle proměnných (Zvára, 2008). S tím rozdílem, že bude omezena v rozsahu počtu hodnot závislé proměnné, kdy predikce pravděpodobnosti jevu nastane nebo nenastane. Proto vypočítaná pravděpodobnost bude nabývat hodnot pro jev, který nastal 1, nebo 0 pro jev, který nenastal. Logistická regrese může být aplikována na jakoukoliv kombinaci spojitých nebo diskretních proměnných, ale vyžaduje znalost, jak závislé proměnné, tak i nezávislé proměnné (Meloun a Militký 2004). Podle Hosmera a Lewenshowa (2000) je pravděpodobnost π závislé proměnné na nezávisle proměnné x modelována logistickým modelem

$$\pi(x) = \frac{e^{\beta_0 + \beta_1 x}}{1 + e^{\beta_0 + \beta_1 x}}$$

Transformace $\pi(x)$, která je základem studie logistické regrese je tzv. *logitová transformace* (Hosmer a Lewenshow, 2000; Burcin a kol. 2010), která je definována z hlediska tzv. *poměru šancí, naděje*, tedy porovnání pravděpodobnost jevu, který nastal a jevu, který nenastal (Zvára, 2008). Transformace je definována z hlediska $\pi(x)$, jako (Hosmer a Lewenshow 2000):

$$g(x) = \ln \left[\frac{\pi(x)}{1 - \pi(x)} \right] = \beta_0 + \beta_1 x.$$

Význam transformace je, že má mnoho žádoucích vlastností lineárního regresního modelu. Neboť pravděpodobnost π leží mezi čísly 0 a 1, tak hodnota *šance* $\pi(x)/1-(x)$ zkoumaného jevu náleží do intervalu $(0, \infty)$. Avšak samotná funkce $g(x)$, nazývaná *logit* pravděpodobnosti π , která představuje logaritmus *šance* $\ln (\pi(x)/1-(x))$, je v rozmezí od $-\infty$ do ∞ , v závislosti na x , jak je tomu právě u lineárního regresního modelu (Hosmer a Lewenshow, 2000; Hendl, 2004; Grafarend, 2006).

K odhadu parametrů v logistickém regresním modelu se využívá metody maximální věrohodnosti, namísto MNČ, jak tomu bylo u modelu lineární regrese. Odhady parametrů jsou vybrány tak, aby maximalizovali funkci maximální věrohodnosti. Pro (x_i, y_i) platí, když y_i je roven 1, pak příspěvek na věrohodnostní funkci je $\pi(x_i)$ a pokud je y_i rovno 0, pak je příspěvek k věrohodnostní funkci $1-\pi(x_i)$, kde $\pi(x_i)$ značí hodnotu $\pi(x)$ vzhledem

k x_i . Definice logaritmovaného výrazu pro vyjádření přínosů x , y pro věrohodnostní funkci je

$$L(\beta) = \ln[l(\beta)] = \sum_{i=1}^n \{y_i \ln[\pi(x_i)] + (1 - y_i) \ln[1 - \pi(x_i)]\}.$$

U interpretace nezávislé proměnné se předpokládá, že bude nabývat hodnot 0 nebo 1, tedy rozdíl logitu s $x = 1$ a $x = 0$ je

$$g(1) - g(0) = [\beta_0 + \beta_1] - [\beta_0] = \beta_1.$$

Výsledek vyšel roven β_1 , kdy za účelem interpretace tohoto výsledku použijeme míru sdružení nazývanou *odds ratio* (poměr šancí). Odds ratio, značený OR, je definován jako poměr šancí mezi $x = 1$ a $x = 0$ a je dán rovnicí

$$OR = \frac{\pi(1)/[1 - \pi(1)]}{\pi(0)/[1 - \pi(0)]} = \frac{\frac{e^{\beta_0 + \beta_1}}{1 + e^{\beta_0 + \beta_1}} / \frac{1}{1 + e^{\beta_0 + \beta_1}}}{\frac{e^{\beta_0}}{1 + e^{\beta_0}} / \frac{1}{1 + e^{\beta_0}}} = \frac{e^{\beta_0 + \beta_1}}{e^{\beta_0}} = e^{(\beta_0 + \beta_1) - \beta_0} = e^{\beta_1}.$$

Pro logistické regrese s dichotomicky nezávislými proměnnými s hodnotami 1 a 0 je vztah mezi poměrem šancí (odds ratio) a regresními koeficienty roven $OR = e^{\beta_1}$. Odds ratio nám udává míru asociace, která nám odhaduje kolikrát je pravděpodobný/nepravděpodobný výskyt nějakého jevu y mezi nezávisle proměnnými x s hodnotou 1 než x s hodnotou 0. Tento vztah je základní důvod, proč je logistická regrese považována za mocný analytický nástroj výzkumu (Hosmer a Lemeshow 2000).

4.2 Softwarové řešení logistické regrese v prostředí RStudio

Predikce kumulativní NT byla vypočítána aplikováním logistického regresního modelu na závisle proměnnou Kumulativní NT a nezávisle proměnnou Rok, vybraných z poskytnutých dat. K výpočtu bylo použito grafické prostředí RStudio s možností využití Zobecněného lineárního modelu (GLM), ve kterém lze nastavit parametry pro výpočet logistické regrese. V podkapitole níže je tento model přiblížen. V příloze práce je možnost nahlédnutí na příkaz pro výpočet predikce, zadávaný do prostředí RStudio. Před samotným výpočtem predikce bylo nutné upravit data v proměnné Kumulativní NT, neboť obsahovala nelogické procentuální a nenumerné hodnoty výše NT, které by nemohly vstupovat do výpočtu. V porostních skupinách, kde došlo k úplnému vytěžení, by měla hodnota NT odpovídat 100%. V některých případech se však v datech objevily hodnoty převyšující hranici 100% nebo proměnná nabývala slovního výroku VYTĚŽENO. Z důvodu chybného zapsání výše NT nebo nenumerné hodnoty, která

nemohla vstupovat do daného výpočtu, byly údaje ručně opraveny a nahrazeny nejvyšší možnou hodnotou 100%.

4.2.1 Zobecnění lineární model (GLM)

Zobecnění lineární model, značen GLM (Generalized Linear Model) nebo GLIM, zahrnuje mnoho běžně používaných modelů. GLM zahrnuje lineární regresi, analýzu rozptylu, analýzu kovariance a zvláštní případ GLIM, do kterého spadají log-lineární modely pro analýzu kontingenčních tabulek, probit/logit regrese, Poissonova regrese a mnoho dalších (Olsson, 2002).

K logistickému regresnímu modelu se lze dostat z GLM, zápisem

$$g[E(Y|x)] = \mathbf{x}^T \beta,$$

ve kterém je funkce g , podmíněná střední hodnotou náhodné závislé proměnné Y , vyjádřené jako lineární funkce vektoru nezávisle proměnné $\mathbf{x}^T = (1, x_1, x_2, \dots, x_s)$ s regresními koeficienty $\beta^T = (\beta_0, \beta_1, \dots, \beta_s)$. K vytvoření logistického regresního modelu $\ln \pi(x)/1 - \pi(x) = \mathbf{x}^T \beta$ dojde, pokud má proměnná Y alternativní rozdělení, tedy $Y \sim A(\pi)$, které, jak je známo, má střední hodnotu $E(Y) = \pi$. A jako spojovací funkce se v GLM zvolí, již výše zmíněný, logit - logit (π) = $\ln \pi(x)/1 - \pi(x)$, který je vyjádřený jako lineární funkce regresorů (Tvrdlík, 2013).

4.3 Shluková analýza

Jak již bylo výše zmíněno, bakalářská práce byla rozšířena o shlukovou analýzu, která rozdělila porostní skupiny na předem daný počet shluků podle vzájemné podobnosti vlastností. Předem nadefinovaný počet shluků vypovídá o využití nehierarchické shlukové analýze. Nejdříve jsou uvedeny obecné informace o dané analýze a samotné řešení analýzy na datech s popisem využití funkce je shrnuto v kapitole 4.4. Daný příkaz pro shlukovou analýzu je součástí přílohy práce.

Analýza shluků (Cluster Analysis, CLU) je metoda zabývající se šetřením podobnosti objektů s velkým množstvím naměřených hodnot (vícerozměrné objekty), která vytváří jejich následnou klasifikaci do shluků neboli clusterů (Meloun a Militký 2004).

Metoda shlukové analýzy využívá jako základ pro tvorbu shluků *míru podobnosti* (resp. *nepodobnosti*) (Meloun a Militký 2004). Jejím cílem je najít v datech/oblastech

podmnožinu objektů, které vykazují míru podobnosti. Obecně máme určitou množinu značenou M obsahující m objektů, kde pro objekty $a, b \in M$ je dáno číslo neboli tzv. *numerická podobnost* $\sigma(a, b)$, která musí splňovat požadavky: $0 \leq \sigma(a, b) \leq 1$; $\sigma(a, a) = 1$; $\sigma(a, b) = \sigma(b, a)$ a $\min(\sigma(a, b), \sigma(b, c)) \leq \sigma(a, c)$. Tzv. *normovanou vzdáleností* objektů a, b je chápána charakteristika $(1 - \sigma(a, b))$. Shluková analýza se pokouší množinu M rozložit na $\{M_i\}_{i=1}^k$, platí $\cup_{i=1}^k M_i = M$, $M_i \cap M_j = \emptyset$ ($i \neq j$) tedy objekty náležící množině M_i jsou si více podobné než s objekty z množiny M_j .

Ve shlukové analýze se hledá rozdělení na „přirozené skupiny“, ke kterým lze dojít pomocí algoritmů, které automaticky provádějí hledání za pomoci počítačového prostředí (Hendl, 2004). Dle způsobu shlukování můžeme algoritmy rozdělit na *hierarchické* a *nehierarchické*.

4.3.1 Hierarchické shlukování

Analýza založena na hierarchickém uspořádání shluků a graficky vizualizovaná pomocí vývojového stromu nebo dendrogramu. Dělí se dle směru přístupu na *aglomerační* a *divizní shlukování*.

Aglomerační shlukování

Na začátku shlukování jsou jednoprvkové objekty, které tvoří nultý rozklad, kdy každý objekt vytváří jeden shluk. Poté se spojí dva jednoprvkové shluky, které jsou si nejpodobnější a vytvoří se nový shluk. Celý tento postup se opakuje, dokud všechny objekty neutvoří jediný shluk (is.muni.cz).

$$\{M_{ij}\}_{j=1}^k \quad (i_1 < i_2) \text{ platí } M_{i_1,j} \subseteq M_{i_2,j}$$

Objekty, které se na začátku spojí, zůstávají stále spojeny. Mezi nejznámější metody spojování shluků patří:

- **Metoda nejbližšího souseda** – minimální vzdálenost. Dva objekty nejbližší k sobě se spojí a vytvoří shluk, ke kterému se přidává třetí nejbližší objekt apod.
- **Metoda nejvzdálenějšího souseda** – maximální vzdálenost.
- **Metoda průměrné vzdálenosti** – průměrná vzdálenost všech objektů v jednom shluku vůči všem objektům ve shluku druhém.

Dalšími metodami patřícími do aglomeračního shlukování jsou *Wardova metoda*, *metoda těžiště*, *metoda mediánová* a jiné (Hendl, 2004).

Divizní shlukování

Opačný postup shlukování než u předešlého shlukování, kdy se vychází z jednoho shluku, tvořeného všemi objekty a jeho postupným rozdělováním dojde až k jednoprvkovým shlukům.

4.3.2 Nehierarchické shlukování

Oproti hierarchickému shlukování metoda nevytváří vývojový strom nebo dendrogram, ale rozděluje objekty do předem zadaných počtů shluků, které se však mohou během výpočtu měnit. Nehierarchické metody se dělí podle přerozdělování shluků na *optimalizační metodu*, která hledá optimální rozklad přesunováním jednotlivých objektů z jednoho shluku do druhého, za účelem minimalizovat/maximalizovat danou charakteristiku rozkladu a na *analýzu módu, neboli medoidů*, hledající rozklad do shluku, z hlediska chápání shluků jako oblasti se zvýšenou koncentrací objektů v m-rozměrném prostoru znaků (Meloun a Militký 2004). Mezi nejznámější metody nehierarchického shlukování patří:

- **Shlukování metodou nejbližších těžišť (K - Means)** – metoda založena na nejbližším těžišti, kdy je objekt přiřazen shluku s nejmenší vzdáleností mezi objektem a těžištěm shluku.
- **Shlukování metodou optimálních středů neboli medoidů** – medoid označuje optimální střed shluku, neboli střední objekt, který má minimální průměrnou vzdálenost k okolním objektům v daném shluku.
 - **Metoda PAM (Partition Around Medoid)** – metoda je zaměřená na minimalizaci celkové vzdálenosti D , kdy se v souboru k objektů hledá objekt s nejkratší vzdáleností představující střed. Poté se $k - 1$ objektů nachází podle nejmenší hodnoty D . Alternativa pozice k objektů je zjištěna iteračním způsobem. Dosud nezařazené objekty se přesouvají do určitého shluku, aby se hodnota D snižovala až do doby, kdy se D již nesnižuje.

Mezi další metody je zahrnuta např.: *Späthova metoda* a další.

4.4 Softwarové řešení shlukové analýzy v prostředí RStudio

Poskytnutá data se vyznačovala velkým počtem naměřených hodnot, které udávaly mnoho vlastností sledovaných porostních skupin v rozmezí let 1999-2008. V tomto kroku byla práce zaměřena na data za rok 2008, jako poslední rok zaznamenávající vlastnosti pro všechny porostní skupiny. Podrobnějším prostudováním významu jednotlivých proměnných bylo zjištěno, že některé proměnné vypovídají charakteristické vlastnosti stanoviště/terénu, kde se porostní skupiny nacházejí, porostu/stromů, které se tam vyskytují, činitelé způsobující NT a procentuální poškození porostních skupin. Proto byla pro potřebu výstižnější interpretace a možnosti zjištění, jak jsou jednotlivé porostní skupiny shodné nebo odlišné z pohledu určité skupiny vlastností popisující dané oblasti zájmu, využita shluková analýza, popsána výše, neboť se zdála být nejvhodnější metodou pro danou interpretaci. Z toho vyplývá, že data byla rozdělena na čtyři oblasti obsahující vlastnosti pro ně charakteristické (Tab. 3), na kterých byla jednotlivě provedena shluková analýza, která dále tyto oblasti rozdělila na shluky podle podobnosti. Stanovení počtu shluků bylo provedeno předem na 5, 7 a 10 shluků. V konečné fázi byl vybrán způsob rozdělení oblastí na 5 shluků, kdy vhodný počet shluků byl zvolen pomocí funkce cluster.Sim (viz. kapitola 4.4.1).

Tab. 3 Definování oblastí podle vlastností (zdroj:vlastní)

STANOVENÍ OBLASTÍ A VLASTNOSTÍ PRO SHLUKOVOU ANALÝZU	
OBLAST	VLASTNOST
stanoviště (terénu) porostu (stromů) NT	HOSP. U., TROF.U., ASPECT, SLOUPE, ELEVATION, ozáření, LVS zast.dr., zakm., přesný věk, vert.str. NT kůrovec, NT tracheomyk, NT živelná, NT ostatní
%pošk.porostu (%)	1.(0,08-0,25), 2.(0,26-0,36), 3.(0,37-0,44), 4.(0,45-0,54), 5.(0,55-0,75)

Pro upřesnění: poslední oblast, poškození porostu, nebyla rozdělena na shluky dle shlukové metody, jak tomu bylo u ostatních oblastí, ale ručně pomocí histogramu četností do určitého počtu intervalů hodnot podle zvoleného počtu shluků. Důvodem bylo, že do dané oblasti vstupuje pouze jedna proměnná/vlastnost a to výsledek bakalářské práce Kudělkové (2014), % poškození porostu kde užití shlukové metody na data s jednou proměnnou pozbývalo smysl.

Pro stanovení optimální shlukové metody a normované vzdálenosti z pohledu zvoleného počtu shluků, byla využita funkce cluser.Sim, jejíž popis je níže a ukázkou příkazu pro analýzu může čtenář nalézt v příloze práce.

4.4.1 Funkce cluster.Sim

Funkce cluster.Sim z knihovny clusterSim odhaduje jakou metodu shlukování a normovanou vzdálenost je z hlediska zvoleného počtu shluků nejlépe vhodné na oblasti využít. Základ je nastavení parametrů, jako je vstupní matice, typ dat vstupujících do analýzy, kde byl zvolen typ 7 - data smíšená, neboť poskytnutá data mají různé typy proměnných. Nastavení minimálního a maximálního počtu shluků, v mém případě 5 až 10 shluků. Index vnitřní kvality shluku, kde byl zvolen index Silhouette ("S"), kde $S \in [-1; 1]$, udává míru úspěšnosti rozdělení objektů do určitého počtu shluků prostřednictvím porovnání vzdáleností uvnitř shluku. Hodnota -1 značí špatnou klasifikaci a hodnota blízko 1 naopak klasifikaci dobrou. Zda je počet shluků nastaven správně, zjistíme pomocí takového počtu, který maximalizuje průměrnou hodnotu indexu značenou SC - hodnota 0,71-1 (silná struktura, úspěšné shlukování), 0,51-0,70 (přijatelná struktura), 0,26-0,50 (slabá, zkusit jinou metodu) a -1,00-0,25 (naprosto nevhodná struktura) (Meloun a Militký 2004). Podle výsledku hodnot indexu Silhouette (Obr. 4) provedeného pro každou z oblastí, bylo rozhodnuto dané oblasti zájmu rozdělit do 5 shluků.

RESULTS OF CLASSIFICATIONS

PATH = 7 (Interval or mixed (ratio & interval) data without normalization)

INDEX = Silhouette

NO. OF CLUSTERS = < 5 ; 10 >

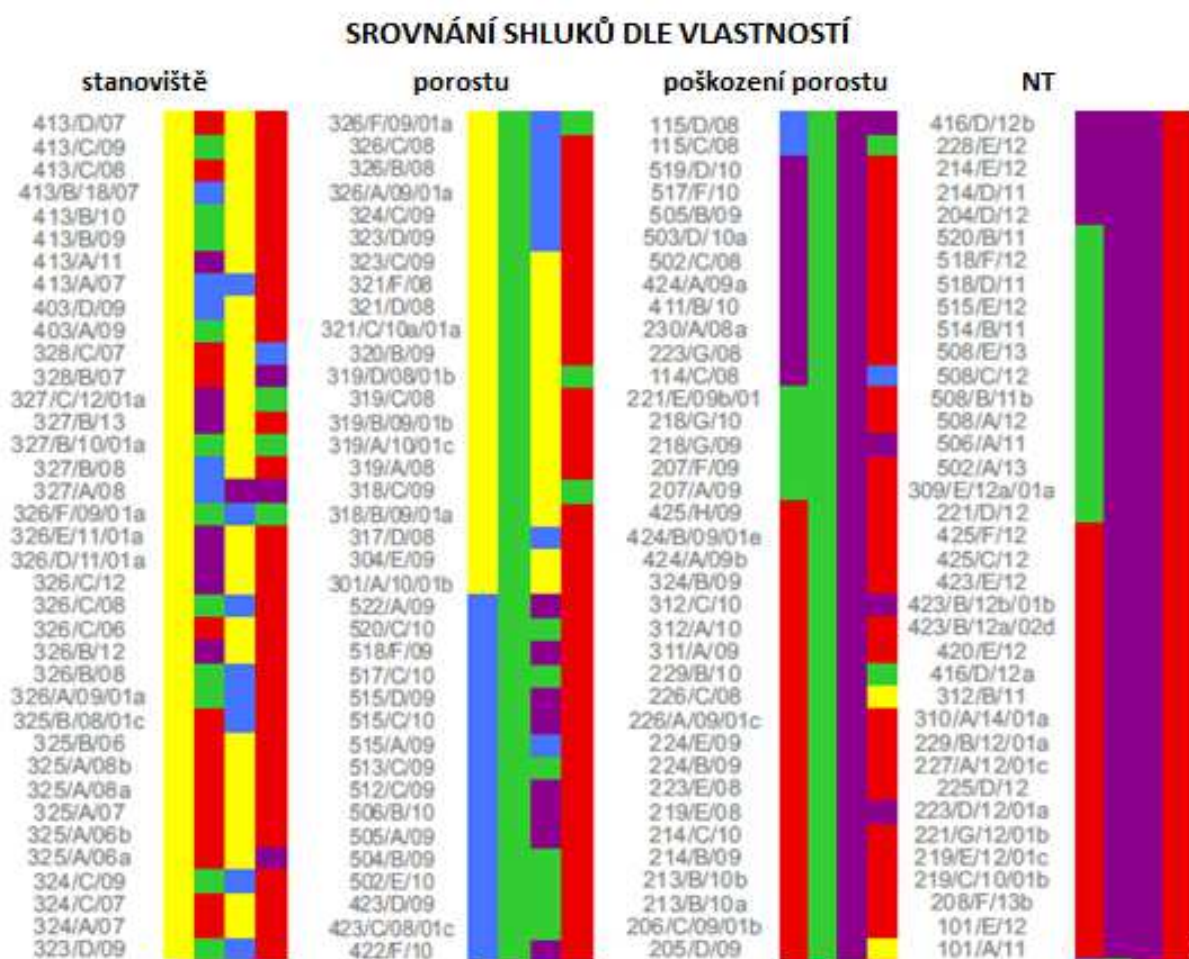
	Rank	No.	No. of clusters	Normalization formula	Distance measure	Clustering method	Silhouette
1	1	163	5	Without normalization	GDM1	pam	0.672680090404592
2	2	165	7	Without normalization	GDM1	pam	0.631532447380247
3	3	164	6	Without normalization	GDM1	pam	0.631284133921341
4	4	116	6	Without normalization	Squared Euclidean	pam	0.629706914257799

Obr. 4 Ukázka výsledků SC s výběrem míry vzdálenosti a metody shlukování pro 5 až 10 shluků
(zdroj: vlastní)

Vhodná míra vzdálenosti a metoda shlukování se vybírala prostřednictvím výsledné hodnoty SC, ze skupin metod a vzdáleností předem nadefinovaných určitým typem dat. V mém případě byla zvolena míra vzdáleností GDM1 a metoda shlukování PAM, která je výše popsána.

Pro vizualizaci a vyhodnocení výsledků shlukové analýzy byla vytvořena tzv. *heatmapa*, která představuje grafické znázornění porostních skupin čtyř oblastí rozdělených na 5 barevně odlišených shluků. V heatmapě řádky reprezentují jednotlivé porostní skupiny a sloupce jednotlivé oblasti zájmu (1. - stanoviště, 2. – porost, 3. – poškození a 4. – NT). Takto byly vytvořeny čtyři heatmapy (Obr. 5), kdy jednotlivě představovaly srovnání shluků podle vlastností stanoviště, porostu, poškození a NT.

Za účelem zjistit, jaké vlastnosti jsou typické pro vytvořené shluky oblastí stanoviště, porostu, poškození a NT, bylo zapotřebí zpětné základní charakteristiky proměnných v jednotlivých vytvořených shlucích.



Obr. 5 Ukázka částí jednotlivých heatmap (zdroj: vlastní)

5 VÝSLEDKY

Prvním dílčím výsledkem bakalářské práce byla základní charakteristika proměnných v datech pomocí popisné statistiky, za účelem bližšího seznámení se s rozsahem a strukturou poskytnutých dat. Tato část byla zaměřena na grafické porovnání vybraných proměnných porostních skupin dle pěti revírů za období deseti let, z hlediska výsledků základní statistiky. Mimo grafy s porovnání revírů byly dané výsledky zaznamenány i do přehledových tabulek, umístěných v příloze práce. Vztah mezi jednotlivými proměnnými byl zaznamenán v grafickém znázornění korelačního koeficientu, dle jeho těsnosti.

Hlavním výsledkem práce byl výpočet predikce kumulativní nahodilé těžby pro všechny sledované porostní skupiny do roku 2025. Zaznamenán byl jako tabulární, grafický výstup i jako krátké video, vystihující průběh kumulativní nahodilé těžby od roku 1999 do roku 2025.

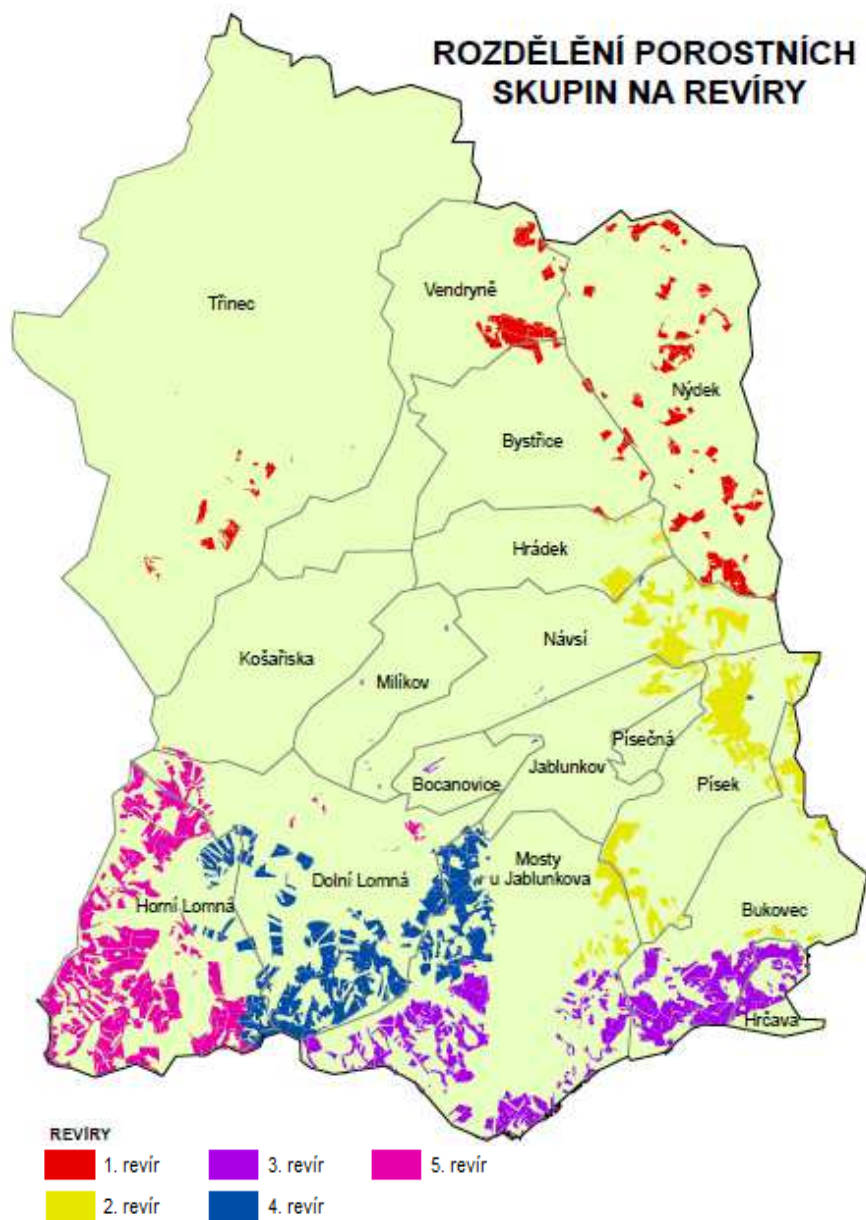
Bakalářská práce byla rozšířena o výsledek rozdělení porostních skupin na oblasti/shluky, které se vyznačovaly podobnými vlastnostmi z hlediska stanoviště, na kterém se nacházely, porostu, který se v dané porostní skupině vyskytoval, procentuálního poškození porostu a činitelů způsobující nahodilou těžbu. Výsledky byly znázorněny pomocí heatmap, podle kterých lze srovnávat vzniklé shluky podle jednotlivých vlastností stanoviště, porostu, poškození a nahodilé těžby.

Výsledkem práce je i poster, který obsahuje mapový výstup predikce kumulativní nahodilé těžby pro rok 2020 i s grafickou vizualizací a dalšími informacemi, zároveň obohacený o znázornění výsledku rozdělení sledovaných porostních skupin na shluky pomocí heatmap.

5.1 Popisná statistika dat

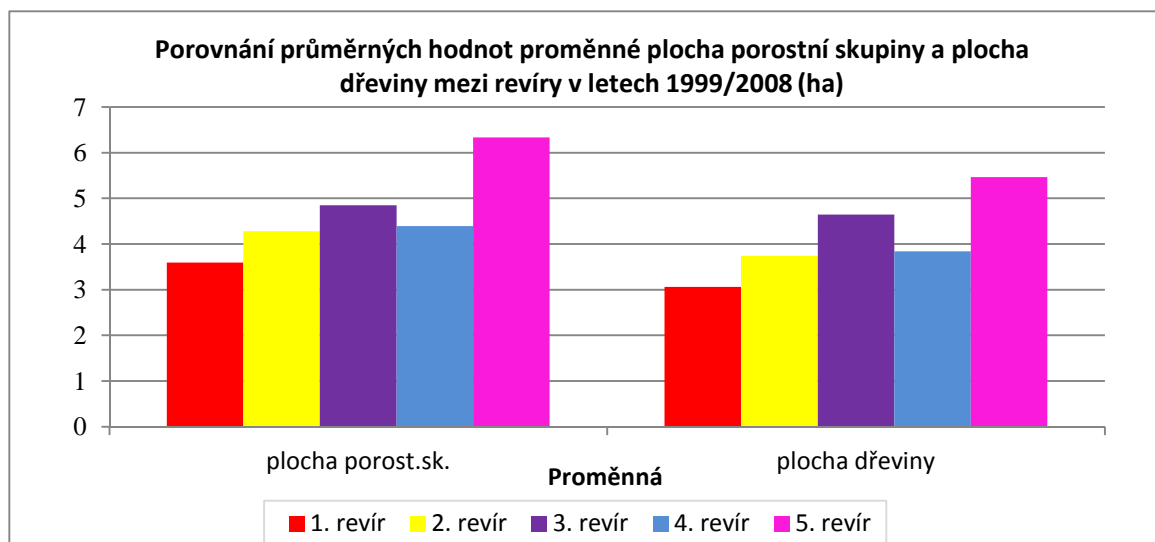
Prostřednictvím popisné statistiky bylo možné porovnat všech pět revírů, do kterých náležely sledované porostní skupiny. Obrázek 6 zachycuje rozdělení a polohu jednotlivých porostních skupin dle pěti revírů. Výpočty byly zaměřeny především na střední hodnoty a rozptýlenost dat s ohledem na různorodost proměnných v datech. Pro data spojitá/metrická byl vypočítán aritmetický průměr, medián, rozptyl a směrodatná odchylka a pro data nominální a ordinální absolutní četnost. Z hodnot, kterých nabývaly, byla možnost srovnání jednotlivých revírů dle dané proměnné. Zhodnocení a grafická vizualizace všech výsledků statistických výpočtů prováděných na datech, by představovali značně rozsáhlé rozpracování, proto bylo grafické porovnání revírů podle

dané proměnné vytvořeno pouze pro aritmetický průměr a četnosti. Ostatní výpočty byly shrnuty do přehledových tabulek. Některé vizualizace výsledků jsou i s popisem uvedeny v následující části, ostatní lze nalézt v příloze této práce.

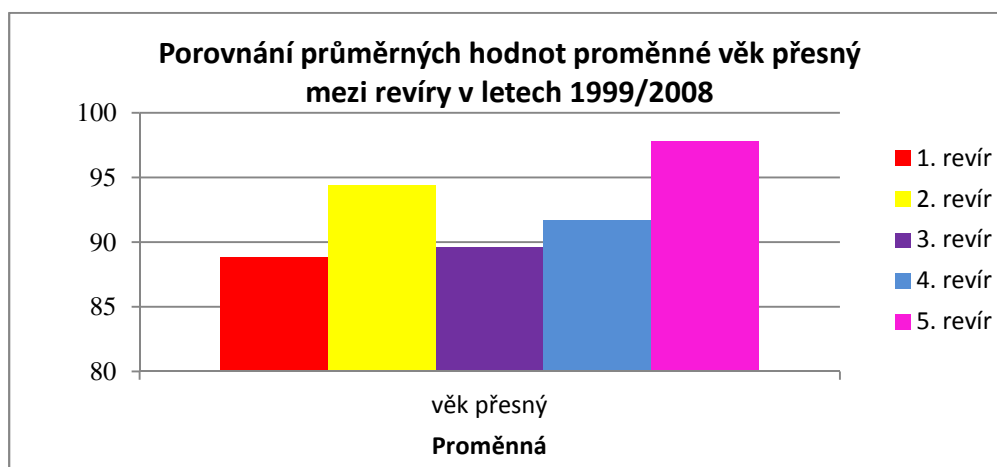


Obr. 6 Ukázka rozdělení porostních skupin do pěti revírů (zdroj: vlastní)

Proměnné v datech vypovídaly velké množství informací o porostních skupinách a revírech, ke kterým se vztahovaly. Plocha, věk, hustota porostních skupin aj. mají velkou váhu, co se týče poškození dřeva náhodnými činiteli. Podle Grafů 1 a 2 nejrozsáhlejší a nejstarší porostní skupiny náleží do pátého revíru a naopak první revír vykazuje dvojnásobně menší rozlohu s nižším věkovým rozhraním porostu.



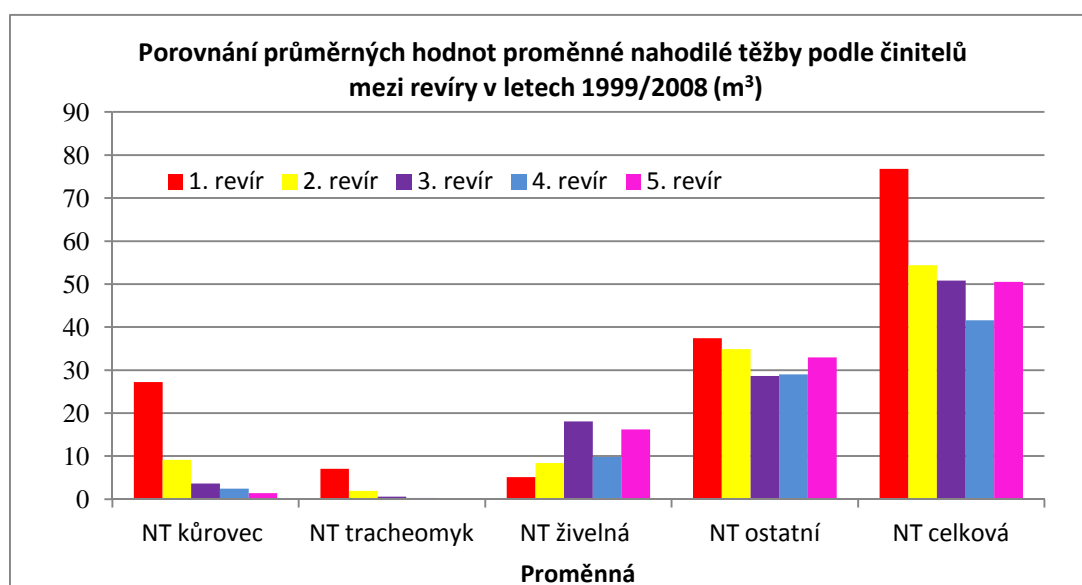
Graf 1 Porovnání průměrného stáří mezi revíry v letech 1999/2008



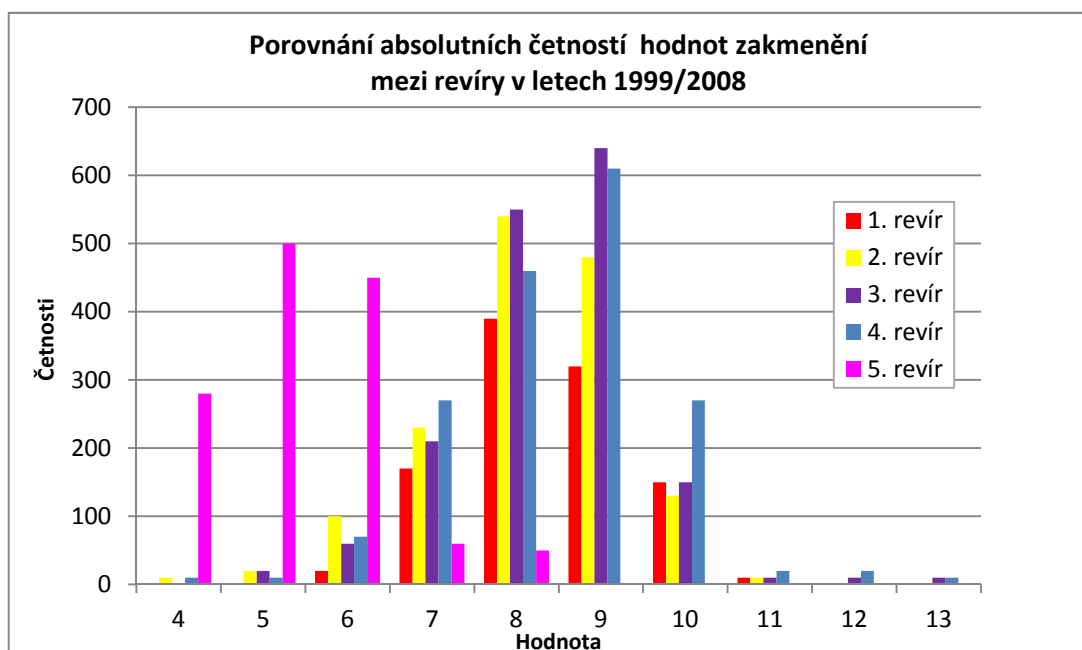
Graf 2 Porovnání ploch porostních skupin a dřevin mezi revíry v letech 1999/2008

Na Grafu 3 jsou průměrné hodnoty nahodilé těžby podle činitelů, kteří ji způsobili během deseti let. Z grafické vizualizace je zřejmé, že pro první revír, ačkoliv je nejmladší ze všech a měl by být i nejsilnější ze všech, co se týče napadení činitelů způsobující nahodilou těžbu, je nejvíce zasahován činností všech zmíněných činitelů, s výjimkou NT živelné. Tento revír je rovněž charakteristický poměrně vysokou hustotou porostu, kde zakmenění sahá k největším četnostem hodnot od 6 do 11 (Graf 4). Proto pokud se zde

vyskytne napadení např.: kůrovce, je velmi pravděpodobné, že rozsah škod bude vyšší než u méně hustých porostů, kde k rozšíření kůrovce ze stromu na strom stěžuje řídký porost. Toto platí pro podstatně méně hustý pátý revír, který oproti prvnímu revíru byl minimálně zasahován všemi činiteli, ovšem s výjimkou NT živelné. Ta oproti biotickým činitelům nejvíce zasahuje porosty málo husté, které nevytvářejí pevný štít proti náporům abiotických činitelů.



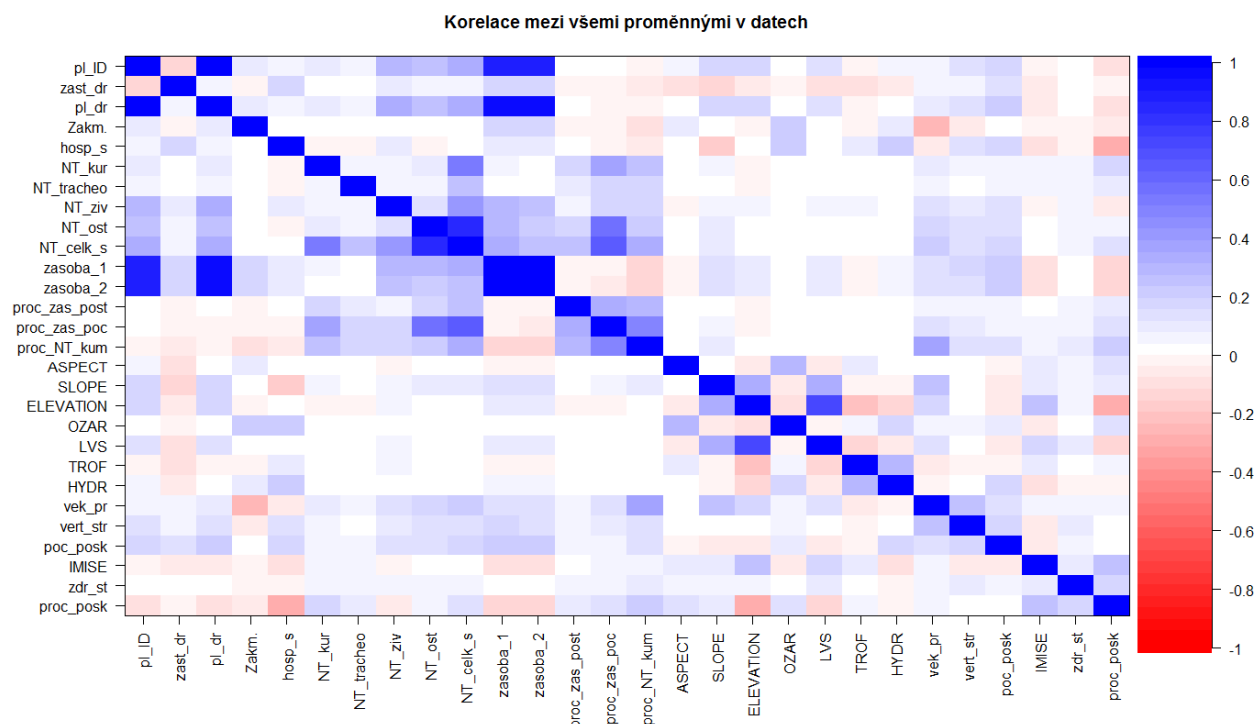
Graf 1 Porovnání průměrných hodnot NT podle činitelů mezi revíry v letech 1999/2008



Graf 2 Porovnání absolutních četností hodnot zakmenění mezi revíry v letech 1999/2008

Je mnoho dalších vlastností, které zde nebyly z důvodu rozsáhlosti vysvětlení jednotlivých proměnných uvedeny a podle kterých lze definovat, jak jsou dané oblasti charakteristické z pohledu zranitelnosti/obranyschopnosti, orientace, nadmořské výšky, aj. Možností vizualizace a tabulky nalézt v příloze k bakalářské práci.

K určení míry vztahů/ovlivnění proměnných mezi sebou v rámci celé sady dat bylo vytvořeno schéma, které barevně odlišuje hodnotu korelačního koeficientu, podle kterého je snadné odhadnout proměnné navzájem se ovlivňující a naopak proměnné nenabývající žádné vztahy. Ze schématu korelačního koeficientu je zřejmé, jak plocha porostních skupin ovlivňuje výši NT způsobenou jakýmkoli činitelem. Stejně tomu je u přesného věku, hodnot imisí a počátečním poškození, které mají také zcela přímou korelaci mezi činiteli NT. Vyskytují se zde i proměnné vykazující antikorelaci neboli zcela žádnou vazbu na ostatní proměnné (Obr. 7).



Obr. 7 Ukázka barevně odlišeného schématu hodnot korelačního koeficientu

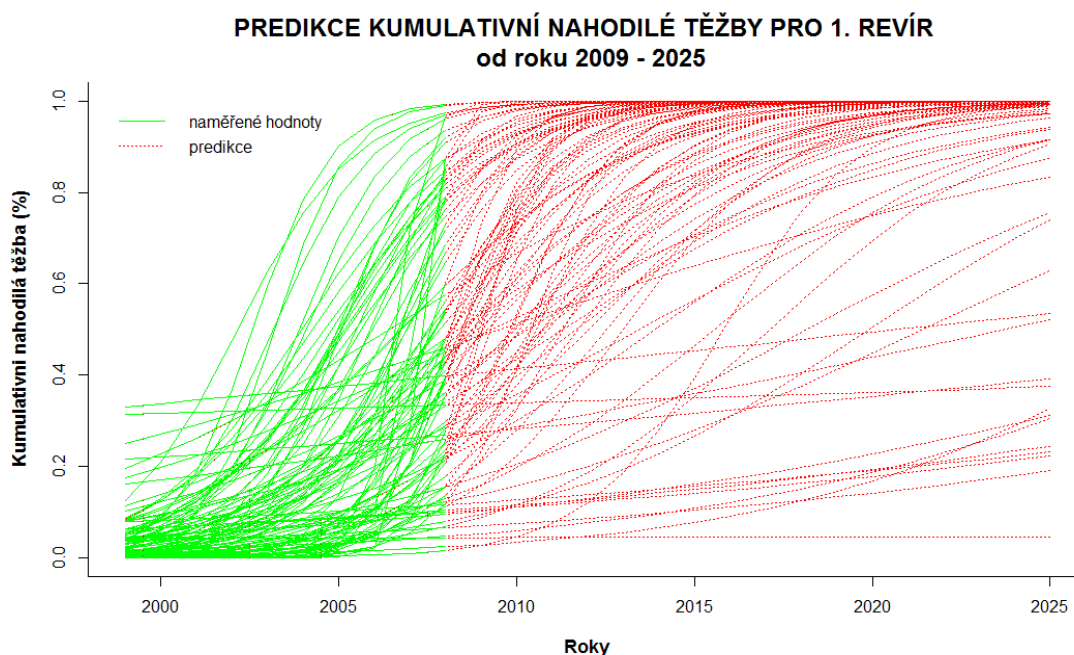
5.2 Predikce kumulativní nahodilé těžby

Hlavním výsledkem bakalářské práce byl výpočet predikce kumulativní nahodilé těžby pro sledované porostní skupiny do roku 2025. Jako hlavní proměnná, podle které byla predikce vypočítána, byla kumulativní nahodilá těžba, která vycházela z načítaných hodnot procentuální nahodilé těžby z počátečních zásob, stanovené LHP. Pomocí logistické regrese, jejíž křivka nejlépe proložila data, byl vypočítán daný predikt kumulativní NT od roku 2009 až do roku 2025.

Před provedením výpočtu bylo předpokládáno, že průběh křivky zachycující NT bude mít esovitý tvar. Toto tvrzení vyšlo z předpokladu, že na začátku máme les, respektive porost, zdravý a silný, jen s běžnými ztrátami – prořezování porostu, kácení z důvodu zásahu do silnic, elektrického vedení apod. – začátek křivky. Následně je lesní porost vystaven stresu způsobeného činiteli - kůrovcem, vichřicí, mrazem, kteří ničí a požírají les i po delší dobu a z toho důvodu je pro zamezení dalšího rozšíření náhlý vzestup nutné NT – vzestup křivky. Po nepříznivém zásahu činitelů způsobující NT dojde k úplnému zničení lesního porostu – dokončení křivky.

Výpočet predikce byl zaznamenán v tabulární podobě, kde je možné zjistit přesné hodnoty kumulativní NT pro všech 736 porostních skupin v rozmezí od roku 1999-2008, hodnoty již naměřeny a od roku 2009-2025 hodnoty představující predikci NT. Náhled části tabulky s hodnotami predikce NT je možný v příloze, celá tabulka je uložena na DVD. Pro srovnání jednotlivých porostních skupin z pohledu průběhu NT v daném časovém rozmezí byly vytvořeny grafy, jak pro porostní skupiny v jednotlivých revírech, tak i pro všechny revíry dohromady. Z grafického znázornění výpočtu predikce, kde na ose x jsou jednotlivé roky od roku 1999 do roku 2025 v intervalu po pěti letech a na ose y jsou v omezeném rozsahu od 0,1-1 zaznamenány procentuální hodnoty dané kumulativní NT porostních skupin, lze zaznamenat, že křivka průběhu NT má opravdu esovitý tvar, jak tomu bylo vysvětleno výše. Pro ukázkou jednoho z výsledků, Obrázek 8 zobrazuje průběh NT pro porostní skupiny 1. revíru, jak pro již naměřené hodnoty od roku 1999-2008 (zeleně značené), tak i pro predikované hodnoty od roku 2009-2025 (červeně značené). Z grafu průběhu NT pro všechny revíry dohromady (viz. příloha) bylo možné, i přes velké množství křivek, vyčíst, že většina porostních skupin dosáhne v roce 2025 hranice 100% NT, tedy úplnému vykácení lesa. Mnohé z porostních skupin dosáhne této hranice ještě před rokem 2025. Z hlediska vizualizace hodnot, ze kterých byla predikce vypočítána, lze zjistit, že některé porostní skupiny nabývaly hranice 100% již

před rokem 2008 a některé měly k této hranici velmi blízko. Vyskytují se zde i takové porostní skupiny, jejichž výše NT nedosáhne ani na 50% hranici, ovšem jejich zastoupení je málo početné. Z výsledků lze konstatovat, že je zapotřebí pro danou oblast vytvořit taková opatření, aby se zamezilo markantnímu zvýšení NT, které by znamenalo ztrátu velkého množství lesního porostu. Pro přehlednější hodnocení výsledků je v příloze bakalářské práce možné nahlédnout na grafickou vizualizaci predikce NT pro jednotlivé revíry zvlášť i pro všechny revíry dohromady a v příloženém DVD lze shlédnout krátké video zaznamenávající průběh NT během let 1999-2025.



Obr. 8 Graf predikce kumulativní nahodilé těžby pro 1. revír

5.3 Rozdělení porostních skupin podle podobnosti

Výsledkem shlukové analýzy bylo rozdělení porostních skupin na pět podoblastí/shluků, vyznačující se určitou podobností, z pohledu vlastností čtyř oblastí - stanoviště, porostu, poškození porostu a NT.

Z výpočtů základní statistické charakteristiky vlastností každého ze shluků dané oblasti, kde pro vlastnosti spojité byl vypočten aritmetický průměr, medián, směrodatná odchylka a pro vlastnosti ordinální a nominální absolutní četnosti, bylo možné zjistit, jak jsou jednotlivé shluky porostních skupin charakteristické.

5.3.1 Charakteristika shluků dle vlastností stanoviště

Oblast vlastností stanoviště se vyznačuje proměnnými: hospodářský soubor, trofická úroveň, aspekt, sloupe, elevation, ozáření a LVS. Z Tabulky 4 a 5 jsou vidět vypočítané průměry, mediány, směrodatné odchylky a jednotlivé četnosti vlastností stanoviště pro všech pět shluků. Z absolutní četnosti LVS je patrné, že skoro všechny shluky náležejí do oblasti jedlovo – bukového porostu, s výjimkou prvního shluku, který má nejvyšší četnost u porostu bukového. Rovněž lze konstatovat velkou podobnost dle trofické úrovně, kde pro všechny shluky, vyjímaje třetího shluku, byla nejvyšší četnost u hodnoty 8, tedy středně dobré podmínky obsahu živin ve vodě. Porovnání vhodnosti smrkové produkce z hlediska ozáření má druhý a čtvrtý shluk o něco horší, kdy je nejvyšší četnost u hodnoty 3, nežli u ostatních shluků s hodnotou 1, která vykazuje oblast velmi vhodnou pro produkci smrkového porostu. Zbylé vlastnosti nabývají rozdílných hodnot mezi shluky a jsou popsány níže v pořadí podle hospodářského souboru, orientace, sklonu svahu a nadmořské výšky.

První vytvoření shluk je podle nejvyšší četnosti hospodářského souboru 451 oblastí tzv. smrkového hospodářství na živných stanovištích středních poloh, kde se jedná o lesy hospodářské s průměrnou orientací svahu 204° , tedy jihozápadní, s průměrným sklonem svahu 14° a průměrnou nadmořskou výškou oblasti 547 m n. m.

Druhý shluk je rovněž oblastí smrkového hospodářství, ale oproti předešlému shluku jde o exponovaná stanoviště vyšších poloh s hospodářskými lesy a navíc ještě s ochrannými pásmy léčivých vod. Oblast má strmější průměrný sklon svahu 19° s jihovýchodní orientací a průměrnou nadmořskou výškou 679 m n.m.

Dle hospodářského souboru je třetí shluk velmi podobný druhému shluku s rozdílem, že tyto hospodářské lesy spadají do tzv. lesů v PHO 1, neboli pásem hygienické ochrany s přímou ochranou zdrojů (50-100m) (Pokorný, 2014). Vyznačuje se také nejstrmějšími svahy ze všech shluků, které jsou až 22° , jižně orientované a průměrnou nadmořskou výškou 693 m n.m.

Čtvrtý shluk je stejně jako první shluk oblastí hospodářských lesů zaměřené na smrkové hospodářství živných stanovišť, ale již ve vyšších polohách, z důvodu nadmořské výšky, která je průměrně až 762 m n.m. Svahy mají sklon průměrně 16° a jsou orientovány jihovýchodně.

Poslední shluk je nejodlišnější od ostatních, neboť se jedná o oblast lesů zvláštního určení, potřebných pro zachování biologické rozmanitosti neboli genové základny. Oblast je dále dle hospodářského souboru zaměřena na smrkové hospodářství živných stanovišť vyšších poloh, kde nadmořská výška je průměrně 645 m n.m. a sklony svahu jsou zde nejmírnější ze všech, přibližně 12°, orientovány jižně stejně jako u třetího shluku.

Tab. 4 Základní statistická charakteristika vlastností stanoviště dle jednotlivých shluků

Základní statistická charakteristika vlastností stanoviště dle shluků									
SHLUK	Aritmetický průměr			Medián			Směrodatná odchylka		
	Aspect (°)	Slope (°)	Elevation (m)	Aspect (°)	Slope (°)	Elevation (m)	Aspect (°)	Slope (°)	Elevation (m)
1.	204,09	13,59	547,13	214	13	547	106,30	5,72	57,85
2.	166,66	19,07	679,00	167	19	692	100,33	6,51	89,22
3.	187,98	22,11	692,69	215	22	687	121,98	6,00	107,80
4.	148,56	16,19	762,12	124	16	743	108,04	5,13	95,01
5.	174,09	11,98	645,91	161	12	606	76,57	5,20	132,48

Tab. 5 Absolutní četnosti vlastností stanoviště pro jednotlivé shluky

Absolutní četnosti vlastností stanoviště dle shluků						Absolutní četnosti vlastností stanoviště dle shluků						
	1. shluk	2. shluk	3. shluk	4. shluk	5. shluk		1. shluk	2. shluk	3. shluk	4. shluk	5. shluk	
HOSPODÁŘSKÝ SOUBOR	451	129	x	x	x	x	5	4	2	x	2	x
	471	3	x	x	x	x	7	x	x	2	x	1
	501	5	x	x	7	x	8	89	27	3	145	52
	516	1	x	x	1	x	10	6	21	84	x	5
	531	4	x	x	2	x	11	30	x	x	11	44
	551	87	x	x	200	x	12	77	18	6	53	1
	556	2	x	x	3	x	13	15	x	x	1	10
	1511	x	x	99	x	x	14	3	2	x	1	1
	2511	x	72	x	x	x	15	5	1	4	x	2
	3506	x	1	x	x	x	17	1	1	x	x	4
4501	x	x	x	x	1	20	1	x	x	x	x	
8501	x	x	x	x	7	21	x	1	x	x	x	
8541	x	x	x	x	112	1	61	20	34	80	11	
LVS	3	x	1	x	x	x	2	61	14	27	48	21
	4	133	6	8	x	x	3	56	21	25	49	49
	5	98	66	91	189	24	4	53	18	13	36	39
	6	x	x	x	24	10						

5.3.2 Charakteristika shluků dle vlastností porostu

Oblast vlastností porostu se vyznačuje proměnnými: zastoupení dřevin, zakmenění, přesný věk a vertikální struktura.

Podle Tabulky 6 lze souhrnně zhodnotit, že průměrná hodnota zastoupení jedné dřeviny je v prvních třech shlucích podobné a pohybuje se kolem 96%, u čtvrtého a pátého shluku je přibližně kolem 70%. Z četností hodnot zakmenění, které určují hustotu daného porostu je zřejmé, že druhý, třetí a pátý shluk je, co do hustoty porostu, řidší, oproti prvnímu a čtvrtému shluku. Vertikální struktura jednotlivých shluků je podle četnosti stejná pro všechny a určuje, že oblast se nevyznačuje strmými svahy, ale spíše rovinným povrchem s mírným sklonem svahu (Tab. 7). Nejrozdílnější hodnot nabývá přesný věk jednotlivých shluků. Podle průměrného věku je třetí (126 let) a pátý shluk (118 let) nejstarší, nejmladším je první shluk (73 let), druhý a čtvrtý shluk se pohybuje v rozmezí 80-90 let.

Tab. 6 Základní statistická charakteristika vlastností porostu dle jednotlivých shluků

Základní statistická charakteristika vlastností porostu dle shluků						
SHLUK	Aritmetický průměr		Medián		Směrodatná odchylka	
	zastoupení dřevin (%)	věk přesný	zastoupení dřevin (%)	věk přesný	zastoupení dřevin (%)	věk přesný
1.	95,30	72,71	97	73	5,06	7,44
2.	96,38	94,53	98	93	4,39	6,13
3.	96,05	126,07	98	124	4,67	12,52
4.	68,24	80,28	70	81	10,02	8,79
5.	71,89	117,75	75	115	9,89	11,32

Tab. 7 Absolutní četností vlastností porostu dle shluků

Absolutní četností vlastností porostu dle shluků						
		1. shluk	2. shluk	3. shluk	4. shluk	5. shluk
ZAKMENĚNÍ	4	x	x	x	1	1
	5	1	x	4	x	x
	6	2	3	15	2	6
	7	24	25	47	10	10
	8	58	58	77	26	25
	9	108	46	44	38	14
	10	35	10	15	10	6
	11	2	2	3	x	3
	12	x	x	2	1	x
	13	x	x	2	x	x
VERTIKÁLNÍ STRUKTURA	0	2	x	1	x	x
	1	215	111	129	78	47
	2	13	33	78	10	18
	11	x	x	1	x	x

5.3.3 Charakteristika shluků dle vlastností NT

Oblast vlastností NT se vyznačuje proměnnými způsobující NT: NT kůrovcem, NT houbou tracheomyk, NT živelná a NT ostatní.

Podle základní statistiky (Tab. 8) lze zjistit, že nejmenší NT způsobenou činiteli výše zmíněnými je vystaven první shluk, který má průměrnou výši poškození všemi činiteli v rozmezí od 1-3 m³, kde nejvyšší průměrná hodnota je 2,58 m³ u NT živelné.

Druhý shluk má oproti předešlému shluku třikrát vyšší poškození kůrovcem, průměrná hodnota poškození houbou tracheomyk je 17,7 m³, živelné pohromy průměrně způsobí NT ve výši 5,6 m³ a ostatní činitelé zvyšující NT ve výši 109 m³.

Třetí shluk má druhou nejvyšší průměrnou hodnotu navýšení NT způsobené kůrovcem a to 15 m³. Zvýšení NT houbou tracheomyk zastupuje nejvyšší průměrnou hranici z činitelů způsobující NT ve výši 235 m³, živly mají míru poškození 26 m³ a ostatní činitelé 35 m³.

Čtvrtý shluk vyniká především v průměrné hodnotě kůrovcem způsobující NT, kdy jeho hodnota dosahuje 118 m³. Hodnota za houbu tracheomyk – 16 m³ je třetí nejvyšší z této proměnné a za NT způsobenou živelnými pohromami je průměrná hodnota 9,2 m³. Nejméně je sužován ostatními činiteli s průměrnou hodnotou 8 m³, které jsou zároveň druhou nejmenší v proměnné ostatní činitelé NT.

Poslední shluk se vyznačuje poměrně nízkými průměrnými hodnotami za všechny činitele, kteří byly zahrnuti do analýzy. Nejméně poškození způsobuje kůrovec, přibližně 8,4 m³, druhou nejnižší je poškození způsobené houbou tracheomyk a to 9,8 m³, nejvyšší hodnotu zastupují ve výši 84 m³ živelné pohromy a pro poškození způsobené ostatními činiteli byla vypočtena průměrná hodnota 16,2 m³.

Tab. 8 Základní statistická charakteristika vlastností způsobujících NT dle jednotlivých shluků (1. část)

Základní statistická charakteristika vlastností způsobujících NT dle shluků								
SHLUK	Aritmetický průměr (m ³)				Medián (m ³)			
	kůrovec	tracheomyk	NT živelná	ostatní	kůrovec	tracheomyk	NT živelná	ostatní
1.	1,06	1,97	2,58	2,09	0	0	0	0
2.	3,30	17,65	5,56	108,70	0	0	0	64
3.	14,85	235,03	26,04	34,90	0	169	5	0
4.	118,27	15,71	9,20	7,87	88	0	0	0
5.	8,41	9,75	83,94	16,15	0	0	65	8

Tab. 8 Základní statistická charakteristika vlastností způsobujících NT dle jednotlivých shluků (2. část)

Směrodatná odchylka (m ³)			
kůrovec	tracheomyk	NT živelná	ostatní
4,53	7,70	5,71	5,69
9,02	42,08	9,13	89,81
61,18	196,29	52,05	130,92
102,07	39,53	15,55	19,29
13,14	19,93	51,64	32,50

5.3.4 Charakteristika shluků dle vlastností poškození porostu

Oblast vlastností poškození porostu je odlišný případ od stávajících oblastí tím, že neobsahuje více proměnných, ale pouze jednu a to procentuální poškození porostu, jehož dopočtem se zabývala ve své bakalářské práci Kudělková (2014), neboť ve stávajícím souboru byl odhad procentuálního poškození určen jen pro určité porostní skupiny. Z dopočítaných hodnot procentuálního poškození porostu byly pomocí histogramu vytvořeny intervaly pro pět shluků (Tab. 9).

Tab. 9 Rozdělení hodnot poškození porostu do intervalů

Intervaly procentuálního poškození dle shluků	
SHLUK	intervaly
1.	0,000 - 0,25
2.	0,26 - 0,36
3.	0,37 - 0,44
4.	0,45 - 0,54
5.	0,55 - 0,75

5.3.5 Vyhodnocení pomocí heatmap

Pomocí heatmap je možné mezi sebou porovnávat vzniklé shluky náležící do dané oblasti. Nejdříve byly porostní skupiny srovnány podle vlastností stanoviště, kdy první sloupec v heatmapě, zastupující vlastnosti stanoviště, byl seřazen podle pěti vytvořených po sobě následujících shluků. Z tohoto uspořádání bylo možné zjistit, zda bylo rozdělení porostních skupin na shluky v jednotlivých oblastech podobné nebo zcela odlišné. Tedy jestli některé porostní skupiny splňovaly podobnost vlastností stanoviště, porostu, poškození porostu i NT dohromady pro stejný shluk. Následně byl celý postup zopakován a porostní skupiny byly seřazeny podle zbylých vlastností porostu, poškození porostu a NT.

V heatmapě srovnání dle vlastností stanoviště bylo zjištěno, že rozdělení porostních skupin do 1. shluku má velkou podobnost v rozdělení porostních skupin do 1. shluku v oblasti vlastností NT. Velké shody nastávají i v případě 5. shluku, kde je velké množství stejného rozdělení pro oblast vlastností poškození. V ostatních shlucích v oblasti vlastností stanoviště nejsou tak značné shody, jak v předešlé situaci.

Vizualizace srovnání shluků dle vlastností porostu ukazuje případ, kdy velká shoda v rozdělení do 1. shluku je zaznamenána v oblastech vlastností stanoviště a těžby. V případě rozdělení do 3. shluku jsou v malé míře shodné shluky pro oblast vlastností stanoviště a vlastností poškození. Pro 4. shluk je několik porostních skupin zařazen stejně i v oblasti vlastností stanoviště, ale již ne tak markantně, jak tomu bylo u shluku prvního. 5. shluk vykazuje podobnost rozdělení na shluk v malém počtu porostních skupin v oblasti vlastností stanoviště.

V případě srovnání dle vlastností poškození lze sledovat shoda u rozdělení na 1. shluk s oblastí vlastností těžby. Stejně zastoupení porostních skupin je sledováno s oblastí vlastností porostu a to pro 2. shluk. Podobnost zde vykazuje o mnoho nižší zastoupení porostních skupin oproti rozdělení na 3. shluk, kde je výše shodných rozdělení větší a je částečně sledována i pro oblast vlastností stanoviště. Jak již bylo výše zmíněno, v situaci srovnání podle vlastností stanoviště, lze sledovat mezi těmito oblastmi velké shody v rozdělení do 5. shluku.

Srovnání dle vlastností těžby (NT) určilo rozsáhlé zastoupení většiny porostních skupin do 1. shluku. Z toho vyplývá, že shody s ostatními oblastmi nastanou právě pro rozdělení na 1. shluk, mezi které patří všechny oblasti vlastností, ale nejvíce podobného rozdělení měly především oblasti vlastností porostu a následně vlastností stanoviště.

Pro rozsáhlost datové sady a s tím spojené vizualizace všech čtyř heatmap, jsou jednotlivé náhledy na dané výstupy umožněny na webových stránkách a posteru práce.

6 DISKUZE

Součástí práce bylo nastudování si problematiky nahodilé těžby, příčiny vzniku a statistické metody využití v minulosti k jejímu modelování. Zdrojem informací byly, jak domácí, tak i zahraniční publikace.

Nezbytným krokem bylo seznámení se s veškerými proměnnými v poskytnutých datech, prostřednictvím konzultace s odborníkem, neboť autorovi znalosti v problematice lesnictví nebyly dostačující.

K vypracování cíle práce byla zvolena proměnná kumulativní nahodilá těžba, kde bylo s překvapením zjištěno, že některé údaje v proměnné se rozcházejí s logickým uvažováním, kdy proměnná nabývala hodnot nad uvedenou hranici. Tyto hodnoty byly autorem opraveny, neboť při jejich zpracování by byl výsledek značně ovlivněn a nenabýval by správných hodnot. Během práce byl také zjištěn fakt, že by bylo zapotřebí do výpočtů zahrnout zvyšování/obměnu ploch lesů během let, způsobený postupným zalesňováním. Tento problém by mohl být řešen analýzou přežívání, ale z důvodů krátkého časového rozpětí na vypracování a nedostatku potřebných dat, nebyla tato metoda ve finále použita a pracovalo se pouze na základě poskytnutých dat. K zlepšení výsledků predikce nahodilé těžby, by bylo vhodné rozšířit danou analýzu o stanovení statisticky významných vlastností, zejména s pohledu porostu a stanoviště, které korelují se stoupající pravděpodobností nahodilých těžeb. Podobné výsledky podává použitá shluková analýza, ovšem aplikování vícenásobné regrese by zajišťovalo přesnější korelaci jednotlivých vlastností porostu a stanoviště na výši nahodilé těžby.

Velice přínosné může být porovnání výsledků shlukové analýzy, se zaměřením na výsledky dle vlastností NT a poškození porostu a hodnot výsledku predikce NT. Aby bylo možné zhodnotit správné/chybné zařazení porostních skupin s predikovanou zvyšující se NT do oblastí dle vlastností poškození porostu a vlastností NT.

7 ZÁVĚR

Úkolem bakalářské práce bylo zpracování poskytnutých dat, zaměřených na desetileté sledování porostních skupin v zájmové oblasti nejvýchodnějšího cípu České republiky. Cílem byl výpočet predikce kumulativní nahodilé těžby do roku 2025. Práce byla rozšířena o rozdělení jednotlivých porostních skupin obsažených v datech na skupiny, které se vyznačují vzájemnou podobností.

První část bakalářské práce se zabývala popisem pojmu nahodilá těžba, činitelé ji vyvolávající a sepsáním krátké rešerše na již využitě matematicko-statistických metody pro modelování nahodilých jevů v lesnictví.

Následující část byla zaměřena na základní informace o zájmové oblasti a datech s podrobným popsáním jednotlivých proměnných. Součástí byl popis základní statistiky aplikované na datech, pro zjištění charakteristiky dat z hlediska střední hodnoty, rozptýlenosti, četnosti a vzájemných vztahů mezi proměnnými.

Praktickou částí byl výpočet predikce nahodilé těžby, se kterou souvisí přiblížení použitých metod, zejména pak logistické regrese, na jejímž základě byla provedena modelace pravděpodobnosti zvýšení nahodilé těžby v jednotlivých letech do roku 2025. Tento výpočet se zaměřuje na vztah mezi nezávisle proměnnou roky a závisle proměnnou udávající hodnotu kumulativní nahodilé těžby omezenou intervalem 0,1;1. Použita byla shluková analýza k rozdělení jednotlivých porostních skupin na 5 shluků, které se vyznačují vzájemnou podobností dle vlastností stanoviště, porostu, poškození a nahodilé těžby.

Prostřednictvím výpočtu predikce nahodilé těžby bylo zjištěno, že většina porostních skupin dosáhne v roce 2025 hranici 100% nahodilé těžby, tedy úplnému vykácení lesního porostu. Mnohé z porostních skupin tuto hranici dosáhnou ještě před rokem 2025 a některé dokonce nabývají této hranice ještě v letech, ze kterých se predikce počítala.

Na základě shlukové analýzy, byly jednotlivé porostní skupiny rozděleny na pět shluků, podle podobnosti vlastností čtyř oblastí – stanoviště, porostu, poškození a nahodilé těžby. Zpětným provedením základní statistiky pro každý shluk dané oblasti bylo zjištěno, jak jsou jednotlivé shluky charakteristické. Následně bylo studováno, zda bylo rozdělení porostních skupin na shluky stejné nebo odlišné v rámci čtyř oblastí.

Výsledky predikce nahodilé těžby mohou sloužit jako podklady pro lesní hospodářství v dané oblasti, k zajištění takových opatření, aby bylo zamezeno

predikovanému zvýšení nahodilé těžby, které by znamenalo velkou ztrátu lesního porostu. Charakteristika porostních skupin dle rozdělení na shluky může být užitečná pro vytipování oblastí náchylnějších k NT z hlediska jejich vlastností stanoviště, porostu a poškození.

POUŽITÁ LITERATURA A INFORMAČNÍ ZDROJE

BONE, Christopher; WULDER, A., Michael; WHITE, C., Joanne; ROBERTSON, Colin; NELSON, A, Trisalyn. *A GIS-based risk rating of forest insect outbreaks using aerial overview surveys and the local Moran's I statistic*. Applied Geography, 2013, 40, s. 161-170. ISSN 0143-6228.

BURCIN, Boris, TESÁRKOVÁ, Klára, ŠÍDLO, Luděk. *Nejpoužívanější metody vyrovnávání a extrapolace křivky úmrtnosti a jejich aplikace na českou populaci*. Demografie, 2010, 53 (2), s. 77-89. ISSN 0011-8265.

Geomorfologické členění České republiky: Subprovincie, provincie, subsystémy a systémy. Treking: portál o horách, trekingu a turistice [online]. 2006 [cit. 2014-04-04]. Dostupné z: <http://www.treking.cz/regiony/subprovincie.htm>

GRAFAREND, W., Erik. *Linear and nonlinear models fixed effects, random effects, and mixed models*. Berlin: de Gruyter, 2006. 752 s. ISBN 31-101-6216-4.

HENDL, Jan. *Přehled statistických metod zpracování dat: analýza a metaanalýza dat*. Praha: Portál, 2004. 584 s. ISBN 80-7178-820-1.

HLÁSNÝ, Tomáš; KRÍSTEK, Štěpán; HOLUŠA, Štěpán; TROMBIK, Jiří; URBAŇCOVÁ, Naděžda. *Snow distribution in secondary Norway spruce forests in Central Europe: Regression modeling and its implications for forest management*. Forest Ecology and Management. 2011, 262, s. 2151-2161. ISSN 0378-1127.

HOSMER, David, W.; LEMENSHOW, Stanley. *Applied Logistic Regression: Second Edition*. New York: John Wiley & Sons, Inc., 2000. 375 s.

Hospodářská úprava lesa: Okruhy vědomostí. Borová šiška: studijní materiály pro studenty lesního inženýrství [online]. 2012 [cit. 2014-03-11]. Dostupné z: http://mezismrky.cz/borova_siska/r_3l.htm#Hospod%C3%A1%C5%99sk%C3%A1_%C3%BAprava_lesa

Charakteristika okresu Frýdek-Místek. Český statistický úřad [online]. ČZSO, 2012 [cit. 2014-04-04]. Dostupné z: http://www.czso.cz/xt/redakce.nsf/i/charakteristika_okresu_frydek_mistek

KAMIMURA, Kana; SHIRASI, Norihiko. *A review of strategies for wind damage assessment in Japanese forest*. Journal of Forestry Research. 2007, 12, s. 162-176. ISSN 1993-0607.

KŘÍSTEK, Jiří. a kol. *Ochrana lesů a přírodního prostředí*. Písek: Matice lesnická s.r.o., 2002. 386 s. ISBN 80-86271-08-0.

KŘÍSTEK, Štěpán; HOLUŠA, Jaroslav; HLÁSNÝ, Tomáš; TROMBIK, Jiří. Poškození alochtonních smrkových porostů sněhem: Analýza a návrh managementu s podporou geoinformatických technologií. In INSPEKTOR, Tomáš; HORÁK, Jiří; RŮŽIČKA, Jan. *Sborník - Symposium GIS Ostrava 2013*. Ostrava: Technická univerzita Ostrava, 2013. ISBN 978-80-248-2951-7.

KUDEĽKOVÁ, Romana. *Pravděpodobnost poškození porostu - využití metod nelineární regrese*. Olomouc: Univerzita Palackého, 2014. Bakalářská práce.

Lesnická typologie. Biogeografie: multimediální výuková příručka [online]. Brno: Masarykova univerzita, 2010 [cit. 2014-03-11]. Dostupné z: http://is.muni.cz/do/rect/el/estud/prif/ps10/biogeogr/web/index_book_5-4-2.html

MARTIŠ, Miroslav. *Člověk versus krajina*. Praha: Horizont, 1988. 262s.

MELOUN, Milan.; MILÍTKÝ, Jiří. *Statistická analýza experimentálních dat*. Praha: Academia, 2004. 953 s. ISBN 80-200-1254-0.

MEZERA, Alois. a kol. *Tvorba a ochrana krajiny*. Praha: SZN, 1979. 467 s.

NEUBAUER, Jiří.; SEDLAČÍK, Marek.; KŘÍŽ, Oldřich. *Základy statistiky: aplikace v technických a ekonomických oborech*. Praha: Grada, 2012. 233 s.

ISBN 978-80-247-4273-1.

OLSSON., U. *Generalized Linear Models: An Applied Approach* [online]. Sweden: Studentlitteratur, 2002 [cit. 2014-04-10]. Dostupné z:
http://nozdr.ru/data/media/biblioteka/kolxo3/M_Mathematics/MV_Probability/MVsa_Statistics%20and%20applications/Olsson%20U.%20Generalized%20linear%20models.%20An%20applied%20approach%20%28Studentlitteratur,%202002%29%28ISBN%209144041551%29%28O%29%28243s%29_MVsa_.pdf

OVERBECK, Marc; SCHMIDT, Matthias. *Modelling infestation risk of Norway spruce by *Ips typographus* (L.) in the Lower Saxon Harz Mountains (Germany)*. Forest Ecology and Management. 2012, 266, s. 155-125. ISSN 0378-1127.

PFEFFER, Antonín. *Kůrovcovití Scolytidae a jádrohlobovití Platypodidae*. Praha: Academia, 1989. 137 s. ISBN 80-20-0089-5.

POKORNÝ, Radek. *Voda a les: funkce hydrická a funkce vodohospodářská*. Brno: Mendelova univerzita, Provozně ekonomická fakulta [online]. 2014 [cit. 2014-05-01]. Dostupné z:
https://akela.mendelu.cz/~xcepl/inobio/inovace/Pesteni_lesu_II/9_Voda%20a%20les.pdf

POLENO, Zdeněk. *Lesnický naučný slovník II.díl, P-Ž*. Praha: Agrospoj, 1995. 683 s. ISBN 0-7084-131-1.

PRŮŠA, Eduard. *Přirozené lesy České republiky*. Praha: SZN, 1990. 246 s.

ISBN 80-209-0095-0.

Přírodní podmínky v CHKO Beskydy. Regionální informační servis [online]. Centrum pro regionální rozvoj České republiky, 2012 [cit. 2014-04-04]. Dostupné z: http://www.risy.cz/Files/Images/beskydy/ochrana-prirody/prir_podm_01.jpg

REICHHOLK, Josef. *Les: Ekologie středoevropských lesů*. Praha: Knížní klub Ikar, 1999. 223 s. ISBN 80-7202-494-9.

ROBERTSON, Colin; WULDER, A., Michael; NELSON, A., Trisalyn; WHITE, C., Joanne. *Risk rating for mountain pine beetle infestation of lodgepole pine forests over large areas with ordinal regression modelling*. Forest Ecology and Management. 2008, 256, s. 900-912. ISSN 0378-1127.

SAMEC, Pavel.; TUČEK, Pavel a kol. *Modelování růstových podmínek lesů v České republice*. Olomouc: Univerzita Palackého, 2012. 312 s. ISBN 978-80-244-3306-6.

Shluková analýza. [online]. Brno: Masarykova univerzita. [cit. 2014-04-10]. Dostupné z: http://is.muni.cz/th/172767/fi_b/5739129/web/web/main.html#Shlukov%C3%A1_anal%C3%BDza

SCHROEDER, Lief, Martin. Colonization of storm gaps by the spruce bark beetle: influence of gap and landscape characteristics. *Agricultural and Forest Entomology*. 2010, 12, s. 29-39. ISSN 1461-9563.

STADELMANN, Golo; BUGMANN, Harald; MEIER, Franz, WERMELINGER, Beat; BIGLER, Christof. Effects of salvage logging and sanitation felling on bark beetle (*Ips typographus* L.) infestations. *Forest Ecology and Management*. 2013, 305, s. 273-281. ISSN 0378-1127.

TVRDÍK, Josef. Analýza vícerozměrných dat: určeno pro vzdělání v akreditovaných studijních programech [online]. Ostrava: Ostravská univerzita, 2013 [cit. 2014-04-10]. Dostupné z: http://www1.osu.cz/~tvrdik/down/files/AVDAT_2013.pdf

ULBRICHTOVÁ, Iva. *Primární a sekundární produkce*. Fakulta lesnická české zemědělské univerzity [online]. 2005 [cit. 2014-03-11]. Dostupné z: http://fle.czu.cz/~ulbrichova/Skripta_EKOL/produkce/produkce.htm

VOŽENÍLEK, Vít. (2002): Diplomové práce z geoinformatiky. Vydavatelství Univerzity Palackého, Olomouc, UP, 31 s.

Vyhláška Ministerstva zemědělství č. 78/1996 Sb., o stanovení pásem ohrožení lesů pod vlivem imisí § 1: Charakteristika pásem ohrožení. EAGRI [online]. Ministerstvo zemědělství, 2009 [cit. 2014-03-11]. Dostupné z: <http://eagri.cz/public/web/mze/legislativa/pravni-predpisy-mze/tematicky-prehled/100051836.html>

ZVÁRA, Karel. Regrese. Praha: Matematicko-fyzikální fakulta Univerzity Karlovy, 2008. 253 s. ISBN 978-80-7378-041-8.

SUMMARY

Aim of this thesis is analysis of available data focused on 10-years lasting monitoring of growth groups in area of interest in eastern part of Czech Republic, and calculate prediction of random logging in between years 2009 and 2025. Growth groups were also divided into units with similar attributes.

First part of this thesis deals with description of random logging, causing agents and also with short theoretical work about used mathematical and statistical methods suitable for modelling random phenomena in forestry. As a source was used czech and foreign literature.

Next part is focused on basic information about area of interest and available data with detail description of particular attributes. Basic statistics were carried out to discover information about mean, variation, distribution and relationships between variables.

Practical part of thesis is formed by calculation of prediction in random logging by logistic regression, which is basis for modelling of probability of increase of random logging in every year until year 2025. This method is focusing on relationship between independent variable (year) and depend variable indicating value of cumulative random logging in interval 0 - 1. Value 0,1 is 10% and 1 means 100%. Some of this numbers had to be corrected, because their value was more than limit of 100%, or wasn't numerical therefore couldn't be used in calculation. Results of prediction revealed fact, that most of the growth groups will probably reach in year 2025 the 100% limit in random logging. Others will reach this limit before 2025 and some have already reached limit in year on which was calculation based.

Work revealed fact that it would be better include in calculation also information about changes of size of areas over the years. This problem could be solved by analysis of surviving, but this method wasn't used because of lack of time and suitable data.

Growth groups were divided into 5 main clusters with similar characteristic from several points of view: location, vegetation type, percentage of damage, and agents causing random logging.

Cluster analysis was used as a suitable statistical method for analysing similarity between groups of big data. On every cluster, basic statistical analysis was performed for description of basic characteristics. Visualisation of clusters in several approaches (spatial, vegetation type, percentage of damage, and agents causing random logging) helped to find out dependence between clusters, thus decide if there is a any system in

classification of growth group, or it is a random process. Using of multiple regression for discovering correlation between attributes would increase quality of prediction in random logging.

Results of this thesis could be used for forestry economy and ensuring of steps for reduction increasing of random logging, which means a lost of forests. Characteristic of growth groups could be useful for selection of areas predisposed to NT with respect to spatial characteristic, vegetation and damage.

PŘÍLOHY

SEZNAM PŘÍLOH

Vázané přílohy:

- Příloha 1 Základní statistická charakteristika dat
- Příloha 2 Kompletní vyhodnocení predikce kumulativní nahodilé těžby
- Příloha 3 Příkazový kód predikce
- Příloha 4 Příkazový kód shlukové analýzy

Volné přílohy

- Příloha 5 Poster predikce nahodilé těžby pro rok 2020 + heatmapy shlukové analýzy
- Příloha 6 DVD – metadata, text práce, vstupní data, výstupní data - predikce, shluková analýza, základní statistika, video – průběh nahodilé těžby, webové stránky, poster.

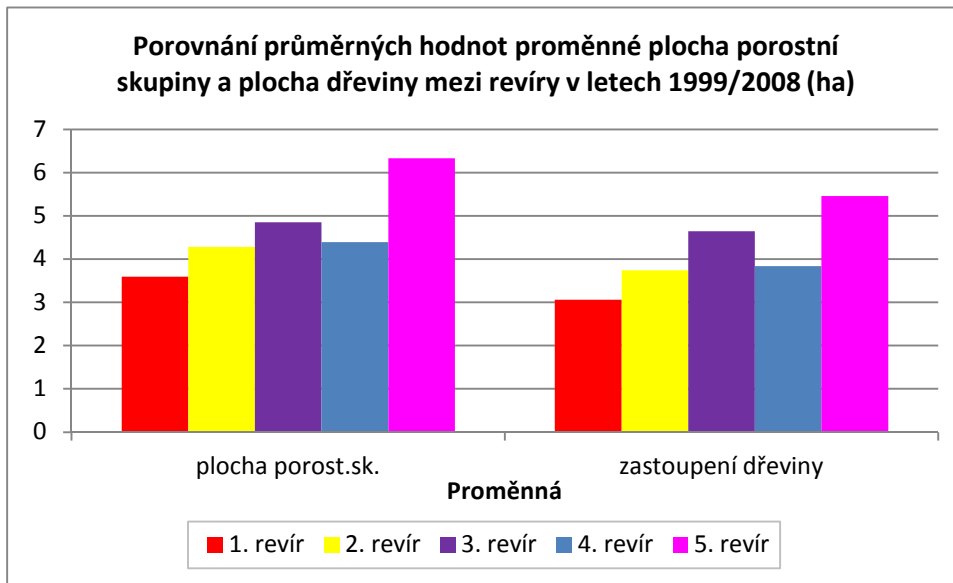
Příloha 1 – Základní statistická charakteristika dat

Tab. 1 Charakteristika střední hodnoty v datech – aritmetický průměr, medián

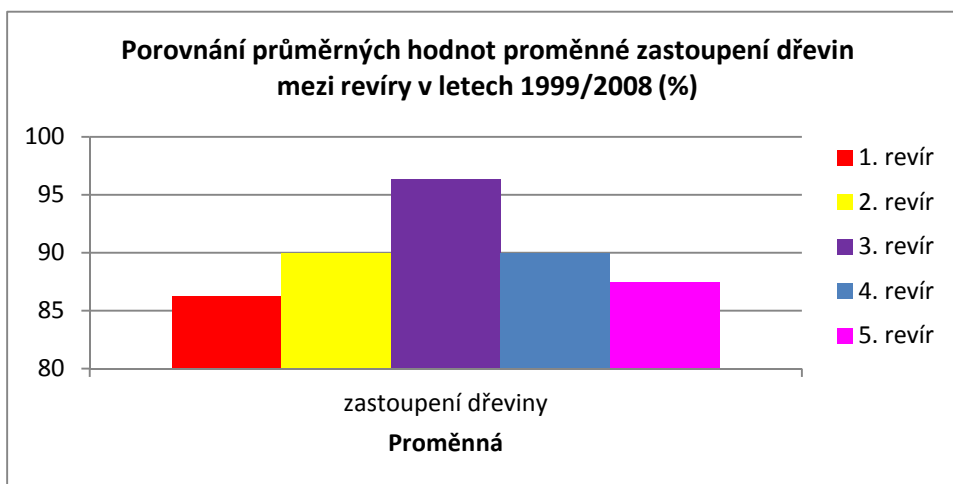
PROMĚNNÁ	Aritmetický průměr - \bar{x}					Medián - \tilde{x}				
	1. revír	2. revír	3. revír	4. revír	5. revír	1. revír	2. revír	3. revír	4. revír	5. revír
pl_ID (ha)	3,6	4,3	4,8	4,4	6,3	2,6	3,2	3,7	3,1	5,1
zast_dr (%)	86,3	90,0	96,3	90,0	87,5	90,0	95,0	100,0	95,0	93,0
pl_dr (ha)	3,1	3,7	4,6	3,8	5,5	2,2	2,8	3,6	2,6	4,3
NT_kur (m ³)	27,2	9,1	3,6	2,5	1,4	0,0	0,0	0,0	0,0	0,0
NT_tracheo (m ³)	7,0	1,9	0,6	0,3	0,0	0,0	0,0	0,0	0,0	0,0
NT_ziv (m ³)	5,1	8,4	18,0	9,8	16,2	0,0	0,0	0,7	0,0	0,0
NT_ost (m ³)	37,4	34,9	28,6	29,0	32,9	1,3	0,0	1,2	0,0	0,0
NT_celk_s (m ³)	76,8	54,3	50,8	41,6	50,5	6,8	3,7	9,0	3,9	6,5
zasoba_1 (m ³)	1168,8	1460,8	2262,4	1688,0	2462,2	854,0	1129,0	1709,0	1094,5	1847,5
zasoba_2 (m ³)	1092,4	1406,6	2211,6	1646,4	2411,7	739,5	1099,0	1667,5	1070,5	1833,5
%NT_zas_post (%)	9,2	3,9	2,6	2,7	2,3	0,8	0,3	0,5	0,3	0,3
%NT_zas_poc (%)	4,9	2,9	2,1	2,3	2,0	0,7	0,2	0,4	0,3	0,3
%NT_kum (%)	19,4	16,5	10,9	13,0	11,5	9,6	10,1	5,0	7,6	7,8
ASPECT (°)	196,2	203,5	164,9	167,6	160,8	204,5	213,0	151,5	114,0	135,0
SLOPE (°)	16,0	16,4	11,5	16,5	19,2	16,0	16,0	12,0	16,0	19,0
ELEVATION (m)	574,2	608,4	618,8	697,1	776,0	546,0	609,0	602,0	691,0	756,0
vek_pr	88,8	94,4	89,6	91,7	97,8	86,0	87,0	84,0	84,0	100,0
%pošk (%)	0,4	0,4	0,0	0,0	0,0	0,4	0,3	0,0	0,0	0,0

Tab. 2 Charakteristika míry rozptýlenosti v datech – rozptyl, směrodatná odchylka

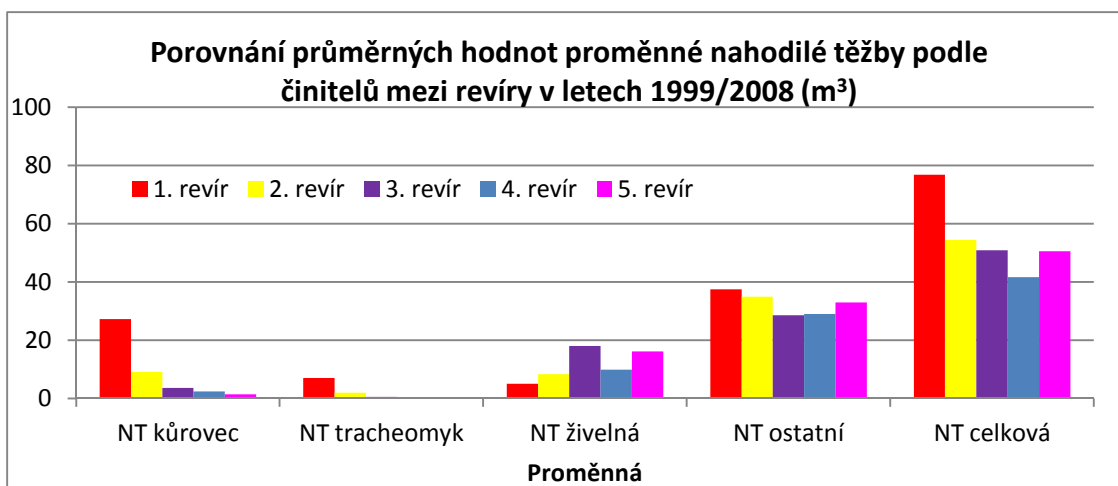
PROMĚNNÁ	Rozptyl - s_{x2}					Směrodatná odchylka - s_x				
	1. revír	2. revír	3. revír	4. revír	5. revír	1. revír	2. revír	3. revír	4. revír	5. revír
pl_ID (ha)	11,3	13,6	16,2	17,5	23,0	3,4	3,7	4,0	4,2	4,8
zast_dr (%)	178,0	162,4	50,3	150,5	174,4	13,3	12,7	7,1	12,3	13,2
pl_dr (ha)	8,5	10,3	14,4	13,3	17,6	2,9	3,2	3,8	3,7	4,2
NT_kur (m ³)	16795,9	2038,1	466,3	142,1	110,8	129,6	45,1	21,6	11,9	10,5
NT_tracheo (m ³)	2867,4	339,4	87,9	23,6	0,0	53,5	18,4	9,4	4,9	0,0
NT_ziv (m ³)	695,2	652,9	3385,7	989,8	2113,7	26,4	25,6	58,2	31,5	46,0
NT_ost (m ³)	11070,9	11217,3	7196,5	8437,1	11123,2	105,2	105,9	84,8	91,9	105,5
NT_celk_s (m ³)	38366,2	16226,0	13169,1	10373,1	13763,0	195,9	127,4	114,8	101,8	117,3
zasoba_1 (m ³)	1575793,2	1709162,3	4316551,2	3371822,4	4949665,2	1255,3	1307,4	2077,6	1836,3	2224,8
zasoba_2 (m ³)	1427660,5	1603422,2	4163529,2	3229779,3	4806606,3	1194,8	1266,3	2040,5	1797,2	2192,4
%NT_zas_post (%)	4361,2	103,5	42,2	54,6	43,1	66,0	10,2	6,5	7,4	6,6
%NT_zas_poc (%)	110,4	48,8	19,5	41,5	29,2	10,5	7,0	4,4	6,4	5,4
%NT_kum (%)	564,3	391,6	240,4	274,9	193,7	23,8	19,8	15,5	16,6	13,9
ASPECT (°)	9808,0	10589,8	6707,6	16041,0	11170,3	99,0	102,9	81,9	126,7	105,7
SLOPE (°)	43,0	30,6	22,5	47,4	32,7	6,6	5,5	4,7	6,9	5,7
ELEVATION (m)	9798,7	6597,6	8899,7	19367,1	9868,4	99,0	81,2	94,3	139,2	99,3
vek_pr	401,1	568,9	635,4	699,7	535,4	20,0	23,9	25,2	26,5	23,1
%pošk (%)	0,0	0,0	0,0	0,0	0,0	0,1	0,1	0,0	0,0	0,0



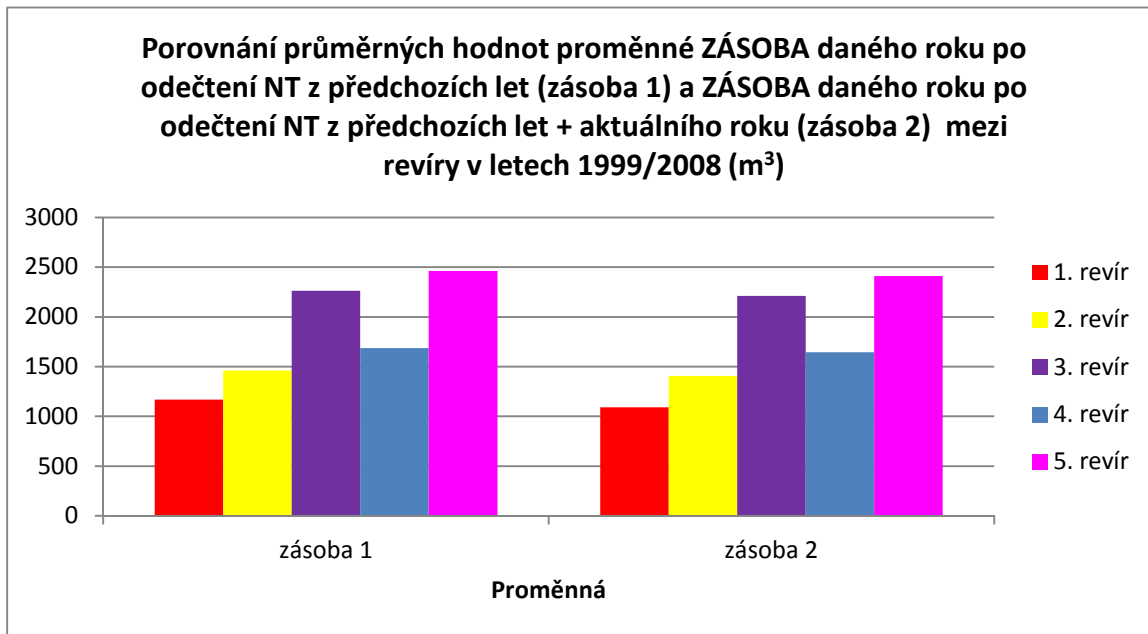
Graf 1 Porovnání průměrných hodnot plochy porostní skupiny a plochy dřevin mezi revíry v letech 1999/2008



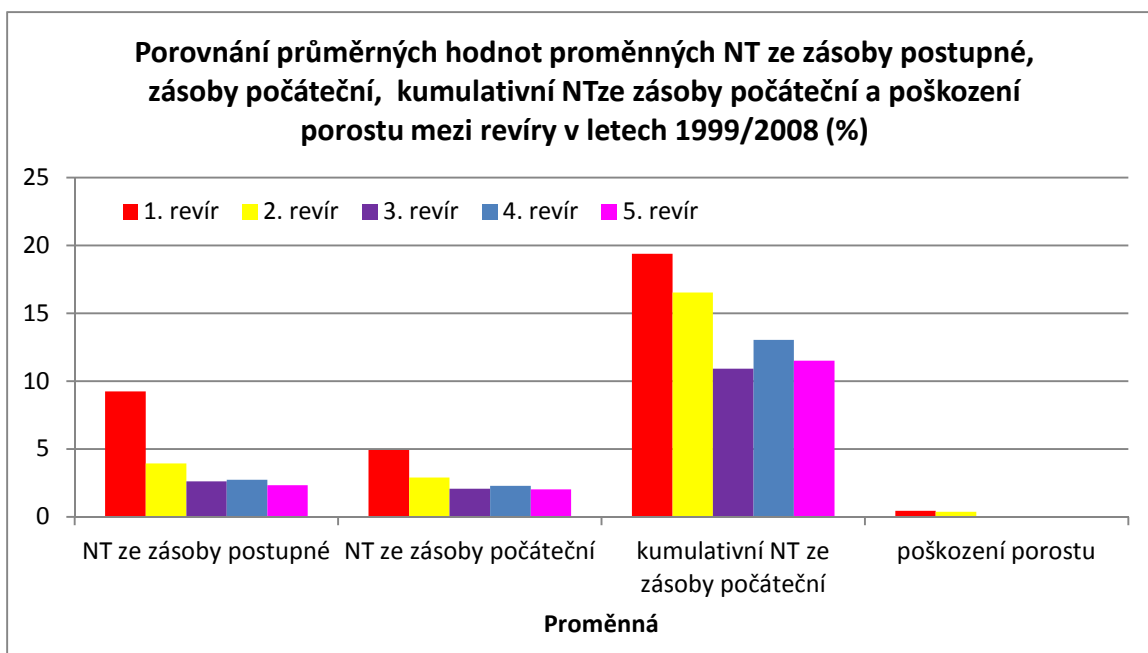
Graf 2 Porovnání průměrného zastoupení dřevin mezi revíry v letech 1999/2008



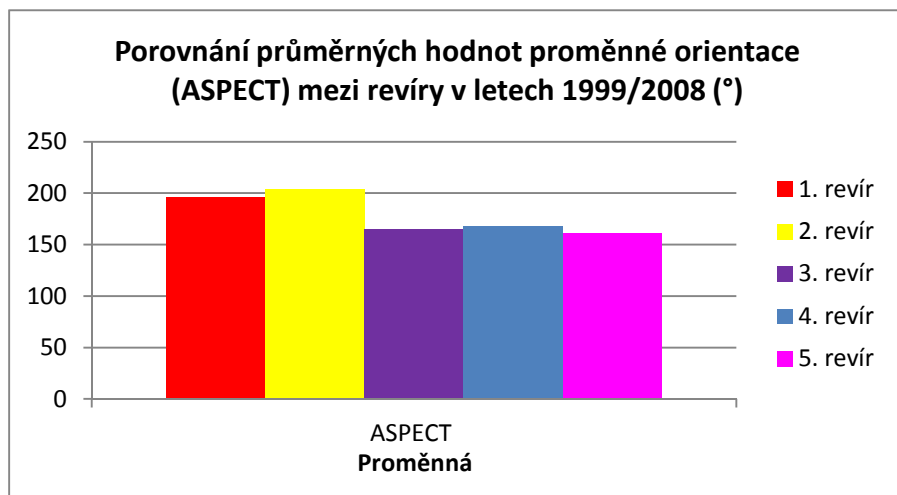
Graf 3 Porovnání průměrné nahodilé těžby podle činitelů mezi revíry v letech 1999/2008



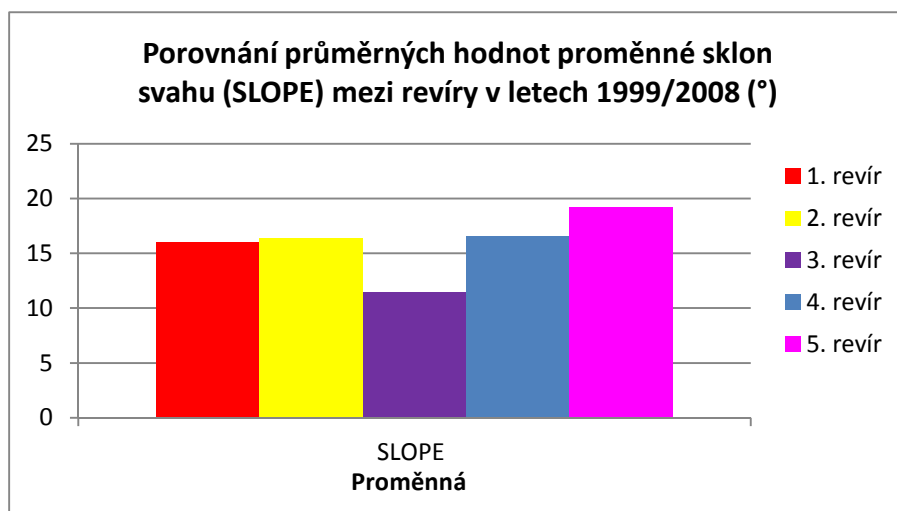
Graf 4 Porovnání průměrných hodnot ZÁSOBY daného roku po odečtení NT z předchozích let (zásoba 1) a ZÁSOBY daného roku po odečtení NT z předchozích let + aktuálního roku mezi revíry v letech 1999/2008



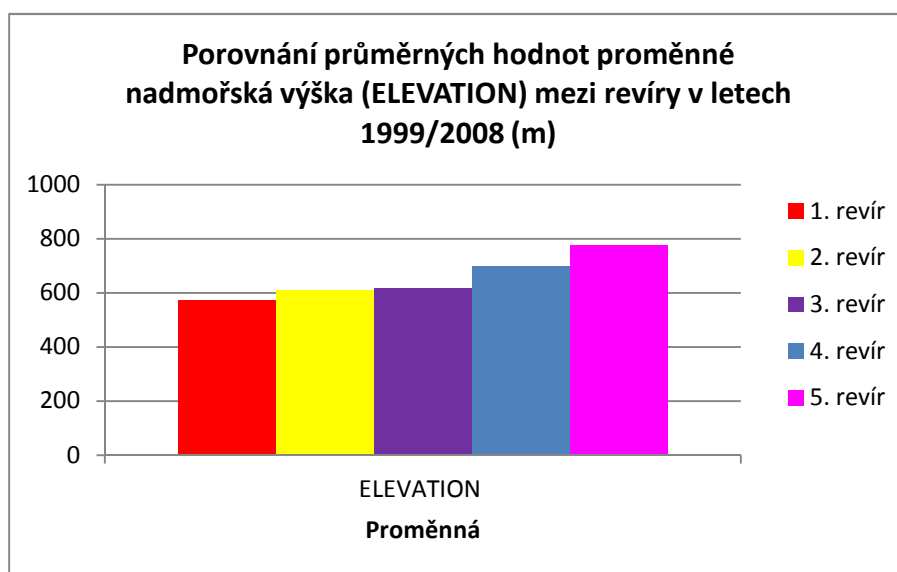
Graf 5 Porovnání průměrných hodnot NT ze zásoby postupné, zásoby počáteční, kumulativní NT ze zásoby počáteční a poškození porostu mezi revíry v letech 1999/2008



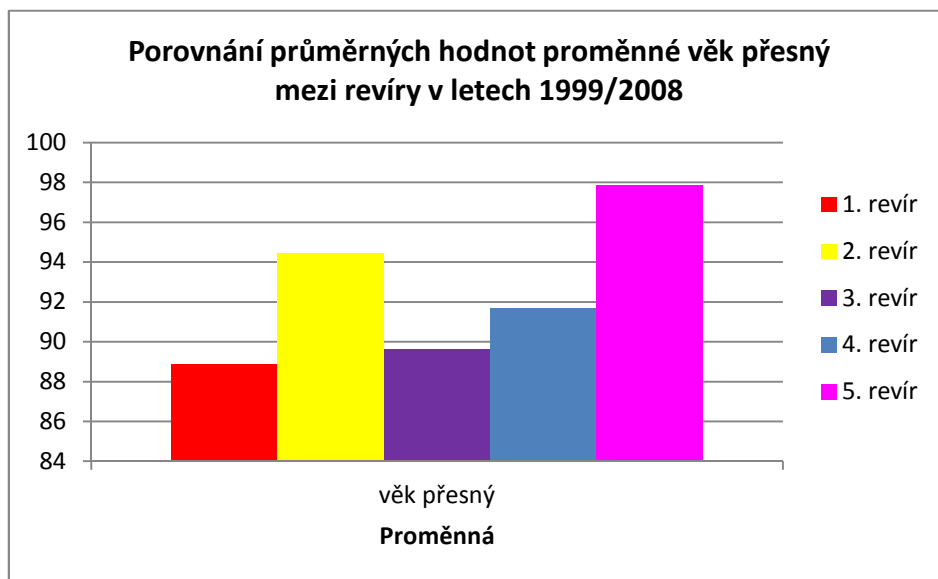
Graf 6 Porovnání orientace mezi revíry v letech 1999/2008



Graf 7 Porovnání sklonu svahu mezi revíry v letech 1999/2008



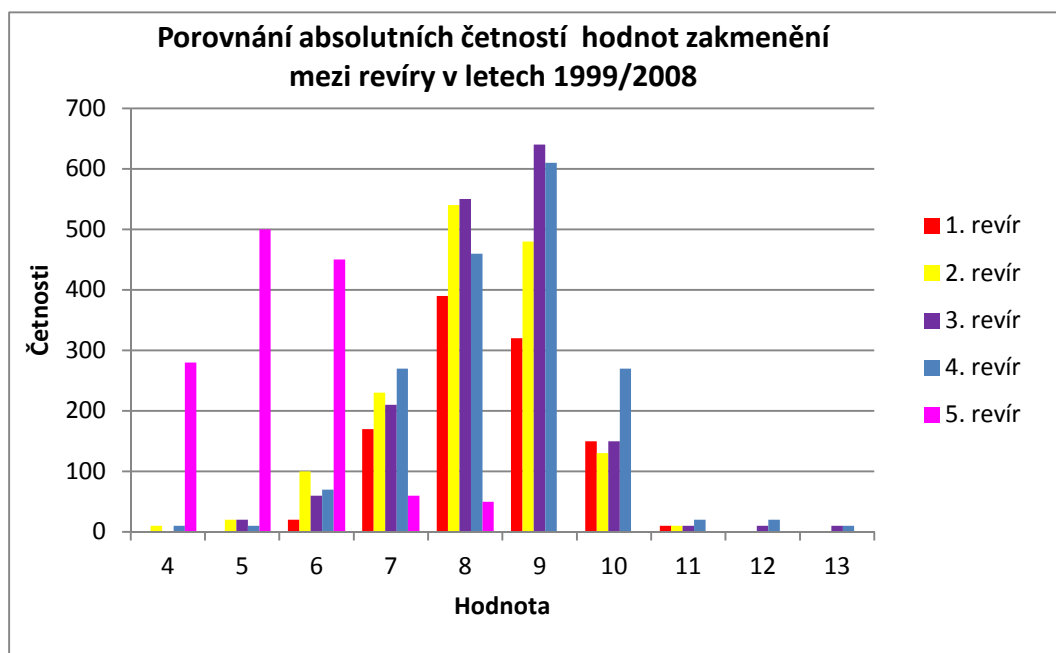
Graf 8 Porovnání nadmořské výšky mezi revíry v letech 1999/2008



Graf 9 Porovnání přesného věku mezi revíry v letech 1999/2008

Tab. 3 Absolutní četnost zakmenění mezi revíry

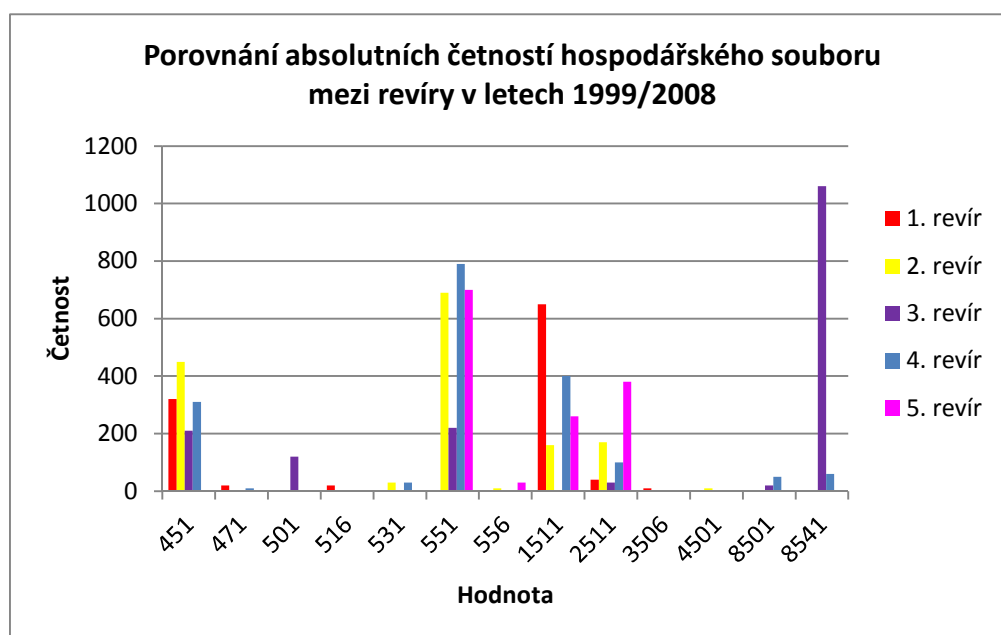
Absolutní četnost hodnot zakmenění mezi revíry					
Hodnota	1. revír	2. revír	3. revír	4. revír	5. revír
4	x	10	x	10	280
5	x	20	20	10	500
6	20	100	60	70	450
7	170	230	210	270	60
8	390	540	550	460	50
9	320	480	640	610	x
10	150	130	150	270	x
11	10	10	10	20	x
12	x	x	10	20	x
13	x	x	10	10	x



Graf 10 Porovnání absolutních četností zakmenění mezi revíry v letech 1999/2008

Tab. 4 Absolutní četnost hospodářského souboru mezi revíry

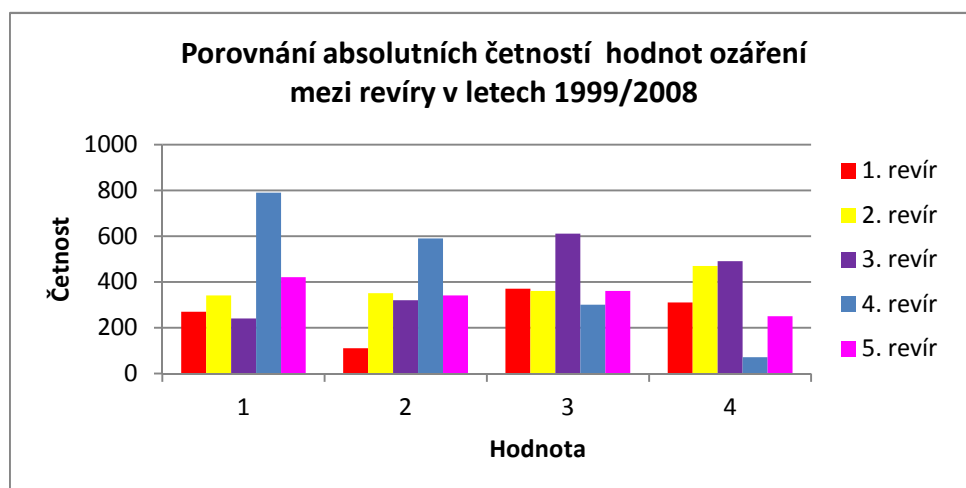
Absolutní četnost hospodářského souboru mezi revíry					
Hodnota	1. revír	2. revír	3. revír	4. revír	5. revír
451	320	450	210	310	x
471	20	x	x	10	x
501	x	x	120	x	x
516	20	x	x	x	x
531	x	30	x	30	x
551	x	690	220	790	700
556	x	10	x	x	30
1511	650	160	x	400	260
2511	40	170	30	100	380
3506	10		x	x	x
4501	x	10	x	x	x
8501	x	x	20	50	x
8541	x	x	1060	60	x



Graf 11 Porovnání absolutních četností hospodářského souboru mezi revíry v letech 1999/2008

Tab. 5 Absolutní četnost ozáření mezi revíry

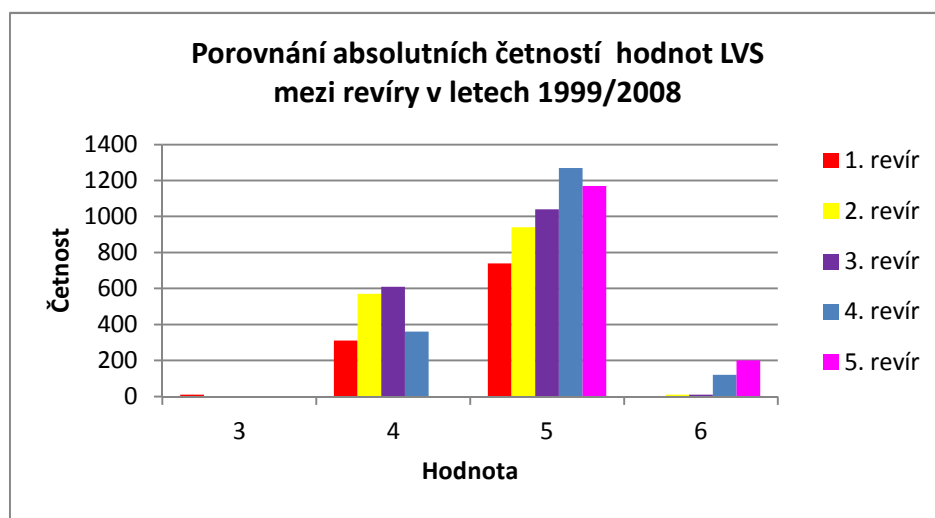
Absolutní četnost hodnot ozáření mezi revíry					
Hodnota	1. revír	2. revír	3. revír	4. revír	5. revír
1	270	340	240	790	420
2	110	350	320	590	340
3	370	360	610	300	360
4	310	470	490	70	250



Graf 12 Porovnání absolutních četností ozáření mezi revíry v letech 1999/2008

Tab. 6 Absolutní četnost LVS mezi revíry

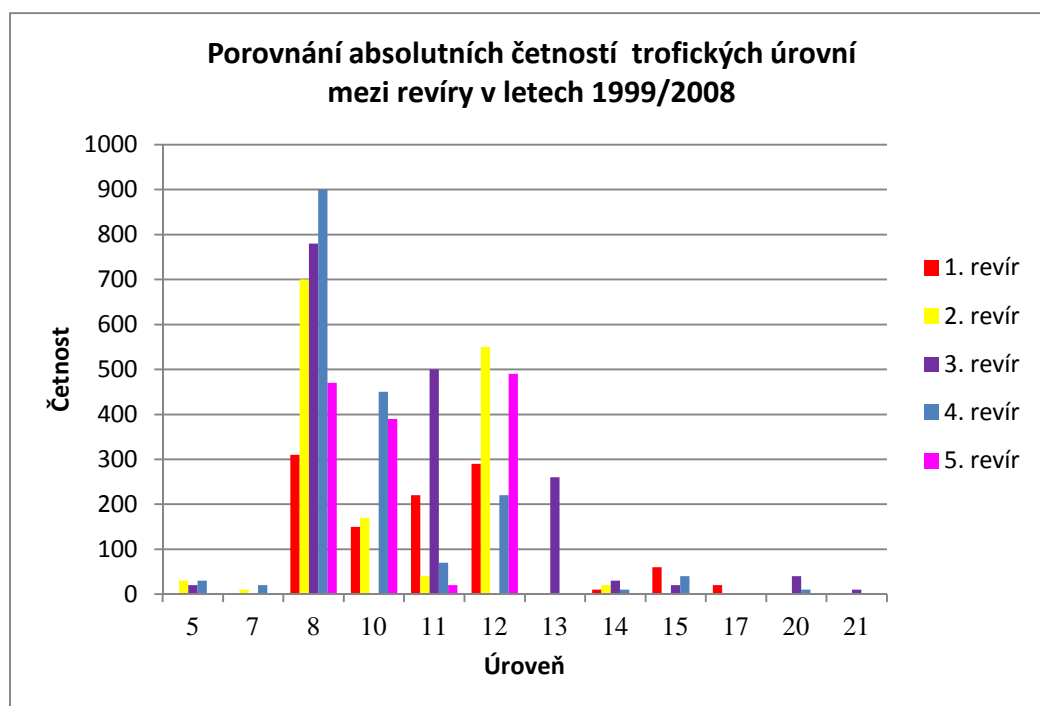
Absolutní četnost hodnot LVS mezi revíry					
Hodnota	1. revír	2. revír	3. revír	4. revír	5. revír
3	10	x	x	x	x
4	310	570	610	360	x
5	740	940	1040	1270	1170
6	x	10	10	120	200



Graf 13 Porovnání absolutních četností LVS mezi revíry v letech 1999/2008

Tab. 7 Absolutní četnost trofických úrovní mezi revíry

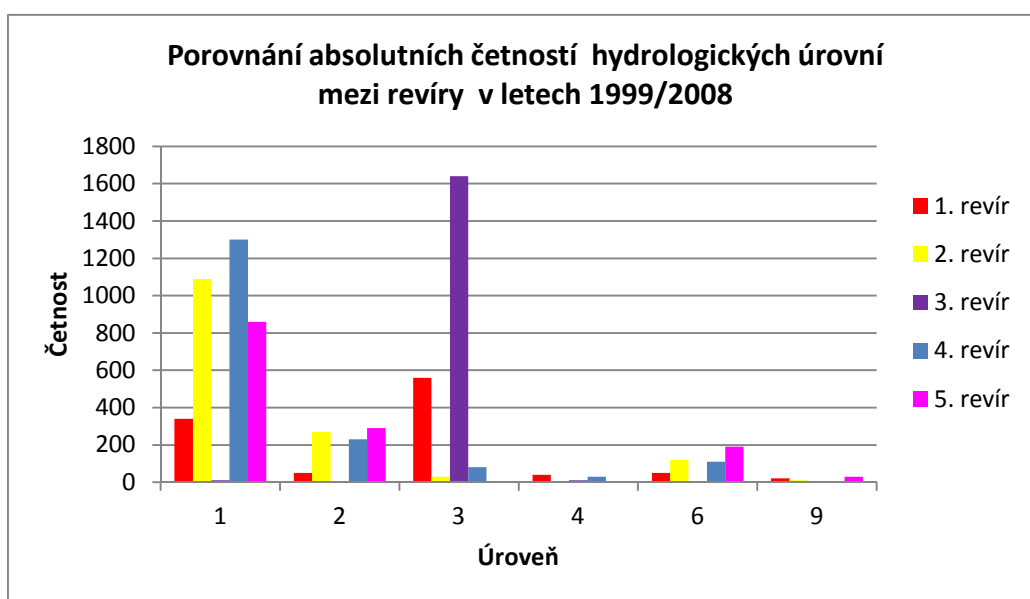
Absolutní četnost trofických úrovní mezi revíry					
Úroveň	1. revír	2. revír	3. revír	4. revír	5. revír
5	x	30	20	30	x
7	x	10	x	20	x
8	310	700	780	900	470
10	150	170	x	450	390
11	220	40	500	70	20
12	290	550	x	220	490
13	0	x	260	x	x
14	10	20	30	10	x
15	60	x	20	40	x
17	20	x	0	0	x
20	x	x	40	10	x
21	x	x	10	x	x



Graf 14 Porovnání absolutních četností trofických úrovní mezi revíry v letech 1999/2008

Tab. 8 Absolutní četnost hydrologických úrovní mezi revíry

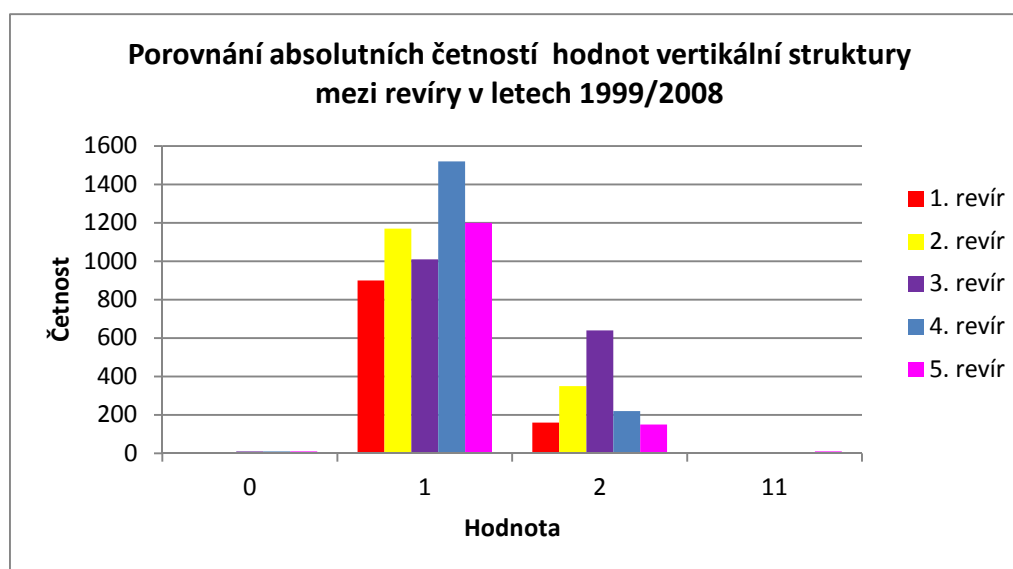
Absolutní četnost hydrologických úrovní mezi revíry					
Úroveň	1. revír	2. revír	3. revír	4. revír	5. revír
1	340	1090	10	1300	860
2	50	270	x	230	290
3	560	30	1640	80	x
4	40	x	10	30	x
6	50	120	x	110	190
9	20	10	x	x	30



Graf 15 Porovnání absolutních četností hydrologických úrovní mezi revíry v letech 1999/2008

Tab. 9 Absolutní četnost vertikální struktury mezi revíry

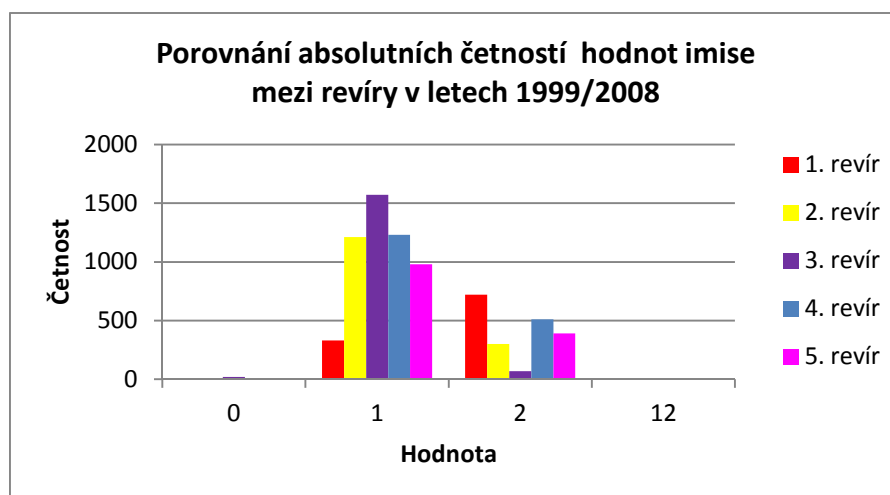
Absolutní četnost hodnot vertikální struktury mezi revíry					
Hodnota	1. revír	2. revír	3. revír	4. revír	5. revír
0	x	x	10	10	10
1	900	1170	1010	1520	1200
2	160	350	640	220	150
11	x	x	x	x	10



Graf 16 Porovnání absolutních četností vertikální struktury mezi revíry v letech 1999/2008

Tab. 10 Absolutní četnost hodnot imise mezi revíry

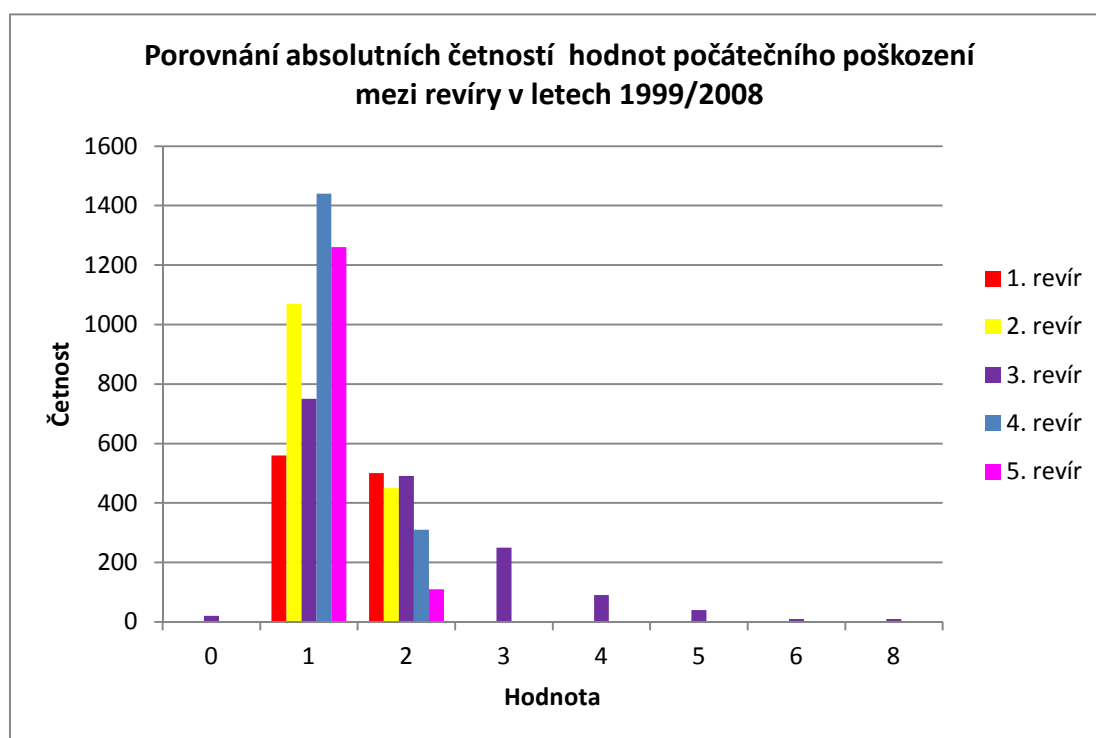
Porovnání absolutních četností hodnot imise mezi revíry					
Hodnota	1. revír	2. revír	3. revír	4. revír	5. revír
0	10	x	20	10	x
1	330	1210	1570	1230	980
2	720	300	70	510	390
12	x	10	x	x	x



Graf 17 Porovnání absolutních četností hodnot imise mezi revíry v letech 1999/2008

Tab. 11 Absolutní četnost počátečního poškození mezi revíry

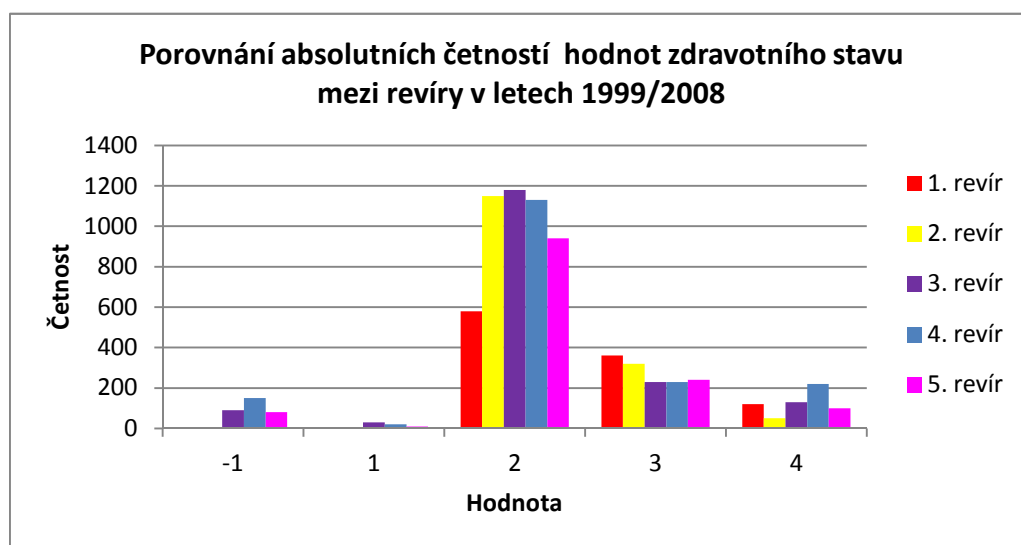
Absolutní četnost hodnot počátečního poškození mezi revíry					
Hodnota	1. revír	2. revír	3. revír	4. revír	5. revír
0	x	x	20	x	x
1	560	1070	750	1440	1260
2	500	450	491	310	110
3	x	x	250	x	x
4	x	x	90	x	x
5	x	x	40	x	x
6	x	x	10	x	x
8	x	x	10	x	x



Graf 3 Porovnání absolutních četností počátečního poškození mezi revíry v letech 1999/2008

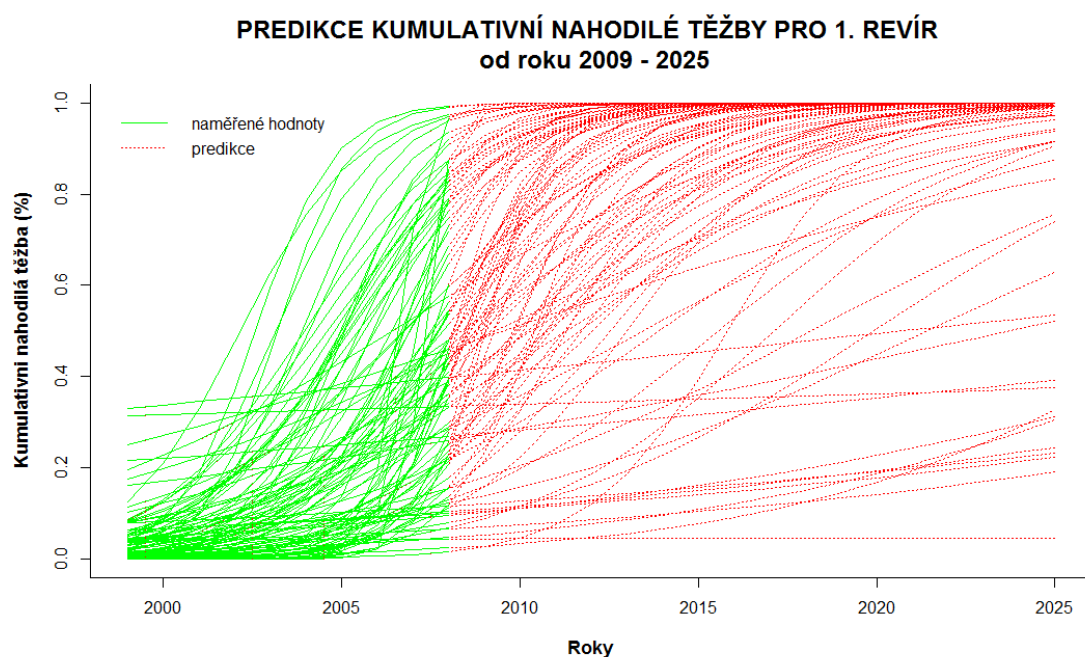
Tab. 12 Absolutní četnost zdravotního stavu mezi revíry

Absolutní četnost hodnot zdravotního stavu mezi revíry					
Hodnota	1. revír	2. revír	3. revír	4. revír	5. revír
-1	x	x	90	150	80
1	x	x	30	20	10
2	580	1150	1180	1130	940
3	360	320	230	230	240
4	120	50	130	220	100

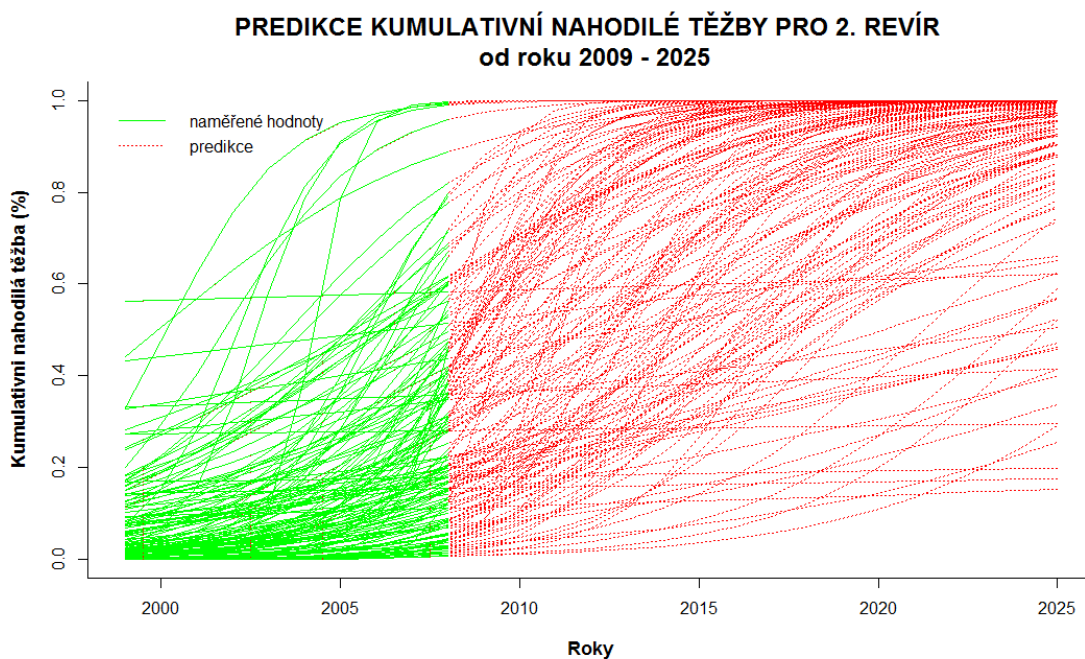


Graf 194 Porovnání absolutních četností zdravotního stavu mezi revíry v letech 1999/2008

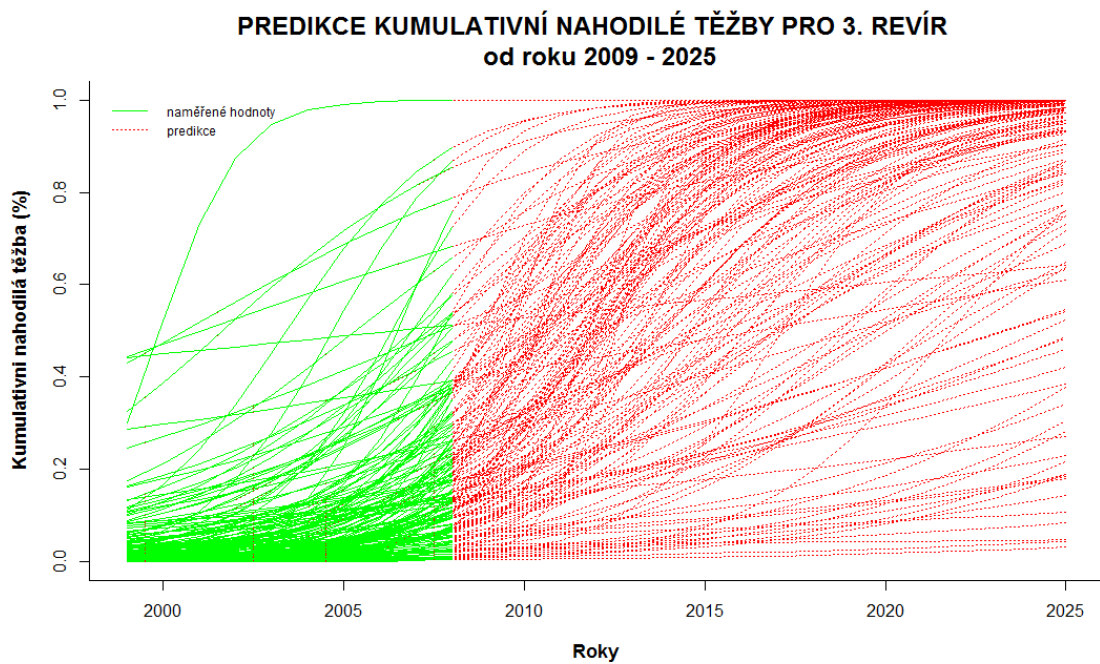
Příloha 2 – Kompletní vyhodnocení predikce kumulativní nahodilé těžby



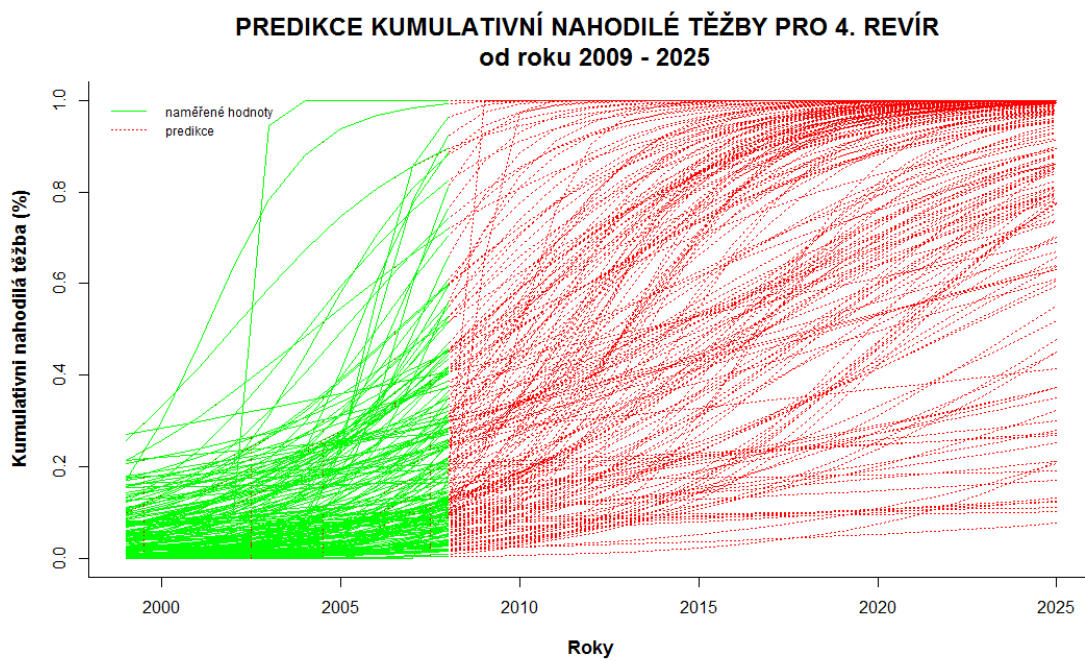
Obr. 1 Graf predikce kumulativní nahodilé těžby pro 1. revír



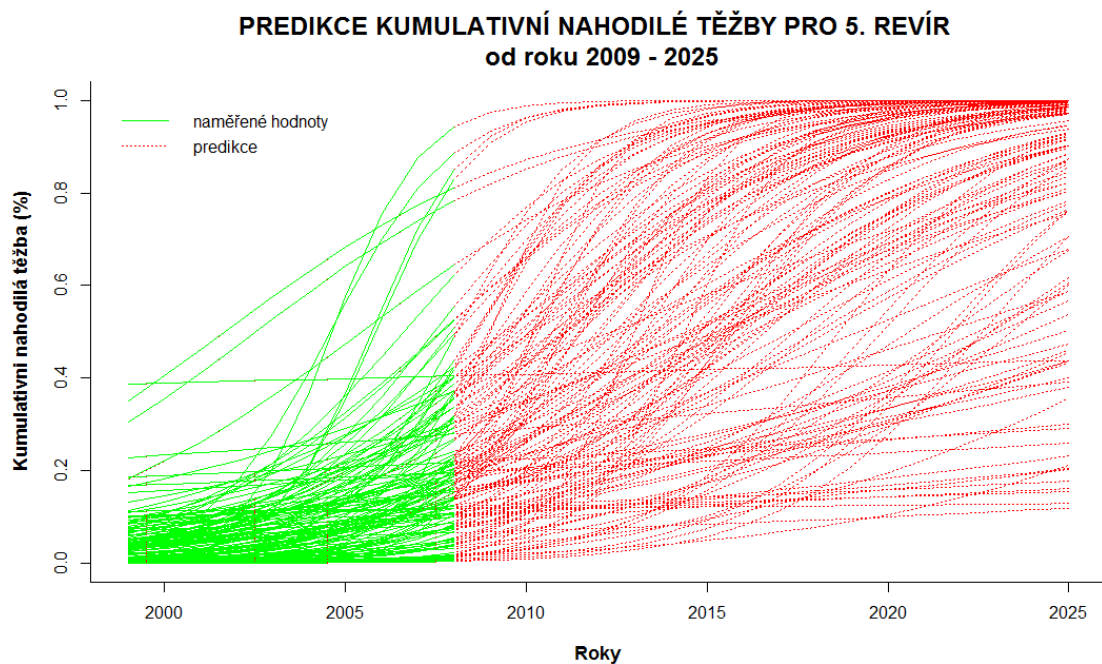
Obr. 2 Graf predikce kumulativní nahodilé těžby pro 2. revír



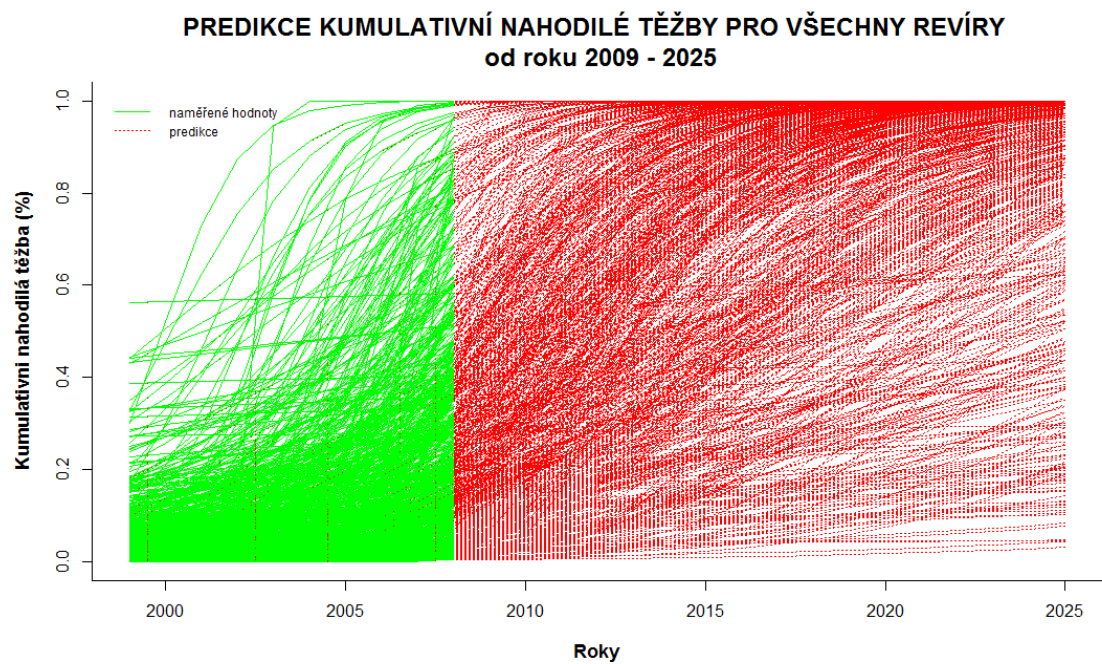
Obr. 3 Graf predikce kumulativní nahodilé těžby pro 3. revír



Obr. 4 Graf predikce kumulativní nahodilé těžby pro 4. revír



Obr. 5 Graf predikce kumulativní nahodilé těžby pro 5. revír



Obr. 6 Graf predikce kumulativní nahodilé těžby pro všechny revíry

Příloha 3 – Příkazové kódy pro predikce kumulativní nahodilé těžby

Příkaz pro predikci kumulativní nahodilé těžby pro 1. revír

```
#nacteni dat
#pro prvni revir
data <-
read.delim("C:/Users/Renata/Desktop/Renata_SLEZAKOVA_bc/Rko/Bakalarska_p
race/prvni_revir.txt")
#nastaveni zaznamu > 100 hodnotou 100
data[data > 100 & data < 1999] <- 100
#vydeleni cele matice 100, aby nebyla nutnost to delat cele dokola
data[,2:ncol(data)] <- data[,2:ncol(data)]/100
#vytvoreni prazdne matice pro predikci
data.predict <- matrix(0, ncol = ncol(data), nrow = 27, dimnames =
list(seq(1999, 2025, by=1), colnames(data)))
data.predict[,1] <- as.numeric(rownames(data.predict))
#definovani roku, ktere chceme
roky <- seq(1999, 2008, by=1)
roky.pred <- data.frame(roky=seq(1999, 2025, by=1))
#prazdny graf pro vykresleni
plot(c(1999, 2025), c(0,1), type = "n", xlab = "Roky", ylab =
"Kumulativni nahodilá těžba (%)", font.lab="2", cex.lab=1.1, bty="l",
#pozice nadpisu a podnadpisu pod sebou
main = "PREDIKCE KUMULATIVNÍ NAHODILÉ TĚŽBY PRO 1. REVÍR \n od roku 2009
- 2025",
#barva, font a velikost pisma nadpisu a podnadpisu
col.main=rgb(0,0,0), font.main="2",cex.main=1.5, col.sub=rgb(0,0,0),
font.sub="2", cex.sub=1.5)
#pro 2. a posledni sloupec nejdrive vezmeme promennou y - kumul,
#pote vypocet glm(zobecneni linearni model) z dat,
#zapsani predikce do vytvorene prazdne matice,
#vykresleni
for(i in 2:ncol(data)){
  kumul <- c(data[,i])

  glm.out <- glm(kumul ~ roky,family=quasibinomial(logit))
  predict <- predict(glm.out, roky.pred, type="response")
```

```
data.predict[,i] <- predict

lines(roky.pred[,1], predict, lty = "dotted", col="red", lwd="1" )
lines(roky, glm.out$fitted, type="l", col="green", lwd="1")
}
#legenda (pozicem, názvy proměnných, typ linie, šířka linie, barva,
ohraničení legendy, vzdálenost řádků, mezery mezi linií a popisem, délka
linie, velikost )
legend("topleft", c("naměřené hodnoty", "predikce"), lty= c("solid",
"dotted"), lwd="1", col=c("green", "red"), bty="n", y.intersp=.5,
x.intersp=0.85, seg.len= 1.5)
#vypsání dat do textového souboru
write.table(data.predict, "predikce_1_revir.txt")
```

Příkaz pro predikci kumulativní nahodilé těžby pro 2. revír

```
#nacteni dat

#pro druhy revir

data <- read.delim("C:/Users/Renata/Desktop/Renata_SLEZAKOVA_bc/Rko/Bakalarska_p
race/druhy_revir.txt")

#nastaveni zaznamu > 100 hodnotou 100

data[data > 100 & data < 1999] <- 100

#vydeleni cele matice 100, aby nebyla nutnost to delat cele dokola

data[,2:ncol(data)] <- data[,2:ncol(data)]/100

#vytvoreni prazdne matice pro predikci

data.predict <- matrix(0, ncol = ncol(data), nrow = 27, dimnames =
list(seq(1999, 2025, by=1), colnames(data)))

data.predict[,1] <- as.numeric(rownames(data.predict))

#definovani roku, ktere chceme

roky <- seq(1999, 2008, by=1)

roky.pred <- data.frame(roky=seq(1999, 2025, by=1))

#prazdny graf pro vykresleni

plot(c(1999, 2025), c(0,1), type = "n", xlab = "Roky", ylab =
"Kumulativni nahodilá těžba (%)",font.lab="2", cex.lab=1.1, bty="l",

#pozice nadpisu a podnadpisu pod sebou

main = "PREDIKCE KUMULATIVNÍ NAHODILÉ TĚŽBY PRO 2. REVÍR \n od roku 2009
- 2025",

#barva, font a velikost pisma nadpisu a podnadpisu

col.main=rgb(0,0,0), font.main="2",cex.main=1.5,col.sub=rgb(0,0,0),
font.sub="1", cex.sub=1.5)

#pro 2. a posledni sloupec nejdrive vezmeme promennou y - kumul,

#pote vypocet gml(zobecneni linearni model) z dat,

#zapsani predikce do vytvorene prazdne matice,

#vykresleni

for(i in 2:ncol(data)){

  kumul <- c(data[,i])
```

```
glm.out <- glm(kumul ~ roky,family=quasibinomial(logit))
predict <- predict(glm.out, roky.pred, type="response")

data.predict[,i] <- predict

lines(roky.pred[,1], predict, lty = "dotted", col="red", lwd="1" )
lines(roky, glm.out$fitted, type="l", col="green", lwd="1")
}

#legenda (pozicem, názvy proměnných, typ linie, šířka linie, barva,
ohraničení legendy, vzdálenost řádků, mezery mezi linií a popisem, délka
linie)

legend("topleft", c("naměřené hodnoty", "predikce"), lty= c("solid",
"dotted"), lwd="1", col=c("green", "red"), bty="n", y.intersp=.5,
x.intersp=0.85, seg.len= 1.5)

#vypsani dat do textoveho souboru

write.table(data.predict, "predikce_2_revir.txt")
```

Příkaz pro predikci kumulativní nahodilé těžby pro 3. revír

```
#nacteni dat
#pro treti revir
data <- read.delim("C:/Users/Renata/Desktop/Renata_SLEZAKOVA_bc/Rko/Bakalarska_p
race/treti_revir.txt")
#nastaveni zaznamu > 100 hodnotou 100
data[data > 100 & data < 1999] <- 100
#vydeleni cele matice 100, aby nebyla nutnost to delat cele dokola
data[,2:ncol(data)] <- data[,2:ncol(data)]/100
#vytvoreni prazdne matice pro predikci
data.predict <- matrix(0, ncol = ncol(data), nrow = 27, dimnames =
list(seq(1999, 2025, by=1), colnames(data)))
data.predict[,1] <- as.numeric(rownames(data.predict))
#definovani roku, ktere chceme
roky <- seq(1999, 2008, by=1)
roky.pred <- data.frame(roky=seq(1999, 2025, by=1))
#prazdny graf pro vykresleni
plot(c(1999, 2025), c(0,1), type = "n", xlab = "Roky", ylab =
"Kumulativni nahodilá těžba (%)",font.lab="2", cex.lab=1.1, bty="l",
#pozice nadpisu a podnadpisu pod sebou
main = "PREDIKCE KUMULATIVNÍ NAHODILÉ TĚŽBY PRO 3. REVÍR \n od roku 2009
- 2025",
#barva, font a velikost pisma nadpisu a podnadpisu
col.main=rgb(0,0,0), font.main="2",cex.main=1.5,col.sub=rgb(0,0,0),
font.sub="1", cex.sub=1.5)
#pro 2. a posledni sloupec nejdrive vezmeme promennou y - kumul,
#pote vypocet glm(zobecneni linearni model) z dat,
#zapsani predikce do vytvorene prazdne matice,
#vykresleni
for(i in 2:ncol(data)){
  kumul <- c(data[,i])

  glm.out <- glm(kumul ~ roky,family=quasibinomial(logit))
  predict <- predict(glm.out, roky.pred, type="response")

  data.predict[,i] <- predict
```

```
lines(roky.pred[,1], predict, lty = "dotted", col="red", lwd="1" )
lines(roky, glm.out$fitted, type="l", col="green", lwd="1")
}
#legenda (pozicem, názvy proměnných, typ linie, šířka linie, barva,
ohraničení legendy, vzdálenost řádků, mezery mezi linií a popisem, délka
linie)
legend("topleft", c("naměřené hodnoty", "predikce"), lty= c("solid",
"dotted"), lwd="1", col=c("green", "red"), bty="n", y.intersp=.5,
x.intersp=0.85, seg.len= 1.5,cex=0.75)
#vypsání dat do textového souboru
write.table(data.predict, "predikce_3_revir.txt")
```

Příkaz pro predikci kumulativní nahodilé těžby pro 4. revír

```
#nacteni dat
#pro ctvrty revir
data <- read.delim("C:/Users/Renata/Desktop/Renata_SLEZAKOVA_bc/Rko/Bakalarska_p
race/ctvrty_revir.txt")
#nastaveni zaznamu > 100 hodnotou 100
data[data > 100 & data < 1999] <- 100
#vydeleni cele matice 100, aby nebyla nutnost to delat cele dokola
data[,2:ncol(data)] <- data[,2:ncol(data)]/100
#vytvoreni prazdne matice pro predikci
data.predict <- matrix(0, ncol = ncol(data), nrow = 27, dimnames =
list(seq(1999, 2025, by=1), colnames(data)))
data.predict[,1] <- as.numeric(rownames(data.predict))
#definovani roku, ktere chceme
roky <- seq(1999, 2008, by=1)
roky.pred <- data.frame(roky=seq(1999, 2025, by=1))
#prazdny graf pro vykresleni
plot(c(1999, 2025), c(0,1), type = "n", xlab = "Roky", ylab =
"Kumulativni nahodilá těžba (%)", font.lab="2", cex.lab=1.1, bty="l",
#pozice nadpisu a podnadpisu pod sebou
main = "PREDIKCE KUMULATIVNÍ NAHODILÉ TĚŽBY PRO 4. REVÍR \n od roku 2009
- 2025",
#barva, font a velikost, pisma nadpisu a podnadpisu
col.main=rgb(0,0,0), font.main="2",cex.main=1.5,col.sub=rgb(0,0,0),
font.sub="1", cex.sub=1.5)
#pro 2. a posledni sloupec nejdrive vezmeme promennou y - kumul,
#pote vypocet glm(zobecneni linearni model) z dat,
#zapsani predikce do vytvorene prazdne matice,
#vykresleni
for(i in 2:ncol(data)){
  kumul <- c(data[,i])

  glm.out <- glm(kumul ~ roky,family=quasibinomial(logit))
  predict <- predict(glm.out, roky.pred, type="response")

  data.predict[,i] <- predict
```

```
lines(roky.pred[,1], predict, lty = "dotted", col="red", lwd="1" )
lines(roky, glm.out$fitted, type="l", col="green", lwd="1")
}
#legenda (pozicem, názvy proměnných, typ linie, šířka linie, barva,
ohraničení legendy, vzdálenost řádků, mezery mezi linií a popisem, délka
linie)
legend("topleft", c("naměřené hodnoty", "predikce"), lty= c("solid",
"dotted"), lwd="1", col=c("green", "red"), bty="n", y.intersp=.5,
x.intersp=0.85, seg.len= 1.5, cex=0.75)
#vypsání dat do textového souboru
write.table(data.predict, "predikce_4_revir.txt")
```


Příkaz pro predikci kumulativní nahodilé těžby pro 5. revír

```
#nacteni dat

#pro paty revir

data <- read.delim("C:/Users/Renata/Desktop/Renata_SLEZAKOVA_bc/Rko/Bakalarska_p
race/paty_revir.txt")

#nastaveni zaznamu > 100 hodnotou 100

data[data > 100 & data < 1999] <- 100

#vydeleni cele matice 100, aby nebyla nutnost to delat cele dokola

data[,2:ncol(data)] <- data[,2:ncol(data)]/100

#vytvoreni prazdne matice pro predikci

data.predict <- matrix(0, ncol = ncol(data), nrow = 27, dimnames =
list(seq(1999, 2025, by=1), colnames(data)))

data.predict[,1] <- as.numeric(rownames(data.predict))

#definovani roku, ktere chceme

roky <- seq(1999, 2008, by=1)

roky.pred <- data.frame(roky=seq(1999, 2025, by=1))

#prazdny graf pro vykresleni

plot(c(1999, 2025), c(0,1), type = "n", xlab = "Roky", ylab =
"Kumulativni nahodilá těžba (%)",font.lab="2", cex.lab=1.1, bty="l",

#pozice nadpisu a podnadpisu

main = "PREDIKCE KUMULATIVNÍ NAHODILÉ TĚŽBY PRO 5. REVÍR \n od roku 2009
- 2025",

#barva, font a velikost pisma nadpisu a podnadpisu

col.main=rgb(0,0,0), font.main="2",cex.main=1.5,col.sub=rgb(0,0,0),
font.sub="1", cex.sub=1.5)

#pro 2. a posledni sloupec nejdrive vezmeme promennou y - kumul,

#pote vypocet gml(zobecneni linearni model) z dat,

#zapsani predikce do vytvorene prazdne matice,

#vykresleni

for(i in 2:ncol(data)){

  kumul <- c(data[,i])
```

```
glm.out <- glm(kumul ~ roky,family=quasibinomial(logit))
predict <- predict(glm.out, roky.pred, type="response")

data.predict[,i] <- predict

lines(roky.pred[,1], predict, lty = "dotted", col="red", lwd="1" )
lines(roky, glm.out$fitted, type="l", col="green", lwd="1")
}

#legenda (pozicem, názvy proměnných, typ linie, šířka linie, barva,
ohraničení legendy, vzdálenost řádků, mezery mezi linií a popisem, délka
linie)

legend("topleft", c("naměřené hodnoty", "predikce"), lty= c("solid",
"dotted"), lwd="1", col=c("green", "red"), bty="n", y.intersp=.5,
x.intersp=0.85, seg.len= 1.5)

#vypsani dat do textoveho souboru

write.table(data.predict, "predikce_5_revir.txt")
```

Příkaz pro predikci kumulativní nahodilé těžby pro všechny revíry

```
#nacteni dat
#po vsechny reviry dohromady
data <- read.delim("C:/Users/Renata/Desktop/Renata_SLEZAKOVA_bc/Rko/Bakalarska_p
race/data_predikt.txt")
#nastaveni zaznamu > 100 hodnotou 100
data[data > 100 & data < 1999] <- 100
#vydeleni cele matice 100, aby nebyla nutnost to delat cele dokola
data[,2:ncol(data)] <- data[,2:ncol(data)]/100
#vytvoreni prazdne matice pro predikci
data.predict <- matrix(0, ncol = ncol(data), nrow = 27, dimnames =
list(seq(1999, 2025, by=1), colnames(data)))
data.predict[,1] <- as.numeric(rownames(data.predict))
#definovani roku, ktere chceme
roky <- seq(1999, 2008, by=1)
roky.pred <- data.frame(roky=seq(1999, 2025, by=1))
#prazdny graf pro vykresleni
plot(c(1999, 2025), c(0,1), type = "n", xlab = "Roky", ylab =
"Kumulativni nahodilá těžba (%)", font.lab="2", cex.lab=1.1, bty="l",
#pozice nadpisu a podnadpisu pod sebou
main = "PREDIKCE KUMULATIVNÍ NAHODILÉ TĚŽBY PRO VŠECHNY REVÍRY \n od
roku 2009 - 2025",
#barva, font a velikost pisma nadpisu a podnadpisu
col.main=rgb(0,0,0), font.main="2",cex.main=1.5,col.sub=rgb(0,0,0),
font.sub="2", cex.sub=1.5)
#pro 2. a posledni sloupec nejdrive vezmeme promennou y - kumul,
#pote vypocet glm(zobecneni linearni model) z dat,
#zapsani predikce do vytvorene prazdne matice,
#vykresleni
for(i in 2:ncol(data)){
  kumul <- c(data[,i])

  glm.out <- glm(kumul ~ roky,family=quasibinomial(logit))
  predict <- predict(glm.out, roky.pred, type="response")

  data.predict[,i] <- predict
```

```
lines(roky.pred[,1], predict, lty = "dotted", col="red", lwd="1" )
lines(roky, glm.out$fitted, type="l", col="green", lwd="1")
}
#legenda (pozicem, názvy proměnných, typ linie, šířka linie, barva,
ohraničení legendy, vzdálenost řádků, mezery mezi linií a popisem, délka
linie, velikost)
legend("topleft", c("naměřené hodnoty", "predikce"), lty= c("solid",
"dotted"), lwd="1", col=c("green", "red"), bty="n", y.intersp=.5,
x.intersp=0.85, seg.len= 1.5, cex=0.75)
#vypsání dat do textového souboru
write.table(data.predict, "predikce_vsechny_reviry.txt")
```

Příloha 4 – Příkazový kód pro shlukovou analýzu

```
#nahrani knihoven
library(cluster)
library(clusterSim)
library (reshape)
library (amap)
library (ggplot2)

#nahrani dat
data <- read.delim("C:/Users/Renata/Desktop/data/Rko/Bakalarska_prace/data_shlukovka.txt")
data <- data_shluk

#zjisteni typu dat v datasetu
str(data)

#zjisteni rozhrani tabulek (kolik sloupcu/radku)
dim(data)

#nadefinovani prvnioho sloupecku, jako id dat
rownames(data) <- data [,1]

#rozdeleni dat podle potreby pro shlukovou analyzu
data_stan <- data[,c(12,24,25,26,27,28,29)]
data_porost <- data[,c(5,9,31,32)]
data_NT <- data [,c(13,15,16,17)]
poskozeni <- data[,37]

#stanoveni optimalniho typu pro shlukovani pomoci cluster.sim
#pro smisene promenne (ratio & interval)
sim_data_stan <- cluster.Sim(data_stan, 7, 5, 10, icq = "S", outputHtml = "sim_data_stan", normalizations = c("n0"), distances = c("d1", "d2", "d3", "d4", "d5"), methods = c("m1", "m2", "m3", "m4", "m5", "m6", "m7", "m8"))
sim_data_porost <- cluster.Sim(data_porost, 7, 5, 10, icq = "S", outputHtml = "sim_data_porost", normalizations = c("n0"), distances = c("d1", "d2", "d3", "d4", "d5"), methods = c("m1", "m2", "m3", "m4", "m5", "m6", "m7", "m8"))
sim_data_NT <- cluster.Sim(data_NT, 7, 5, 10, icq = "S", outputHtml = "sim_data_NT", normalizations = c("n0"), distances = c("d1", "d2", "d3", "d4", "d5"), methods = c("m1", "m2", "m3", "m4", "m5", "m6", "m7", "m8"))
```

```

# matice vzdalenosti pro smisena data (brana jako klasicka cisla)
data_stan_dist <- GDM1(data_stan, weightsType = "equal")
data_porost_dist <- GDM1(data_porost, weightsType = "equal")
data_NT_dist <- GDM1(data_NT, weightsType = "equal")
# vlastni shlukovani smisenych dat pomoci pam
pam_data_stan <- pam(data_stan_dist, 5, diss = TRUE)
pam_data_porost <- pam(data_porost_dist, 5, diss = TRUE)
pam_data_NT <- pam(data_NT_dist, 5, diss = TRUE)
# tvorba dat se skupinami
skupiny_smisene_data_5_cluster <-
as.data.frame(cbind(pam_data_stan$clustering,
pam_data_porost$clustering, poskozeni, pam_data_NT$clustering), file =
"data_5_clusteru.txt", col.names = c("pam_data_stan",
"pam_data_porost", "poskozeni", "pam_data_NT"), row.names =
rownames(data_stan))
colnames(skupiny_smisene_data_5_cluster) <- c("pam_data_stan",
"pam_data_porost", "poskozeni", "pam_data_NT")
#vypsání shluků do textového souboru
write.table(skupiny_smisene_data_heat, file = "data_5_clust.txt",
row.names = rownames(data_stan))
SROVNÁNÍ DLE STANOVIŠTĚ
#heatmaps
skupiny_smisene_data_heat <-
cbind(rownames(skupiny_smisene_data_5_cluster),
skupiny_smisene_data_5_cluster)
colnames(skupiny_smisene_data_heat) <- c("JPRL",
"stanoviště", "porost", "poškození", "těžba")
skupiny_smisene_data_heat$JPRL <- with(skupiny_smisene_data_heat,
reorder(JPRL, stanoviště))
skupiny_smisene_data_melt <- melt(skupiny_smisene_data_heat)
#nastavení palety barev
paleta5 <- c("#EE0000", "#32CD32", "#8B008B", "#4876FF", "#FFFF00")
#nastaveni paramteru vizualizace pomoci heatmaps
g <- ggplot(skupiny_smisene_data_melt, aes(variable, JPRL, fill =
factor(value) )) + geom_tile()
gg <- g + labs(title = "SROVNÁNÍ SHLUKŮ PODLE VLASTNOSTÍ STANOVIŠTĚ,
POROSTU, POŠKOZENÍ POROSTU A NAHODILÉ TĚŽBY V ROCE 2008 \nsrovnání dle

```

```

stanoviště") + scale_fill_manual(name = "Označení shluku", values =
paleta5)
ggg <- gg + theme_grey(base_size = 4) + labs(x = "", y = "") +
theme(legend.position = "bottom", axis.ticks = element_blank(),
axis.text.x = element_text(size = 5))
ggg
SROVNÁNÍ DLE POROSTU
#heatmaps
skupiny_smisene_data_heat <-
cbind(rownames(skupiny_smisene_data_5_cluster),
skupiny_smisene_data_5_cluster)
colnames(skupiny_smisene_data_heat) <- c("JPRL",
"stanoviště", "porost", "poškození", "těžba")
skupiny_smisene_data_heat$JPRL <- with(skupiny_smisene_data_heat,
reorder(JPRL,porost))
skupiny_smisene_data_melt <- melt(skupiny_smisene_data_heat)
#nastavení palety barev
paleta5 <- c("#EE0000", "#32CD32", "#8B008B", "#4876FF", "#FFFF00")
#nastavení paramteru vizualizace pomoci heatmaps
g <- ggplot(skupiny_smisene_data_melt, aes(variable, JPRL, fill =
factor(value) )) + geom_tile()
gg <- g + labs(title = "SROVNÁNÍ SHLUKŮ PODLE VLASTNOSTÍ STANOVIŠTĚ,
POROSTU, POŠKOZENÍ POROSTU A NAHODILÉ TĚŽBY V ROCE 2008 \nsrovnání dle
porostu") + scale_fill_manual(name = "Označení shluku", values =
paleta5)
ggg <- gg + theme_grey(base_size = 4) + labs(x = "", y = "") +
theme(legend.position = "bottom", axis.ticks = element_blank(),
axis.text.x = element_text(size = 5))
ggg
SROVNÁNÍ DLE POŠKOZENÍ POROSTU
#heatmaps
skupiny_smisene_data_heat <-
cbind(rownames(skupiny_smisene_data_5_cluster),
skupiny_smisene_data_5_cluster)
colnames(skupiny_smisene_data_heat) <- c("JPRL",
"stanoviště", "porost", "poškození", "těžba")
skupiny_smisene_data_heat$JPRL <- with(skupiny_smisene_data_heat,
reorder(JPRL,poškození))
skupiny_smisene_data_melt <- melt(skupiny_smisene_data_heat)

```

```

#nastavení palety barev
paleta5 <- c("#EE0000", "#32CD32", "#8B008B", "#4876FF", "#FFFF00")
#nastaveni paramteru vizualizace pomoci heatmaps
g <- ggplot(skupiny_smisene_data_melt, aes(variable, JPRL, fill =
factor(value) )) + geom_tile()
gg <- g + labs(title = "SROVNÁNÍ SHLUKŮ PODLE VLASTNOSTÍ STANOVIŠTĚ,
POROSTU, POŠKOZENÍ POROSTU A NAHODILÉ TĚŽBY V ROCE 2008 \nsrovnání dle
poškození porostu") + scale_fill_manual(name = "Označení shluku", values
= paleta5)
ggg <- gg + theme_grey(base_size = 4) + labs(x = "", y = "") +
theme(legend.position = "bottom", axis.ticks = element_blank(),
axis.text.x = element_text(size = 5))
ggg
SROVNÁNÍ DLE POŠKOZENÍ NAHODILÉ TĚŽBY
#heatmaps
skupiny_smisene_data_heat <-
cbind(rownames(skupiny_smisene_data_5_cluster),
skupiny_smisene_data_5_cluster)
colnames(skupiny_smisene_data_heat) <- c("JPRL",
"stanoviště", "porost", "poškození", "těžba")
skupiny_smisene_data_heat$JPRL <- with(skupiny_smisene_data_heat,
reorder(JPRL, těžba))
skupiny_smisene_data_melt <- melt(skupiny_smisene_data_heat)
#nastavení palety barev
paleta5 <- c("#EE0000", "#32CD32", "#8B008B", "#4876FF", "#FFFF00")
#nastaveni paramteru vizualizace pomoci heatmaps
g <- ggplot(skupiny_smisene_data_melt, aes(variable, JPRL, fill =
factor(value) )) + geom_tile()
gg <- g + labs(title = "SROVNÁNÍ SHLUKŮ PODLE VLASTNOSTÍ STANOVIŠTĚ,
POROSTU, POŠKOZENÍ POROSTU A NAHODILÉ TĚŽBY V ROCE 2008 \nsrovnání dle
nahodilé těžby") + scale_fill_manual(name = "Označení shluku", values =
paleta5)
ggg <- gg + theme_grey(base_size = 4) + labs(x = "", y = "") +
theme(legend.position = "bottom", axis.ticks = element_blank(),
axis.text.x = element_text(size = 5))
ggg

```



```

#průměr, směrodatná odchylka a medián pro jednotlivé shluky
#sjednoceni dat
data_stan_stat <- cbind(data_stan, pam_data_stan$clustering)
data_porost_stat <- cbind(data_porost, pam_data_porost$clustering)
data_NT_stat <- cbind(data_NT, pam_data_NT$clustering)
#statistika vlastnosti stanoviste
prum_stan <- aggregate(data_stan_stat, by=
list(pam_data_stan$clustering), mean)
median_stan <- aggregate(data_stan_stat, by=
list(pam_data_stan$clustering), median)
sd_stan <- aggregate(data_stan_stat, by= list(pam_data_stan$clustering),
sd)
#statistika vlastnosti porostu
prum_porost <- aggregate(data_porost_stat, by=
list(pam_data_porost$clustering), mean)
median_porost <- aggregate(data_porost_stat, by=
list(pam_data_porost$clustering), median)
sd_porost <- aggregate(data_porost_stat, by=
list(pam_data_porost$clustering), sd)
#statistika vlastnosti NT
prum_NT <- aggregate(data_NT_stat, by= list(pam_data_NT$clustering),
mean)
median_NT <- aggregate(data_NT_stat, by= list(pam_data_NT$clustering),
median)
sd_NT <- aggregate(data_NT_stat, by= list(pam_data_NT$clustering), sd)

```