

Univerzita Palackého v Olomouci

Přírodovědecká fakulta

Katedra geoinformatiky

**ANALÝZA GEOGRAFICKÉHO ROZŠÍŘENÍ
MEDICAGO TRUNCATULA**

Bakalářská práce

Jiří KOMÍNEK

Vedoucí práce RNDr. Jan BRUS, Ph.D.

Olomouc 2017

Geoinformatika a geografie

Anotace

Cílem bakalářské je popsat stanovištní podmínky *Medicago truncatula*. Výstupy jsou mapové produkty, statistické přehledy a datová sada nálezových dat včetně environmentálních faktorů.

Samotné vlastní řešení je ukázkou postupu výběru dat, problémů a možných řešení při jejich zpracování. V rešeršní části jsou popsány volně dostupné zdroje environmentálních dat. Dále je zde představena problematika prostorového modelování druhů. Praktická část je zaměřena na zpracování získaných dat a přiřazení environmentálních faktorů k bodové vrstvě výskytu *Medicago truncatula*. Zpracování rozsáhlých datasetů a špatná kvalita nálezových dat přináší celou řadu problémů. Mezi ně patří zdlouhavé manuální postupy, nebo nepřesnost v poloze nálezových dat. Je zde popsáno několik možných způsobů řešení prostřednictvím nástrojů v programu *ArcGIS for Desktop*.

Výsledek práce bude dále použit při výzkumu dormance semen *Medicago truncatula* prováděného Katedrou botaniky Univerzity Palackého v Olomouci.

KLÍČOVÁ SLOVA

Medicago truncatula, SDM, environmentální faktory, nálezová data

Počet stran práce: 46

Počet příloh: 7 (z toho 1 vázaná a 6 volných)

Anotation

The aim of this bachelor thesis is to describe the habitat conditions of *Medicago truncatula*. The outputs consist of map products, statistical reports and the data set of occurrence data including environmental factors.

The solution represents an example of the process of choosing data, problems, and possible solutions gained by processing them. The search section describes the free environmental data sources. In the next section, the problem of species distribution modeling is presented.

The practical part focuses on the processing of the acquired data and the assignment of the environmental factors to the point of occurrence of *Medicago truncatula*. The processing of large datasets and the poor quality of the occurrence data brings a number of problems. These include the lengthy manual procedures and also inaccuracies in the location of the occurrence data. There are several possible ways how to solve this problem, for example through *ArcGIS for Desktop* tools. The result of this work will be further used in the research of the dormancy of *Medicago truncatula* seeds by the Department of Botany of Palacky University in Olomouc.

KEYWORDS

Medicago truncatula, SDM, environmental factors, occurrence data

Number of pages: 46

Number of appendixes: 7

Prohlašuji, že

- bakalářskou práci včetně příloh, jsem vypracoval samostatně a uvedl jsem všechny použité podklady a literaturu.

- jsem si vědom, že na moji bakalářskou práci se plně vztahuje zákon č.121/2000 Sb. - autorský zákon, zejména § 35 – využití díla v rámci občanských a náboženských obřadů, v rámci školních představení a využití díla školního a § 60 – školní dílo,

- beru na vědomí, že Univerzita Palackého v Olomouci (dále UP Olomouc) má právo nevýdělečně, ke své vnitřní potřebě, bakalářskou práci užívat (§ 35 odst. 3),

- souhlasím, aby jeden výtisk bakalářské práce byl uložen v Knihovně UP k prezenčnímu nahlédnutí,

- souhlasím, že údaje o mé bakalářské práci budou zveřejněny ve Studijním informačním systému UP,

- v případě zájmu UP Olomouc uzavřu licenční smlouvu s oprávněním užít výsledky a výstupy mé bakalářské práce v rozsahu § 12 odst. 4 autorského zákona,

- použít výsledky a výstupy mé bakalářské práce nebo poskytnout licenci k jejímu využití mohu jen se souhlasem UP Olomouc, která je oprávněna v takovém případě ode mne požadovat přiměřený příspěvek na úhradu nákladů, které byly UP Olomouc na vytvoření díla vynaloženy (až do jejich skutečné výše).

V Olomouci dne 15. 8. 2017

Jiří Komínek

Děkuji vedoucímu práce RNDr. Janu Brusovi, Ph.D. za podněty a připomínky při vypracování práce. Děkuji také rodině za podporu a trpělivost při mém studiu. V neposlední řadě všem přátelům, kteří mi pomáhali dokončit práci. Tato práce byla podpořena studentským projektem IGA_PrF_2017_024 Univerzity Palackého v Olomouci.

OBSAH

SEZNAM POUŽITÝCH ZKRATEK	10
ÚVOD	11
1 CÍLE PRÁCE	12
2 METODY A POSTUPY ZPRACOVÁNÍ.....	13
2.1 Použitá data.....	13
2.2 Použitý software.....	13
2.3 Postup zpracování.....	14
3 SOUČASNÝ STAV ŘEŠENÉ PROBLEMATIKY	15
3.1 <i>Medicago truncatula</i>	15
3.1.1 Biologické studie <i>Medicago truncatula</i>	17
3.2 Modelování prostorové distribuce druhů	18
3.2.1 Typy dat vstupujících do SDM.....	19
3.2.2 Dálkový průzkum a SDM	19
3.2.3 Vliv měřítka na výsledek SDM	20
3.3 Datové zdroje	21
3.3.1 WorldClim.....	21
3.3.2 Harmonised Word Soil Database v 1.2.....	22
3.3.3 SoilGrids.....	23
3.3.4 Shuttle Radar Topography Mission.....	23
3.3.5 ASTER Global Digital Elevation Model (GDEM2)	23
3.3.6 CCI Land cover 2015.....	23
3.3.7 500 meter MODIS Land Cover Maps.....	23
3.3.8 Copernicus Global Land Service (GLS).....	23
3.3.9 Corine Land Cover (CLC)	24
3.3.10 GlobCover 2009	24
4 VLASTNÍ ŘEŠENÍ.....	26
4.1 Prvotní zpracování dat	26
4.1.1 Nálezová data <i>Medicago truncatula</i>	26
4.1.2 WorldClim.....	29
4.1.3 ASTER GDEM 2.0	29
4.1.4 Harmonised Word Soil Database v 1.2.....	30
4.1.5 SoilGrids.....	31
4.1.6 GlobCover 2009	31
4.1.7 CCI Land cover 2015.....	31
4.2 Finální úprava dat	31
4.2.1 Detekce potencionálně rizikových bodů	32
4.2.2 Testování proměnlivosti hodnot	33
4.3 Statistické výpočty	34
4.3.1 Environmentální data.....	35
5 VÝSLEDKY.....	36
5.1 Výsledná bodová vrstva.....	36
5.2 Statistické přehledy	36

5.2.1	Klimatické charakteristiky.....	37
5.2.2	Nadmořská výška, sklon, orientace	39
5.2.3	Půdní data	40
5.2.4	Krajinný pokryv	41
5.2.5	Souhrnná statistika	42
5.3	Mapové výstupy	42
6	DISKUZE	44
7	ZÁVĚR.....	46
	POUŽITÁ LITERATURA A INFORMAČNÍ ZDROJE	
	PŘÍLOHY	

SEZNAM POUŽITÝCH ZKRATEK

Zkratka	Význam
CIAT	Center for Tropical Agriculture
CLC	Corine Land Cover
CLINO	WMO climatological normals
DPZ	Dálkový průzkum země
EEA	European Environment Agency
EPSG	Geodetic Parameter Set
ESA	European Space Agency
ESD	European Soil Database
FAO	Food and Agriculture Organization of the United Nations
FTP	File Transfer Protocol
GAČR	Grantová agentura České republiky
GDEM	Global Digital Elevation Model
GHCN	Global Historical Climatology Network
GOFC-GOLD	Global Observation of Forest and Land Cover
HWSD	Harmonised World Soil Database
CHMU	Český Hydrometeorologický ústav
IGBP	International Geosphere-Biosphere Programme
ILASA	Institute for Landscape Architecture in South Africa
INRA	International Research Associates
ISRIC	International Soil Reference and Information Centre
ISSCAS	Institute of Soil Science, Chinese Academy of Sciences
JRC	Joint Research Centre
LAI	Leaf Area Index
LISA	Local Indicators of Spatial Association
NASA	National Aeronautics and Space Administration
NDVI	Normalized Difference Vegetation Index
NGA	National Geospatial-Intelligence Agency
PCA	Principal Component Analysis
SDM	Species distribution modeling
SOTER	Soil Terrain Database
SRTM	Shuttle Radar Topographic Mission
TEM	Toolbox extract multivalues
UNEP	United Nations Environment Programme
WISE	World Inventory of Soil Emission Potentials

ÚVOD

Medicago truncatula je po mnoho dekád v zájmu biologů a evolučních genetiků z celého světa. V dosavadních studiích byla zkoumána především z pohledu genetiky, kdy se podařilo sekvenovat její genom a dokonale popsat vnitřní strukturu. V poslední době se začíná zkoumat i z pohledu adaptace na klimatické podmínky nebo využití v zemědělství.

Velkým odborníkem na řád bobovité je olomoucký botanik doc. Ing Petr Smýkal, PhD. z Katedry botaniky Univerzity Palackého v Olomouci. V současné době probíhá pod jeho vedením výzkum za účelem zjištění vlivu přírodních faktorů na dormanci semen bobovitých rostlin. Výzkum je realizován za spolupráce botaniků, geoinformatiků a pracovníků Výzkumného ústavu rostlinné výroby Dormance představuje stav, kdy jsou semena ve stavu klidu a nekličí. Nejčastější příčinou může být vysoký obsah hormonů inhibiční povahy (Procházka a Šebánek, 1997). Semena ve stavu dormance mohou vyklíčit díky vlivu působení environmentálních faktorů, jako jsou například teplota a srážky, které mohou za určitý čas ukončit stav dormance (Korčáková, 2011).

Výsledek bakalářské práce je součástí výzkumu dormance semen prováděným společně s Katedrou botaniky Univerzity Palackého v Olomouci v rámci projektu GA16-21053S *Využití přístupů ekologické genomiky k poznání adaptivního významu dormance semen u bobovitých rostlin* financovaného Grantovou agenturou ČR.

1 CÍLE PRÁCE

Cílem práce je provést analýzu geografického rozšíření *Medicago truncatula* v oblastech kolem Středozemního moře (Evropa, severní Afrika a Blízký východ) s využitím dostupných datových zdrojů a GIS analýz.

V rešeršní části budou blíže popsány všechny datové sady vstupující do analýz. Dále bude popsána problematika modelování prostorové distribuce druhů a dat potřebných k jejímu vyhotovení.

Práce si dále klade za cíl shromáždit dostupné datové zdroje vhodné pro modelování geografického výskytu druhu na základě dat o současném výskytu. Dále provést porovnání těchto dat s ohledem na prostorové rozlišení, pokryvnost, vhodnost dat pro výzkumné účely a metadata. Na základě uvedených dat proběhne analýza environmentálních podmínek výskytu studovaného taxonu. Výsledky budou vizualizovány a prezentovány ve formě tabelárních a grafických výstupů. Bude vytvořena datová sada obsahující nálezová data s doplněnými environmentálními faktory.

2 METODY A POSTUPY ZPRACOVÁNÍ

Po studiu literatury zabývající se prostorovou statistikou a seznámení s vědeckými publikacemi věnující se *Medicago truncatula*, bylo možné začít vyhledávat a zpracovávat datové sady a následně vytvořit datovou sadu, přehledovou statistiku a mapové výstupy. V následující kapitole budou popsána použitá data a postupy celé práce.

2.1 Použitá data

Datový obsah práce lze rozdělit do 3 hlavních skupin. Hlavní data byla tvořena sadou náleзовých dat *Medicago truncatula*, která pochází z výzkumného zemědělského institutu INRA Montpellier ve Francii. Druhou velkou skupinu tvořila volně dostupná environmentální data, která byla ve formě rastrových datasetů. Třetí skupina obsahovala data získaná analýzou primárních dat (sklon, orientace). V příloze 1 bakalářské práce se nachází kompletní přehled použitých dat. V tabulce 1 je uveden základní přehled použitých dat. Blíže popis lze nalézt v kapitole 3.3 nebo v příloze 3i, kde je uveden bližší popis jednotlivých atributů.

Tab. 1 Základní přehled použitých dat

název	WorldClim v2.0	ASTER GDEM 2.0	SoilGrids	HWSD 1.2	CCI land cover 2015	GLC 09
téma	srážky/teplota /bioklimatické charakteristiky	digitální výškový model	půdní charakteristiky	půdní charakteristiky	krajinný pokryv	krajinný pokryv
prostorové rozlišení	1 km/pixel	30 m/pixel	1 km/pixel	1 km/pixel	300 m/pixel	300 m/pixel
časové rozlišení	1970–2000	2011	-	-	2015	2009
jednotky	mm/°C	m n. m.	*	*	22 kategorií	23 kategorií

*podrobnější informace (Příloha 3i)

2.2 Použitý software

Hlavní náplň práce představovalo zpracování dat. Pro zpracování bylo potřeba použít větší množství použitých software k jejich zpracování.

Správa dat

Velká část práce spočívala ve vyhledávání, stahování a zpracování obsáhlých rastrových datasetů. Rastrů bylo častokrát velké množství a pro usnadnění práce a zrychlení stahování ze serverů, kde to bylo možné, byla data stažena pomocí FTP protokolu. K tomuto účelu byl použit program FileZilla ve verzi 3.23.01. Rastry byly dávkově zpracovány pomocí programu Total Commander ve verzi 9.0. Výškový model (ASTERGDEM v2.0) byl stažen pomocí software Bulk Download tool, který pracuje na podobném principu jako FileZilla.

Práce s daty

Bodová data byla prvotně zpracována v tabulkovém procesoru Microsoft Excel 2007. Pro zpracování rastrových a bodových dat bylo použito programu *ArcGIS for Desktop 10.4*, ve kterém byla prováděna vizualizace získaných dat, úprava a filtrace náleзовých

dat a v neposlední řadě samotná extrakce environmentálních faktorů pomocí funkce *Extract multi values to point*. Několikrát byl použit nástroj *Modelbuilder*, k vytvoření toolboxů pro zautomatizování postupů. Bodová data byla prvotně zpracována v tabulkovém procesoru Microsoft Excel 2007.

Statistické a grafické výstupy

Pro výpočet statistických charakteristik bylo použito především softwaru STATISTICA 12 a GeoDa 1.8. Oba software nabízejí potřebné nástroje k vytvoření statistických přehledů, STATISTICA disponuje více možnostmi přizpůsobení výstupů, změnu popisků barev apod. GeoDa je více zaměřena na geografickou stránku a nabízí možnost propojení statistické a prostorové informace a je volně dostupná ke stažení. Oproti programu STATISTICA nabízí možnost importu GIS formátů jako Esri Shapefile, GeoJSON a současně propojení statistických výstupů s mapovým podkladem. Její velká výhoda je tzv. „linking“, díky kterému je možné při výběru prvků v jednom okně zároveň zvýraznit prvky v ostatních oknech. Oproti jiným statistickým software nenabízí tak bohaté možnosti nastavení výstupů, proto byla použita pouze k vyšetření autokorelace a jako doplňkový nástroj k okamžité identifikaci prvků v prostoru.

Grafické výstupy a mapy byly dotvářeny v softwaru INKSCAPE 0.92 a GIMP 2.8.20. Pro tvorbu webových stránek s výsledky práce byl použit PSPad Editor verze 4.6.2. ve kterém byl také vytvářen online vizualizační nástroj za pomoci JavaScript knihovny Leaflet 0.7.3.

2.3 Postup zpracování

Na začátku práce byla provedena rešerše dostupné literatury a vědeckých publikací, které by mohly být v průběhu práce užitečné. Rešerše byla zaměřena jak na popsání samotného taxonu, tak na studium článků zabývajících se alespoň částí postupu aplikovaných v bakalářské práci.

Po obdržení náleзовých dat proběhla jejich validace a oprava chyb. Během řešení práce nebylo dlouhou dobu zcela jasné, která konkrétní data budou nakonec považována za finální pro potřebu pracovníků z Katedry botaniky. Proto bylo pracováno s celým souborem a veškerá práce byla zaměřena na zautomatizování celého procesu získávání informací k bodům nálezu. Bylo tak možné z nově dodaných míst výskytu průběžně extrahovat potřebné hodnoty.

Po úpravě náleзовých dat, byla po konzultaci s botaniky vyhledána environmentální data. Následně probíhalo vyhledání několika zdrojů ke každému tématu (půda, klima, nadmořská výška). Poté bylo několik datasetů vybráno a staženo. Během práce vzniklo několik nástrojů, které řešily mnohdy zdlouhavé manuální postupy. Prvním výsledkem byla datová sada míst sběru rostlin, spolu s environmentálními faktory. Z těchto dat byla vytvořena nová datová sada, ze které byly zpracovány statistické a mapové výstupy. V závěru práce, na základě komunikace s doc. Ing. Petrem Smýkalem, Ph.D, byl vyhotoven online nástroj pro vizualizaci získaných dat.

3 SOUČASNÝ STAV ŘEŠENÉ PROBLEMATIKY

Práce je primárně zaměřena na analýzu geografického rozšíření rostliny *Medicago truncatula*. Tato rostlina je předmětem zájmu biologů a genetiků, kteří si ji vybrali jako modelový organismus pro zástupce čeledi bobovitých především pro její krátký diploidní genom a krátkou regenerační dobu (Buchtová, 2016). Žádná z publikovaných prací se detailně nevěnuje geografickému popisu stanovišť *Medicago truncatula*. Výzkumy a odborné práce se zaměřením na *Medicago truncatula* se věnují primárně zkoumání jejího genomu a případně adaptaci genů na klimatické podmínky. V této kapitole jsou popsány principy modelování prostorové distribuce druhů (SDM), charakteristika rozšíření a popis zkoumaného taxonu. V závěru je přehled volně dostupných environmentálních dat, která jsou často využívána v odborných studiích s geograficko-biologickým zaměřením.

3.1 *Medicago truncatula*

Tolice (*Medicago truncatula*) je jednou z modelových rostlin. Mezi další modelové rostliny patří například houseníček rolní (*Arabidopsis thaliana*) nebo kejklířka skvrnitá (*Mimulus guttatus*).

Základní charakteristika

Medicago truncatula je jednoletý druh patřící do rodu *Medicago* čítajícího celkem 87 druhů (Smýkal a kol. 2014). Jedná se o malou, jednoletou, samosprašnou rostlinu vyskytující se v oblasti Středomoří. Je známá především díky svému využití v genetickém výzkumu. Má trojčetné listy a dosahuje menšího vzrůstu mezi 10–60 cm. Listy jsou oválné s rozměry 1–2 cm. Zpravidla má 1–4 světle žluté květy viz (obr. 1). Lusky obsahují 4–12 žlutých semen. Přibližná váha 1000 semen je 3,9 gramu (FAO, 2017).

Charakteristika geografického rozšíření

Rostlina má původní rozšíření ve Středomoří a západní Asii (Lesins a Lesins, 1979). Vzhledem k využití jako pícnina, se *Medicago truncatula* podobně jako kulturní a více rozšířená příbuzná vojtěška (*Medicago sativa*), rozšířilo po celém světě, a to do oblastí, kde se vyskytuje středomořský typ klimatu. Rostlina se vyskytuje např. v jižní Austrálii, Jižní Africe nebo v Chile (Piano and Talamucci, 1996).

Středomořský typ klimatu se dá obecně charakterizovat jako oblast s mírnými vlhkými zimami a horkými suchými léty. Nachází se mezi 28° a 45° zeměpisné šířky a je obvykle přechodovou zónou mezi mírným a tropickým klimatem (Dallman, 1998). Budedenhagen (1990) definuje středomořské klima, jako oblast s průměrným ročním úhrnem srážek 250-900 mm, kde se nejméně 65 % vyskytuje v zimním období. Aschmann (1973) vyčleňuje středomořské klima, jako oblast, kde je alespoň jede měsíc s průměrnou teplotou pod 15 °C, ale počet hodin za rok s teplotou nižší než 0 °C by neměl překročit 3 % roční sumy teplot. Středomořské klima zabírá přibližně 2 % celosvětové souše (Thrower and Bradbury, 1973).

Na světě je pět regionů, mající pravé středomořské klima (Dallman, 1998).

- Středomoří (jižní Evropa, severní Afrika, východní Asie)
- Část jižní Austrálie
- Nejjižnější část Jižní Afriky
- Kalifornie
- Centrální oblast Chile

Vegetační podmínky

Rostlina se vyskytuje na různých druzích půdy, od písčitých až po jílovité, ale především upřednostňuje dobře odvodněné, neutrální až zásadité půdy s pH 6–8 (Evans, Dear a O'Connor, 1990). Nemá ráda půdy kyselé s pH pod 6 a půdy které nejsou dostatečně vlhké, pokud není dobře zakořeněná. Nejlépe roste ve středomořském klimatu s horkými suchými léty a mírnými vlhkými zimami, se srážkami okolo 250–600 mm za rok. Není odolná vůči zimním mrazům (Brandsaeter a kol., 2002). Životní cyklus je adaptovaný, aby se vyhnul letním suchům. Po vykvetení a vytvoření semen rostlina v létě uhynie a lusky zůstávají ležet na zemi, než na podzim semena znovu vyklíčí.

Význam

Medicago truncatula pro svoji příbuznost s tolicí vojtěškou (*Medicago sativa*) byla vybrána vědci jako modelová rostlina zastupující čeleď bobovité. Vyznačuje se malým (500 Mbp) a dobře popsaným genem, samosprašností a rychlým růstem. Rostlina je diploidní s $x=2n=16$ chromozomy. Pro své vlastnosti je často využívána biologií a genetikou (Colebatch a kol., 2002). Zejména její schopnost fixovat vzdušný dusík a zlepšovat půdní vlastnosti je z pohledu zemědělství velmi důležitá. Francouzský institut pro zemědělský výzkum (INRA) vybral 4 kultivary pro možné využití ve 3 oblastech: zlepšení kvality a kvantity pastvin, zvětšení půdní odolnosti vůči erozi nebo v oblasti bezorebního zemědělství (INRA, 2017).



Obr. 1 Květ *Medicago truncatula* (zdroj: <http://www.fao.org>)

3.1.1 Biologické studie *Medicago truncatula*

Žádná z nalezených studií se dosud podrobně nezabývala geografickým rozšířením *Medicago truncatula*. V tomto směru je výsledek práce unikátní. Svým pojetím představuje podrobnou studii, zaměřenou na získávání dat environmentálních podmínek daného taxonu. *Medicago truncatula* je objektem zájmu v několika oblastech, například jako modelová rostlina v genetickém inženýrství, rostlina pro zkoumání vlivu environmentálních podmínek na její genom a v zemědělství díky své schopnosti vázat dusík. Je často využívána v bezorebním zemědělství. Má schopnost fixovat až 120 kg/ha a tím zvyšovat úrodnost půdy (Uzun a Aydin, 2004).

Nejdále v analýze geografického rozšíření došla studie Bounejmate, Beale a Robson (1992), která se zaměřovala na vztah rozmístění stanovišť rostlin rodu *Medicago* a klimatických a půdních faktorů v oblasti Maroka. Půdní data byla sesbírána ručně a klimatická data byla použita ze zveřejněných zdrojů, např. záznamů meteorologických stanic. Bylo zjištěno, že na rozmístění stanovišť mají větší vliv půdní než klimatické faktory.

Stručně charakterizované prostředí, ve kterém se *Medicago* nachází, se dá nalézt v několika pracích (Clarkson a Russell, 1975; Moreau a kol. 2007; Pierre a kol. 2008), které jsou primárně zaměřeny na klimatická data. V článku (Burgarella a kol. 2016) je podrobně popsána studie, která zkoumá, jakým způsobem se odráží klimatické podmínky na genetické vybavě rostlin. Zdroj klimatických dat představovala databáze WorldClim a nálezová data pocházela z genobanky INRA Montpellier. Na vzorku 174 rostlin se podařilo zachytit dvě genově odlišné skupiny, jedna byla z východní části Středomoří (Maroko, jižní část Španělska), druhá představovala zbytek středomořské oblasti. Jeden ze zajímavých výsledků bylo zjištění, jaká je doba kvetení v závislosti na poloze a genetické skupině. Korelace mezi dobou kvetení a zeměpisnou šířkou byla pozitivní, u rostlin s vyšší zeměpisnou šířkou byl počátek posunut na později. Naopak korelace se zeměpisnou délkou se nepotvrdila. Obzvláště silná korelace doby kvetení byla zjištěna s jarními a letními srážkami. Vztah mezi nadmořskou výškou nebyl průkazný. Tato studie navazovala na předchozí studii s podobným zaměřením (Yoder a kol. 2014), která se také věnovala studii genetického založení adaptace na klimatické podmínky, především z pohledu identifikace možných kandidátních genů.

Zemědělství

Jednou z důležitých vlastností, zvláště pokud se jedná o plodinu, která je používána jako krmivo, je klíčivost případně dormance semen. Zajímavá studie na téma klíčivost a překonání dormance semen je práce (Uzun a Aydin, 2004). V této studii je na několika rostlinách z rodu tolíce (*Medicago*) a jetel (*Trifolium*) testováno, jaká metoda úpravy semene dokáže nejvíce překonat stav dormance. Semena byla sesbírána na volné pastvině v Turecku. Pro zvýšení klíčivosti bylo použito několik metod: mechanické narušení obalu semena, máčení ve vodě o teplotě 75 °C, zahřátí na 40 °C po dobu pěti minut, zchlazení na 3 °C a další. Po provedení těchto kroků byla semena vystavena klíčicímu cyklu se střídáním světla a tmy po 12 hodinách při stálé teplotě 25 °C. Klíčení semen bez provedení úpravných opatření bylo jen v rozmezí 11–43 % v závislosti na konkrétní rostlině. Největší vliv napříč všemi druhy mělo mechanické narušení obalu semena, pro většinu druhů se klíčivost zvedla na více než 70 %. Výsledky studie ukázaly, jakým způsobem je možné překonat dormanci a zlepšit tak zemědělské výnosy.

V práci (Moriuchi a kol. 2016) bylo testováno, jakým způsobem se adaptuje *Medicago truncatula* na salinitu půdy z hlediska dvou faktorů: množství biomasy

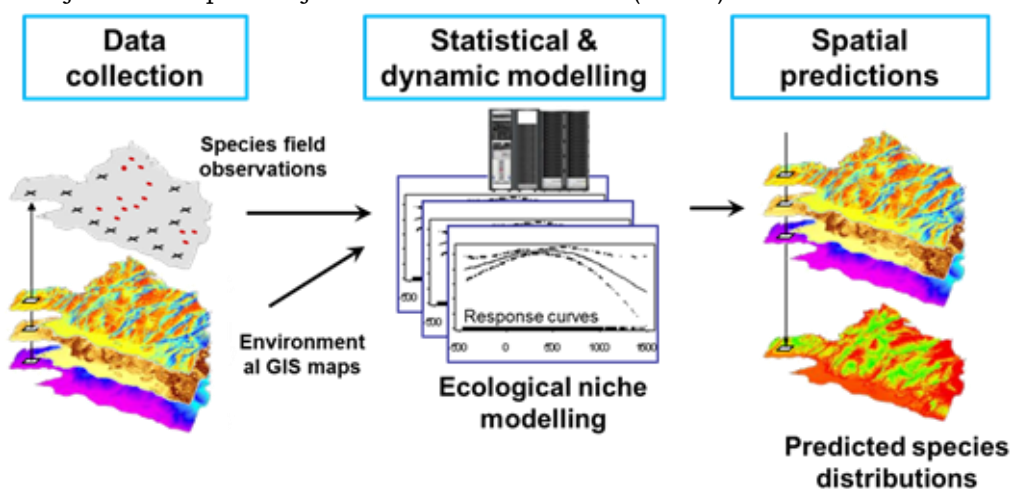
a reprodukce. Pro testování bylo použito několik populací rostlin z oblasti severního Tuniska z oblasti se slanou a neslanou půdou. Pro každou populaci bylo náhodně vybráno 10 genotypů pro další testování. Ze sesbíraných semen byly vypěstovány mateřské rostliny, v přibližně stejně slané půdě, z jaké byly sesbírány původní semena. Ve výsledku bylo prokázáno, že adaptaci dalších potomků ovlivňuje prostředí matečních rostlin. Potomstvo vykazovalo větší nárůst biomasy ve slaném prostředí, pouze pokud matečné rostliny vyrůstaly také ve slaných podmínkách. Další prokazatelný výsledek byl, že rostliny pěstované ve slaném prostředí s původem ve slaném prostředí klíčily a plodily dříve než rostliny, které se vyskytovaly v neslaném prostředí a byly taktéž pěstovány ve slaných podmínkách. Závěrem vyplynulo, že pro populace adaptované na slanou půdu, pěstování mateřských rostlin ve slaném prostředí urychluje životní cyklus, zatímco pěstování potomstva z těchto rostlin ve slaném prostředí, je příčinou vyšší tolerance. Předpoklad je, že s rostoucím počtem dalších generací a vystavování se slanému prostředí tolerance stoupá.

3.2 Modelování prostorové distribuce druhů

Tato práce je zaměřena především na popsání stanovištních podmínek a způsobu, jak tato data získat a zpracovat. Nálezovým datům jsou přiřazovány environmentální faktory na základě globálních datasetů. Výsledná datová sada může být brána jako základ pro další zkoumání problematiky a především může sloužit jako podklad pro prostorovou modelaci druhů (SDM).

Principy modelování prostorové distribuce druhů

Modelování prostorové distribuce druhů (SDM) je hlavním principem řešení mnoha problémů v biogeografii, evoluční ekologii, rozšiřování populací a biologii ochrany životního prostředí (Drake, 2014). Velmi často je používáno také synonymum modelování ekologických nik. Ekologická nika představuje interakce mezi druhem a biotickými a abiotickými faktory prostředí a řadí se k základním ekologickým pojmům (Polechová a Storch, 2008). SDM modely mohou být rozděleny do dvou druhů na modely korelativní a mechanistické (Dorman, 2012). Korelativní modely propojují nálezová data s environmentálními vrstvami a výsledkem jsou mapy, kde se daný druh vyskytuje, nebo by se mohl potenciálně vyskytovat (Kumar a kol. 2014). Environmentální vrstvy nejčastěji představují rastry například úhrn srážek, průměrná teplota a jiné. Princip SDM je znázorněn na obrázku (obr. 2).



Obr. 2 Schéma SDM (zdroj: <https://www.unil.ch>)

3.2.1 Typy dat vstupujících do SDM

Jak již bylo zmíněno výše, SDM využívá kombinaci jak nálezových, tak i environmentálních dat (kapitola 3.2.). Z pohledu výsledků je výběr environmentálních vrstev zcela klíčový. Kniha Franklin (2010) představuje ucelený „návod“ a obsahuje popsané postupy a metody SDM od výběru dat až po příklady statistických modelů. Tento text může sloužit jako první publikace pro seznámení s SDM. Podle Franklin (2010) lze mezi nejčastěji používaná data v SDM zařadit:

Výškopisná data

- nadmořská výška: ovlivňující teplotu a srážkový režim,
- sklon: ovlivňující vodní režim, vlhkost a texturu půdy,
- orientace: sklon ke světovým stranám ovlivňuje míru ozáření povrchu,

Půdní data

- pH,
- objemová hmotnost půdy,
- obsah dusíku, minerálů,
- vlhkost půdy,

Vegetační data

- land cover,
- land use,
- albedo,
- LAI (Leaf area index),
- NDVI (NormalizedDifference Vegetation Index).

3.2.2 Dálkový průzkum a SDM

S rozvojem moderních technologií se ve sběru dat stále více prosazuje Dálkový průzkum Země (DPZ). Využitím DPZ při modelování prostorové distribuce se objevuje v mnoha studiích (Rushton a kol. 2004; Bradley a Fleishman, 2008; Adams a kol. 2008). Díky dálkovému průzkumu můžeme získat data pro rozsáhlá území za krátký čas. Velkou výhodou je, že mohou být pořizována opakovaně, navíc i na nepřístupném území. Zpravidla se snímací zařízení nachází na družici.

Velká část dat, se kterými bylo pracováno v rámci bakalářské práce, byla získána právě pomocí metod DPZ.

Využití dat DPZ

V článku He a kol. (2015) jsou shrnuty aktuální trendy v pořizování dat pomocí DPZ pro účely SDM. Jsou zde uvedeny příklady satelitů a druhů dat, které je možné s nimi pořídit. Tato práce tvoří dobrý základ pro seznámení se s aktuálními možnostmi využití DPZ.

Data získaná pomocí DPZ mají široké možnosti využití. Od monitoringu krajiny až po sledování migrace zvířat. Výzkumem rostlinných druhů pomocí DPZ se v poslední době zabývala celá řada prací (Huang, 2009; Zhang a Jia, 2013; Reddy, 2015). Zejména díky hyperspektrálním snímkům je možné rozpoznat mnoho rostlinných i živočišných druhů a také zjistit další velmi podrobné informace, jako je například obsah vody v listech nebo množství dusíku (He a kol, 2015). Přes všechny výhody má DPZ své omezení. Jeho největší síla se projeví teprve v kombinaci s měřením v místě výskytu. Pomocí dálkového průzkumu je možné monitorovat i vodní prostředí včetně benthické zóny, ale pouze do určité hloubky, hlubší oblasti již vyžadují jiné formy monitoringu (Geller a kol, 2017).

Reddy (2015) se ve své práci zmiňuje, jakým způsobem mohou být využita DPZ data při monitoringu rozšiřování invazivních rostlinných druhů. Využití dálkového průzkumu může být velmi finančně a časově efektivní. Zejména pozemní průzkum ve spojení s DPZ může přinést výrazně lepší výsledky. Xie a kol. (2008) se zabývají mapováním vegetace za pomoci dat DPZ. Práce pokrývá široký rozsah témat. Od popisu používaných satelitů až po způsoby klasifikace hyperspektrálních dat.

Jedním z hlavních přínosů DPZ pro výzkum rostlin jsou vegetační indexy. Na základě spektrálního chování vegetace mohou pomoci zvýraznit pokrytí území vegetací, nebo určit druh vegetace. Mezi nejpoužívanější patří NDVI nebo LAI.

3.2.3 Vliv měřítka na výsledek SDM

Před samotnou analýzou rozšíření jakéhokoli druhu, ať už se jedná o rostlinu či živočicha, je nutné si určit, v jaké podrobnosti a k jakému účelu má být studie provedena. Požadované měřítko ovlivní data, která budou do analýzy vstupovat, a do jisté míry může mít vliv na celkový výsledek.

Při zvolení hrubého měřítka se dobře zachycují trendy ve velkých oblastech. Hrubé měřítko je vhodné pro vytvoření výskytu biomů. Mohou se využít snáze dostupná globální data.

Lokální měřítko naopak dobře zachycuje místní činitele jako vliv reliéfu apod. Klimatická data jsou často dostupná pouze v měřítku, které je vhodné pro výzkum oblastí o rozloze např. 1 000–10 000 km², zatímco růst rostlin je obvykle zkoumán v mnohem menším měřítku např. na ploše 100–10 000 m² (Peterson a kol. 1998). Proto je někdy nutné při absenci podrobných dat a požadavku na správný model přistoupit k tzv. downscalingu. Downscaling je proces přenosu dat z modelů o hrubém měřítku do modelů s jemnějším měřítkem (Flint, 2011).

V kopcovitém terénu je důležité mít data s jemným měřítkem. Hrubé měřítko zde nemusí dobře zachycovat terénní vlivy. Velmi podrobné měřítko v rozmezí (0,5–10 m) můžeme použít pro popsání klimatu blízko zemského povrchu.

Randin a kol. (2009) ukázal, že SDM založené na datech s jemným měřítkem mohou poskytnout výrazně odlišné odhady dopadů klimatické změny na rozšíření taxonu než při použití hrubého měřítka. Při použití jemného měřítka bylo odhadováno vyšší přetrvání druhu na daném stanovišti oproti použití hrubého měřítka. Nicméně tato studie byla zaměřena pouze na alpské rostliny. Práce Trivedi a kol. (2008) ukázala opačné výsledky. Zaměřovala se na to, jaký bude v budoucnu vliv klimatu na populaci horských rostlin na území Skotska. Bylo využito dat s jemným (50 m/pixel) a hrubým rozlišením (50 km/pixel). Výsledek modelu při použití podrobnějšího měřítka předpokládal úhyn 7 z 10 testovaných druhů. Při použití dat s hrubým rozlišením byl předpokládán úhyn jen jednoho druhu.

3.3 Datové zdroje

V dnešní době se na Internetu nachází celá řada volně dostupných environmentálních dat, která jsou vhodná ke geografickým analýzám nebo k SDM modelování. V této kapitole budou popsány volně dostupné zdroje, které se nejčastěji vyskytují v odborných studiích s geograficko-biologickým zaměřením, a velká část z nich byla použita také v této práci. Nejpodrobnější globální volně dostupná data o výškopisu představují ASTER GDEM2 (30 m/pixel), SRTM (30 m/pixel) nebo méně známý ALOS World 3D (30 m/pixel). Za jedny z nejpodrobnějších globálních půdních dat lze považovat například Harmonised World Soil Database (HWSD) nebo SoilGrids. Alternativou k velkým rastrovým půdním databázím v určitých případech, kdy není potřeba velké podrobnosti, může být digitalizovaná FAO-UNESCO Soil Map of the World v měřítku 1 : 5 000 000. Pro získání klimatických dat je v drtivé většině případů používána databáze WorldClim. Jako rozcestník při hledání volně dostupných dat může sloužit webová stránka <https://freegisdata.rtwilson.com>.

3.3.1 WorldClim

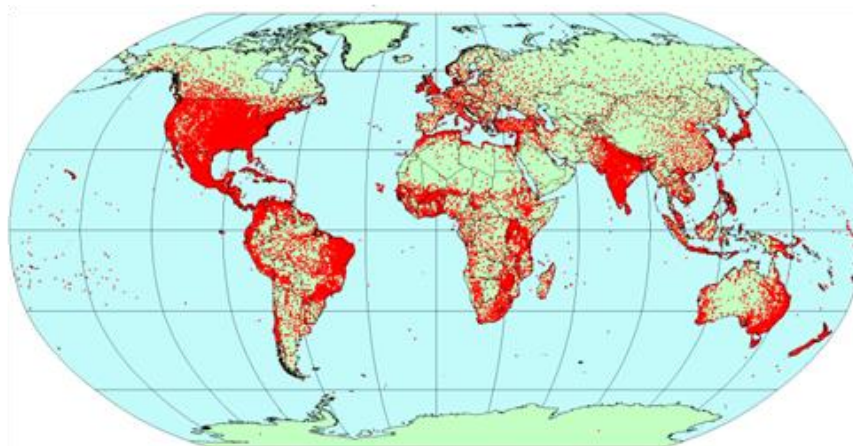
WorldClim databáze slouží jako základní zdroj environmentálních dat v mnoha odborných studiích například v práci, která se zabývá adaptací *Medicago truncatula* na klimatické podmínky (Yoder a kol, 2014). WorldClim se vyskytuje jako základní zdroj dat i v dalších pracích např. (Burgarella a kol, 2016), (Estrella a kol, 2012), (Cao a kol, 2016) a mnoha dalších.

WorldClim představuje rozsáhlou sadu rastrů, které pokrývají celý svět. V době psaní této práce byly dostupné 3 verze (1; 1,4; 2.0). Verze se od sebe liší datovým obsahem. U verze 2.0 jsou dostupná data pouze tzv. current conditions, která byla interpolovaná za období 1970–2000. Předcházející verze využívaly data interpolovaná za období 1960–1990. Rastry mají nejpodrobnější rozlišení 1 km² na rovníku. Dostupná prostorová rozlišení (30 arc-seconds~1km, 2,5 minutes~4,5 km, 5 minutes~9 km, 10 minutes~18,5 km) na rovníku. Souřadnicový systém je WGS84 (EPSG: 4326).

Veškerá data dostupná na www.worldclim.org jsou k dispozici zdarma ke stažení. Obsah je chráněn licencí *Attribution-ShareAlike 4.0 International (CC BY-SA 4.0)*.

Rastry je možné stáhnout ve formátech *Generic grid*, *Esri grid*, *GeoTIFF*. Při rozlišení 30 úhlových vteřin a formátu GeoTIFF mají srážky pro celý svět velikost přibližně 1,5 GB. Obsah je možné stáhnout po dlaždicích o velikosti 30 × 30°.

Za projektem stojí řada odborníků a institucí např. University of California, Berkeley a další subjekty. Gridy byly vytvářeny pomocí interpolace několika databází měřících stanic (obr. 3) z celého světa (GHCN, FAOCLIM, CLINO, CIAT a dalších). Každý bod představoval jednu měřicí stanici. Data o srážkách pochází ze 47 554 stanic, průměrné teplotě (24 542 stanic) minimální a maximální měsíční teplotě (14 835 stanic). Metodika tvorby dat je podrobněji popsána v publikaci (HIJMANS a kol, 2005).



Obr. 3 Stanice s hodnotami pro srážky (zdroj: <http://www.worldclim.org/>)

Tematické členění dat

- **Současný stav** (*current conditions*): data byla interpolována z měření za období 1960-1990 (ver.1.4) Najdeme zde údaje o teplotě (min., max., prům.) srážkách (mm) a bioklimatické charakteristiky. Jedná se o 19 odvozených údajů z teploty a srážek.
- **Budoucí stav** (*future climate data*) nabízí data z 20 klimatických modelů (GCM) počítaných po celém světě pro 4 scénáře vývoje skleníkových plynů (RCP). Data jsou dostupná v rozlišení (10; 5; 2,5) úhlových minut, a 30 vteřin. Tematicky obsahují průměrné maximální a průměrné minimální měsíční teploty a měsíční srážky. Časové rozlišení představuje období 2041–2060 a 2061–2080.
- **Minulý stav:** (*paleoclimate data*) data získaná z globálních klimatických modelů (GCM) dostupná pro dvě období: Střední holocén (6000 let př. n. l.) a poslední Glaciál (před 22 000 lety). Data jsou dostupná v rozlišení od 30 vteřin do 10 úhlových minut. Tematicky obsahují průměrné maximální a průměrné minimální měsíční teploty a měsíční srážky.

3.3.2 Harmonised World Soil Database v 1.2

HWSD představuje jeden z mála globálních zdrojů půdních dat a je velmi často využíván. Příkladem je práce (Cao a kol, 2016), která se zabývá modelováním distribuce habitatu *dřínu lékařského*.

Harmonised World Soil Database (HWSD) je výsledkem spolupráce organizací (FAO, IIASA, ISRIC-World Soil Information, Institute of Soil Science, Chinese Academy of Sciences (ISSCAS) a the Joint Research Centre of the European Commission (JRC). HWSD obsahuje velké množství atributů, které jsou blíže popsány v práci (Nachtergaele a kol. 2009).

Jedná se o databázi půdních typů s pokrytím celého světa kromě Antarktidy. Databáze je ve formátu Microsoft Access dbf. Obsahuje přes 15 000 unikátních půdních jednotek. Pro použití v rámci GIS se dá databáze připojit na rastr s půdními jednotkami o rozlišení 30 úhlových vteřin (1 km²) ve formátu GeoTIFF. Databáze kombinuje několik zdrojů půdních dat (SOTER, ESD, Soil Map of China, WISE). Data pocházejí z několika časových období v závislosti na databázi. Celá databáze je poskytována zdarma ke stažení, pod podmínkou nekomerčního užití pro vzdělávací účely (FAO, 2017).

3.3.3 SoilGrids

Jedná se o unikátní projekt poskytující mnoho půdních informací (pH, texturu, zrnitost a další) v několika hloubkách (0, 5, 15, 30, 60, 100, 200 cm). Gridy jsou dostupné v rozlišení 250 a 1000 m/pixel. Data, ze kterých jsou SoilGrids generována, představují půdní profily (150 000), data ze senzoru MODIS a výškopisná data ze SRTM. Data je možné stáhnout po jednotlivých dlaždicích 200x200 km nebo přes FTP server (ISRIC, 2017).

3.3.4 Shuttle Radar Topography Mission

Model byl vytvořen v rámci spolupráce organizací NASA a NGA. Historicky se jedná o první celosvětový digitální model terénu s vysokým prostorovým rozlišením. Vytvořený byl v roce 2000 za pomoci radaru umístěného na raketoplánu Endeavour. Data pokrývají povrch mezi 60° s. š. a 56° j. š. V roce 2000 byla dostupnost modelu omezena pouze na území USA. V roce 2003 byl zveřejněn méně podrobný model s rozlišením 90 m/pixel pro velkou část světa. Poslední zveřejněná verze pro celý svět má rozlišení 30 m/pixel (NASA, 2017).

3.3.5 ASTER Global Digital Elevation Model (GDEM2)

GDEM2 je vyvíjen ve spolupráci Americké NASA a METI (japonské ministerstvo ekonomiky, obchodu a průmyslu). Byl snímán pomocí senzoru ASTER na satelitu Terra. Data je možné pro osobní využití stáhnout zdarma.

Výškový model ASTER GDEM druhé generace byl vytvořen v říjnu roku 2011. Ve druhé verzi GDEM přibýlo při zpracovávání 260 tis. stereo-dvojic. Díky tomu je vylepšeno nejen pokrytí zemského povrchu, ale i prostorové rozlišení a horizontální a vertikální přesnost. Vodní plochy v modelu jsou zobrazeny reálněji než v první verzi a zachovány jsou jen ty s průměrem o minimální velikosti 1 km. Data pokrývají zemský povrch mezi 83° s. š. a 83° j. š. Model je dostupný v rozlišení 1 úhlová vteřina ve formátu GeoTIFF (GISAT, 2017).

3.3.6 CCI Land cover 2015

ESA (European Space Agency) nabízí mapy krajinného pokryvu derivované pro každý rok v období 1992–2015. Data jsou ve formě rastru o rozlišení 300 m/pixel s dostupností pro celý svět. Nejnovější verze obsahuje 22 tříd krajinného pokryvu klasifikovaných ze snímků pořízených družicí Proba-V (ESA, 2017).

3.3.7 500 meter MODIS Land Cover Maps

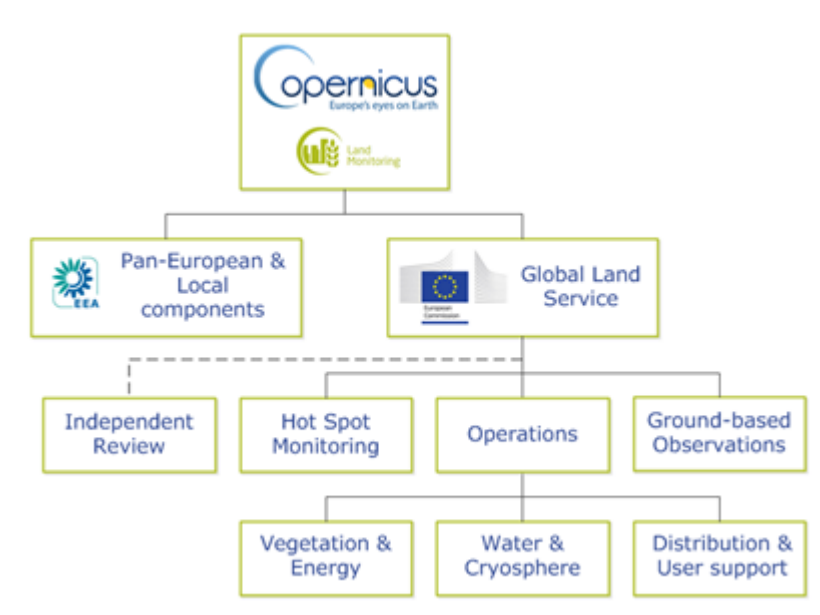
Na stránkách USGS (United States Geological Survey) je možné nalézt land cover klasifikovaný na základě dat ze satelitu MODIS v rozlišení 500 m/pixel. Data byla derivována za časové období 2001–2010. Produkt obsahuje 17 kategorií klasifikace krajinného pokryvu na základě dat 5.1 MCD12Q1 (USGS, 2016).

3.3.8 Copernicus Global Land Service (GLS)

Global Land Service je součástí Evropského projektu Copernicus, který je hlavním evropským programem pro monitoring planety Země. Struktura programu Copernicus je znázorněna na obrázku 4. Program využívá kombinaci satelitů Sentinel a pozemních senzorů. Program zahrnuje 6 jádrových oblastí. Data poskytovaná v rámci celého světa jsou spravována European Commission Joint Research Centre (EC JRC).

Produkty GLS jsou členěny do 4 základních kategorií: energie, vegetace, voda a kryosféra. Data jsou dostupná ve formě rastrů s prostorovým rozlišením 300 m a 1 km na pixel. V této práci byl použit Normalized Difference Vegetation Index v rozlišení 1 km/pixel. Data jsou dostupná od roku 1998 do současnosti. Je možné je bezplatně stáhnout pro vlastní potřebu. Údaje jsou zaznamenávány pomocí senzorů SPOT-VGT a PROBA-V v 10 denní aktualizaci. Rastry jsou dostupné ve formátech GeoTIFF, HDF5, NETCDF a ENVI (LAND COPERNICUS, 2017).

V době psaní této práce byla zhruba polovina nejnovějších dat s vysokým rozlišením ve fázi vývoje a nebyla dostupná.



Obr. 4 Schéma programu Copernicus (zdroj: <http://land.copernicus.eu/global>)

3.3.9 Corine Land Cover (CLC)

CLC je produkt programu Pan-European v rámci programu Copernicus, spravovaného European Environment Agency (EEA).

Jedná se o databázi s typy krajinného pokryvu členěnou do 44 kategorií. V rámci ČR se vyskytuje pouze 28 těchto kategorií. CLC je dostupná pro roky 1990, 2000, 2006, 2012.

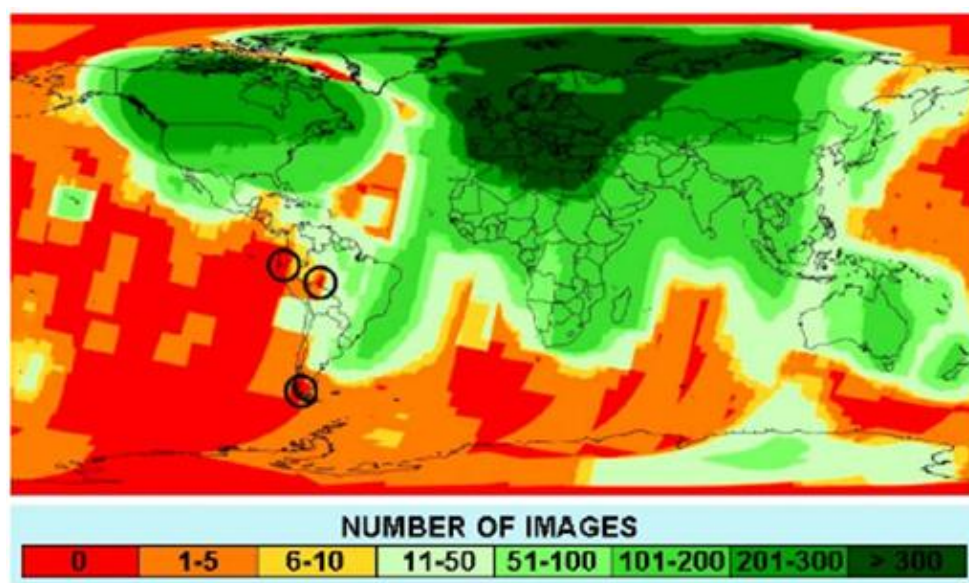
Databáze je k dispozici ve formě rastrů ve formátu GeoTIFF s rozlišením 100 nebo 250 m na pixel. Dostupná je i vektorová databáze (Esri Geodatabase, SQLite database). Verze 2012 byla vytvořena z dat sesbíraných v roce 2011. CLC je možné zdarma stáhnout ze stránek <http://land.copernicus.eu>. Doba zpracování po zveřejnění CLC byla u poslední verze 2 roky. Nejmenší jednotka pro mapování plošných prvků je 25 ha a 100 m šířky pro liniové prvky (LAND COPERNICUS, 2017).

3.3.10 GlobCover 2009

GlobCover je digitální mapa krajinného pokryvu pro celý svět. Jedná se o výsledek spolupráce mezi Evropskou vesmírnou agenturou (ESA) a dalšími organizacemi jako jsou například JRC, EEA, FAO, UNEP, GOFC-GOLD a IGBP. V době psaní práce byla dostupná již verze 2.3, jež navazovala na předchozí verzi z roku 2005, která posloužila jako referenční.

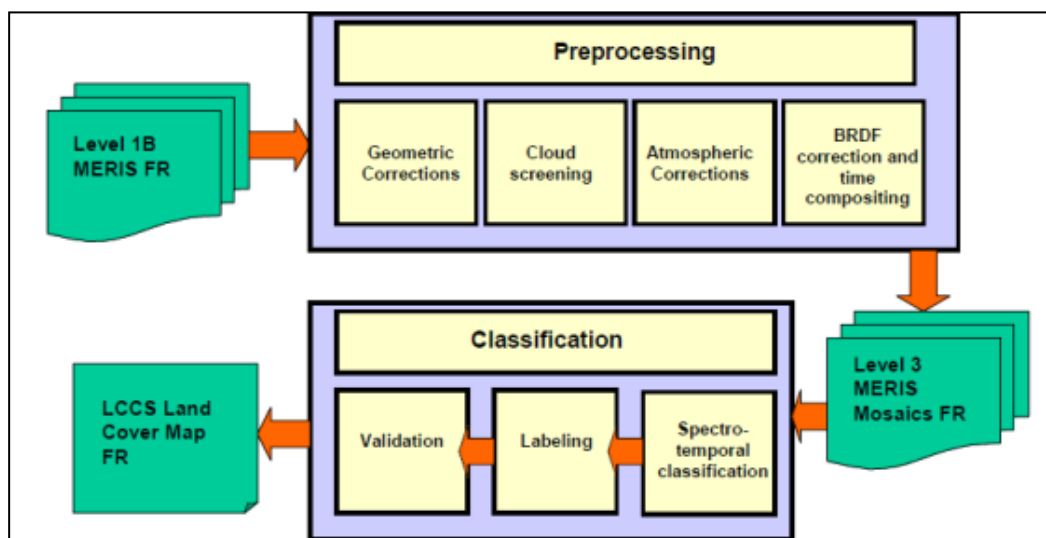
Data jsou dostupná ve formě GeoTIFF včetně souboru lyr. pro stylování v *ArcGIS for Desktop*. Rastr má prostorové rozlišení 300 m/pixel. Záznamy byly získány pomocí senzoru MERIS neseného družicí ENVISAT za celé období roku 2009. Celkem GlobCover obsahuje 23 kategorií členění krajinného krytu, které vycházejí ze systému klasifikace organizace FAO.

I přes výborné rozlišení pro celý svět, si tato databáze nese řadu problémů. Senzor MERIS nebyl navržen pro tak rozsáhlý úkol, jako je pořizování časových řad z celého světa. Pokrytí snímky tak není rovnoměrné a také kvůli oblačnosti je na některých místech málo dat (obr. 5).



Obr. 5 Pokrytí snímky ze senzoru Meris (zdroj: <http://www.gisat.cz>)

Zpracování snímků ze senzoru MERIS až po výslednou mapu je rozděleno do dvou částí (obr. 6). Pre-processing (geometrické a atmosférické korekce) a samotná klasifikace. Před klasifikací je svět rozdělen do 22 oblastí, které jsou z hlediska dálkového průzkumu podobné. Klasifikace se provádí pro každou oblast zvlášť (GISAT, 2017).



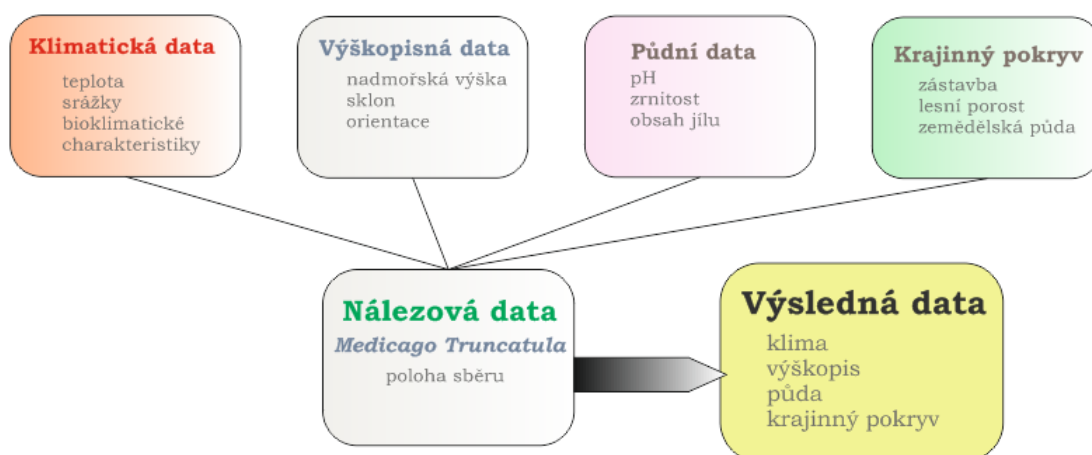
Obr. 6 Schéma zpracování dat GlobCover (BONTEMPS, 2011)

4 VLASTNÍ ŘEŠENÍ

Hlavní cíl práce představuje analýza stanovištních podmínek. K dílčím cílům patří vytvoření nové datové sady a následná prezentace pomocí mapových výstupů. Pro analýzu podmínek bylo vybráno několik datových sad, které byly popsány v podkapitole 3.3 Datové zdroje. V kapitole 4 Vlastní řešení jsou uvedeny postupy, jakým způsobem bylo s daty pracováno. Dále jsou zde popsány problémy, které nastaly během zpracování dat a popis statistického zpracování.

4.1 Prvotní zpracování dat

Po výběru dat bylo možné začít s jejich zpracováním. Data obsažená v práci je možné rozdělit do několika skupin. První skupinu tvoří nálezová data, která jsou jedinou vektorovou vrstvou v celé práci. Další skupinu představují environmentální rastry a poslední skupina je tvořena sekundárními např. (sklon, orientace). Na obrázku 7 je znázorněno schéma práce s daty.



Obr. 7 Schéma zpracování dat

4.1.1 Nálezová data *Medicago truncatula*

Hlavní data byla tvořena několika textovými soubory ve formátu xls. Data byla poskytnuta prostřednictvím doc. Ing. Petra Smýkala, Ph.D z Katedry botaniky Univerzity Palackého v Olomouci.

Struktura

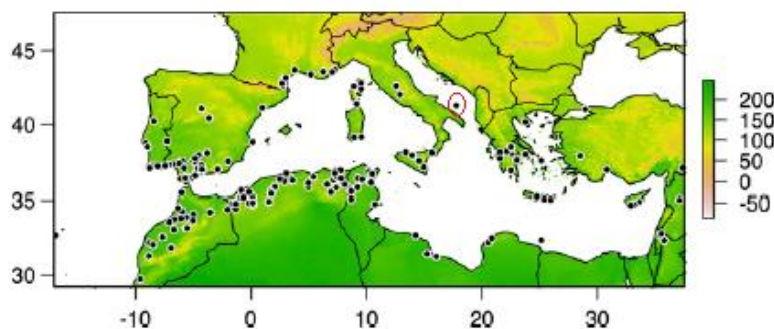
Získané záznamy obsahovaly údaj o poloze, ID vzorku, název státu případně oblasti, kde byly sesbírány (viz tabulka 2). U některých dat byla také uvedena nadmořská výška. Všem datům bohužel zcela chyběla jakákoliv metadata, největší problém nálezových dat tak představovala jejich kvalita.

Tab. 2 Struktura poskytnutých dat

Collection	country	HapMap	lat	lon
L000163	France	HM001	43,571	6,224
L000544	Syria	HM002	36	37
L000162	Spain	HM003	38	-3,816
L000651	Cyprus	HM004	34,783	33,167
L000235	Italy	HM005	35,75	14,517
L000555	Italy	HM006	35,56666667	13,66666667
L000666	Morocco	HM007	31,254	-8,5

Na první pohled byla v datech vidět značná nehomogenita. Poloha místa sběru byla místy uvedena v souřadnicích ve formátu desetinného čísla (17,5; 4,3). V některých případech byla poloha ve stupních (17° 30' s. š., 4° 30' v. d.). Pro následnou práci v *ArcGIS for Desktop 10.4* byly souřadnice převedeny na desetinné číslo. U některých záznamů nebyla poloha zaznamenána vůbec. Tyto body byly vyfiltrovány a po konzultaci s botaniky k nim byla poloha dohledána, nebo byly zcela odstraněny.

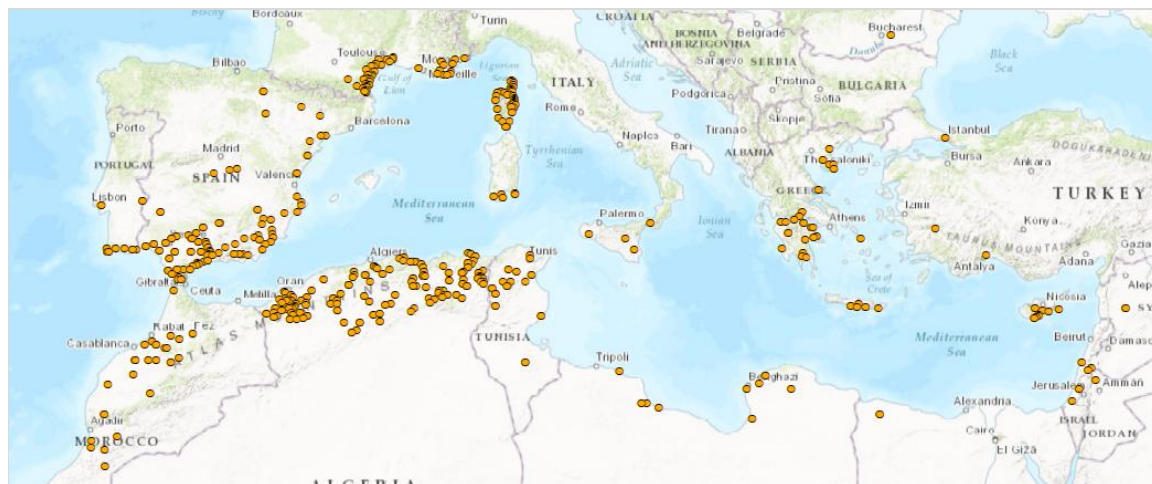
Po vyfiltrování záznamů s chybějící polohou a sjednocení všech souřadnic bylo možné data zobrazit v mapě a odhalit případnou chybu v poloze. Záznamy byly převedeny do bodové vrstvy a zobrazeny v *ArcGIS for Desktop*. Některé body ležely uprostřed vodní hladiny, tyto údaje byly označeny jako chybné. Každý byl zkontrolován a bylo ověřeno, zda se nejedná například o chybu v záznamu (např. místo 34,784 bylo uvedeno 34,00784). Pokud u takového případu souhlasil textový údaj o poloze s možnou lokalizací, byla souřadnice opravena. Některé záznamy obsahovaly chybnou polohu, která nemohla být zaměněna se špatným zadáním souřadnice. Tyto případy byly zcela odstraněny. Část bodů ležela mimo sledované území. Každý byl zkontrolován. Na základě dalších atributů, ze kterých bylo možné vyčíst údaje o poloze, byly také opraveny prohozené souřadnice délky a šířky. V několika případech byla západní délka uvedena bez znaménka (-). Postup při manuální kontrole byl poměrně zdoluhavý a jen těžko nahraditelný automatizovanou metodou. Například v článku (Yoder a kol., 2009) byla použita stejná datová sada jako v této práci, ale bez zjevné hlubší kontroly polohy záznamů (viz obr. 8).



Obr. 8 Chybná souřadnice upraveno podle (Yoder a kol. 2009)

Po prvním zpracování bylo v sadě 1 179 záznamů. Nicméně velké množství bodů mělo totožné souřadnice a na jedné lokaci leželo několik bodů. Zde bylo otázkou, zda je tento výskyt dán vyšší koncentrací taxonu nebo jinými faktory. Pokud by se nejednalo o vyšší výskyt v daném místě a byly by záznamy ponechány, mohlo by to zásadním způsobem ovlivnit statistické charakteristiky daného souboru. Po konzultaci s vedoucím práce a docentem Smýkalem byla pro potřebu statistického vyhodnocení

environmentálních podmínek vytvořena zredukováná sada. K redukci bodů byla použita funkce *Dissolve*, na základě zeměpisných souřadnic. Výsledná bodová vrstva obsahovala 418 záznamů. Téměř všechna data ležela v oblasti Středozemního moře (obr. 9). Další místa výskytu byla: Jihoafrická republika (2 lokality), Rumunsko (1 lokalita) a Írán (1 lokalita).



Obr. 9 Rozložení bodů po úpravě souřadnic
Snímek obrazovky z ArcGIS for Desktop 10.4

Kvalita dat

Polohová přesnost poskytnutých dat byla velmi rozdílná. U některých záznamů byla uvedena poloha na několik desetinných míst, u jiných pouze na celé stupně. Z tohoto důvodu byl do dat zaveden atribut kvality polohové informace (polohová přesnost). Pro klasifikaci byly zvoleny 3 stupně (viz tab. 3), kdy 1. představoval body s nejvyšší kvalitou a polohou určenou u obou souřadnic minimálně na 3 desetinná místa. Kategorie 2 zastupovala body, kdy alespoň jeden měl souřadnici danou méně než 3 desetinnými místy. Do kategorie 3 byly zařazeny body, které měly alespoň v jedné souřadnici méně než 2 desetinná místa.

Tab. 3 Kategorie polohové přesnosti

Kategorie	1	2	3
Příklad souřadnic	43,73376465; 4,048339844	37,45; -1,990666747	43; 2
Počet bodů¹	961	73	145
Počet bodů²	330	29	59

¹ hlavní datová sada 1179 bodů ² redukováná sada pro statistické výpočty

Počet desetinných míst je pro určení polohy zcela zásadní, zpřesnění jen o 1 setinu v jedné souřadnici může zcela změnit polohu bodu v řádu kilometrů. Vzdálenost posunu při změně desetinných míst v jedné souřadnici je ukázána v tabulce 4.

Tab. 4 Posun při změně jedné souřadnice

Změna souřadnice	42,9 ; 42,8	42,95 ; 42,96	42,951 ; 42,952
Řád desetinných míst	0,1	0,01	0,001
Vzdálenost [m]	11 000	1 100	110

Chyba v poloze v řádu setin může pouze v rámci jedné souřadnice představovat značný posun v rozsahu stovek metrů a může představovat značný problém při aplikaci možných výsledků.

4.1.2 WorldClim

WorldClim umožňoval poměrně snadné stažení datasetů bez jakékoliv registrace či složitějšího vyhledávání. Pro účely práce byla použita databáze ve verzi 2.0. (viz kapitola 3.3.1).

V práci bylo použito 79 rastrů (průměrné, maximální a minimální měsíční teploty, průměrné srážky, solární radiace a bioklimatické charakteristiky). Jednalo se o data s rozlišením 30 úhlových vteřin. Všechna data v archivovaném formátu zabírala okolo 7 GB. WorldClim data nebylo nutné příliš upravovat. Po stažení a následné úvaze použití těchto souborů pro extrakci hodnot, byla data hromadně rozbalena a přejmenována v programu Total Commander.

Příklad přejmenování: maximální teplota za měsíc leden *wc2.0_30s_tmax_01* byl přejmenován na *tmax_01*. Obdobný postup byl aplikován i u dalších rastrů.

4.1.3 ASTER GDEM 2.0

Stažení druhé generace výškopisného modelu ASTER GDEM je možné více způsoby. V práci bylo nutné stáhnout velké množství dlaždic a z tohoto důvodu se jako nejlepší a nejrychlejší možnost ukázalo využití stránek <https://reverb.echo.nasa.gov/reverb/>. Velmi často využívaná stránka <https://earthexplorer.usgs.gov/> nebyla zvolena z důvodu omezení stahování. Zde bylo možné vybrat pouze 100 rastrů na jedno hledání, oproti stránkám NASA, kde bylo možné stáhnout libovolně velké území.

Po vytvoření účtu je k dispozici mapové okno, ve kterém se nachází možnost vybrat oblast zájmu pomocí polygonu a období, za které mají být data vyhledávána. V dolní části jsou zobrazeny dostupné datasety. V případě bakalářské práce byl zvolen (*ASTER Global Digital Elevation Model V002*). V dalším kroku následovala možnost *Search by granules*. Zde byly vybrány požadované dlaždice a vloženy do košíku. Ve chvíli kdy byla data připravena, na email zadaný přeregistrací přišel email s odkazem ke stažení. V případě stahování zip archivu, nastala chyba po rozbalení. Tento problém byl vyřešen přepsání koncovky .zip na .tar a stažení v tomto formátu.

Po stažení a rozbalení všech rastrů bylo nutné spojit jednotlivé dlaždice do jedné mozaiky. Velké množství rastrů a výsledného datového objemu 40 GB si vyžádalo značné časové i hardwarové nároky. Soubory bylo nutné spojovat po menších částech, např. část Evropy a následně je spojovat k sobě. Ve výsledku se nepodařilo v programu *ArcGIS for Desktop 10.4* ani v *QGIS 2.18* spojit všechny rastry do jednoho. V konečné fázi bylo pracováno se dvěma rastry.

Před vstupem rastrů do funkce *Extract multivalues to point* byl vytvořen čtvercový buffer okolo každého bodu. Pro zrychlení práce byly pomocí funkce *Clip* extrahovány pouze okolní části digitálního výškového modelu. Poté již byly rastry spojeny do jednoho pomocí funkce *Mosaic to new raster*.

Výpočet sklonu a orientace svahů

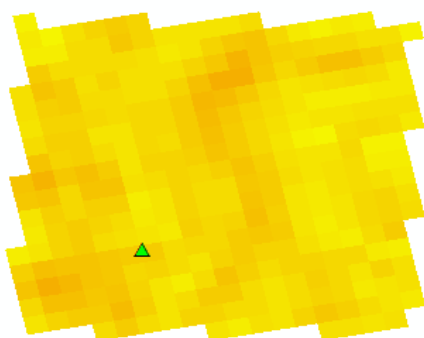
Vzhledem k tomu, že výškový model ASTER GDEM2 je v souřadnicovém systému WGS 1984, bylo nutné pro správné výpočty v *ArcGIS for Desktop* před vstupem do analýz rastr upravit. Nálezová data se rozprostírala na velkém území od 17° z. d. až po 60° v. d. Nebylo tak možné celý rastr transformovat do jednoho souřadnicového

systému bez výraznějšího zkreslení. Bylo zvoleno řešení v podobě transformace jednotlivých částí rastru do Coordinate system UTM WGS84, v zónách 28N–37N a 40N. Pro automatizaci části postupu byl vytvořen toolbox, dostupný v příloze 3e, který celý rastr rozdělil na části podle zóny.

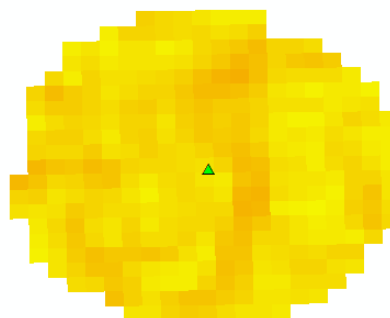
Vstupní data tvoří rastr a dále vrstva, podle které se rastr rozdělí na jednotlivé díly. V tomto případě byla použita vektorová vrstva UTM zón ve formátu SHP (dostupná na http://earth-info.nga.mil/GandG/coordsys/grids/universal_grid_system.html#zzd1), ve které byl klíčový atribut ZONE, označující číslo zóny. Pomocí iterátoru byla vždy vybrána příslušná zóna a ta následně vstupovala do ořezové funkce *Clip*. Pro pojmenování rastru podle zóny, ve které se nachází, je nutné nastavit relativní jméno výstupu. Například ve tvaru *%value%.tiff*. Hodnota *value* představuje hodnotu atributu ZONE použitého v *Iterate Feature selection*. Obdobný postup byl proveden i s bodovou vrstvou. Zde byla použita funkce *Copy feature*, která ukládala jednotlivé výběry do nových souborů.

Po rozdělení obou vrstev na jednotlivé části byla každá část převedena do Coordinate system UTM WGS84. Na transformovaných rastrech byl spočítán sklon a vypočtena orientace. Funkce *Extract Multi Values to Points* přiřadila hodnoty k jednotlivým bodům. V závěru byly extrahovány atributové tabulky do formátu xls nástrojem *Table To Excel* a spojeny do jednoho souboru následně připojeny do základní bodové vrstvy.

Pro porovnání bylo vybráno jedno zobrazení pro celý výškový model. Konkrétně Europe Lambert Conformal Conic. Zpracování proběhlo mnohem rychleji bez dalších manuálních postupů. Vybrané oblasti si rozložením dat odpovídaly, ale výsledný rastr byl značně deformovaný (obr. 10) a poloha bodu již byla rozdílná od rastru v UTM projekci (obr. 11). Správným řešením se ukázalo rozdělení rastru po jednotlivých zónách.



Obr. 10 Bod v souřadnicovém systému Europe Lambert Conformal Conic



Obr. 11 Bod v souřadnicovém systému UTM 30 N

4.1.4 Harmonised World Soil Database v 1.2

Harmonised World Soil Database představuje poměrně dobrý zdroj informací o půdě. Velkou výhodou je globální pokrytí celého světa. HWSD není příliš známá a mnoho prací ji nevyužívá. Data je možné stáhnout ze stránek FAO. Stažení probíhalo jednoduchým způsobem, stačilo pouze zvolit možnost *download data*. Problém byl, že data nebyla v rastrovém formátu, ale ve formě databáze mdb. Pro možnost práce s databází v GIS bylo nutné ji nejprve spojit s referenčním rastrem. Na stránkách FAO

však bohužel tento rastr není dostupný a pro účely práce byl stažen z archivu stránek IIASA <http://webarchive.iiasa.ac.at/Research/LUC/External-World-soil-database>.

Po stažení rastru a databáze byly oba soubory spárovány. V dokumentaci k HWSD bylo popsáno, podle kterého atributu je možné data spárovat, ale už zde nebyl uveden postup. Proto bylo nutné potřebné informace dohledat v diskuzních fórech a v manuálu k produktům Esri.

Postup připojení Microsoft Access databáze na rastr

Celý postup je zaznamenán na videu v příloze 3m na DVD

- V samostatně spuštěném ArcCatalogu bylo přidáno nové spojení OLE database
- Byla vybrána a připojena požadovaná databáze (v této práci. Microsoft Jet 4 OLE)
- V ArcGIS for Desktop byl spuštěn nástroj *Build Raster Attribute Table*
- Po kliknutí na rastr byla vybrána funkce *join* a databázové spojení mdb. databáze
- Pro spojení sloužil atribut *MU global* ve složce HWSD data.

Při použití funkce *Extract values to point* se objevil problém. Tato funkce dokáže uložit jen jednu informaci z buňky a nedokáže pracovat s připojenou databází. Proto byl nejdříve extrahován primární klíč z rastru do bodové vrstvy a následně se pomocí tohoto atributu připojily ostatní atributy.

4.1.5 SoilGrids

Půdní data SoilGrids byla stažena pomocí FTP serveru. Oproti HWSD každou půdní charakteristiku zastupuje samostatný rastr. Spolu se soubory GeoTIFF byla v adresáři legenda v podobě textového souboru, ve které byly uvedeny zkratky názvů rastrů a jejich význam. Před extrakcí hodnot byly jednotlivé atributy přejmenovány.

4.1.6 GlobCover 2009

Databáze pokrytí GlobCover je dostupná pro celý svět v jednom GeoTIFF souboru. Ke stažení nebyla nutná registrace. Zabalená data obsahovala GeoTIFF, legendu ve formátu xls, soubor lyr pro stylování rastru v *ArcGIS for Desktop*.

4.1.7 CCI Land cover 2015

Stažení dat Land Cover Maps proběhlo pomocí přímého odkazu ke stažení. Na portálu se nacházela možnost zadat vlastní souřadnici nebo ji nahrát pomocí textového souboru a extrahovat informace pouze ke zvoleným bodům, pro potřeby bakalářské práce byl stažen celý rastr ve formátu GeoTIFF.

4.2 Finální úprava dat

Informace ze všech rastrů byla extrahována do bodové vrstvy pomocí funkce *Extract multivalues to point*, ve které bylo možné nastavit jméno extrahovaných atributů, které se defaultně pojmenovaly podle jména vstupujících rastrů. Během extrakce nastala nespecifikovaná chyba a na dataset bioklimatických charakteristik (WorldClim) nešla tato funkce aplikovat. Z tohoto důvodu byl vytvořen toolbox (Příloha 3f) v prostředí

ModelBuilder, který je založen na modifikaci funkce *Extract value to point* a přináší stejné výsledky.

Body v pobřežních oblastech

Rastry vykazovaly nestejně překrytí a ne zcela dobře reprezentovaly pobřežní hranici. Po první extrakci hodnot, do některých bodů nebyla zapsána hodnota rastru. Ve většině případů ležely v těsné blízkosti hranice pevniny, ale z důvodu nepřesnosti v poloze nebo přesnosti rastru spadly na vodní hladinu. Většinou se tyto body nacházely do 1 km od pevniny a byly manuálně posunuty v průměru o 500 m na nejbližší pixel. Vzhledem k největšímu počtu dat z databáze WorldClim byly posunuty právě na tyto rastry. Jednalo se o 39 bodů. U rastru HWSD několik bodů bylo lokalizováno na vodní hladině, pomocí funkce *Buffer* bylo oříznuto okolí bodů, rastr byl převeden na body a následně byla extrahována nejbližší hodnota. Tyto hodnoty byly poté zpět připojeny k původnímu souboru.

Po konzultaci s vedoucím práce byl zvolen postup, který by měl porovnat, jaký může mít vliv nepřesnost v poloze bodu na jeho hodnotu. Vztažení bodu pouze k danému pixelu může být zavádějící. Zejména u bodů, které leží ve výrazněji členitém terénu, mohou mít okolní pixely značné rozpětí. V případě kvalitativních dat (land cover) může mít každý pixel zcela jinou hodnotu a zásadně měnit význam hodnoty pro daný bod (lesní porost/zastavěné území). U kvantitativních spojitých dat (srážky, teplota) se nepředpokládá velký rozdíl hodnot v blízkém okolí pixelů a menší nepřesnost v poloze nemusí znamenat zcela odlišné výsledky.

4.2.1 Detekce potenciálně rizikových bodů

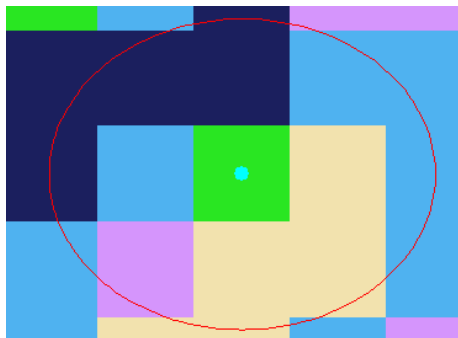
V této části bude vyšetřeno okolí pro kvalitativní data. Zejména v případě land cover každá hodnota pixelu může představovat zcela jinou kategorii. Například se může jednat o kombinaci pokryvu v okolí (lesní porost/zástavba/zemědělská půda). V případě nepřesnosti v poloze bodu může být jeho hodnota interpretována nesprávně. V této části bude představen způsob, jak je možné detekovat potenciálně rizikové body.

Pro testování byla vybrána vrstva Globcover09 a 330 testovacích bodů, které byly označeny kategorií 1 s nejlepší kvalitou prostorové informace. K testování na datech land coveru byl vytvořen toolbox (Příloha 3g) s možností nastavení vzdálenosti okolí, pro které mají být hodnoty vyšetřovány. Za pomoci funkce *Zonal statistic as table* byla spočítána statistika pro každé okolí bodu. Jako statistické charakteristiky byly vybrány: hodnota pixelu, počet vztažných pixelů, počet unikátních hodnot, hodnota s nejmenší/největší četností. Následně probíhal výpočet manuálně. Nejprve byl vytvořen nový atribut (*původní hodnota pixelu-hodnota s nevyšší četností*). Pokud byla hodnota atributu nulová, pixel odpovídal převládající hodnotě okolí a bylo možné zanedbat chybu v poloze a brát tuto hodnotu za správnou. Pokud ale hodnota pole nebyla nulová, byly tyto body označeny, jako „rizikové“. K poli rizikové bylo přidáno pole potenciálně „nejvíce rizikových“ bodů. Nejvíce rizikové body byly určeny jako ty, kde počet unikátních hodnot v oblasti byl větší než 1 a zároveň se shodovala hodnota pixelu s hodnotou nejmenšího počtu pixelů v oblasti. Tudiž zde byl předpoklad, že v okolí jsou převážně pixely s jinou hodnotou.

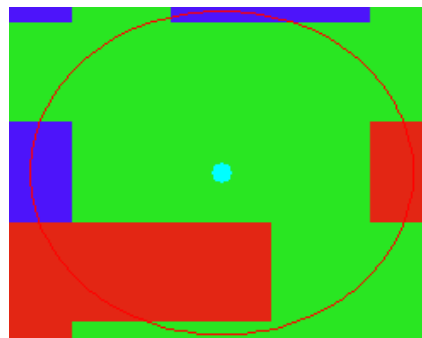
Výsledek testování

Výše zmíněný způsob byl otestován na 330 bodech, ze kterých vyšlo 241 bodů málo rizikových a 89 rizikových. Z rizikových bodů bylo 39 velmi rizikových (obr. 12). 241 bodů mělo hodnotu pixelu stejnou s převažující hodnotou okolních pixelů. Tyto body byly označeny jako málo rizikové (obr. 13).

Tímto způsobem je možné odhalit místa, kde může poloha bodu hrát velkou roli. Jak je vidět, tak v 500m okolí se může vyskytovat až pět různých hodnot. Zejména u dat land cover je důležité, aby byl bod přiřazen ke správnému pixelu. Vzdálenost 500 m představuje chybu v poloze reprezentovanou v řádu setin stupně (viz kapitola 4.1.1), která může znamenat zařazení bodu do zcela jiné kategorie.



Obr. 12 Velmi rizikový bod



Obr. 13 Málo rizikový bod

4.2.2 Testování proměnlivosti hodnot

Jak již bylo zmíněno v předchozí kapitole, v okolí bodu se může vyskytovat značné rozpětí hodnot. Proto bylo přistoupeno k praktickému testování. Cílem bylo odhalit, jak moc by se mohly změnit hodnoty celého souboru v případě náhodné nepřesnosti v poloze jednotlivých bodů.

Test byl proveden na souboru 330 bodů. Jako testovací rastr byl zvolen *GlobeCover 09* v rozlišení 300 m/pixel. Po úvaze autora byly vybrány pouze sousední pixely. Sousední pixel představuje nejbližší pixel v přímém směru (S, V, J, Z) od centrálního pixelu. V programu Microsoft Excel byly vygenerovány kombinace hodnot bodů ze všech směrů. Pro každý bod byla náhodně vybrána hodnota z jednoho směru. Z nově vzniklých kombinací byla vytvořena tabulka četností jednotlivých hodnot.

Postup

Nejprve byly vytvořeny čtyři bodové vrstvy v každém směru. Funkce *Bearing Distance to Line* umožňuje vytvořit úsečku vedenou od vybraného bodu. Je možné zvolit vzdálenost a azimut, pod kterým bude vedena. V tomto případě byla zvolena vzdálenost 300 m, která představuje možnou chybu v řádu tisícín stupně v rámci jedné souřadnice. Větší vzdálenost by mohla zapříčinit přeskočení sousední buňky v případě, že by bod ležel na hranici pixelu. Pomocí funkce *Feature Vertices To Points* byl převeden konec úsečky na bod. Po vytvoření bodových vrstev byla do každé extrahována hodnota rastru. V posledním kroku byly atributové tabulky sloučeny. Následně byla tabulka exportována do Microsoft Excel. Vzorcem ve tvaru `=Z1 & INDEX(N2:R2;1;RANDBETWEEN(1;5))` byla náhodně vybrána hodnota pro každý řádek ze všech směrů. Proces byl pětikrát iterován pro více kombinací.

Výsledky

Po analýze všech okolních bodů může být konstatováno, že u testovaného rastru a bodové vrstvy není v těsné blízkosti v hlavních směrech příliš velká variabilita hodnot pixelů (viz tab. 5) a zastoupení jednotlivých kategorií se výrazně nemění.

Tímto způsobem je možné odhalit, u kterých kategorií může nastat největší změna v počtu hodnot, při náhodné chybě rozprostřené na celý soubor. Tabulka všech hodnot je k dispozici v příloze 3j.

Tab. 5 Porovnání četností kategorií krajinného pokryvu mezi náhodnými výběry

Kategorie*	Soubor výběru					
	Původní	Náhodný 1	Náhodný 2	Náhodný 3	Náhodný 4	Náhodný 5
20	88	86	86	100	86	95
210	1	5	0	2	0	3
110	8	10	10	10	10	4

*názvy jednotlivých kategorií v příloze 3i

4.3 Statistické výpočty

Po zpracování bodové vrstvy a potřebných úpravách byly vypočítány statistické charakteristiky environmentálních faktorů. Do statistické části vstupovala redukováná sada 418 bodů a část s nejvyšší prostorovou přesností 330 bodů. Před výpočtem zpracováním jednotlivých tematických částí byla otestována korelace mezi vstupními daty.

Téměř všechny body se nacházely v oblasti Středomoří, 2 body ležely na území Jihoafrické republiky, jeden bod v Íránu a v Rumunsku.

Prostorová korelace

Prvním testem bylo zjištění prostorové korelace mezi body. Důležité před samotnou statistikou bylo zjistit, zda body s nízkou kvalitou polohové informace jsou rovnoměrně rozprostřeny po celém sledovaném území, nebo vytvářejí shluky na určitých místech, proto bylo přistoupeno k otestování prostorové korelace na základě parametru kvality dat.

Prostorová autokorelace je velmi významný způsob zkoumání existence a míry prostorového shlukování, zejména identifikace prostorových shluků může pomoci se zaměřit na konkrétní oblasti a identifikovat působení lokálních faktorů.

Rozlišujeme globální a lokální míru a také zda je pozitivní či negativní. Jako pozitivní můžeme označit místa, kde sousední oblasti mají stejné či podobné hodnoty. Naopak negativní se vyznačuje stav, kdy jsou hodnoty odlišné.

Bylo přistoupeno k výpočtu lokálních indikátorů prostorové asociace (LISA) v programu GeoDa. Pro výpočet bylo použito schéma prostorových vah typu věž s faktorem 1. Hodnota Moranova I byla 0,157 a nenaznačovala tendenci k vytváření shluků napříč daty. Pomocí indikátoru LISA byly identifikovány oblasti Blízkého východu a severovýchodního pobřeží Afriky jako tzv. hot spots, kde se vyskytují body s nejmenší kvalitou prostorové informace. V oblasti Blízkého východu většina bodů spadala do kategorie 3 s nejmenší kvalitou, tento fakt mohl být zapříčiněn mnoha faktory, například chybou při zpracování. Naopak v oblasti Korsiky byla většina bodů v kategorii 1 s nejlepší kvalitou, tato skutečnost může ukazovat na odlišnou metodiku sběru jednotlivými institucemi a zejména pak na pečlivější práci s daty.

Korelace

Pro zjištění míry závislosti mezi vstupními daty bylo přistoupeno k výpočtu korelace v software STATISTICA 12 pomocí korelační matice, do které vstupovala všechna kvantitativní data. Kompletní tabulka je dostupná v příloze 3l. Korelační matice vybraných faktorů je uvedena v kapitole 5 Výsledky.

4.3.1 Environmentální data

Klimatická data

Pro testování klimatických charakteristik byly vybrány atributy průměrná roční teplota, roční úhrn srážek a několik bioklimatických charakteristik. Z průměrných měsíčních teplot a měsíčních úhrnů srážek byly vypočteny souhrnné měsíční hodnoty pro celý soubor, ze kterých byl sestaven klimatogram v programu Microsoft Excel 2007. Veškeré výpočty probíhaly v Software STATISTICA 12 za použití histogramů, základních statistických charakteristik a krabicových grafů pro detekci odlehlých hodnot. Výsledky korespondovaly s typem Středomořského klimatu. Podrobné výsledky v kapitole 5.

Nadmořská výška, sklon, orientace

Pro charakteristiku nadmořské výšky, sklonu a orientace byla použita základní popisná statistika v programu STATISTICA 12 a zejména funkce tabulky četností, pomocí které byly identifikovány nejvíce zastoupené intervaly každé veličiny.

Půdní data

Půdní data vstupující do analýzy představovaly dvě databáze SoilGrids a HWSD obě s rozlišením 1 km/pixel. V případě HWSD u některých záznamů chyběla data a soubor tak nebyl zcela kompletní. Pro analýzu bylo vybráno několik charakteristik pro vzájemné porovnání. V rámci HWSD se hodnoty vztahovaly k vrchním 30 cm půdního horizontu. U SoilGrids byly hodnoty vztaženy k hloubce 5 cm pod povrchem. K výpočtu byl použit software STATISTICA 12 a funkce základních statistik a tabulky četností spolu s vizualizací četností pomocí histogramu.

Krajinný pokryv

Krajinný pokryv podává informace o krajinném krytu, ve kterém rostlina roste. Zda se jedná o lesní porost, zemědělskou půdu, pastviny apod. Pro zachycení možné změny v krajině byly vybrány databáze z roku 2009 a 2015 obě s rozlišením 300 m/pixel.

Pro přehled zastoupení jednotlivých kategorií byly vytvořeny tabulky četností podle unikátních hodnot v programu STATISTICA 12. Do přehledu byly vybrány 3 nejpočetnější kategorie z každé databáze. Kompletní přehled všech kategorií je uveden v příloze 3k.

5 VÝSLEDKY

5.1 Výsledná bodová vrstva

Jedním z hlavních výstupů práce je samotná bodová vrstva nálezových dat, kterou je možné stáhnout a využít pro další výzkum. Samotný datový soubor čítá 1179 záznamů a 165 atributů klimatických, výškopisných a půdních faktorů. Primárním atributem pro identifikaci vzorků pro další využití je Catalog number. Některé atributy byly vytvořeny následnou analýzou primárních dat (sklon, orientace, polohová přesnost).

Data jsou přiložena ve formě GIS formátů Esri Shapefile, Esri Feature Class a v textovém formátu xls (Přílohy 3a, 3b, 3c). Součástí je také tabulka (Příloha 3i) s detailním popisem jednotlivých atributů ve formátu xls, která se skládá z hlavního a vedlejších listů, na hlavním listu jsou názvy atributů, jednotky a význam atributů. Vedlejší listy obsahují bližší informace o attributech. Hlavní list je propojen s vedlejšími pomocí hyperlinku, díky kterému je uživatel po kliknutí odkázán ke konkrétní informaci. Veškeré atributy i soubor s popisem jsou provedeny v anglickém jazyce pro širší okruh možných uživatelů.

5.2 Statistické přehledy

Vypočtené statistické charakteristiky vzhledem ke kvalitě vstupních dat a různému rozlišení environmentálních rastrů jsou spíše orientační a nelze je považovat za zcela prokazatelné, do jisté míry však korespondují s charakteristikou prostředí uvedenou u popisu *Medicago truncatula* (FAO, 2017).

Korelace

Výsledky korelace shrnuje tabulka 6, ve které jsou vybrány hodnoty s nejvyšší mírou korelace v rámci celého souboru 418 bodů, čím intenzivnější odstín, tím vyšší míra korelace. Z tabulky vyplynulo několik závěrů, kromě předpokládané korelace mezi průměrnou roční teplotou a nadmořskou výškou, je zde prokazatelná závislost mezi srážkami a půdními veličinami jako je pH a objemovou hmotností. Čím vyšší jsou srážky, tím nižší je pH. Zejména stanoviště v pouštních oblastech se vyznačují vysokým pH půdy. Závislost byla vyznačována i mezi zeměpisnou šířkou a ročním úhrnem srážek. S přibývajícím šířkou přibývá i srážek. Tabulka korelace vstupních dat je uvedena v příloze 3l.

Tab. 6 Korelační matice vybraných faktorů

Proměnná								
	z. d.	z. š.	Sklon °	Prům. roč. teplota	Roč. srážky	Nadm. výška	pH H ₂ O	Objm. hmotnost
z. š.	1,000000	-0,071363	0,049404	-0,344968	0,352276	-0,191164	-0,264083	-0,370679
z. d.	-0,071363	1,000000	0,019318	0,215315	0,039491	-0,198610	-0,005707	-0,111627
Sklon °	0,049404	0,019318	1,000000	-0,193612	0,332960	0,160834	-0,373147	-0,397836
Prům. roč. teplota	-0,344968	0,215315	-0,193612	1,000000	-0,321555	-0,520464	0,270218	0,362982
Roč. srážky	0,352276	0,039491	0,332960	-0,321555	1,000000	-0,217450	-0,719683	-0,671268
Nadm. výška	-0,191164	-0,198610	0,160834	-0,520464	-0,217450	1,000000	0,281135	0,246082
pH H ₂ O	-0,264083	-0,005707	-0,373147	0,270218	-0,719683	0,281135	1,000000	0,851956
Objm. hmotnost	-0,370679	-0,111627	-0,397836	0,362982	-0,671268	0,246082	0,851956	1,000000

5.2.1 Klimatické charakteristiky

Teplota a srážky

Do výpočtu byly zahrnuty oba soubory, jak celkový tak pouze výběr s nejvyšší kvalitou prostorové informace. Vzhledem k rozložení bodů s nejnižší kvalitou nebyly prokázány velké rozdíly mezi těmito soubory. Všechny údaje o srážkách a teplotě jsou vztaženy k období 1970–2000.

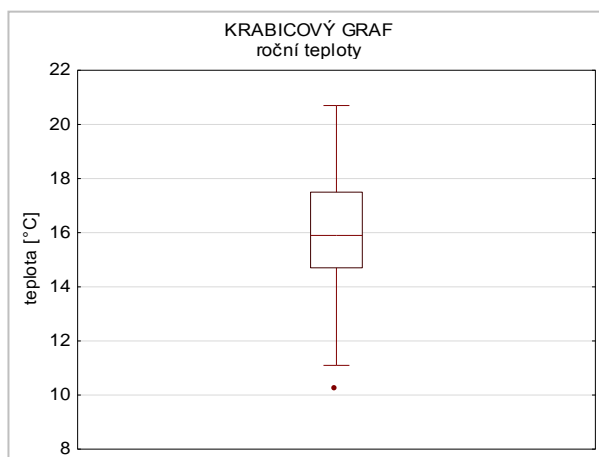
Průměrná roční teplota stanovišť zkoumaných vzorků byla 16,1 °C. Nejnižší teplota 10,2 °C byla zjištěna u stanoviště v severní části Alžírsko v pohoří Atlas, tato nízká hodnota byla dána především nadmořskou výškou, ve které se stanoviště nacházelo (1822 m n. m.). Nejvyšší teplota 20,7 °C byla zjištěna při pobřeží Galilejského jezera v Izraeli. Směrodatná odchylka 1,8 byla poměrně nízká a značila malou variabilitu hodnot kolem průměru. Přehled základních charakteristik je zachycen v tabulce 7.

Tab. 7 Základní charakteristiky ročních klimatických faktorů

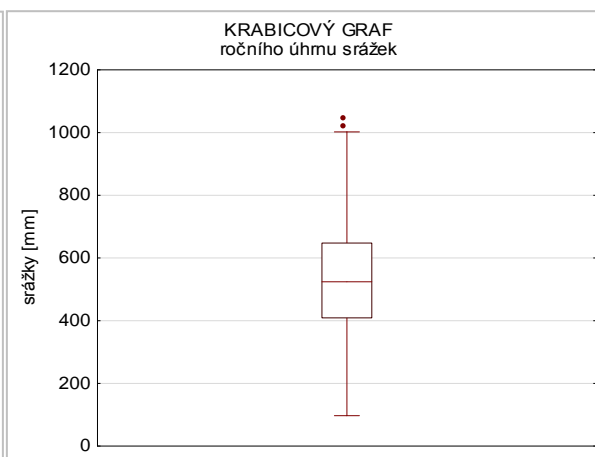
	N platných	Průměr	Medián	Minimum	Maximum	Sm. odch.
Roční teplota	418	16,1	15,9	10,2	20,7	1,8
	330	15,9	15,8	12,1	20,7	1,5
Roční úhrn srážek	418	531,5	524	97	1041	181,5
	330	553,7	550	171	1041	174,4

V rámci srážek byl zřejmý velký rozdíl mezi nejvyšším a nejnižším ročním úhrnem. Habitat rostliny s nejvyšším ročním úhrnem se nacházel na pobřeží Alžírsko na návětrné straně pohoří Atlas. Naopak stanoviště s nejmenším ročním úhrnem srážek bylo zjištěno v severní části Libye. Směrodatná odchylka 181,5 představovala poměrně značnou variabilitu hodnot v okolí průměru.

Pro lepší představu o rozložení dat byly sestaveny box ploty (obr. 14, obr. 15), pomocí kterých je možné detekovat extrémní hodnoty. Při zobrazení box plotů byly zvoleny 25% a 75% kvantil a medián a odlehlé hodnoty.



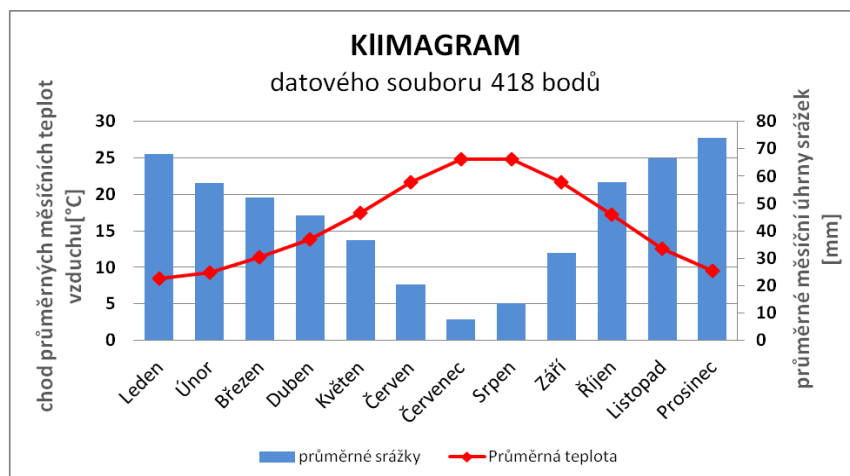
Obr. 14 Krabicový graf průměrné roční teploty 418 bodů



Obr. 15 Krabicový graf ročního úhrnu srážek 418 bodů

Z grafu roční teploty je vidět, že se vyskytují odlehlé hodnoty při dolní hranici souboru, naopak data se srážkami vykazují odlehlé hodnoty v blízkosti maximálních hodnot. Poloha mediánu ukazuje, zda jsou hodnoty blíže koncentrovány k minimu nebo

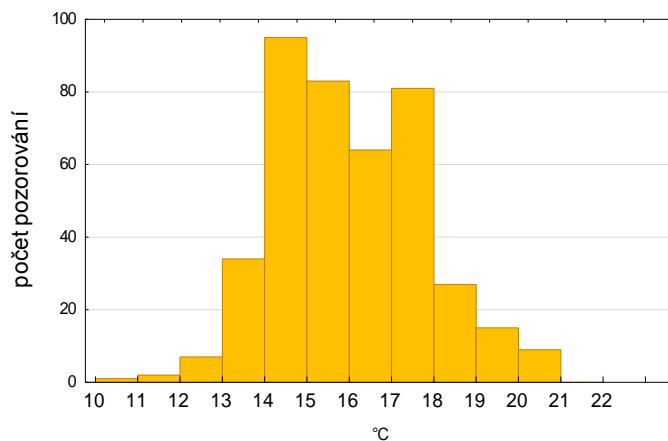
maximu, u obou datových souborů leží medián přibližně uprostřed a data se vyznačují normálním rozdělením. Po analýze ročních hodnot bylo přikročeno ke zkoumání chodu teploty a srážek v ročním cyklu. Pro zobrazení ročního chodu byl sestaven klimagram (obr. 16) celého souboru z průměrných měsíčních hodnot.



Obr. 16 Klimagram datového souboru 418 bodů

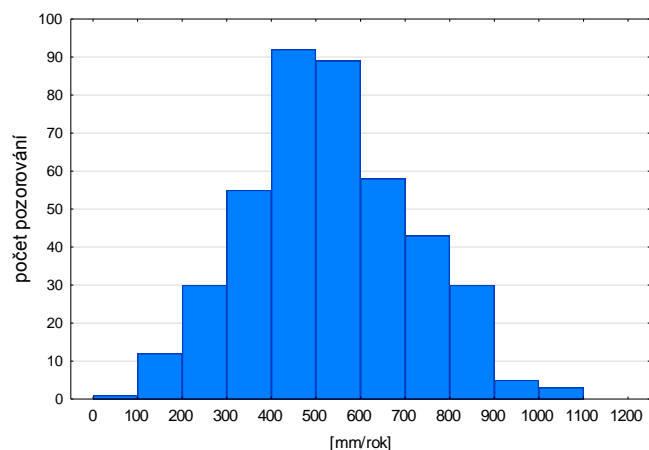
Z grafu je jasně patrný pokles srážek v letním období, v měsících (květen–září) průměrné srážky nedosahují ani 40 mm/měsíc a zejména v červenci a srpnu jsou průměrné srážkové úhrny do 10 mm/měsíc. V takřka inverzním chodu se projevují teploty, které dosahují svého maxima v červenci a srpnu, naopak nejnižší jsou v zimním období, kdy se v lednu a únoru dostanou těsně pod hranici 10 °C.

Rozložení teploty a srážek napříč intervaly ukazují krabicové grafy (obr. 17, obr. 18).



Obr. 17 Histogram průměrné roční teploty

Histogram průměrné roční teploty ukazuje rozložení hodnot napříč intervaly, nejpočetnější je interval mezi 14–18 °C. V levé části grafu jsou patrné odlehle nízké hodnoty, které byly detekovány krabicovým grafem.



Obr. 18 Histogram průměrných ročních úhrnů srážek

U rozložení srážek se jedná téměř o normální rozdělení, v pravé části grafu jsou rovněž patrné odlehle vysoké hodnoty. Nejpočetnější je interval 400–600 mm. Alternativou k histogramu je tabulka četností, která nabízí přesné počty jednotlivých záznamů, ale nepřináší již tak dobrou vizuální interpretaci.

5.2.2 Nadmořská výška, sklon, orientace

Ze základního přehledu (tab. 8) je patrný rozdíl mezi hodnotou mediánu a průměru, medián je o 117 menší a naznačuje, že se většina hodnot nalézá pod vypočteným průměrem. Průměr zde proto není zcela vypovídající, neboť je ovlivněn extrémními hodnotami. Velmi vysoké je i variační rozpětí a to 2003, spolu s velkou směrodatnou odchylkou 383,82 ukazují na značné kolísání hodnot a možnou přítomnost odlehlých hodnot (outliers). U souboru 418 bodů se vyskytuje i extrémní hodnota -181 m n. m. u stanoviště poblíž Mrtvého moře. Maximum bylo určeno u stanoviště ležícího v pohoří Atlas.

Tab. 8 Základní charakteristiky nadmořské výšky

	N platných	Průměr	Medián	Minimum	Maximum	Rozpětí	Sm. odch.
Nadmořská výška	418	414,7	297,5	-181	1822	2003	383,82
	330	413,5	277,5	0	1730	1730	383,14

Z tabulky 9 je patrné rozložení četnosti a je možné identifikovat nejvíce zastoupené intervaly. Nejpočetnější je kategorie 0–200 m n. m., která tvoří přibližně 40 % všech sledovaných hodnot. Podle hodnot ve sloupci relativní kumulativní četnosti je patrné, že přibližně 80 % všech rostlin se nachází do 800 m n. m. Intervaly v rozmezí 1200–2000 m n. m. jsou velmi málo početné a značí výskyt odlehlých hodnot

Tab. 9 Četnost nadmořské výšky pro soubory 330 418

Nadmořská výška m n. m.	Četnost		Kumulativní četnost		Kumulativní relativní četnost	
	330	418	330	418	330	418
-200 <x<=0	2	3	2	3	0,6061	0,7177
0 <x<=200	134	164	136	167	41,2121	39,9522
200 <x<=400	52	73	188	240	56,9697	57,4163
400 <x<=600	50	58	238	298	72,1212	71,2919
600 <x<=800	25	38	263	336	79,6970	80,3828
800 <x<=1000	30	38	293	374	88,7879	89,4737
1000 <x<=1200	27	32	320	406	96,9697	97,1292
1200 <x<=1400	6	7	326	413	98,7879	98,8038
1400 <x<=1600	3	3	329	416	99,6970	99,5215
1600 <x<=1800	1	1	330	417	100,0000	99,7608
1800 <x<=2000	0	1	330	418	100,0000	100,0000

Sklon a orientace

V tabulce 10 jsou dva záznamy v intervalu (-5–0), interval představuje povrch, který byl klasifikovaný jako plochý. Zcela nepočtenější jsou intervaly 0–5° a 5–10°, ve kterých se dohromady nachází téměř 70 % všech záznamů. V intervalech od 30° do 45° se nachází jen velmi málo bodů a představují odlehlé hodnoty.

Tab. 10 Četnost sklonu svahů pro soubory 330 418 bodů

Sklon [°]	Četnost		Kumulativní četnost		Kumulativní relativní četnost	
	330	418	330	418	330	418
-5 <x<=0	1	2	1	2	0,3030	0,4785
0 <x<=5	135	175	136	177	41,2121	42,3445
5 <x<=10	84	115	220	292	66,6667	69,8565
10 <x<=15	46	52	266	344	80,6061	82,2967
15 <x<=20	32	37	298	381	90,3030	91,1483
20 <x<=25	16	19	314	400	95,1515	95,6938
25 <x<=30	11	12	325	412	98,4848	98,5646
30 <x<=35	1	1	326	413	98,7879	98,8038
35 <x<=40	2	3	328	416	99,3939	99,5215
40 <x<=45	2	2	330	418	100,0000	100,0000

Orientace

Při analýze orientace svahu pomocí nástroje tabulky četností se neprojevil výraznější trend, který by naznačoval preferenci rostliny k určité orientaci svahu. Relativní četnosti napříč stranami byly rovnoměrně zastoupené s nepočtenější skupinou jihovýchodní orientace.

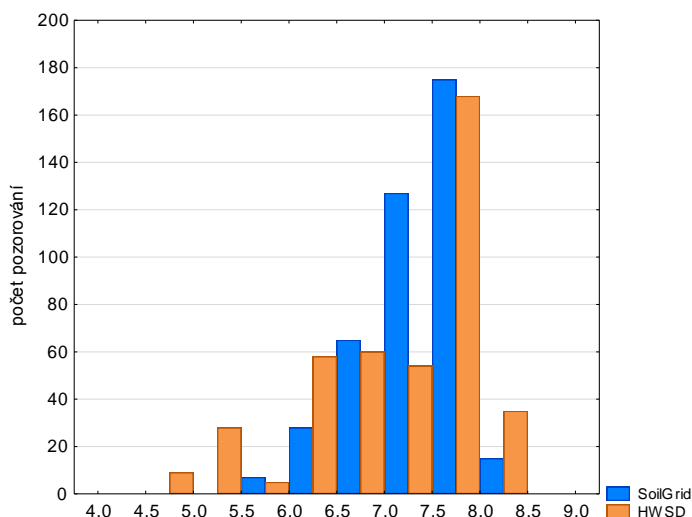
5.2.3 Půdní data

U obou souborů vyšly podobné hodnoty pH viz tabulka 11. Průměr byl kolem 7,3, což ukazuje na výskyt stanovišť v zásaditých půdách. Směrodatná odchylka neprokazuje větší výkyv kolem vypočteného průměru. Pro lepší vizualizaci výsledků byl spočítán histogram pH pro soubor 418 bodů, který srovnává data z databáze SoilGrids a HWSD.

Tab. 11 půdní reakce (H₂O)

Datový soubor	Průměr	Minimum	Maximum	Sm. odch.
330	7,2	4,7	8,3	0,90
418	7,35	4,7	8,3	0,94

Z histogramu pH půdy (obr. 19) je jasně patrné, že hodnoty jsou koncentrovány nejvíce v intervalu 7,5–8 jak u dat ze SoilGrids tak z HWSD. Poměrně početný interval zejména u dat pocházejících ze SoilGrids je 7–7,5, oproti datům HWSD, které v tomto intervalu vykazují poloviční počet hodnot. Půdy s pH vyšším než 7 se označují za alkalické půdy. Na půdách s pH pod 6 se vyskytuje jen malé množství dat.



Obr. 19 Histogram pH půdy 418 bodů

V tabulce 12 je základní statistický přehled podíl jílovitých částic v půdě. U obou souborů vyšel průměr s hodnotou kolem 24 %, který značí, že se rostliny nacházejí spíše na lehčích půdách. Maximum bylo pouze 38 %, tento výsledek naznačuje, že sledovaný soubor se nevyskytoval na těžkých jílovitých půdách.

Tab. 12 Podíl jílovitých částic [%]

Datový soubor	Průměr	Medián	Minimum	Maximum	Sm. odch.
330	24,53	24	11	38	3,95
418	24,74	24	6	38	4,32

5.2.4 Krajinný pokryv

Nejpočetnější kategorií povrchu u GLC09 byla kategorie 20 (Mosaic cropland (50-70%)/vegetation (20-50%)), která tvořila přibližně 25 %. Druhá nejpočetnější byla kategorie 14 (Rainfed croplands) a třetí kategorie 30 (Mosaic vegetation (50-70%)/cropland (20-50%)). V rámci Landcover 15 nebyla už kategorie s tak vysokým procentem četnosti jako CLC 09. Nepočetnější zde byla kat. 11 (Herbaceous cover) s relativní četností téměř 18,5 %. Druhou nejzastoupenější kategorií byla kat. 10 (Rainfed croplands) která je totožná s kategorií 20 u CLC09, avšak u CCI-LC 15 je zastoupena o 10 % méně. Třetí nejvíce zastoupená kategorie byla s 13,4 % (Tree or shrub cover). Srovnání obou klasifikací land cover přineslo rozličné výsledky dané rozdílnou klasifikací každé databáze. U obou se mezi nepočetnějšími typy krajinného pokryvu

vyskytovala kategorie bezzávlahová pole (Rainfed croplands). Mezi zastoupením nejpočetnějších kategorií napříč soubory nebyly výraznější rozdíly (viz tab. 13).

Tab. 13 Zastoupení nejpočetnějších kategorií krajinného pokryvu 330 a 418 bodů

Global Land cover 2009			CCI-LC 15		
Kategorie	Relativní četnost [%]		Kategorie	Relativní četnost [%]	
	330	418		330	418
20	26,67	25,36	11	17,87	18,42
14	17,27	15,31	10	16,36	15,55
30	13,94	14,11	12	16,97	13,40

*kompletní přehled všech kategorií (Příloha 3k)

5.2.5 Souhrnná statistika

Vypočtené důležité charakteristiky shrnuje tabulka 14. I přes rozporuplnou kvalitu nálezových dat bylo dosaženo obdobných hodnot, jaké byly uvedeny u popisu růstového prostředí *Medicago truncatula* (FAO, 2017).

Tab. 14 Přehled převládajících stanovištních podmínek

Veličina	Hodnota
Průměrná roční teplota	14–16 °C
Roční úhrn srážek	400–600 mm
Nadmořská výška	0–200 m n. m.
Sklon svahu	0–5°
pH půdy	7,5–8

5.3 Mapové výstupy

Pro vizualizaci nálezových dat bylo vytvořeno několik mapových výstupů ve formě tištěných produktů a online vizualizační aplikace.

Analogové mapy

Byly vytvořeny 4 tištěné mapy s vizualizací nálezových dat nad několika druhy podkladových vrstev (fyzicko-geografická, roční úhrn srážek, průměrná roční teplota, vizualizace nejistoty). Veškeré tištěné mapy byly vyhotoveny na výkresy o formátu A3, samotné mapové pole je v měřítku 1 : 15 500 000. Pro obecnou mapu a pro vizualizaci polohové přesnosti byl použit rastr Natural Earth Data 1. Všechny mapy jsou přiloženy jako volné přílohy. Vzhledem k absenci potřebných metadat nálezových dat a nemožnosti dohledání časových údajů, nesplňují mapové výstupy kartografické zásady z pohledu časového určení.

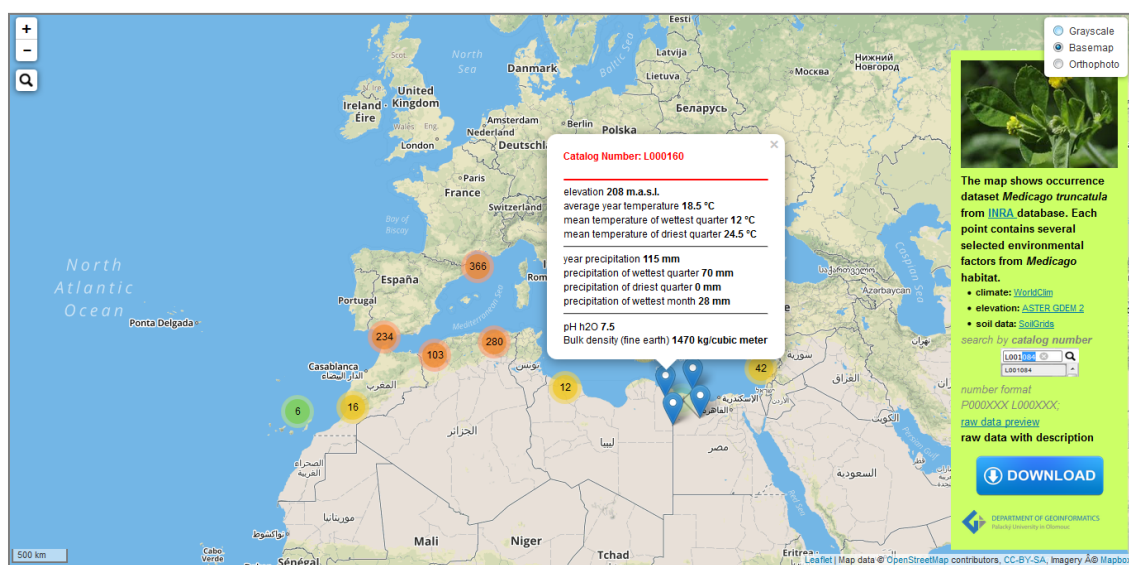
Online vizualizační nástroj

Pro konstrukci bylo zvoleno řešení pomocí open-source JavaScript knihovny Leaflet ve verzi 0.7.3 (Leaflet, 2017), která umožňuje vytvářet webové mapové aplikace. Zobrazovaná data tvoří soubor typu JSON.

Nástroj umožňuje volbu 3 podkladových map a jednu tematickou vrstvu *Medicago truncatula*. Pokladové mapy tvoří ortofoto, fyzicko-geografická mapa a mapa v odstínech

šedi. Možnost volby podkladové mapy je umístěna v kontrolním panelu v pravém horním rohu. Zobrazení informací samotných bodů probíhá prostřednictvím Pop-up okna s několika vybranými charakteristikami z každého tématu. Na informačním panelu v pravé části se nachází možnost stažení dat v csv. formátu se všemi extrahovanými atributy včetně souboru s jejich popisem.

V levém horním rohu je umístěno vyhledávací pole (viz obr. 20), ve kterém je možné vyhledat bod pomocí jeho katalogového čísla. Základní knihovna Leaflet neumožňuje vyhledávání, proto byl použit volně dostupný plugin Leaflet Control Search z webových stránek <http://labs.easyblog.it/maps/leaflet-search/>. Vzhledem k velkému počtu bodů a urychlení načítání byl použit plugin Leaflet.markercluster, pomocí kterého při různých úrovních zoomu body vytvářejí shluky. Po kliknutí na vybraný bod se zobrazí pop-up okno s vybranými atributy. Všechny popisky jsou provedeny v anglickém jazyce pro širší možnosti využití nástroje. Dostupný na webových stránkách k bakalářské práci <http://www.geoinformatics.upol.cz/dprace/bakalarske/kominek17/>.



Obr. 20 Prostředí mapové aplikace

6 DISKUZE

Práce byla obsáhlá jak na použitá data, tak na software, metody a postupy zpracování. Cílem bylo vytvoření statistických přehledů stanovištních podmínek, vhodné vizualizace pomocí mapových výstupů a nové datové sady nálezových dat s environmentálními faktory. Pro extrakci environmentálních faktorů byla použita data o různých měřících. Pokud by se jednalo o podrobnější studii, bylo by zapotřebí vybrat menší zkoumaný vzorek a environmentální data se stejnými měřící, a hlavně přesněji lokalizovaná nálezová data. Na výsledky práce je proto nutné nahlížet z pohledu snahy o zachycení určitých trendů.

Při řešení nastalo mnoho problémů, na které bylo nutné najít řešení. Již při výběru dat nastala otázka, jaká data budou nejvhodnější a jakým způsobem je vybrat. Vzhledem k tomu, že výsledná datová sada by měla sloužit pro výzkum dormance *Medicago truncatula*, byla vybrána data, která by na to mohla mít největší vliv. Jejich výběr proběhl na základě dostupné literatury a konzultace s vedoucím práce a doc. Ing. Petrem Smýkalem, Ph.D., který uvedl, že všechna vybraná data mohou mít vliv na dormanci semen. Nejen zaměření datových vrstev, ale i výběr jejich zdroje a měřítka může hrát roli. U klimatických dat byl výběr velmi omezen a nejvhodnější datová sada, která by nejlépe pokrývala vymezené území v jednotném měřítku, byla pouze sada WorldClim. U zdrojů dat krajinného pokryvu se nabízelo několika možností. Pro účely práce bylo proto vybráno hned několik zdrojů z důvodu jejich odlišné klasifikace. Prakticky tak byly pokryty všechny globální zdroje v požadovaném měřítku, které byly v době psaní práce dostupné. Zaměření bylo především na popis stanovištních podmínek, proto byl kladen důraz na výběr půdních dat. Nálezová data se rozkládala na rozsáhlém území (především v okolí Středozemního moře, v Iránu a v Jihoafrické republice). Z tohoto důvodu nebylo možné unifikovat měřítko pro všechny datasety. Zde bylo otázkou, zda zvolit méně podrobné měřítko rastru např. 5–10 km/pixel nebo zvolit pro některá data jemnější rozlišení. Při příliš hrubém měřítku by mohly být informace značně zkreslené, a proto raději byla vyhledávána data v měřítku 1 km/pixel a podrobnějším. Pomocí několika metod bylo otestováno, jak moc se mohou data v okolí měnit a jaký vliv tato skutečnost může mít na celkovou statistiku souboru. Zde je prostor k rozvoji tohoto tématu např. v rámci bakalářské práce, která by se zaměřovala na nejistotu v datech.

Při zpracování bylo provedeno několik kroků, které by mohly přinést zkreslený výsledek. Prvním bylo odstranění bodů se stejnou souřadnicí a ponechání pouze jednoho záznamu na lokaci. V tomto případě nebylo zcela jasné, zda se jedná o nadbytek rostlin v daném místě nebo je tato skutečnost dána pouze větším množstvím sběrů, tato důležitá metadata, zde bohužel chyběla. Po konzultaci s vedoucím práce a docentem Smýkalem bylo zvoleno řešení v podobě redukce bodů se stejnou souřadnicí.

Dalším krokem bylo zavedení dodatečného atributu prostorové přesnosti. Z pohledu obsahu je tento atribut subjektivní, nicméně může být dobrým ukazatelem rozdílu kvality dat. Na první pohled je zřejmé, že souřadnice udaná pouze na celé stupně nebude zcela přesná a velmi pravděpodobně došlo k jejímu znehodnocení v průběhu práce. Vzhledem k možnému posunu polohy bodu při zpřesnění souřadnic o desetiny nebo setiny (viz podkapitola 4.1.1) je nutné s těmito daty pracovat obezřetně.

Vzhledem k různým datovým zdrojům, odlišným měřítkům a také ne příliš přesné poloze nálezových dat, mnohdy nastal problém, že body ležely svojí polohou při okraji pevniny na hladině moře. U rastrů, kde bod ležel na vodní hladině v blízkosti pevnin,

byla informace extrahována z nejbližšího pixelu. Z tohoto důvodu bylo také přistoupeno k otestování, jak moc se mohou měnit hodnoty v nejbližším okolí pixelu. Následně bylo zjištěno, že v nejbližším okolí pixelu není příliš velký rozsah hodnot. Proto byl postup extrakce hodnoty z nejbližšího bodu vyhodnocen jako dostatečně přesný. Alternativní řešení by mohlo spočívat ve filtraci a odstranění těchto bodů před statistickým zpracováním.

Při výpočtu orientace a sklonu svahů byla použita celá řada kroků, která vedla ke konečnému výsledku. Vzhledem k velké datové náročnosti digitálního modelu nebylo reálně možné spočítat sklon či orientaci pro celé území v rámci jednoho výstupního rastru. Z tohoto důvodu byl zvolen způsob extrakce nejbližšího okolí každého bodu a tvorby těchto charakteristik zvlášť. Aster GDEM je v souřadnicovém systému WGS 84 a před výpočtem orientace a sklonu svahů je nutné ho převést na projekční systém v metrových souřadnicových jednotkách. Vzhledem k faktu, že se body rozkládají na velkém území, není možné použít jeden souřadnicový systém, aniž by došlo k výraznějšímu zkreslení některých oblastí (viz podkapitola 4.1.3). Proto bylo zvoleno řešení převodu rastru a bodů po částech do UTM zón a následného výpočtu. Jelikož výsledné rastry orientace a sklonu jsou konečným procesem mnoha dílčích kroků, mohlo během těchto kroků dojít k posunu a částečnému zkreslení rastrů a výsledek nemusí být zcela korektní.

Při zpracování mapových výstupů nebylo možné zjistit časové vymezení nálezových dat. Tato informace proto nemohla být vložena do podtitulu mapy. Mapy tím nesplňují časové vymezení a nejsou z hlediska kartografie zcela v pořádku. Po uvážení a bylo rozhodnuto místo časových odhadů toto vymezení raději neuvádět.

Výsledná datová sada obsahuje mnohem více atributů, než bylo zpracováno ve statistických přehledech. Zde byly zpracovány jen některé, zejména s ohledem na celkové popsání geografického rozšíření. Finální datová sada svojí obsáhlostí environmentálních faktorů z velkého množství zdrojů skýtá velký potenciál pro její další využití a pozdějšímu zpracování. Skvělým příkladem by bylo její využití pro některou bakalářskou práci z oblasti biologie či botaniky. Vždy ale musí být brán zřetel na ne příliš kvalitní nálezová data a různá měřítka environmentálních vrstev a další faktory uvedeny výše.

Vzhledem k rozsáhlé pokryvnosti nálezových dat, práce nenabízela mnoho alternativních řešení nebo možnost použití více různých datových sad. Zajímavý výsledek by mohla přinést studie se zaměřením na nálezová data v rámci ČR, při které by bylo možné porovnat, jak velký je rozdíl hodnot odvozených globálních nebo lokálních dat. Vytvořený online nástroj přinesl pozitivní reakce ze strany lidí z negeografického prostředí, kteří nejsou zvyklí pracovat s GIS. Přináší nový pohled na zkoumaná data obohacený o vymezení v prostoru a v budoucnu by mohl být rozšířen o další funkce, jako je dynamické vykreslování grafů, výběr bodů v oblasti apod.

Práce s daty obsahovala mnoho úkolů k jejich řešení, které s sebou přinesly mnohdy velmi zajímavé postupy a nové poznatky z oblasti GIS, které předčily očekávání autora a ukazují průřez napříč celou řadou dovedností, které autor získal během svého bakalářského studia.

7 ZÁVĚR

Cílem práce bylo popsání stanovištních podmínek *Medicago truncatula*. Mezi další cíle patřilo vytvoření statistických přehledů a mapových výstupů.

Během průzkumu volně dostupných dat byly vyhledány a popsány volně dostupné datové zdroje včetně problémů, které mohou nastat při jejich zpracování. Vstupní data nebyla příliš kvalitní a v případě jejich souřadnic bylo nutné uvažovat i o možné nepřesnosti. Před samotným statistickým zpracováním musela být data značně upravena. Z důvodu ne příliš kvalitních nálezových dat a nejednotného měřítka rastrových vrstev, nelze považovat statistické charakteristiky za zcela vypovídající. I přes různou kvalitu dat bylo možné vypořizovat určité trendy a v rámci výzkumu lze výsledek považovat jako vypovídající.

Výsledek neslouží jen jako statistický přehled. Celá práce popisuje velké množství datových sad, možných problémů a řešení při jejich zpracování a nabízí možnost jednoduché online vizualizace výsledných dat pro široké spektrum odborníků. Díky provázanosti napříč obory je zajímavým příkladem využití geoinformatiky v praxi.

Vytvořená datová sada a mapové výstupy budou dále sloužit pro potřeby botaniků z Katedry botaniky Univerzity Palackého v Olomouci, kteří je využijí pro své potřeby při výzkumu dormance semen *Medicago truncatula*.

POUŽITÁ LITERATURA A INFORMAČNÍ ZDROJE

BONTEMPS, Sophie, et al. GLOBCOVER 2009-Products description and validation report. 2011.

ADAMS, John B.; GILLESPIE, Alan R. Remote sensing of landscapes with spectral images: A physical modeling approach. Cambridge University Press, 2006.

ASCHMANN, Homer. Distribution and peculiarity of Mediterranean ecosystems. In: Mediterranean type ecosystems. Springer Berlin Heidelberg, 1973. p. 11-19.

Bounejmate, M., A. D. Robson, and P. E. Beale. "Annual Medicago species in Morocco. II. Distribution in relation to soil and climate." *Australian Journal of Agricultural Research* 43.3 (1992): 751-763.

BRANDSAETER, Lars Olav, et al. Freezing resistance of winter annual and biennial legumes at different developmental stages. *Crop Science*, 2002, 42.2: 437-443.

BRADLEY, Bethany A.; FLEISHMAN, Erica. Can remote sensing of land cover improve species distribution modelling?. *Journal of Biogeography*, 2008, 35.7: 1158-1159.

BURGARELLA, Concetta, et al. Adaptation to climate through flowering phenology: a case study in *Medicago truncatula*. *Molecular ecology*, 2016, 25.14: 3397-3415.

BUCHTOVÁ, Tereza. Analýza mutantních linií vojtěšky pro studium funkce genů aktivních během vývoje osemení. Olomouc, 2016. Bakalářská práce. Univerzita Palackého v Olomouci. Vedoucí práce Ing. Petr Smýkal, Ph. D.

BUDDENHAGEN, I. W. Legumes in farming systems in Mediterranean climates. The role of legumes in the farming systems of the Mediterranean areas, 1990, 3-29.

CAO, Bo, et al. Modeling habitat distribution of *Cornus officinalis* with Maxent modeling and fuzzy logics in China. *Journal of Plant Ecology*, 2016, 9.6: 742-751.

Colebatch, G., B. Trevaskis, and M. Udvardi. "Symbiotic nitrogen fixation research in the postgenomics era." *New Phytologist* 153.1 (2002): 37-42.

Clarkson, N. M., and J. S. Russell. "Flowering responses to vernalization and photoperiod in annual medics (*Medicago* spp.)." *Australian Journal of Agricultural Research* 26.5 (1975): 831-838.

DALLMAN, Peter R. Plant life in the world's Mediterranean climates: California, Chile, South Africa, Australia, and the Mediterranean basin. Univ of California Press, 1998.

DORMANN, Carsten F., et al. Correlation and process in species distribution models: bridging a dichotomy. *Journal of Biogeography*, 2012, 39.12: 2119-2131

DRAKE, JOHN M., CHRISTOPHE RANDIN a ANTOINE GUISAN. Modelling ecological niches with support vector machines. *Journal of Applied Ecology*. 2006, 43(3), 424-432. DOI: 10.1111/j.1365-2664.2006.01141.x. ISSN 0021-8901

DE LA ESTRELLA, Manuel, et al. Legume diversity patterns in West Central Africa: influence of species biology on distribution models. *PloS one*, 2012, 7.7: e41526.

EVANS, J.; DEAR, B.; O'CONNOR, G. E. Influence of an acid soil on the herbage yield and nodulation of five annual pasture legumes. *Animal Production Science*, 1990, 30.1: 55-60.

FLINT, Lorraine E.; FLINT, Alan L. Downscaling future climate scenarios to fine scales for hydrologic and ecological modeling and analysis. *Ecological Processes*, 2012, 1.1: 2.

FRANKLIN, Janet. Mapping species distributions: spatial inference and prediction. Cambridge University Press, 2010.

GELLER, Gary N., et al. Remote Sensing for Biodiversity. In: *The GEO Handbook on Biodiversity Observation Networks*. Springer International Publishing, 2017. p. 187-210.

He, Kate S., et al. "Will remote sensing shape the next generation of species distribution models?." *Remote Sensing in Ecology and Conservation* 1.1 (2015): 4-18.

HIJMANS, Robert J., et al. Very high resolution interpolated climate surfaces for global land areas. *International journal of climatology*, 2005, 25.15: 1965-1978.

HUANG, Cho-ying; ASNER, Gregory P. Applications of remote sensing to alien invasive plant studies. *Sensors*, 2009, 9.6: 4869-4889.

KORČÁKOVÁ, Marta. Studium vegetativní a generativní reprodukce vybraných druhů plevelů. Praha, 2011. Disertace. Česká zemědělská univerzita v Praze. Vedoucí práce Doc. Ing. Evžen Prokinová, CSc.

KUMAR, Sunil, Lisa G. NEVEN a Wee L. YEE. Evaluating correlative and mechanistic niche models for assessing the risk of pest establishment. *Ecosphere*. 2014, 5(7), art86. DOI: 10.1890/ES14-00050.1. ISSN 2150-8925."

LESINS, Karlis Adolfs; LESINS, Irma. Genus *Medicago* (Leguminosae): a taxogenetic study. Springer Science & Business Media, 2012.

MOREAU, Delphine; SALON, Christophe; MUNIER-JOLAIN, NATHALIE. A model-based framework for the phenotypic characterization of the flowering of *Medicago truncatula*. *Plant, cell & environment*, 2007, 30.2: 213-224.

MORIUCHI, Ken S., et al. Salinity adaptation and the contribution of parental environmental effects in *Medicago truncatula*. *PloS one*, 2016, 11.3: e0150350.

NACHTERGAELE, Freddy, et al. Harmonized world soil database. Wageningen: ISRIC, 2009

PETERSON, David L.; SCHREINER, Edward G.; BUCKINGHAM, Nelsa M. Gradients, vegetation and climate: spatial and temporal dynamics in the Olympic Mountains, USA. *Global ecology and biogeography Letters*, 1997, 7-17.

PIERRE, Jean-Baptiste, et al. Detection of QTLs for flowering date in three mapping populations of the model legume species *Medicago truncatula*. *Theoretical and Applied Genetics*, 2008, 117.4: 609-620

PIANO, E.; TALAMUCCI, P. Annual self-regenerating legumes in Mediterranean areas. In: G. Parente, J. Frame and S. Orsi (eds), " Grassland and land use systems". Proc. XVI EGF General Meeting, Grado (Italia). 1996. p. 895-909.

POLECHOVÁ, Jitka; STORCH, David. Ecological niche. *Encyclopedia of ecology*, 2008, 2: 1088-1097.

PROCHÁZKA, Stanislav a Jiří ŠEBÁNEK. Regulátory rostlinného růstu: Physiology of vegetative propagation of woody species : monografie. Praha: Academia, 1997. *Folia Universitatis Agriculturae et Silviculturae Mendelianae Brunensis*. ISBN 80-200-0597-8.

RANDIN, Christophe F., et al. Climate change and plant distribution: local models predict high-elevation persistence. *Global Change Biology*, 2009, 15.6: 1557-1569

REDDY, C. Sudhakar. Applications of remote sensing in plant sciences: an overview. In: *Plant Biology and Biotechnology*. Springer India, 2015. p. 713-727.

RUSHTON, S. P.; ORMEROD, S. J.; KERBY, G. New paradigms for modelling species distributions?. *Journal of applied ecology*, 2004, 41.2: 193-200.

SMÝKAL, Petr, et al. Legume crops phylogeny and genetic diversity for science and breeding. *Critical Reviews in Plant Sciences*, 2015, 34.1-3: 43-104

THROWER, Norman JW; BRADBURY, David E. The physiography of the Mediterranean lands with special emphasis on California and Chile. Mediterranean Type Ecosystems, Origin and Structure, 1973, 37-52.

TRIVEDI, Mandar R., et al. Spatial scale affects bioclimate model projections of climate change impacts on mountain plants. *Global Change Biology*, 2008, 14.5: 1089-1103.

UZUN, Ferat; AYDIN, Ibrahim. Improving germination rate of *Medicago* and *Trifolium* species. *Asian Journal of Plant Sciences*, 2004, 3.6: 714-717.

XIE, Yichun; SHA, Zongyao; YU, Mei. Remote sensing imagery in vegetation mapping: a review. *Journal of plant ecology*, 2008, 1.1: 9-23.

YODER, Jeremy B., et al. Genomic signature of adaptation to climate in *Medicago truncatula*. *Genetics*, 2014, 196.4: 1263-1275.

ZHANG, Anzhi; JIA, Gensuo. Monitoring meteorological drought in semiarid regions using multi-sensor microwave remote sensing data. *Remote Sensing of Environment*, 2013, 134: 12-23

Internetové zdroje

Copernicus [online]. 2017 [cit. 2017-05-08]. Dostupné z: <http://land.copernicus.eu/>

Dálkový průzkum země. *Copernicus* [online]. 2017 [cit. 2017-05-08]. Dostupné z: <http://copernicus.gov.cz/dalkovy-pruzkum-zeme>

ESA/CCI viewer. *Esa* [online]. [cit. 2017-05-08]. Dostupné z: <http://maps.elie.ucl.ac.be/CCI/viewer/download.php>

Free Gis Datasets. *Free Gis Datasets* [online]. [cit. 2017-05-08]. Dostupné z: <https://freegisdata.rtwilson.com/>

Gisat / ASTER GDEM. *GISAT* [online]. [cit. 2017-05-08]. Dostupné z: <http://www.gisat.cz/content/cz/produkty/digitalni-model-terenu/aster-gdem>

Gisat/zprávy ze světa/Mapa GlobCover 2009. *Gisat* [online]. 2011 [cit. 2017-05-08]. Dostupné z: <http://www.gisat.cz/content/cz/novinky/zpravy-ze-sveta/mapa-globcover-2009>

Harmonised World Soil Database v.1.2. *Food and Agriculture Organization of the United Nations* [online]. [cit. 2017-05-08]. Dostupné z: <http://www.fao.org/soils-portal/soil-survey/soil-maps-and-databases/harmonized-world-soil-database-v12/en/>

Welcome to the internet site on Medicago truncatula. *INRA* [online]. [cit. 2017-07-14].
Dostupné z: <http://www1.montpellier.inra.fr/BRC-MTR/accueil.php?menu=medicago&page=menu11>

Leaflet - a JavaScript library for interactive maps [online]. 2017 [cit. 2017-05-08].
Dostupné z: <http://leafletjs.com/>

Medicago truncatula. *FAO* [online]. [cit. 2017-05-06]. Dostupné z:
<http://www.fao.org/ag/agp/agpc/doc/gbase/data/meditrunc.htm>

Shuttle Radar Topography Mission. *NASA Jet Propulsion Laboratory (JPL)* [online]. [cit. 2017-05-08]. Dostupné z: <https://www2.jpl.nasa.gov/srtm/>

SoilGrids: global gridded soil information. *ISRIC* [online]. 2017 [cit. 2017-05-08].
Dostupné z: <http://www.isric.org/explore/soilgrids>

The USGS Land Cover Institute. *U.S. Geological Survey* [online]. 2017 [cit. 2017-05-08].
Dostupné z: https://landcover.usgs.gov/global_climatology.php

Shuttle Radar Topography Mission [online]. 2017 [cit. 2017-05-08]. Dostupné z:
<https://www2.jpl.nasa.gov/srtm/>

WorldClim [online]. 2017 [cit. 2017-05-08]. Dostupné z: <http://worldclim.org>

PŘÍLOHY

SEZNAM PŘÍLOH

Vázané přílohy

Příloha 1 Tabulka: Přehled zdrojů použitých dat

Volné přílohy

Příloha 2a Mapa nálezů *Medicago truncatula*
Příloha 2b Mapa ročního úhrnu srážek a nálezů *Medicago truncatula*
Příloha 2c Mapa průměrné roční teploty a nálezů *Medicago truncatula*
Příloha 2d Mapa kvality prostorové informace nálezů *Medicago truncatula*
Příloha 3 DVD
Příloha 4 Poster

Popis struktury DVD

Adresáře:

- Data:
 - Příloha 3a Výsledná datová sada xls
 - Příloha 3b Výsledná datová sada Esri Shapefile
 - Příloha 3c Výsledná datová sada Esri Feature Class
 - Příloha 3d Datová sada pro statistické přehledy xls.
- Toolboxy:
 - Příloha 3e Rozdělení rastrů
 - Příloha 3f Extrakce hodnot rastrů
 - Příloha 3g Detekce rizikových bodů
- Tabulky:
 - Příloha 3h Přehled zdrojů použitých dat
 - Příloha 3i Popis atributů výsledné sady
 - Příloha 3j Proměnlivost hodnot náhodných výběrů
 - Příloha 3k Zastoupení kategorií krajinného pokryvu
 - Příloha 3l Korelace vstupních dat
- Video:
 - Příloha 3m Připojení HWSD databáze
- Text práce:
 - Text práce
- Přílohy:
 - Tištěné mapy, poster
- Webové stránky
- Metadata:
 - Metadatový záznam xml.

Příloha 1

	Worldclim 2.0	Aster GDEM 2.0.	HWSD v 1.2	SoilGrids	GlobCover 2.3	CCI Land Cover 2015
formát	GeoTIFF	GeoTIFF	GeoTIFF + mdb.	GeoTIFF	GeoTIFF	GeoTIFF
datum publikace	2016	2011	2012	2014	2010	2016
prostorové rozlišení	1 km/pixel	30 m/pixel	1 km/pixel	1 km/pixel	300 m/pixel	300 m/pixel
časové rozlišení	1970–2000	2009–2011	–	–	2009	2015
epsg	4326	4326	4326	4326	4326	4326
pokrytí	celý svět	83° s. š. - 83° j. š.	celý svět	celý svět	celý svět	celý svět
obsahová složka	teplota, srážky, solární radiace, bioklimatické charakteristiky	digitální model terénu	13 půdních charakteristik	17 půdních charakteristik	23 tříd krajinného pokryvu	22 tříd krajinného pokryvu
poskytování dat	1 rastr (celý svět) - 1 charakteristika	dlaždice 1x1 °	1 rastr (celý svět)	1 rastr (celý svět) - 1 charakteristika	1 rastr (celý svět)	1 rastr (celý svět)
velikost	1-2 GB	24 MB.	rastr 20MB (zip) mdb.databáze 45 MB.	30–250 MB	363 MB	258 MB
distribuce	přímé stažení	přímé stažení, ftp	přímé stažení	ftp	přímé stažení	přímé stažení
nutnost registrace	ne	ano	ne	ne	ne	ne
zdroje dat	meteo stanice, klimatické databáze (GHCN, CLINO, FAOCLIM..)	senzor ASTER na satelitu Terra	SOTER, ESD, Soil Map of China, WISE	12 databází půdních profilů (110 000 bodů)	envisat satellite	Proba-V
licence	creative commons 3.0	pouze osobní použití	pouze pro vědecké a nekomerční účely	Open Data Commons Open Database License	pouze pro vědecké a nekomerční účely	pouze pro vědecké a nekomerční účely

