

Univerzita Palackého v Olomouci

Přírodovědecká fakulta

Katedra geoinformatiky

**PROSTOROVĚ ZALOŽENÝ MODEL PRO
PREDIKCI POTENCIÁLNÍHO VÝSKYTU SRŠNĚ
ASIJSKÉ V ČESKÉ REPUBLICE**

Diplomová práce

Bc. Ondřej BIEMANN

Vedoucí práce RNDr. Jan BRUS, Ph.D.

Olomouc 2023

Geoinformatika a kartografie

ANOTACE

Sršeň asijská *Vespa velutina* je v současné době v Evropě rychle se rozšiřující druh invazivního hmyzu. Pro své četné dopady na biodiverzitu se stává vážnou zdravotní, ekologickou, ale i ekonomickou hrozbou. Ve velkém decimuje včelstva, a působí tak velké škody nejen včelařům, ale i celé společnosti. Tento druh je sice o něco menší než sršeň obecná, zato ale agresivitou připomíná včely útočné. Proto je potřeba další šíření tohoto druhu náležitě monitorovat, k čemuž mohou pomoci prostorové modely. Cílem této diplomové práce je sestavit prostorově založený model pro predikci výskytu sršně asijské v České republice a zjištěné výsledky spolu s doprovodnými informacemi prezentovat formou map a informační brožury pro laickou veřejnost. Na základě poznatků z rešerše a statistického testování byly stanoveny nejdůležitější převážně klimatické proměnné, které ovlivňují výskyt sršně asijské na Evropském kontinentu. Získané znalosti poté byly uplatněny v praxi při modelování míry vhodnosti habitatu pro sršeň asijskou. Výsledkem práce je sestavený predikční model, soubor map prezentujících vhodnost podmínek pro šíření sršně asijské v České republice a brožura, která na několika stranách poskytuje nejdůležitější informace o sršni asijské, jejím dopadům na přírodu a možnostech obrany proti jejím negativním vlivům. Brožura, na jejímž podkladu je vytvořen i krátký informační leták, zároveň prezentuje vzniklé mapy a poskytuje vizualizace výskytu sršně v takové podobě a podrobnosti, v jaké je v Česku ani jinde v Evropě ještě nikdo neprezentoval. Digitální verze této brožury i informačního letáku jsou dostupné v PDF na webu i pro širokou veřejnost.

KLÍČOVÁ SLOVA

Prostorové modelování; sršeň asijská; Maxent; Wallace; bioklimatické podmínky

Počet stran práce: 75

Počet příloh: 11 (z toho 5 volných)

ANNOTATION

The Asian hornet *Vespa velutina* is currently a rapidly spreading species of invasive insect in Europe. Due to its numerous impacts on biodiversity, it is becoming a serious health, ecological and economic threat. It decimates bee colonies on a large scale, causing great damage not only to beekeepers but also to society. This species is slightly smaller than the common hornet, but its aggressiveness is similar to that of the European wasp. Therefore, the further spread of this species needs to be properly monitored, for which spatial models can help. This thesis aims to build a spatially based model for the prediction of the distribution of the Asian hornet in the Czech Republic and to present the results together with accompanying information in the form of maps and an information brochure for the public. Based on the findings of the research and statistical testing, the most influential predominantly climatic variables affecting the occurrence of Asian hornets on the European continent were determined. The knowledge gained was then put into practice in modelling habitat suitability rates for the Asian hornet. The results of the work are a prediction model, a set of maps presenting the suitability of the conditions for the distribution of the Asian hornet in the Czech Republic and a brochure that provides the most important information on the Asian hornet, its impacts on nature and the possibilities of defence against its negative effects. The brochure, which is also the basis for a short information leaflet, also presents the resulting maps and provides visualisations of the hornet's range in a form and detail that has never been presented in the Czech Republic or elsewhere in Europe. Digital versions of this brochure and the information leaflet are available in PDF format on the web and for the public.

KEYWORDS

Spatial modelling; Asian hornet; Maxent; Wallace; bioclimatic conditions

Number of pages: 75

Number of appendixes: 11

Prohlašuji, že

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

- jsem si vědom, že na moji diplomovou 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 nevydělečně, ke své vnitřní potřebě, diplomovou práci užívat (§ 35 odst. 3),

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

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

- použít výsledky a výstupy mé diplomové 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 4. 5. 2023

Bc. Ondřej BIEMANN
podpis autora

Chtěl bych velmi poděkovat vedoucímu své práce RNDr. Janu Brusovi, Ph.D. za trpělivost, vstřícnost, přínosné konzultace, věcné připomínky a rady během vypracovávání této diplomové práce. Velké poděkování patří i mé rodině za neustálou nejen psychickou podporu v době studia a psaní této práce.

UNIVERZITA PALACKÉHO V OLMOUCI

Přírodovědecká fakulta

Akademický rok: 2021/2022

ZADÁNÍ DIPLOMOVÉ PRÁCE

(projektu, uměleckého díla, uměleckého výkonu)

Jméno a příjmení: Bc. Ondřej BIEMANN
Osobní číslo: R21849
Studijní program: N0532A330009 Geoinformatika a kartografie
Téma práce: Prostorově založený model pro predikci potenciálního výskytu Sršně asijské v České republice
Zadávající katedra: Katedra geoinformatiky

Zásady pro vypracování

Cílem diplomové práce je vytvoření prostorově založeného modelu a následné vizualizace potenciálního výskytu Sršně asijské (*Vespa velutina*) v České republice. Student na základě současných přístupů a dat o výskytu v Evropě a Asii provede modelování potenciálních lokalit výskytu v ČR s využitím GIS. Modelování bude založeno na environmentálních charakteristikách a stanovištních podmínkách. Součástí práce bude také predikce výskytu založená na předpokládaných změnách klimatu. Výstupem práce bude výsledný model a soubor map, který bude sloužit pro informování včelařů. Dalšími výstupy práce budou brožura a informační leták pro zvyšování povědomí o Sršni asijské mezi odbornou a laickou veřejností.

Student vyplní údaje o všech datových sadách, které vytvořil nebo získal v rámci práce do Metainformačního systému katedry geoinformatiky a současně vytvoří zálohu údajů ve formě validovaného XML souboru. Celá práce (text, přílohy, výstupy, zdrojová a vytvořená data, XML soubor) se odevzdá v digitální podobě na CD (DVD) a text práce s vybranými přílohami bude odevzdán ve dvou svázaných výtiscích na sekretariát katedry. O diplomové práci student vytvoří webovou stránku v souladu s pravidly dostupnými na stránkách katedry. Práce bude zpracována podle zásad dle Voženílek (2002) a závazné šablony pro diplomové práce na KGI. Povinnou přílohou práce bude poster formátu A2.

Rozsah pracovní zprávy: max. 50 stran
Rozsah grafických prací: dle potřeby
Forma zpracování diplomové práce: tištěná

Seznam doporučené literatury:

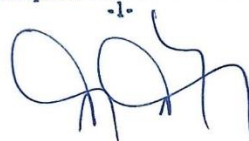
- BESSA, Ana S., João CARVALHO, Alberto GOMES, Frederico SANTARÉM, Raphael DIDHAM a Seán BRADY. Climate and land-use drivers of invasion: predicting the expansion of *Vespa velutina nigrithorax* into the Iberian Peninsula. *Insect Conservation and Diversity* [online]. 2016, 9(1), 27-37 [cit. 2021-12-16]. ISSN 1752458X. Dostupné z: doi:10.1111/icad.12140
- DRAKE, John B. *Climate Modeling for Scientists and Engineers*. Knoxville: University of Tennessee, 2014. ISBN 978-1-611973-53-2.
- HERRERA, Cayetano, José JURADO-RIVERA a Mar LEZA. Ensemble of Small Models As a Tool For Alien Invasive Species Management Planning: Evaluation of *Vespa Velutina* (Hymenoptera: Vespidae) Under Mediterranean Island Conditions. *Research Square* [online]. Universitat de les Illes Balears, 2021 [cit. 2021-12-16]. Dostupné z: doi:10.21203/rs.3.rs-809089/v1
- MCSWEENEY, Robert a Zeke HAUSFATHER. Q&A: How do climate models work? *Carbon Brief* [online]. United Kingdom, 2018 [cit. 2021-12-16]. Dostupné z: <https://www.carbonbrief.org/qa-how-do-climate-models-work>
- SHEN, Samuel S. P. and Richard C. J. SOMERVILLE, 2019. *Climate Mathematics: Theory and Applications* [online]. Cambridge: Cambridge University Press. Retrieved z: doi:10.1017/9781108693882
- VERDASCA, Maria João, Hugo REBELO, Luísa G. CARVALHEIRO a Rui REBELO. Invasive hornets on the road: motorway-driven dispersal must be considered in management plans of *Vespa velutina*. *NeoBiota* [online]. 2021, 69, 177-198 [cit. 2021-12-16]. ISSN 1314-2488. Dostupné z: doi:10.3897/neobiota.69.71352

Vedoucí diplomové práce: **RNDr. Jan Brus, Ph.D.**
Katedra geoinformatiky

Datum zadání diplomové práce: **9. prosince 2021**
Termín odevzdání diplomové práce: **5. května 2023**

UNIVERZITA PALACKÉHO V OLOMOUCI
PŘÍRODOVĚDECKÁ FAKULTA
KATEDRA GEOINFORMATIKY
17. listopadu 50, 771 46 Olomouc

-1-



L.S.

doc. RNDr. Martin Kubala, Ph.D.
děkan

prof. RNDr. Vít Voženílek, CSc.
vedoucí katedry

OBSAH

SEZNAM POUŽITÝCH ZKRATEK	9
ÚVOD	10
1 CÍLE PRÁCE.....	11
2 SOUČASNÝ STAV ŘEŠENÉ PROBLEMATIKY	12
2.1 Sršeň asijská (<i>Vespa velutina</i>)	12
2.1.1 Základní charakteristika.....	12
2.1.2 Evropské výzkumy modelující výskyt sršně asijské.....	16
2.2 Vybrané přístupy k modelování distribuce druhů.....	18
2.2.1 Bioklimatické obálky	19
2.2.2 Umělé neuronové sítě (ANN).....	19
2.2.3 Bioclim	20
2.2.4 Analýza hlavních komponent	20
2.2.5 Maxent.....	21
3 METODY A POSTUP ZPRACOVÁNÍ	22
4 VLASTNÍ ŘEŠENÍ.....	27
4.1 Seznámení s aplikací Wallace	27
4.2 Stažení nálezočných dat	28
4.3 Predikční proměnné	30
4.3.1 Primární výběr bioklimatických proměnných.....	31
4.3.2 Generování rastrů bioklimatických proměnných	32
4.3.3 Finální výběr proměnných statistickými metodami.....	41
4.3.4 Bioklimatické proměnné pro normál 2041–2060	45
4.4 Zpracování modelu v aplikaci Wallace.....	47
4.4.1 Získání dat výskytu druhu	49
4.4.2 Získání environmentálních dat	49
4.4.3 Zpracování dat výskytu	50
4.4.4 Zpracování environmentálních dat.....	51
4.4.5 Charakteristika environmentálního prostoru.....	53
4.4.6 Segmentace/Rozdělení dat výskytu.....	53
4.4.7 Vytvoření a vyhodnocení modelu niky	55
4.4.8 Vizualizace výsledků modelu	61
4.4.9 Přesunutí modelu.....	64
4.4.10 Reprodukce modelu.....	65
5 VÝSLEDKY	66
6 DISKUZE	72
7 ZÁVĚR	75
POUŽITÁ LITERATURA A INFORMAČNÍ ZDROJE	
PŘÍLOHY	

SEZNAM POUŽITÝCH ZKRATEK

Zkratka	Význam
V. v.	Vespa velutina
AUC	Area Under the Curve
CBI	Continuous Boyce Index
CSV	Comma Separated Values
DBF	dBASE Table File
DMR	Digitální model reliéfu
DTP	Desktop Publishing
EEA	European Environment Agency
IDW	Inverted Distance Weighing
MAP	Mean Annual Precipitation
MAT	Mean Annual Temperature
OR	Ommision Rater
PCA	Principle Component Analysis
PPT	Precipitation
RM	regularizační multiplikátor
SHP	Esri Shapefile
SRS	Souřadnicový referenční systém
Tave	Average Temperature
Tmax	Maximal Temperature
Tmin	Minimal Temperature
USGS	United States Geological Survey

ÚVOD

Současná míra globalizace, zvýšená mobilita a klimatické změny představují významnou hrozbu při introdukci invazivních druhů. Sršeň asijská *Vespa velutina* se přesně z těchto důvodů dostala do Evropy – první hnízdo bylo nalezeno ve Francii již v roce 2004, kam byla zavlečena v kontejneru z Asie. Díky své velké schopnosti přizpůsobovat se prostředí se byla schopna v evropských podmínkách udržet, ba dokonce se jí tu velmi daří (Díaz et al., 2022). Do jisté míry za to může i lhostejnost a nekoordinovaný postup při likvidaci prvních ohnisek sršně asijské v roce 2004 ve Francii – systematická likvidace hnízd ve Francii začala teprve *osm let poté*, kdy byl úřadům oficiálně nahlášen první nálezný sršně asijské. Tyto důvody vedly k tomu, že dnes je tento agresivní druh hmyzu výraznou zdravotní, ekologickou, ale i ekonomickou hrozbou. z jediné sršní královny se situace během téměř dvaceti let vyvinula v záplavu desítek tisíc hnízd napříč Evropou, z nichž každé čítá i několik tisíc jedinců, a další rychle přibývají, jak sršně kolonizují nové oblasti (Proková, 2023). Ve velkém decimuje včelstva, a působí tak velké škody nejen včelařům, ale i celé společnosti. I z toho důvodu je potřeba se připravit na budoucí možnost rozšíření této sršně i v České republice.

Práce navazuje na četná modelování výskytu sršně asijské v Evropě a využívá data o jejím současném celoevropském rozšíření. Cílem této diplomové práce je sestavit prostorově založený model, který by výskyt sršně asijské v České republice mohl predikovat a informovat občany, do jaké míry je potřeba se bát katastrofických scénářů, které jsou leckdy prezentovány v médiích. Šíření je predikováno na základě klimatických podobností se současnými lokalitami nálezu tohoto invazivního hmyzu jinde v Evropě.

Uživatelsky nejpřínosnějším materiálem je výsledná brožura, která na několika stranách poskytuje nejdůležitější informace o sršni asijské, jejím dopadům na přírodu a možnostech obrany proti jejím negativním vlivům. Brožura zároveň prezentuje vzniklé mapy a poskytuje informace o vhodnosti životního prostředí pro šíření sršně asijské, a to v takové podobě a podrobnosti, v jaké je v Česku ani jinde v Evropě ještě nikdo zřejmě neprezentoval.

1 CÍLE PRÁCE

Cílem diplomové práce je vytvoření prostorově založeného modelu a následné vizualizace potenciálního výskytu Sršně asijské (*Vespa velutina*) v České republice. Modelování potenciálních lokalit výskytu v ČR bude provedeno na základě současných přístupů a dat o výskytu v Evropě a Asii. Modelování bude založeno na environmentálních charakteristikách a stanovištních podmínkách. Součástí práce bude také predikce výskytu založená na předpokládaných změnách klimatu, a to do roku 2050.

Výstupem práce bude výsledný model a soubor map, který bude sloužit pro informování včelařů. Výstupy práce budou pro včelaře shrnuty do formy **brožury s informacemi o sršni asijské a vizualizacemi potenciálního výskytu v České republice**. Dále bude vytvořen informační leták pro zvyšování povědomí o sršni asijské mezi odbornou a laickou veřejností.

Dílčí cíle byly rozděleny na část teoretickou a část praktickou:

Teoretická část:

- **odborná rešerše** existujících prací – rešerše biologických charakteristik sršně asijské včetně bližšího popisu poddruhu rozšířeného v Evropě (*Vespa velutina nigrithorax*), popis stanovištních podmínek, vhodnost současných a budoucích klimatických podmínek na rozšíření sršně asijské v ČR, metody a přístupy k modelování výskytu živočišných druhů,
- analýza dostupných **datových sad** (klimatická, využití území atd.) a popis zvolených dat pro potřeby modelování.

Praktické cíle:

- zpracování dat, vytvoření predikčních rastrů pro modelování,
- vytvoření predikčního modelu,
- **zpracování** pravděpodobnostních map, brožury, letáku.

Po konzultaci s vedoucím práce a po prvotní analýze dostupných dat bylo stanoveno, že **data o výskytu sršně asijské budou brána pouze z Evropy**, nikoli i z Asie, jak je v původním zadání. Důvody jsou malá dostupnost dat o výskytu sršně asijské v Asii (která je zapříčiněna pravděpodobně faktem, že v Asii se jedná o běžný druh, zatímco v prostředí Evropy jde o druh vysoce monitorovaný pro jeho výrazné dopady na ekonomiku a společnost) a důraz na natrénování modelu na datech z Evropy, poněvadž evropské klimatické a krajinné charakteristiky mají logicky České republice mnohem blíže než ty asijské.

Součástí zadání práce je také tvorba webových stránek a posteru ve formátu A2 shrnujících celou práci. Konečné verze map, brožury a dalších materiálů budou umístěny spolu s posterem a textem práce na stránkách Katedry geoinformatiky UP.

Výsledky této práce umožní zejména včelařům, ale i ostatní odborné a laické veřejnosti nahlédnout do problematiky možného rozšiřování sršně asijské v České republice. Vzhledem k faktu, že se momentálně jedná o výraznou biologickou hrozbu (z důvodů, které budou dále rozebrány v práci), je potřeba veřejnost informovat o potenciálním nebezpečí, které její šíření v České republice přináší, a způsobech, jak se jí bránit, případně jak se na ni připravit. v neposlední řadě má tato práce přinést sofistikovanější a přehlednější vizualizace výskytu sršně asijské, neboť vizualizační část je i v evropském měřítku nedostatečná.

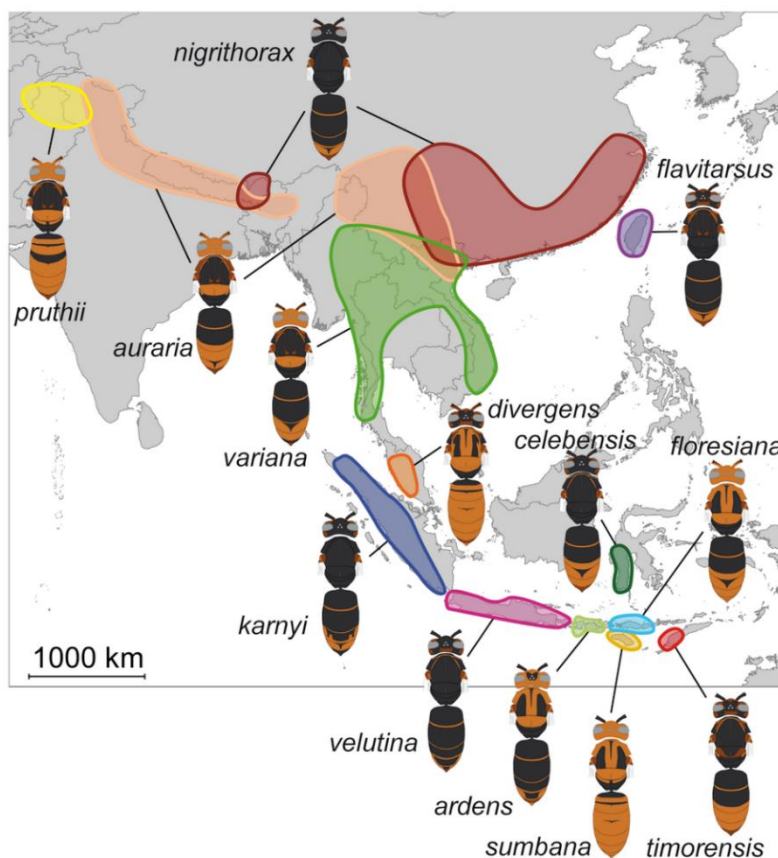
2 SOUČASNÝ STAV ŘEŠENÉ PROBLEMATIKY

Kapitola se zabývá charakteristikou sršně asijské, popsáním důvodů jejího negativního vlivu na prostředí Evropy a mechanismy, jak se jí bránit. Dále jsou popsány stanovištní charakteristiky, v nichž se sršeň vyskytuje a na závěr kapitoly jsou popsány vybrané modelovací přístupy.

2.1 Sršeň asijská (*Vespa velutina*)

2.1.1 Základní charakteristika

Sršeň asijská (*Vespa velutina*) je invazní druh sociálního blanokřídlého hmyzu z čeledi sršňovitých původně pocházející z Jihovýchodní Asie. v roce 2004 však byl poddruh *Vespa velutina nigrithorax* náhodně zavléčen v kontejneru z Číny na území Francie, odkud se začal rychle šířit do okolních oblastí – v Evropě je v současnosti rozšířen na území Itálie, Španělska, Portugalska, Německa, Belgie, Nizozemska, Velké Británie a Švýcarska (Obr. 2) (Proková, 2023).



Obr. 1 Rozšíření různých barevných forem sršně asijské (*Vespa velutina*) na území Asie. (zdroj: <https://cs.wikipedia.org/>)

Jedná se o menší až středně velký hmyz – královny dosahují velikosti okolo 32 mm, dělnice asi 22 mm, je tak asi o 10–20 % menší než v České republice běžná sršeň obecná (*Vespa crabro*). Jde o vysoce agresivního a efektivního predátora, jenž se specializuje na lov hmyzu, zejména včel a dalších druhů opylovačů. Včely tvoří až 70 % potravy sršně asijské, která tak výrazně decimuje včelí kolonie a vyvolává obavy včelařů. Navíc způsobuje

i výrazné škody v ovocných sadech a na vinicích. v Evropské unii je proto označena jako vysoce rizikový cizokrajný druh kvůli ohrožení původní biodiverzity (Diaz, et al., 2023).

Sršeň asijská se vyskytuje v několika barevných variantách, záleží na konkrétním poddruhu či regionální barevné variantě. Některé formy jsou zbarveny spíše do žluta či žlutooranžova (*Vespa velutina divergens*), jiné mají vyrovnaný poměr černého a žlutého zbarvení (*Vespa velutina variana*). Do Evropy byla zavlečena tmavě hnědá až černá varianta sršně asijské *Vespa velutina nigrithorax* (což doslova znamená „černá hrud“). Původní areály výskytu jednotlivých barevných forem jsou znázorněny na Obr. 1.

V rámci druhu je patrný pohlavní dimorfismus (Diaz et. al., 2023) – samce a samice navzájem odlišuje rozdílná velikost tykadel (samičí tykadla jsou kratší a tenčí než samčí) a proporce těla (samec má oválnější hrud a je mohutnější než samice). Dalším poznávacím znakem odlišujícím jednotlivá pohlaví je žihadlo – stejně jako u všech blanokřídlých platí, že samice jej mají, zatímco samci nikoli.

Odlišnosti od sršně obecné (*Vespa crabro*)

V České republice se původně vyskytuje jen jeden druh sršně, a to sršeň obecná (*Vespa crabro*). Ač je možné si sršeň asijskou se sršní obecnou splést, poměrně výrazně se liší velikostí a zbarvením. Níže je podrobnější výčet hlavních vizuálních odlišností (*pozn.: porovnávány jsou V. crabro s V. v. nigrithorax pro její evropské rozšíření, ostatní barevné varianty porovnávány nejsou*):

Sršeň obecná (*V. crabro*)

- délka těla: dělnice cca 25 mm, královna 35 mm
- hlava i hrud rezavá
- kompletně žluté nohy
- zadeček žlutý s černými pruhy

Sršeň asijská (*V. v. nigrithorax*)

- délka těla: dělnice cca 22 mm, královna 32 mm
- hlava shora černá, čelo žluté
- nohy černé, ale chodidla výrazně žlutá
- zadeček černý, jen 4. segment oranžový

(*Pozn.: Porovnání jednotlivých sršní je zobrazeno v brožuře přiložené k této práci.*)

Kromě vizuálních odlišností jsou však výrazné rozdíly v chování jednotlivých druhů.

Sršně obecné na člověka nikdy nezaútočí bez příčiny. Pokud nejsou rozdrážděné či jim není ničeno hnízdo, vůbec si lidí nevšímají. Jsou mírumilovné, a dokonce je možné se přiblížit do těsné blízkosti jejich hnízda, pokud jim není bráněno ve výletu. v případě, že se zabydlí na nevhodném místě a lidem překáží, lze to řešit i jiným způsobem než zničením jejich hnízda – pokud je hnízdo dostupné, je možné přestěhovat celou kolonii na vzdálenější lokalitu (Diaz et al., 2023).

Včelaři i přesto v některých případech sršně obecné likvidují, neboť v případech, kdy je vlivem chemických postřiků na ochranu rostlin na polích, lesích a zahradách nedostatek hmyzu v krajině, jsou včely medonosné na včelnicích leckdy jedinou potravou sršní obecných. Včely jsou pro sršeň obecnou zdrojem potravy spíše v období, kdy je nedostatek hmyzu – hlavně na konci léta a začátkem podzimu, kdy populace sršní v hnízdě vrcholí a množství hmyzu v přírodě přirozeně ubývá. Včela medonosná tak tvoří jen asi 10–30 % z celkové skladby potravy sršně obecné. Sršně obecné se navíc nikdy nesnaží včelstvo vyhubit, loví včely jednotlivě, chytají jen staré, převážně letní včely, které by stejně nepřežily zimu (Diaz et al., 2023).

Přítomnost sršní obecných v blízkosti úlů má ale i své pozitivní stránky. Dokonce by se dalo hovořit o jisté formě skrytého partnerství. Sršně obecné totiž zbavují včelstvo

nežádoucího hmyzu, například vos útočných (*Vespula germanica*) a zavíječe voskového (*Galleria mellonella*). I tak se ale včely přirozeně brání. Na česně (hlavním vchodu do úlu) se včely brání chytáním sršně za nohy, aby jí zabránily odletět do bezpečí. Několik včel útočí na sršeň naráz, koušou ji, případně ji bodají žihadly, následkem čehož sršeň uhynie. Je důležité zmínit, že včely medonosné a sršně obecné mají tento symbiotický vztah už tisíce let a oba druhy jsou na něj adaptované (Diaz et al., 2023).

Oproti tomu **sršeň asijská** má na biodiverzitu Evropy dopady zejména negativní. Je schopna se velmi dobře adaptovat na nové prostředí. v Evropě se jí velmi daří se rychle šířit a kolonizovat nová území, jsou-li na to vhodné klimatické podmínky. Například ve Francii se během prvních tří let sršeň asijská rozšířila na území asi 120 000 km² (Vita Bee Health, 2023) a nové území nyní kolonizuje rychlostí asi 78 km ročně (Robinet et al., 2017). Na konci sezóny žije v hnízdě 1 500–2 000 jedinců (Diaz, et. al., 2023), nicméně v extrémních případech může být v hnízdě až 13 000 jedinců (Skuhrovec, 2023). Oproti sršni obecné, jejíž kolonie čítá většinou jen okolo 400–700 jedinců (Drbohlavová, 2013), jsou tedy kolonie sršni asijských mnohem početnější, a tedy i nebezpečnější a pro své okolí mnohem větší zátěží. Jedna kolonie *Vespa velutina* zkonsumuje v průměru milion jedinců hmyzu za sezónu. Ve Francii ukázaly studie, že *Vespa velutina* se živí jak hmyzem, tak mrtvolami savců a ptáků. Její kořist tvořil z 59 % blanokřídlý hmyz, z nichž včely zastupovaly více než 35 %, z 32 % dvoukřídlý hmyz a z 9 % jiné druhy (řády *Hemiptera*, *Orthoptera*, *Lepidoptera*, *Heteroptera* a další). Složení kořisti se zásadně lišilo podle prostředí – ve městech sršně asijské loví především včely medonosné, které jsou vlivem rostoucího trendu včelaření v těchto oblastech velmi hojné a tvořily někdy až 70 % kořisti sršni asijských. Naopak v lesních oblastech je podíl včel medonosných v potravě sršni na úrovni okolo 30 % (Diaz, et. al., 2023). Sršeň asijská je při svých útocích mnohem vynalézavější a agresivnější než sršeň obecná, agresivitou se spíše blíží vosám (Proková, 2023).

Roční cyklus hnízd a sociální život sršně asijské

Po skončení své hibernace (trvajících od prosince až do března) královna začne stavět tzv. primární hnízdo. Většinou toto hnízdo postaví rychle a na krytém místě, nebojí se stavět si hnízda v těsné blízkosti lidí, naopak urbánní oblasti vyhledávají – často žijí a staví si hnízda v kůlnách, dětských prolézačkách, zahradních domcích a dalších místech blízko lidí (Proková, 2023). Například v Jižní Koreji žije až 70 % kolonií sršně asijské *Vespa velutina* v městském prostředí (Diaz et al., 2023), v roce 2010 počet odchytnutých jedinců v urbánních oblastech (městské parky, centra měst) převyšoval odchyt všech ostatních druhů sršni, *V. velutina* tvořila 170 z 245 odchytnutých jedinců (Choi et al., 2012). Mezi roky 2007–2014 bylo ve Francii v lokalitě Andernos-les-Bain 30 z 36 (77 %) hnízd, identifikovaných jako primárních, postaveno v urbánních oblastech (Franklin et al., 2017).

Primární hnízda jsou postavená většinou z rozžvýkaného dřeva, často jsou velká jen několik centimetrů a obsahují 10–20 dělničích buněk. Dělnice vychovává v primárním hnízdě královna sama, tento vývoj trvá v mírném pásu asi 30–50 dní (Choi et al., 2012).

Pokud není místo primárního hnízda vhodné pro rozšíření, dělnice začnou stavět tzv. sekundární hnízdo (většinou v korunách stromů), které může mít výšku až 100 cm, šířku až 70 cm a může být vybudováno i 200 m daleko od primárního hnízda. od primárního hnízda se též kromě velikosti liší i vchodem – u primárního hnízda je zespodu, u sekundárního je z boku (Diaz et al., 2023).

Královna v tomto období snižuje výlety z hnízda, čímž se vyhýbá riziku úhynu a omezuje se pouze na kladení vajíček. S rostoucím počtem dělnic se hnízdo rychle zvětšuje (zhruba od května do září). Kolonie v této době obsahuje pouze dělnice, a to do doby, dokud

královna nenaklade na konci srpna vajíčka, z nichž se vyvíjejí samci a samice schopní páření. od října do listopadu jsou v hnízdě zejména pohlavně dospělí jedinci a několik set dělnic. z některých oplodněných samic se poté stávají nové královny a zakladatelky hnízd (Proková, 2023).

Od konce léta královny a samci opouští hnízdo a páří se, načež samci umírají a nové oplozené královny přechází do stadia hibernace. Celý cyklus tak začíná nanovo. Doposud není známa míra úmrtnosti královen přes zimu a není známa ani minimální teplota, kterou jsou schopny královny přežít. Velké množství hnízd se přes zimu vyprázdní, neboť většina jedinců zimu nepřežije a nové královny se do stejných hnízd již nevrací (Proková, 2023).

Potrava sršní asijských a způsob lovu

Potravu v podobě cukerných a bílkovinných látek do hnízda donáší jen dělnice. Kromě včel jsou zdrojem bílkovin ještě další druhy hmyzu, které se stanou kořistí dělnic. Sršně jsou ve skutečnosti všežravé, dospělci se živí zejména nektarem, zralým ovocem a sekrety produkovanými larvami v hnízdě. Masožravé jsou v kolonii sršní larvy – živí se hmyzem, masem z mršin, ale i masem z prodejních stánků.

Pokud sršeň asijská nalezne úl, včelstvo napadne. Útočí na různá včelstva bez ohledu na to, zda jde o zdravé nebo nemocné včelstvo. *Vespa velutina* je v lovu včel velmi obratná, loví je za letu chycením za křídla. Na rozdíl od sršní obecných, které pomalu poletují a čekají na včely medonosné, asijská sršeň se před úlem pohybuje nahoru a dolů a čeká na přilétající včely. Takto útočících sršní může být před úlem i 50 (Proková, 2023). Chycené včele sršeň ihned odstraní hlavu, zadeček, křídla a nohy a využívá pouze svalovinu ze včelí hrudi, kterou do hnízda odnáší ve formě kaše, již krmí larvy. Tyto sršní ataky jsou pro včely medonosné velmi stresující, leckdy včely ani neopouští úly a kolonie zanikají kvůli nedostatku potravy (Diaz et al., 2023).

Obranné mechanismy včel

Asijské včely jsou na sršně druhu *Vespa velutina* přizpůsobeny, neboť se vyvíjely společně s nimi a vyvinuly si účinné obranné strategie. Vydávají speciální skupinové vibrace na česnech úlů, které dle různých pozorovatelů slouží buď k rušení sršní asijských, aby byly dezorientované a nemohly včely chytat, nebo k varování ostatních včel, aby se nevracely do úlu, kde je predátor (Darrouzet, 2019 In: Proková, 2023). Dělnice asijských včel využívají ještě další strategii – sršeň obalí více dělnic, čímž vytvoří tzv. termickou kouli a sršeň zabije horkem (včely dokáží zvýšit teplotu a v centru této koule je asi 47 °C, což je pro sršně fatální). Důležité však je, že asijské včely jsou této obrany schopny. U evropské včely medonosné bohužel neexistují důkazy, že by se byly schopné sršním ubránit (Diaz et al., 2023).

Monitoring a likvidace sršně asijské

Způsobem, jak neprofesionálně chytat sršně asijské, je lákat je do pastí jako jsou láhve s návnadou. Tento postup však není účinným monitorovacím ani likvidačním opatřením, neboť jednotky odchycených sršní nemají na život celé kolonie téměř žádný vliv. Takové pastě zabijí i jiné druhy hmyzu včetně chráněných druhů (nejsou tedy selektivní) a navíc se uhynulá sršeň nedá využít ke sledování a hledání hnízda (Diaz et al., 2023).

Dnes jsou již známy samičí feromony, na něž jsou samci sršně asijské velmi citliví. Tato znalost může pomoci při vývoji selektivních feromonových pastí na odchyt sršních samců, čímž by se redukovala možnost spáření se sršními samicemi. Problémem však může být to, že samci sršní asijských potřebují kromě feromonových signálů i vizuální podněty, aby královnu následovali (Cheng et al., 2022).

Monitoring primárních hnízd je náročný, neboť hnízda jsou velmi malá a snadno přehlédnutelná, sekundární hnízda jsou poměrně velká a lehce vzbudí lidskou pozornost. Mnohdy jsou však zpozorována jen náhodně. Nalezená primární hnízda by měla být likvidována ještě předtím, než dělnice začnou stavět sekundární hnízdo, případně na začátku léta, kdy je kolonie teprve ve vývoji. Běžně užívaná opatření pro redukci počtu sršní jsou: zničení hnízd, odchyt královen a dělnic, ochrana včelstev pomocí elektrických zábran před úly a česnové nástavce, podpora včelstev dokrmováním a případný přesun včelstev na jiné místo a do vyšších nadmořských výšek.

Primární hnízda se dají díky své malé velikosti odchyťovat do nádob. Sekundární hnízdo je dobré zničit například odsáním dělnic a dalších jedinců pomocí speciálního vysavače se sběrnou nádobou. Použít lze i různé chemické metody likvidace (biocidy, insekticidy apod.) nebo odchyt královen a dělnic do selektivních pastí.

Více podrobných informací nejen k tematice odchytu sršní je dostupných v brožuře *Sršeň asijská (Vespa velutina), její biologie, monitoring, kontrola a prevence šíření* (Diaz et al., 2023).

Hlášení pozorování sršně asijské mohou nálezci hlásit písemně na Agenturu ochrany přírody a krajiny České republiky (AOPK ČR).

2.1.2 Evropské výzkumy modelující výskyt sršně asijské

Autor ve své práci vycházel z několika různých výzkumů, které se zabývaly modelováním výskytu sršně asijské *Vespa velutina*. Autor se ve své práci opírá o výzkumy z Mallorcy (Herrera et al., 2022), Španělska (Bessa et al., 2016) a Portugalska (Verdasca et al., 2021).

Výzkum na Mallorce byl podnícen poměrně rychlým vývojem sršní populace (v roce 2015 bylo nalezeno jedno hnízdo, v roce 2017 už jich bylo 20) a byla potřeba modelovat případnou možnost šíření do dalších částí ostrova. Vědci při posuzování vhodnosti prostředí vycházeli ze 42 environmentálních proměnných spadajících do pěti hlavních kategorií: klimatické, topografické, hydrografické a antropogenní proměnné a proměnné založené na metodách DPZ. Ostrov Mallorca byl rozdělen do buněk o rozloze 1 km² a všechny proměnné byly do těchto buněk agregovány (zprůměrovány). Byla posouzena multikolinearita mezi všemi 42 proměnnými a z těchto statistických analýz vzešlo osm nejdůležitějších určujících proměnných ovlivňujících výskyt sršně asijské, a to:

- **krajinný pokryv,**
- **lidská stopa (human footprint, %),**
- **sklon svahu,**
- **úhrn srážek v nejteplejším čtvrtletí (bio 18),**
- **úhrn srážek v nejchladnějším čtvrtletí (bio 19),**
- **srážková sezónnost (bio 15)** – rozptyl v měsíčních úhrnech srážek v průběhu celého roku (O'Donnel et al., 2012),
- **izotermalita (bio 3)** – kvantifikuje, jak velké je kolísání denních a nočních teplot v porovnání s kolísáním letních a zimních (ročních) teplot (O'Donnel et al., 2012),
- **kontinentalita** – souhrn vlastností klimatu podmíněných působením pevniny na procesy tvorby klimatu, vzrůstá směrem od oceánu do nitra pevniny (Česká meteorologická společnost, 2023).

Modelování proběhlo vlastním sestaveným modelem, jenž byl sestaven z několika dílčích modelů, jmenovitě například Maxent, Generalizovaný lineární model (GLM),

klasifikační stromy, neuronové sítě a další. Model byl natrénován na třiceti datech výskytu hnízd se zapojením 25 % ostatních nenapadených buněk na Mallorce (více v podkapitole 4.4.4). S ohledem na použitá data bylo vygenerováno 28 jednoduchých bivariačních modelů kombinujících osm výše zmíněných proměnných. Každý jednoduchý model byl natrénován na 2/3 dat a testován na zbylé třetině dat. Po 100násobném přizpůsobení a vyhodnocení každého jednotlivého bivariačního modelu z 28 různých kombinací prediktorů a za použití 10 různých metod vznikl konečný soubor 28 000 jednotlivých bivariačních modelů. Poté bylo odstraněno celkem 694 jednoduchých bivariačních modelů, jejichž AUC bylo nižší než 0,5. Modely byly hodnoceny metrikami AUC, Somersovo D, Boyce Index a TSS (True Skill Statistics) (vybrané statistické ukazatele výkonu modelů jsou vypsány v kapitole 4.4.7). Výsledný složený model výzkumníkům poskytl poměrně velikou hodnotu AUC (0.9165 ± 0.032), což značí vynikající model s velkou diskriminační (rozhodovací) schopností. Tímto modelem byla Mallorca vyhodnocena jako méně vhodná lokalita pro šíření sršně asijské.

Další v této práci zkoumaný přístup k modelování výskytu sršně asijské je **ze Španělska** (Bessa et al., 2016). Ještě před samotným modelováním výzkumníci poukazují na fakt, že vzhledem k mnohem početnějším koloniím sršní asijských *Vespa velutina* oproti sršni obecné *Vespa crabro* jsou tyto invazivní sršně daleko větším nebezpečím pro včelaře. Větší úhyny včel medonosných mj. ovlivňuje i nízká efektivita evropských včel se bránit sršním asijským, což výrazně zvyšuje predaci (Monceau et al., 2013). I tento výzkum pracoval s lokalitami hnízd, do modelování jich vstupovalo 275 a byly pomocí GPS získány na severu Portugalska.

Tento výzkum pracuje s poměrně velkými buňkami – míra vhodnosti habitatu je určena pro hexagony o ploše 100 km², čímž se dle autora této práce potírají lokální charakteristiky a anomálie, které mohou hrát velkou roli ve výskytu sršní (například úzké údolí, v němž je výskyt sršní možný, ač na vrcholcích okolních hor nikoliv). Tuto skutečnost potvrzuje i Dormann et al. (2012).

Rozšíření sršní určuje mnoho klimatických a ekologických faktorů (Choi et al., 2012). Bylo vybráno celkem 22 proměnných, které mohou vysvětlit rozšíření *V. v. nigrithorax*. Devatenáct z nich se týká environmentálních proměnných: klimatických (10), topografických (3), struktury stanoviště (4), produktivity vegetace (1) a hydrografické sítě (1); zatímco další tři odpovídají antropogenním aktivitám. Klimatické proměnné byly získány z databáze BioClim s rozlišením 30 úhlových vteřin.

Pro zamezení korelace mezi proměnnými byla provedena Spearmannova neparametrická korelační analýza a byly vypuštěny prediktory s velkou mírou korelace ($>0,7$). Jako nejlivnější proměnné byly stanoveny:

- **NDVI (Normalized Difference Vegetation Index)**
- **průměrná roční teplota vzduchu (Bio 1)**
- **roční teplotní rozsah (Bio 7)**
- **izotermalita (Bio 3)**
- **úhrn srážek v nejvlhčím měsíci (Bio 13)**
- **průměrná vzdálenost k vodním tokům**
- **využití krajiny (urbánní plochy, zemědělské plochy, křoviska)**

Nejjednodušší model (AIC =2041,50) je zároveň velmi dobře natrénovaný (vysvětluje 72 % rozptylu). Validace ukázala, že finální model vykazuje velmi vysokou diskriminační výkonnost (AUC =0,92).

Výzkum v Portugalsku též využíval dat o lokalitách hnízd, celkem jich do modelování vstupovalo téměř 8 000. Jako ekologické pozadí druhu bylo zvoleno okolí do 5 km kolem hnízd, který by měl představovat maximální schopnost doletu a založení nových hnízd. Toto sdělení dosti kontrastuje s výzkumem ve Francii, kde byla jako maximální rychlost šíření sršňů za rok zjištěna hodnota 78 km (Robinet et al., 2017).

Proměnných, ovlivňujících výskyt sršňů asijských, bylo stanoveno celkem 11 – tři klimatické a osm antropogenních a krajinných proměnných (Land cover). Proměnné v tomto výzkumu byly převzaty z Bessa et al. (2016) s následujícími změnami:

- 1) byly vyloučeny NDVI a izotermalita,
- 2) byly požity vzdálenosti ke každé konkrétní třídě krajinného pokryvu namísto jejího procentuálního zastoupení, protože použití vzdáleností poskytovalo lepší výsledky,
- 3) byly začleněny proměnné *vzdálenost k lesu* a *index lidského vlivu* a
- 4) byly zahrnuty prediktory týkající se vzdálenosti k liniovým stavbám (dálnicím a železnicím).

Aby byla redukována přílišná kolinearita, byly zkontrolovány vysoce korelující proměnné.

K posouzení, které proměnné ovlivňují distribuci *Vespa velutina*, byly použity zobecněné lineární smíšené modely (GLMM). Byly spuštěny úplné modely s klimatickými, půdními pokryvy a antropogenními faktory současně. Jelikož je uznáváno, že klimatické proměnné jsou hlavními faktory ovlivňujícími rozšíření druhů v různých prostorových měřítkách, byly spuštěny další modely pouze s krajinným pokryvem a antropogenními proměnnými ve snaze najít další možné prediktory v regionálním měřítku.

Z modelů vyšly následující nejvlivnější proměnné:

- **vzdálenost od křovisek a přírodních luk**
- **vzdálenost od dálnic**
- **roční teplotní rozsah (Bio 7)**
- **úhrn srážek v nejvlhčím měsíci (Bio 13)**

Z výsledků modelování tohoto výzkumu vyplývá mimo jiné poměrně výrazný vliv liniových staveb (dálnic a silnic) na šíření sršně asijské. Autoři odhalili velkou korelaci v počtu nalezených hnízd a vzdálenosti od dálnic. Toto zjištění by podle nich mělo podnítit více monitorovacích programů v 6km okolí dálnic.

2.2 Vybrané přístupy k modelování distribuce druhů

Modelování distribuce druhů, případně někdy nazýváno také jako modelování ekologických nik, je klíčovým přístupem pro mnoho problémů v populační dynamice, ochraně přírody, biogeografii a evoluční ekologii. v poslední době bylo dosaženo metodologického pokroku převzetím modelovacích principů ze statistického a strojového učení (Drake, 2014). Ekologická nika charakterizuje postavení druhu v ekosystému a zahrnuje jak požadavky na stanoviště (faktory životního prostředí, přičemž faktory jsou biotické a abiotické), tak funkční úlohu druhu v prostředí, v němž se vyskytuje (Polechová et al., 2019). Gutiérrez-Hernández a García (2021) prezentují modelování ekologické niky jako optimální nástroj pro zkoumání vztahů mezi geografickým výskytem druhů a proměnnými prostředí.

Dormann et al. (2012) rozdělují modely distribuce druhů na dvě skupiny: nejčastěji používané přístupy k popisu rozšíření druhů a biodiverzity jsou známé jako **korelační modely rozšíření druhů**. Tyto metody se zaměřují na popis zákonitostí, nikoli mechanismů, ve vztahu mezi výskytem druhů a environmentálními daty (především

klimatickými údaji). Korelační modely rozšíření druhů statisticky vztahují proměnné prostředí přímo k výskytu nebo početnosti druhů. Přinesly užitečné poznatky pro ochranu biodiverzity a ekologické pochopení velkoplošných vzorců.

Procesně založené distribuční modely (také nazývané jako **mechanistické modely**) formulují ekologii druhu jako matematické funkce a definují kauzalitu; výskyt nebo početnost druhu je nepřímým, z vlastností jednotlivých složek prostředí nelehce odvoditelným důsledkem. Modely založené na procesech často vyžadují velký počet parametrů, přičemž mnohé z nich vyžadují omezeně dostupná data s často vysokým časoprostorovým rozlišením. Proto byly dosud tyto modely použity pro mnohem méně druhů než korelační modely a dále nebudou rozebírány.

Pro efektivní modelování korelační přístup vyžaduje tři typy vstupních dat: data výskytu (nálezová data), environmentální data a specifické nastavení parametrů modelu.

V případě nálezových dat se jedná konkrétně o data prezence (data potvrzeného výskytu druhu) a data absence (potvrzeného nevýskytu druhu, přístup využívající i data absence je nazývá *presence-background* přístup). v případě modelování distribuce druhů ale většinou neexistují skutečná absenční data, v procesu modelování se tedy generují náhodně vybraná pseudo-absenční data (nelze říct, že když se v nějakém prostředí druh nevyskytuje, není pro něj prostředí vhodné; prostředí pro něj vhodné být může, pouze se do něj druh ještě nedostal; Lobo et al. 2008). Tato data však vedou k celkovému snížení přesnosti modelu (De Souza Muñoz et al., 2009). Data environmentálního pozadí slouží ke zjištění podmínek v lokalitách výskytu druhu a natrénování modelu.

Podrobněji se pozadí v procesu modelování věnuje ve své bakalářské práci Vlčková (2016). **Následující výběr a popis modelovacích přístupů vychází z bakalářské práce Vlčkové (2016) a Štěrby (2010).**

2.2.1 Bioklimatické obálky

Nezákladnějším nástrojem pro výzkum distribuce druhů jsou bioklimatické obálky. Pomocí nich lze určit vhodnost klimatických podmínek pro rozšíření určitého druhu nebo celého společenstva organismů. Obálky tvoří vícerozměrný prostor, jenž je určen klimatickými proměnnými. Neberou však v potaz např. typ půdy, topografii či biotické interakce mezi druhy. Mají dvě hlavní využití: zjištění, které oblasti mají příhodné podmínky pro výskyt daného druhu, a odhad změny těchto podmínek do budoucna (Beaumont et al., 2005).

Základním přístupem pro stanovení bioklimatické obálky je určení jejího středu průměrem a jejích limitů jako maxima/minima zkoumané proměnné. Může být zvolen i přístup, jenž stanoví hranici pouze mezi 5 až 95 percentilem hodnot.

Bioklimatické obálky jsou využívány i různými modely, např. Bioclim, Habitat a Domain.

2.2.2 Umělé neuronové sítě (ANN)

Artificial Neural Network (či ANN) je skupina umělých vzájemně propojených neuronů, jejichž snahou je napodobit činnosti lidského mozku ve smyslu získání poznatků pomocí procesu učení. Neuronové sítě využívají matematický nebo výpočetní model pro zpracování informací na základě spojitěho přístupu k výpočtu. Základní jednotkou neuronové sítě je neuron. Každý neuron je napojen na jiný a má přidruženou hodnotu váhy a prahu. Pokud je výstup jakéhokoli samostatného neuronu vyšší než práh, tento neuron se aktivuje a posílá informace další vrstvě neuronů (IBM, 2023). Ve většině případů se jedná o adaptivní systém, který mění svou strukturu na základě informací, které jím proudí.

Používá se k modelování vztahu mezi vícerozměrnou vstupní proměnnou x a vícerozměrnou výstupní proměnnou y .

Neuronové sítě nepotřebují ke svému fungování znát vztahy mezi vstupy a výstupy, základní vzorce chování se naučí v průběhu zpracovávání dat. Jako ve spoustě jiných modelovacích přístupů, množina všech dostupných dat je rozdělena na trénovací a testovací data. Trénovací data ovlivňují rychlost a kvalitu učení, testovací data ověřují výkonnost neuronové sítě po ukončení učícího procesu. Pro vyhodnocení kvality predikčních schopností neuronové sítě slouží křížová validace, pro niž je nutné stanovit validační data.

2.2.3 Bioclim

Bioclim je korelativní modelovací přístup založený na přístupu presence-only (Johnson et al., 2023) schopný interpolovat až 35 environmentálních proměnných. Implementuje algoritmus bioklimatických obálek – pro každou zadanou proměnnou algoritmus nalezne průměr a směrodatnou odchylku. Každá proměnná má také navíc maximální a minimální hodnotu stanovenou ze vstupních dat výskytu druhu. Následně algoritmus porovná hodnoty environmentálních proměnných ve zkoumané lokalitě s percentilovým rozložením hodnot ve známých místech výskytu. Čím blíže je místo výskytu k mediánu, tím je vhodnější.

V tomto modelu lze každý bod výskytu klasifikovat jako: vhodný, pokud se všechny hodnoty přidružené environmentální proměnné nachází v rámci stanovené obálky, okrajový, pokud jedna nebo více env. hodnot nachází mimo obálku, ale stále v rozsahu maximálních a minimálních hodnot, nebo nevhodné, pokud se jedna nebo více proměnných nachází mimo limitní hodnoty. Nabývá hodnot od 0 do 1 (openModeller, 2015).

Hodnota 1 bude velmi vzácná, neboť pro její dosažení by musel mít hledaný bod výskytu pro všechny uvažované proměnné hodnotu mediánu trénovacích dat. Oproti tomu hodnota 0 bude velmi častá, protože je přiřazena ke všem buňkám s hodnotou environmentální proměnné mimo percentilový rozsah pro alespoň jednu proměnnou (openModeller, 2015).

2.2.4 Analýza hlavních komponent

Metoda PCA (Principal Component Analysis) bývá velmi často používána při modelování ekologických nik a rozšíření druhů (Herrera et al., 2022; Janžekovič, 2012; Marchi, 2015). Metoda se používá na transformaci původního vícerozměrného prostoru na nový dvourozměrný souřadný systém tak, aby hlavní osa pokrývala co největší variabilitu dat – jde tedy o zjednodušení původních proměnných do menšího počtu proměnných (hlavních komponent) s co nejmenší ztrátou informace. Převedení do dvourozměrného systému výrazně zjednodušuje interpretaci. Podává informace o proměnných, které nejvíce přispívají do jednotlivých komponent, a tedy mají největší vliv na variabilitu dat.

Ve vícerozměrném prostoru lze ohraničit oblast, kde se modelovaný druh vyskytuje. Do stejného prostoru se přidávají lokality (environmentální pozadí) pseudo-absence, tedy takových, u nichž není známo, zda jsou pro výskyt druhu vhodné. Takový postup využil ve výzkumu šíření sršně asijské např. Herrera et al. (2022).

2.2.5 Maxent

Označení Maxent pochází z anglického *Maximum* Entropy, tj. maximální míra neuspořádanosti. Maxent je algoritmus strojového učení, který využívá prezenční data a odhaduje odezvu druhu na životní prostředí tak, aby byla v celé studované oblasti co nejjednodušší vůči vstupním údajům (Elith et al., 2011). Maxent využívá tzv. *presence-background* přístupu, což znamená, že porovnává hodnoty predikční proměnné v bodech pozadí s hodnotami v bodech výskytu (na rozdíl od technik *presence-absence*, které vyžadují data o nepřítomnosti druhu; Johnson et al., 2023). Zjednodušeně řečeno porovnává podmínky v místech, kde se druh vyskytuje, s místy, kde (zatím) není.

Model Maxent obsahuje několik nastavení, které umožňují snížit nebo zvýšit složitost modelu prostřednictvím dvou klíčových faktorů: funkcí, jimiž se model prokládá (v algoritmu Maxent jsou tyto funkce nazvány jako *Feature Classes*) a regularizačních multiplikátorů (RM), které penalizují komplexnost modelu a shlazují výsledné predikce, aby nebyly tak unputé na trénovací data. Různé funkce (*Feature Classes*) určují tvar dostupných modelovaných vztahů v prostoru prostředí. Standardní funkce jsou *linear (L)*, *quadratic (Q)*, *hinge (H)* a *product (P)*. Více kombinací funkcí (např. LQ, LQHP) vede k potenciálně vyšší složitosti modelu. Podrobnějšímu popisu jednotlivých funkcí se dále věnuje Elith et al. (2011).

Modul Maxent automatizuje dva pracovní postupy, a to sestavení několika modelů s odlišnou mírou složitosti podle určených funkcí a regularizačních multiplikátorů a kvantifikace výkonnosti modelu několika metrikami.

Ukázalo se, že Maxent patří mezi nejvýkonnější techniky modelování nik/distribuce pro širokou škálu prostředí a druhů (Elith et al., 2006; Carnaval a Moritz, 2008). I z tohoto důvodu byl modelovací algoritmus Maxent zvolen jako modelovací přístup v této práci.

Vzhledem k tomu, že práce nemá porovnávat jednotlivé modely a vyhodnocovat jejich přesnost, nýbrž aplikovat modelovací postupy na reálný problém šíření sršně asijské, byl zvolen Maxent (Steven et al., 2006). jakožto jeden z nejpoužívanějších a nejlépe hodnocených modelovacích přístupů pro hodnocení distribuce druhů. Příjemným bonusem je přehledné pracovní prostředí aplikace Wallace, v rámci něhož je možné model Maxent spustit. I tak ale bylo potřeba nastudovat spoustu literatury pro pochopení jednotlivých kroků modelování a aplikování na vlastní studii. Celému modelování a jednotlivým nastavením algoritmu Maxent se podrobně věnuje kapitola 4.4.

Podrobnějšímu dělení a dalším algoritmům modelování se dále věnují Vlčková (2016) a Štěrba (2010).

3 METODY A POSTUP ZPRACOVÁNÍ

Ze zadání práce a požadovaných cílů byl stanoven harmonogram dílčích cílů a činností realizovaných v průběhu řešení práce. Bylo zpracováno několik odborných článků, jejichž cílem bylo modelování šíření sršně asijské v různých státech Evropy (zejména Francie, Španělsko, Portugalsko). Některé výzkumy používaly data získaná z terénních prací, které byly cíleny na hledání hnízd sršní asijských. Lokality nalezených hnízd byly následně použity při modelování. Na základě rešerše výzkumů byly stanoveny hlavní prediktory nejvíce ovlivňující výskyt sršně asijské *Vespa velutina nigrithorax*. Pod pojmem prediktory se zde rozumí zejména bioklimatické charakteristiky, jejichž kompletní seznam je zmíněn dále v textu, a charakteristika krajinného pokryvu, nadmořská výška, sklon svahu apod. Tyto prediktory byly následně použity při sestavování a natrénování modelu Maxent, který na základě nich dokáže stanovit míru vhodnosti habitatu pro výskyt sršně asijské. Model byl natrénován na datech současného výskytu sršně asijské na území Evropy a následně aplikován na území České republiky, a to pro současný klimatický normál 1991–2020 a následně do budoucna pro normál 2041–2060. Vznikly tak vizualizace pro dvě různá období, které byly spolu s informacemi o sršni zpracovány do formy map, brožury a letáku pro včelaře, případně širokou veřejnost.

Použité metody

V teoretické části diplomové práce je popisována **sršeň asijská** (konkrétně poddruh *Vespa velutina nigrithorax*, která je momentálně jediná zdokumentovaná invazivní asijská sršeň v Evropě), její charakteristika, odlišnost od sršně obecné (*Vespa crabro*), dopady na společnost a hospodaření. Dále jsou popisovány charakteristiky habitatu, v němž se sršeň v Evropě vyskytuje, a které byly popsány v několika výzkumech. Kapitola se též věnuje popisu několika možných řešení pro modelování distribuce druhů, přičemž největší důraz je kladen na model Maxent (obsažený v aplikaci Wallace), jenž byl vybrán jako finální nástroj pro modelování vhodnosti habitatu pro sršeň asijskou.

V rámci praktické části byly využity poznatky z rešerše – byly analyzovány dostupné datové sady a vybrány zdroje dat vhodné pro potřeby této diplomové práce. Na podkladu těchto dat byly vygenerovány rastry bioklimatických a dalších stanovištních charakteristik pro klimatický normál 1991–2020, které byly použity pro modelování výskytu sršně asijské v Evropě a následně aplikovány na prostředí České republiky, a pro klimatický normál 2041–2060, jenž byl použit pro predikci výskytu v ČR. Vzhledem k poměrně velkému množství potenciálně použitelných proměnných (19 proměnných), jejichž výběr byl stanoven na základě rešerše, byly použity různé statistické metody pro vytipování nejdůležitějších proměnných, které poskytují komplexní, ale přesto nepříliš složitý obraz o prostředí, v němž se šíří sršeň asijská. Výsledné proměnné byly vybrány statistickými metodami PCA a podobnosti jednotlivých bioklimatických proměnných byly spočítány pomocí korelační matice a následně redukovány, aby model neobsahoval příliš korelujících proměnných, které by mohly výsledný model ovlivnit (více ke statistickému zpracování v podkapitole 4.3.3).

Výzkum prováděný v této práci se od ostatních výzkumů popsaných v rešerši liší rozsahem zkoumaného území, počtem zpracovaných nálezových dat (bodů výskytu) a prostorovou podrobností rastrů bioklimatických a dalších stanovištních proměnných. Například ve Španělsku stanovovali výzkumníci míru vhodnosti habitatu v hexagonové síti o velikosti 100 km² na buňku (Bessa et al., 2016), na Mallorce byly pro modelování použity čtvercové pixely o velikosti 1 km² (Herrera et al., 2022). Autor v této práci používá rastry s velikostí pixelu 0,0625 km² (pixel 250×250 m, celkem je v jednom rastru přes 36 milionů pixelů), což by mělo vést k větší přesnosti výsledného modelu

a podrobnosti predikce. Vytvoření rastrů probíhá z bodových vrstev pomocí interpolace **IDW** (Inverse Distance Weighing). Jde o interpolační techniku, která vychází z předpokladu, že věci, které jsou si blízké, jsou si podobnější než ty, které jsou od sebe vzdálenější. K předpovědi hodnoty pro libovolné neměřené místo používá IDW naměřené hodnoty v okolí místa předpovědi. Naměřené hodnoty, které jsou nejbližší předpovídanému místu, mají na předpovídanou hodnotu větší vliv než hodnoty vzdálenější. IDW předpokládá, že každý měřený bod má lokální vliv, který se vzdáleností klesá (Esri, 2023a). Tato metoda interpolace byla použita jako nejrychlejší pro generování rastrů – vzhledem k tomu, že bodová data byla v síti 250×250 m a bylo potřeba vyexportovat rastr o stejném rozlišení, metoda IDW zjednodušeně řečeno sloužila spíše k převodu bodové vrstvy do rastrové.

Takto podrobné rastry byly vyexportovány pro poměrně velké zájmové území (viz sekce *Použitá data* níže), což z celého souboru rastrů dělá poměrně komplexní databázi proměnných, kterou je možné použít pro další výzkumy šíření živočišných druhů a sestavování modelů s velkou prostorovou přesností.

Vzniklo několik různých modelů, jejichž použitelnost byla vyhodnocena na základě vypočítaných statistik, a na těchto podkladech byl vybrán neoptimálnější model. S jeho použitím byla vygenerována **míra vhodnosti habitatu** (ta je klíčová pro předpovídání šíření sršně asijské na další území) na území Evropy a tento model byl poté aplikován na území České republiky, a to pro normál 1991–2020 (tedy pro zjištění, zda již v současnosti jsou v České republice podmínky vhodné pro šíření sršně asijské) a dále normál 2041–2060 (tedy predikce do budoucnosti, kdy je očekávána změna teploty vzduchu, množství srážek a dalších klimatických proměnných, která pravděpodobně povede ke zvyšování vhodnosti habitatu pro tento invazivní druh hmyzu). **z povahy modelovacího přístupu modelu Maxent, jenž byl použit pro predikci v této práci, je výstupem predikční model založený na určování míry vhodnosti habitatu pro sršně asijskou *Vespa velutina*.**

Další část práce, tedy příprava mapových layoutů a kompletace brožury pro veřejnost, probíhala v programech desktop publishing (**DTP**). Zde byly z GIS vytvořené a exportované mapové výstupy upravovány, sepsán informační text a doplněna infografika. Brožura včetně map je hlavním vizualizačním výstupem celé práce.

Použitá data

Diplomová práce byla realizována za využití několika zdrojů dat. Žádná data nebyla ručně měřena, ani digitalizována, veškerá data byla stažena z volně dostupných zdrojů. Vzhledem k tomu, že model byl trénován na datech ze západní Evropy, byla pro dílčí kroky zvolena veřejně dostupná data pokrývající v co největším rozlišení celé zájmové území.

Zájmové území pro následné analýzy bylo stanoveno polygonovou vrstvou všech států Evropy, v nichž byla získána data o výskytu sršně. Jedná se o **Portugalsko, Španělsko, Itálii, Švýcarsko, Velkou Británii, Nizozemsko, Belgie, Lucembursko a Německo**, k těmto státům byly přidány **Česká republika a Rakousko** (Rakousko pro případnou vizualizaci tranzitních koridorů pro šíření sršně asijské). Vrstva zájmového území vznikla kombinací dvou vrstev – polygonové vrstvy států, která byla ve vysokém rozlišení stažena z **Open Street Map** pomocí doplňku QuickOSM v programu QGIS a zahrnovala i území výsoštných vod přímořských států, a polygonové vrstvy **EEA coastline for analysis** (Evropa ořezaná pobřežní čarou), tato vrstva byla vytvořena v roce 2017 a pochází od Evropské agentury pro životní prostředí (European Environment Agency, 2018).

Stěžejním rastrovým datasetem, na jehož podkladu byla následně vygenerována veškerá klimatická data, je **European Digital Elevation Model (EU-DEM) ve verzi 1.1**. Jedná se o aktualizovanou verzi EU-DEM 1.0, tuto aktualizaci koordinovala Evropská

agentura pro životní prostředí (EEA) v rámci programu EU Copernicus. Data jsou poskytována po dlaždicích o velikosti 1 000×1 000 km ve formátu GeoTIFF 32-bit a v rozlišení 25×25 m. EU-DEM ve verzi 1.1 byl vydán v dubnu roku 2016 (Copernicus, 2016). Přístup k datům je založen na zásadě úplného, otevřeného a bezplatného přístupu, jak je stanoveno v nařízení o datové a informační politice programu Copernicus (EU) č. 1159/2013 ze dne 12. července 2013 (European Environment Agency, 2017).

Zdroj klimatických dat byl zvolen s ohledem na fakt, že ve výsledku bylo potřeba vygenerovat rastry o rozlišení 250×250 m, jehož většina volně dostupných dat nedosahuje.

Pro historické hodnoty klimatických charakteristik byl použit program **ClimateEU**, jenž pracuje s datasetem **Climate Research Unit Time Series (CRU-TS) ve verzi 4.05**. Jedná se o data vytvořená oddělením Climate Research Unit (CRU) pro výzkum klimatu na University of East Anglia. Data jsou poskytována ve vysokém rozlišení, a to pro období od ledna 1901 do prosince 2020. Na těchto datech proběhlo natrénování modelu. Data CRU-TS ve verzi 4.05 byly vydány v roce 2021 a jsou poskytovány pod otevřenou licencí Open Government Licence (University of East Anglia Climatic Research Unit, 2021).

Pro klimatické predikce byl použit dataset **European CLimate Index Projections (ECLIPS) ve verzi 2.0**, který obsahuje rastrová data pro 80 ročních, sezónních a měsíčních klimatických proměnných pro dva předchozí normály (1961–1990, 1991–2010) a pět budoucích normálů (2011–2020, 2021–2140, 2041–2060, 2061–2080, 2081–2100). Data jsou poskytována pod licencí Creative Commons Attribution 4.0 International. ECLIPS ve verzi 2.0 byl vydán v červenci roku 2020 (Chakraborty et. al., 2020).

Tato data jsou sice v rozlišení 30×30 úhlových vteřin (tedy asi 1×1 km), nicméně po předchozí konzultaci s vedoucím práce byly vyhodnoceny jako přesnější (použitými algoritmy výpočtu predikovaných hodnot, nikoli rozlišením) v porovnání s daty poskytovanými programem ClimateEU. Predikovaná klimatická data z datasetu ECLIPS 2.0 byla použita pouze pro území České republiky, kde byla provedena dodatečná interpolace hodnot na požadované rozlišení 250×250 metrů (metodou co-kriging zahrnující i digitální model reliéfu). Data sloužila pouze pro aplikaci modelu na nové časové období (normál 2041–2060), nikoli jeho natrénování.

Kromě klimatických charakteristik, které výrazně ovlivňují rozšíření živočišných druhů, byla stažena podrobná data o krajinném pokryvu a využití krajiny **Land Cover Map of Europe 2017** (Malinowski, R., et. al., 2020) od Evropské kosmické agentury (European Space Agency, ESA). Dataset je výsledkem plně automatizované klasifikace, která byla provedena nad obrazovými daty mise Copernicus Sentinel-2 se zapojením nových cloud-computing nástrojů. Rozlišení těchto dat je 10 m na pixel pro celou Evropu a krajinný pokryv je rozdělen do 13 kategorií (Malinowski et al., 2020).

Posledním, ale zato stěžejním zdrojem informací jsou data o výskytu sršně asijské. Zatímco výzkumy, které byly rozebírány v rešerši, pracovaly s bodovými lokalitami výskytu hnízd, autor této práce k těmto datům o výskytu hnízd neměl přístup, a tak byly zvoleny zdroje dat o výskytu živočichů, které jsou postaveny na volně dostupných datech využívajících citizen science přístupu (tedy přístupu, kdy s výzkumy – např. sběrem dat – pomáhají dobrovolníci z řad nadšenců o danou problematiku). v tomto případě byly použity nálezkové databáze **GBIF.org** (Global Biodiversity Information Facility, 2023) a **iNaturalist.org** (iNaturalist, 2023a), do nichž mohou registrovaní uživatelé vkládat nálezková data různých druhů, doplněná o fotografie a další informace ohledně druhu a nálezu. Nejedná se tedy vždy o data hnízd, nýbrž i o pouhé záznamy pozorovaných jedinců. Přesnost těchto dat je ovlivněna přesností zadání dat do systému nálezem, nicméně obě výše zmíněné organizace mají supervisory, kteří na kvalitu zadávaných dat dohlíží a záznamy kontrolují.

Použité programy

Práce na řešeném problému se z programového vybavení dělila na práci s GIS a jiným softwarem na zpracování dat a práci s programy DTP. Jako hlavní program pro práci s prostorovými daty byl zvolen **ArcGIS Pro** ve **verzi 3.1** od americké společnosti Esri, v něm probíhaly veškeré práce s rastry (jejich generování, rastrová algebra i počítání buňkových statistik) i s vektorovými daty. Byl použit jak na zpracování analytickými nástroji, tak na vytváření datových náhledů a prostě prohlížení dat.

Pro generování klimatických dat byl využit program **ClimateEU** ve **verzi 4.63**. Tento software pracuje se soubory CSV (Comma Separated Values) a existuje ve dvou variantách, z nichž každá pracuje s jiným oddělovačem položek a desetinných míst v CSV souborech – jedna varianta pojmenovaná **ClimateEU v4.63** předpokládá, že desetinná místa odděluje desetinná tečka a oddělovačem položek je čárka (anglosaský svět), druhá varianta s názvem **ClimateEU v4.63 (experimental EU number format)**, která byla použita v této práci, pracuje s desetinnou tečkou a oddělovačem položek je středník (např. čeština). U druhé varianty, ač nazvané jako „experimental“, autor práce nezaznamenal jediný problém se zápisem číselných hodnot, který by mohla tato experimentální verze přinést.

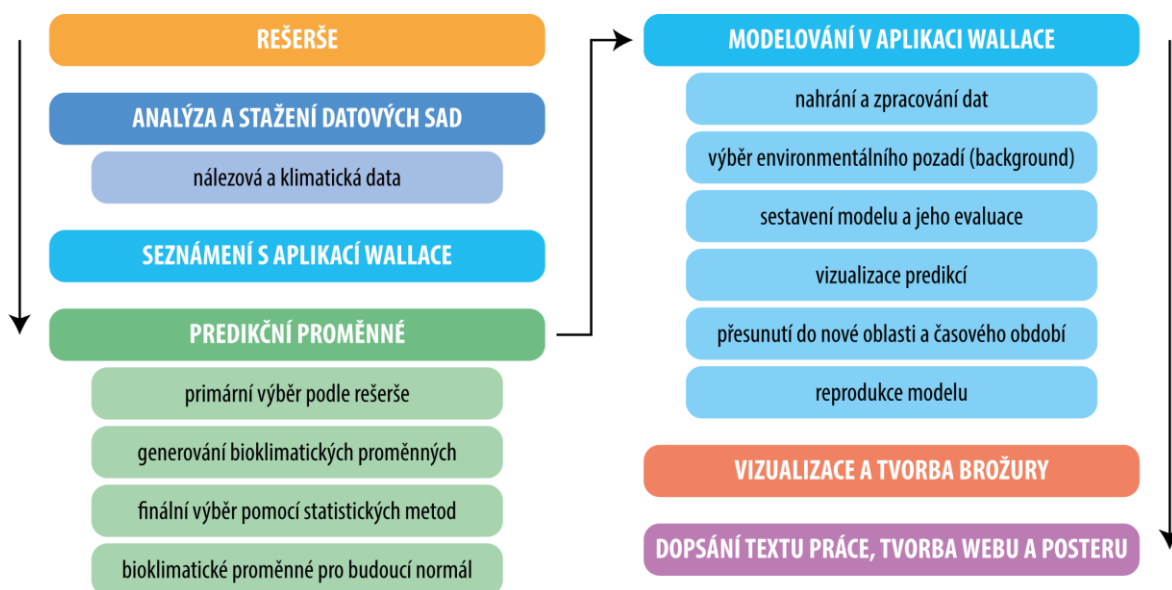
Pro úpravu dat (např. generování velkých CSV souborů, které byly nahrávány do programu ClimateEU), specifické rastrové operace a výpočty (např. počítání standardních odchylek za účelem počítání bioklimatických charakteristik) a pro spuštění modelu byl použit program **RStudio** s předinstalovaným jazykem **R** původně ve verzi 4.1.2, později v aktualizované verzi 4.2.3.

Na statistické analýzy a vizualizace (např. PCA, korelace) byl využit program **Orange** ve **verzi 3.34.0**.

Nahlížení na tabelární data a upravování malých souborů formátu CSV probíhalo v programu **Microsoft Excel** z balíčku Microsoft Office 365.

Vzhledem k omezeným možnostem grafických úprav a nástrojů v programu ArcGIS Pro byl pro grafické dokončení map použit program **Adobe Illustrator** z programové sady Adobe Creative Cloud, kterou může autor práce používat v rámci své studentské licence. v Adobe Illustrator byly nejprve vytvořeny mapy spolu s legendami, které byly následně vloženy do programu **Adobe InDesign** (také ze sady Adobe Creative Cloud), kde byl následně dořešen grafický styl stran, vloženy texty, infografika, podtituly a další kompoziční prvky, finálně zde byla realizována předtisková příprava.

Postup zpracování



Obr. 2 Postup zpracování všech kroků diplomové práce.

Postup práce je patrný z diagramu (Obr. 2).

4 VLASTNÍ ŘEŠENÍ

V této kapitole je podrobně popsána příprava dat, tvorba modelu a následné vytvoření informační brožury, která je stěžejním vizualizačním výstupem celé diplomové práce.

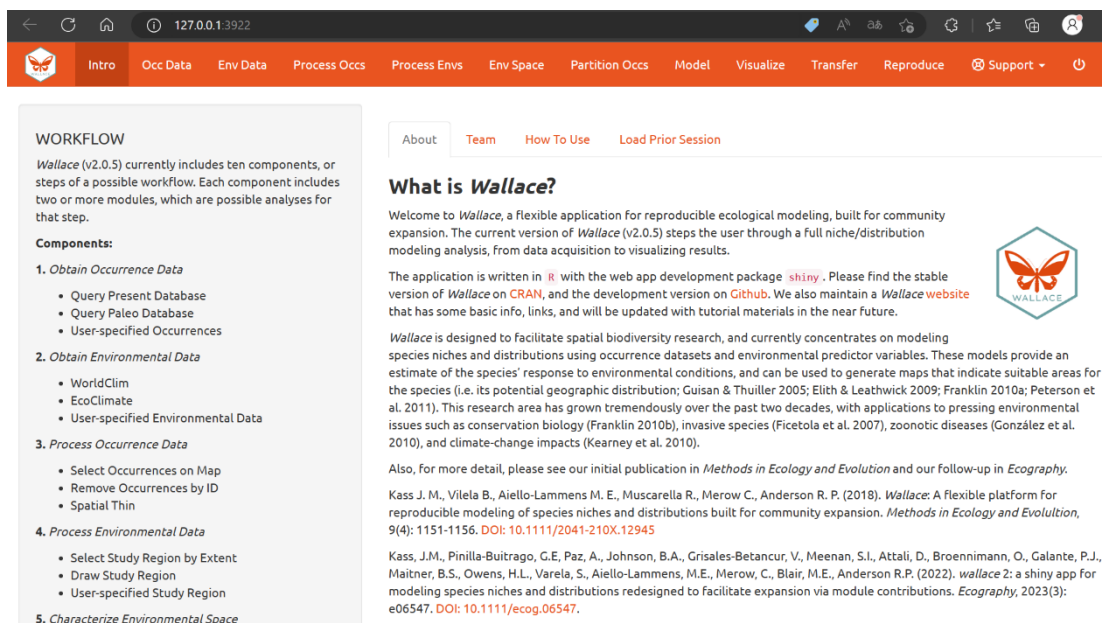
Po získání teoretických poznatků z rešerše (*kapitola 2*) a po analýze dostupných datových sad byly vygenerovány rastry bioklimatických a ostatních charakteristik, stažena a zpracována data o výskytu sršňů. Na podkladě těchto vrstev byl vytvořen model algoritmem Maxent v aplikaci Wallace a aplikován na území České republiky pro současný a budoucí normál. Výstupy v podobě klimatických map a mapy se znázorněnou vhodností habitatu pro sršně asijskou byly spolu s informačním textem a infografikou zpracovány do formy informační brožury pro včelaře a širokou veřejnost se zájmem o zkoumanou problematiku.

4.1 Seznámení s aplikací Wallace

Vzhledem k tomu, že modelování probíhalo v aplikaci Wallace, bylo nejprve provedeno seznámení se s prostředím této aplikace za účelem zjištění veškerých dat a komponent, které jsou nezbytnými vstupy do modelovacího procesu. Podrobný popis práce s Wallace je popsán v podkapitole 4.4.

Aplikace se spouští přes program RStudio (případně pouze přes konzoli jazyka R) a otevře se ve výchozím prohlížeči jako internetová stránka hostovaná na localhost (Obr. 3). Dílčí kroky jsou v aplikaci Wallace pojmenovány jako komponenty a v aplikaci jsou přehledně zařazeny do záložek nacházejících se na liště při horním okraji okna. Každý komponent obsahuje několik tzv. modulů, tedy konkrétních nastavení, která je možné v rámci komponentu spustit (více v podkapitole 4.4). Modelovací postup zahrnuje následující komponenty:

- 1) **Obtain Occurrence Data (Získání dat výskytu druhu)** – lze stáhnout, či nahrát vlastní data ve formátu CSV či TXT,
- 2) **Obtain Environmental Data (Získání environmentálních dat)** – lze stáhnout, či nahrát vlastní rastry ve formátu TIF či ASC, uživatelsky nahraná data musí mít stejné rozlišení, stejný rozsah (extent) a stejné umístění Null hodnot,
- 3) Process Occurrence Data (Zpracování dat výskytu)
- 4) Process Environmental Data (Zpracování environmentálních dat)
- 5) **Characterize Environmental Space (Charakteristika environmentálního prostoru)** – tento *nepovinný* komponent slouží k mezidruhovému srovnání environmentálních podmínek pomocí PCA (Principle Component Analysis, Analýza hlavních komponent; *pozn. komponenta v PCA není programový komponent programu Wallace, nýbrž název pro nekorelovanou proměnnou, která je výsledkem analýzy hlavních komponent*),
- 6) Partition Occurrence Data (Segmentace/Rozdělení dat výskytu) – rozdělení souboru dat výskytu na trénovací a testovací data,
- 7) Build and Evaluate Niche Model (Vytvoření a vyhodnocení modelu niky) – samotné modelování pomocí modelu Maxent,
- 8) **Visualize Model Results (Vizualizace výsledků modelu)** – vytvoření datového náhledu na data vhodnosti habitatu pro zkoumaný druh,
- 9) **Model Transfer (Přesunutí modelu)** – přesunutí natrénovaného modelu do nové oblasti nebo časového úseku (např. predikce do budoucna)
- 10) **Reproduce (Reprodukce modelu)** – stažení kódu pro reprodukci modelovacího postupu a stažení metadat



Obr. 3 Úvodní obrazovka aplikace Wallace, v horní oranžové liště jsou vidět jednotlivé komponenty modelovacího postupu. Vlevo je šedý sloupec WORKFLOW, který obsahuje výčet komponentů (např. 2. Obtain Environmental Data) a v rámci nich jednotlivé moduly (např. WorldClim, EcoClimate).

Po seznámení se s aplikačním prostředím a nastudování jednotlivých komponent modelování byla práce na tvorbě modelu rozčleněna na několik dílčích částí: 1) stažení a úprava náleзовých dat, 2) generování rastrů bioklimatických charakteristik, 3) vymezení zájmového území a 4) zpracování všech dat v aplikaci Wallace. Tyto části budou v následujících několika kapitolách podrobně rozebrány i s postupem a popisem jednotlivých kroků.

4.2 Stažení náleзовých dat

Z vybraných volně dostupných náleзовých databází GBIF.org a iNaturalist.org byla stažena bodová data o výskytu sršně asijské (nejedná se pouze o záznamy hnízd, nýbrž i nálezy jedinců), přičemž většina záznamů pocházela z Evropy. Zadání diplomové práce hovoří o zapojení dat z Asie, nicméně data z Asie nakonec použita nebyla – důvody jsou poměrně malé množství stažených náleзовých dat (z výše uvedených dostupných zdrojů) o výskytu sršně asijské v Asii (zapříčiněno pravděpodobně faktem, že v Asii se jedná o běžný druh, zatímco v prostředí Evropy jde o druh invazivní, a tedy i vysoce monitorovaný pro jeho výrazné dopady na evropskou biodiverzitu) a důraz na natrénování modelu na datech z Evropy, poněvadž evropské klimatické a krajinné charakteristiky mají České republice mnohem blíže než ty asijské.

Data z GBIF.org byla stažena pomocí komponentu **Obtain Occurrence Data** v aplikaci Wallace. v tomto komponentu byl vybrán modul **Query Database (Present)**, v něm nastavena databáze GBIF a staženo 8 000 bodů výskytu. Následně přes záložku **Save** byla data stažena ve formátu CSV. Data z iNaturalist.org (asi 13 000 bodů) byla stažena přímo z webových stránek této databáze také ve formátu CSV.

Data z obou náleзовých databází byla upravena v programu Microsoft Excel tak, aby obsahovala pouze sloupce scientific_name (latinské jméno druhu, tedy *Vespa velutina*), latitude (zeměpisná šířka) a longitude (zeměpisná délka), a jako tabulka nahrána do programu ArcGIS Pro. Pomocí nástroje **Display XY Data** byly oba CSV soubory

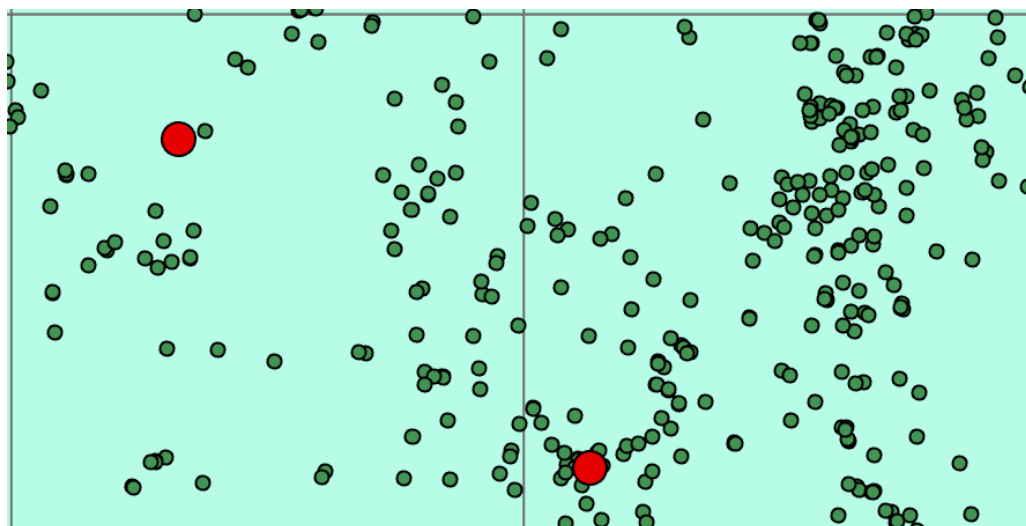
převedeny do podoby bodové vrstvy. Vzhledem k faktu, že obě nálezové databáze jsou propojené a navzájem si sdílí některá nálezová data (iNaturalist, 2023b), byla provedena kontrola duplicitních bodových záznamů (stejných mezi oběma soubory), ty byly pročištěny a následně nástrojem **Append** spojeny do jedné bodové vrstvy.

Redukce hustoty nálezových bodů

Tato bodová vrstva výskytů obsahovala více než 17 000 bodů výskytu, jejichž hustota byla zredukována, a to z několika (na vlastní zkušenosti autora) založených důvodů:

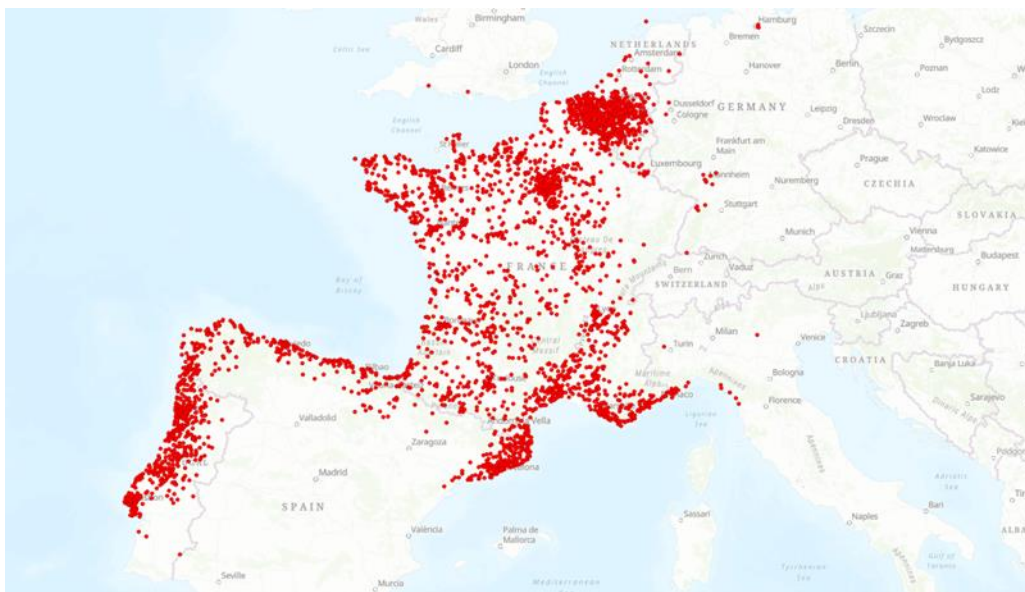
- 1) uvedené množství bodů je **příliš velké** a **výpočetní výkon počítače nestačil** na jejich efektivní zpracování při modelování,
- 2) vzhledem k velké hustotě bodů například v Belgii je pravděpodobné, že někteří jedinci, případně členové jedné kolonie mohli být zaznamenáni vícekrát, **data tedy nemusí mít vypovídací hodnotu při odlišování jednotlivých kolonií**,
- 3) v návaznosti na předchozí důvod, velké množství bodů výskytu na jednom místě (např. Belgie) může způsobit **výraznou autokorelaci** (vzájemnou závislost) hodnot při modelování, čímž bude zvyšována váha hodnot proměnných v místech velkého výskytu bodů, což výrazně ovlivní výstupy výsledného modelu.

Pro zředění hustoty bodů výskytu byl použit nástroj **Generate Tessellation**. Tím byla vygenerována síť čtvercových polygonů („pixelů“) pokrývající bodovou vrstvu výskytů, přičemž jednomu polygonu byla nastavena plocha 25 km² (Obr. 4).



Obr. 4 Ukázka redukce hustoty bodů výskytu sršně asijské – zeleně jsou původní nezředěné body, červeně jsou zředěné body vyexportované tak, aby v každém polygonu mřížky o velikosti 25 km² byl jeden záznam.

Následně byla pomocí nástroje **Add Spatial Join** bodová vrstva napojena do vygenerované sítě tak, aby v každém polygonu byl napojen pouze jeden bod výskytu a do atributů polygonu se přidalo ID tohoto připojeného bodu. Mřížka s napojenými ID bodů byla vyexportována jako nová vrstva. Poté došlo nástrojem **Add Join** ke zpětnému napojení této mřížky do původní bodové vrstvy (přes atribut ID bodu) a bodové záznamy, jejichž ID se shodovaly s ID z mřížky, byly vyexportovány jako samostatná vrstva. Tímto postupem došlo ke zředění bodů z cca 17 000 na 3 119 záznamů. Datový náhled výsledné bodové vrstvy, která byla nahrána do aplikace Wallace, je na Obr. 5.



Obr. 5 Náhled na výsledná bodová data výskytu sršně asijské (mapový podklad World Topographic Map od Esri)

4.3 Predikční proměnné

Stěžejními vstupy do aplikace Wallace, resp. modelu Maxent, bez nichž by nebylo možné modelovat vhodnost habitatu, jsou bioklimatické proměnné a charakteristiky využití území, případně nadmořská výška a její deriváty (sklon). Při procesu modelování výskytu druhů a modelování ekologických nik se používá 19 bioklimatických proměnných. Jsou odvozeny od teploty vzduchu a množství srážek a popisují klimatické podmínky, které nejvíce souvisí s fyziologií druhů (O'Donnell et al., 2012). Bioklimatické proměnné jsou označeny Bio1–Bio 19 a jedná se o tyto charakteristiky (**tučně označené proměnné** byly použity při modelování v této práci, bližší informace k výběru podkapitola 4.3.1 a 4.3.3):

- Bio 1 – Annual Mean Temperature (Průměrná roční teplota vzduchu)
- **Bio 2 – Annual Mean Diurnal Range (Průměrný roční rozsah denních teplot)**
- **Bio 3 = Isothermality (Izotermalita)**
- **Bio 4 = Temp. Seasonality (Teplotní sezónnost)**
- Bio 5 = Max Temp. of Warmest Month (Max. teplota nejteplejšího měsíce)
- Bio 6 = Min Temp. of Coldest Month (Min. teplota nejchladnějšího měsíce)
- Bio 7 = Annual Temp. Range (Roční teplotní rozsah)
- Bio 8 = Mean Temp. of Wettest Quarter (Prům. teplota nejvlhčího čtvrtletí)
- Bio 9 = Mean Temp. of Driest Quarter (Prům. teplota nejsuššího čtvrtletí)
- **Bio 10 = Mean Temp. of Warmest Quarter (Prům. teplota nejtepl. čtvrtletí)**
- **Bio 11 = Mean Temp. of Coldest Quarter (Prům. teplota nejchlad. čtvrtletí)**
- Bio 12 = Annual Precipitation (Roční úhrn srážek)
- Bio 13 = Prec. of Wettest Month (Úhrn srážek v nejvlhčím měsíci)
- Bio 14 = Prec. of Driest Month (Úhrn srážek v nejsušším měsíci)
- **Bio 15 = Prec. Seasonality (Srážková sezónnost)**
- Bio 16 = Prec. of Wettest Quarter (Úhrn srážek v nejvlhčím čtvrtletí)
- Bio 17 = Prec. of Driest Quarter (Úhrn srážek v nejsušším čtvrtletí)
- **Bio 18 = Prec. of Warmest Quarter (Úhrn srážek v nejteplejším čtvrtletí)**
- **Bio 19 = Prec. of Coldest Quarter (Úhrn srážek v nejchladnějším čtvrtletí)**

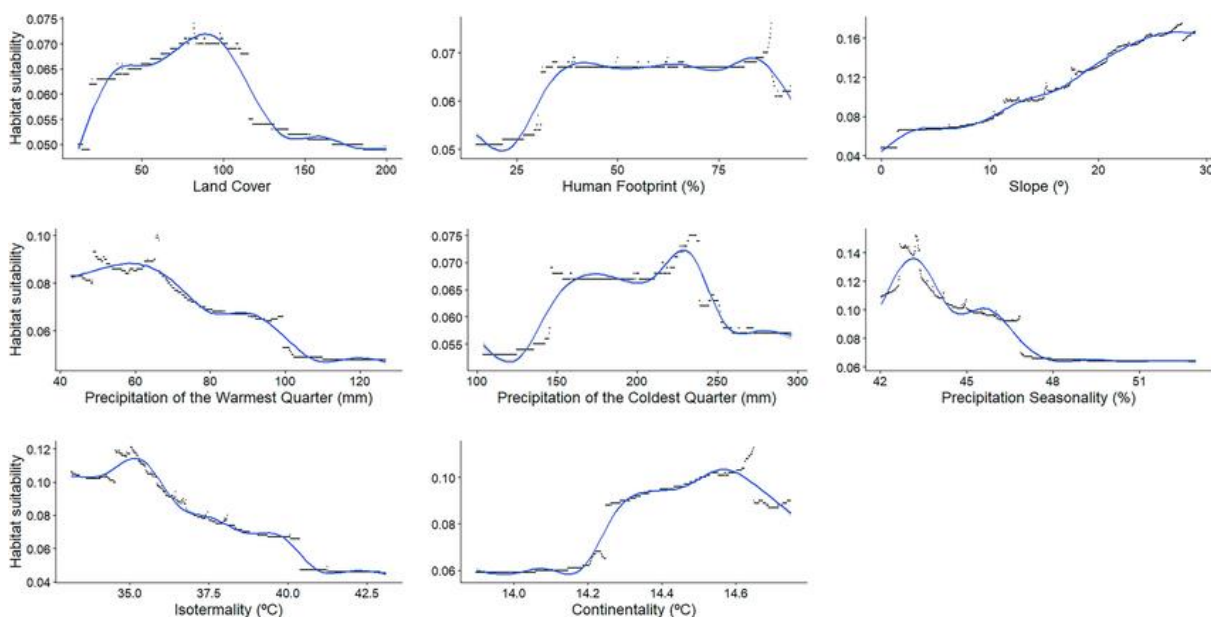
Podrobné informace a postup výpočtu jednotlivých proměnných, z něhož vycházel i autor této práce, jsou obsaženy v metodické příručce ***Bioclimatic Predictors for Supporting Ecological Applications in the Conterminous United States*** (O'Donnell et al., 2012) publikované americkou výzkumnou vládní agenturou USGS (United States Geological Survey), která tyto proměnné vyvinula.

K bioklimatickým proměnným byly přidány ještě nadmořská výška v podobě **digitálního modelu reliéfu**, z něj odvozený **sklon svahu a kategorická proměnná krajinného pokryvu** (se 13 kategoriemi pokryvu). Zdroje dat, z nichž byly generovány jmenované proměnné, jsou popsány v kapitole 3 v sekci *Použitá data*.

4.3.1 Primární výběr bioklimatických proměnných

Primární výběr rastrových proměnných proběhl v rámci rešerše výzkumů v Evropě, v nichž již několik vědeckých týmů (Bessa et al., 2016; Herrera et al., 2022; Verdasca et al., 2021) popsalo proměnné, které mají největší vliv na šíření sršně asijské. v této práci proto nebyla provedena nová kompletní testování použitelných proměnných, nýbrž z již výzkumy stanovených nejlivnějších proměnných byly několika statistickými metodami vybrány ty nejvíce určující pro použitý vzorek dat výskytu v této práci.

Proměnné, které vyplývají z výše zmíněných prací jako nejvíce vysvětlující, jsou vypsány v rešerši v podkapitole 2.1.2.



Obr. 6 Nejlivnější proměnné na výskyt sršně asijské v případě výzkumu na Mallorce podle Herrera et al. (2022). Grafy znázorňují hodnoty sledované proměnné a zjištěnou míru vhodnosti habitatu. (zdroj: <https://www.researchgate.net/>)

Po nastudování nejlivnějších proměnných bylo úkolem vygenerovat rastrová data těchto bioklimatických a stanovištních charakteristik a podrobit je statistickému testování v rámci celoevropského rozsahu nálezových dat. Predikční proměnné vyhodnocené na malém vzorku nálezových dat (např. pouze 30 hnízd zahrnutých do výpočtu na Mallorce) nemusí být totiž obecně platným ukazatelem pro výskyt v celoevropském měřítku.

4.3.2 Generování rastrů bioklimatických proměnných

V následující kapitole bude podrobněji rozebrána tvorba rastrových bioklimatických a stanovištních proměnných, neboť až s jejich použitím bylo přistoupeno ke statistickému vyhodnocení a sekundárnímu (výslednému) výběru nejlivnějších proměnných na výskyt sršně asijské (sekundární výběr proměnných je popsán v kapitole 4.3.3). Podrobnosti o použitých datech a vrstvách jsou popsány v kapitole 3, v sekci *Použitá data*.

Stanovení zájmového území

Zájmové území muselo pro potřeby modelování pokrývat všechny (zředené, viz podkapitola 4.2) body výskytu sršně asijské, které byly do modelování zahrnuty. Pro tyto účely byly staženy polygony následujících států: **Portugalsko, Španělsko, Itálie, Švýcarsko, Velká Británie, Nizozemsko, Lucembursko, Belgie a Německo**, k těmto státům byly přidány **Česká republika a Rakousko** (Rakousko pro případnou vizualizaci tranzitních koridorů pro šíření sršně asijské). Pomocí doplňku QuickOSM v programu QGIS byly z Open Street Map staženy veškeré administrativní hranice těchto států, ty však obsahovaly i území výsoštných vod. Pro získání pouze pevninských bodů byly všechny státy oříznuty polygonovou vrstvou **EEA coastline for analysis** (Evropa ořezaná pobřežní čarou). k tomu posloužil nástroj **Clip v** ArcGIS Pro. Následně pomocí nástroje **Dissolve** byly všechny státy spojeny do jednoho velkého polygonu – ten vymezoval požadované zájmové území.

Digitální model reliéfu a jeho použití pro přípravu dat

Do programu ArcGIS Pro byly nahrány jednotlivé dlaždice digitálního modelu reliéfu (DMR) z celoevropského DMR EU-DEM a pomocí nástroje **Mosaic To New Raster** byly spojeny do jedné bezešvé vrstvy s rozlišením 250 m na pixel. SRS datasetu EU-DEM je ETRS89 LAEA (Lambert Azimutal Equal Area; EPSG: 4258) a jelikož klimatické vrstvy vznikly na podkladu EU-DEM, byl tento SRS použit jako hlavní SRS celoevropských dat.

Následně byl výsledný bezešvý rastr převeden do bodové vrstvy pomocí nástroje **Raster To Point** čítající přes 36 milionů bodů. Body byly v pravidelné mřížce, vzdálenost mezi nimi činila 250 m (tedy shodná s rozlišením rastru). v atributové tabulce je důležitý atribut *grid_code* obsahující informaci o nadmořské výšce v každém bodu. Nástrojem **Add XY Coordinates** byly do atributů přidány XY souřadnice v systému ETRS89 (s nímž bylo pracováno v programu ArcGIS Pro), a do dvou ručně přidáných atributů *Latitude* a *Longitude* byly nástrojem **Calculate geometry** dopočítány souřadnice ve WGS84 (se kterými pracuje program ClimateEU).

Takto připravená bodová vrstva byla vyexportována ve formátu Esri Shapefile (SHP). Nejdůležitější je exportovaný soubor dBASE Table File (DBF), který byl vstupem do programu RStudio. v něm byl DBF soubor zpracován a vyexportován jako CSV, s nímž pracoval ClimateEU. Důvod exportu CSV přes RStudio je ten, že výsledné CSV je příliš velké na to, aby šlo zpracovat například v tabulkovém procesoru Microsoft Excel, na což tento program upozorní chybovou hláškou.

Programový kód pro export CSV z DBF souboru v programu RStudio (verze jazyka R 4.2.3) vypadá následovně:

```
#instalace a spuštění důležitých knihoven
remotes::install_version("SDMTools", "1.1-221")
library(SDMTools)
install.packages("foreign")
library(foreign)

setwd("set_working_directory") #nastaví absolutní cestu k adresáři
table=read.dbf("nazev_souboru.dbf") #z adresáře načte DBF soubor
```

Po načtení souboru DBF musí být opraveny názvy sloupců a pořadí sloupců tak, aby odpovídaly struktuře viditelné v atributové tabulce na Obr. 7:

```
head(table) #zobrazí hlavičku dbf souboru, slouží pouze pro kontrolu
fix(table) #umožní opravit názvy sloupců v hlavičce
table2=table[,c(1,2,5,4,3)] #mění pořadí sloupců
```

Poslední příkaz, jenž generuje tabulku *table2*, slouží ke změně pořadí sloupců. Sloupce totiž musí být pro program ClimateEU připraveny v pořadí *y*, *x*, *Latitude*, *Longitude*, *Elevation*. Příkaz s těmito konkrétními hodnotami `[,c(1,2,5,4,3)]` se použije tehdy, jsou-li sloupce v původní tabulce v pořadí *y*, *x*, *Elevation*, *Longitude*, *Latitude*.

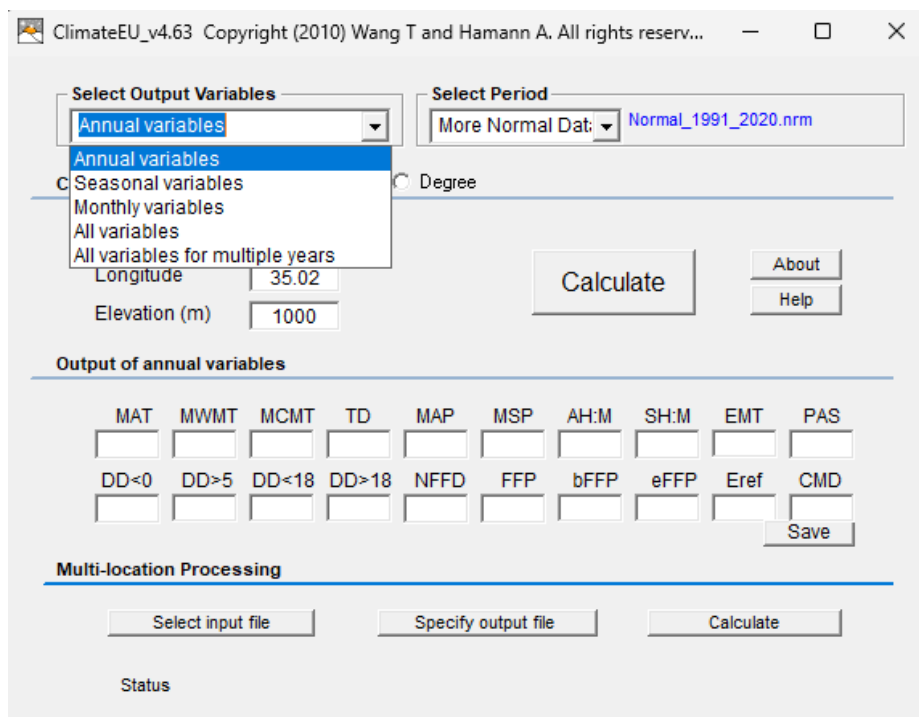
Poté těchto nezbytných úpravách se tabulka vyexportuje ve formátu CSV pomocí následujícího příkazu:

```
write.csv(table2, "nazev_souboru.csv", row.names = F, quote = F)
```

y	x	Latitude	Longitude	Elevation
3999875	3228125	57,80911	-8,603987	29,7608
3999875	3228375	57,80969	-8,599917	174,935
3999875	3228625	57,81027	-8,595847	225,657
3999875	3228875	57,810849	-8,591777	169,634
3999875	3229125	57,811429	-8,587707	161,381
3999875	3229375	57,812008	-8,583637	241,444
3999875	3229625	57,812588	-8,579566	148,921
3999875	3229875	57,813167	-8,575496	93,2513

Obr. 7 Náhled na soubor CSV v ArcGIS Pro obsahující souřadnice v systému ETRS89 (*y*, *x*) a WGS84 (*Latitude*, *Longitude*) doplněný o nadmořskou výšku (*Elevation*).

Vyexportovaný soubor CSV byl jako vstup nahrán do programu ClimateEU (Obr. 8). Byl specifikován výstupní soubor, typ období (měsíční/sezónní/roční) a klimatický normál 1991–2020 a program sám spočítal všechny nastavené proměnné. v případě použitého vstupního CSV o více než 36 milionech bodů trvalo generování jednoho výsledného souboru i více než 30 hodin. Výsledkem je vygenerovaný soubor CSV, v němž jsou každému bodu (řádku) na základě souřadnic a nadmořské výšky přiřazeny hodnoty několika klimatických proměnných.



Obr. 8 Náhled na jednoduché prostředí programu ClimateEU.

Generování bioklimatických proměnných

Soubory CSV byly vygenerovány celkem tři – s ročními (Obr. 9), měsíčními i sezónními proměnnými, to vše pro **klimatický normál 1991–2020**. Vytvořené CSV soubory byly nahrány (drag-and-drop) do programu ArcGIS Pro a pomocí **Display XY Data** byly za použití atributů x a y (souřadnice v ETRS89) převedeny na bodovou vrstvu.

OBJECTID*	Shape*	y	x	Latitude	Longitude	Elevation	MAT	MWMT	MCMT	TD	MAP	MSP	AHM	SHM	DD_0	DD5	DD_18	DD18	NFFD	bFFP	eFFP	FFP	PAS	EMNT	Eref	CMD
1	Point	3999875	3228125	57,80911	-8,603987	29,7608	8,6	13,4	4,7	8,7	1406	475	13,2	28,2	19	1459	3478	0	324	64	322	258	36	-11,6	361	0
2	Point	3999875	3228375	57,80969	-8,599917	174,935	7,9	13	3,9	9,1	1419	479	12,6	27,1	31	1291	3721	0	313	78	315	237	52	-13,3	351	0
3	Point	3999875	3228625	57,81027	-8,595847	225,657	7,7	12,8	3,6	9,2	1431	483	12,4	26,5	36	1235	3811	0	308	83	312	229	59	-14	348	0
4	Point	3999875	3228875	57,810849	-8,591777	169,634	7,9	12,9	3,9	9	1443	487	12,4	26,5	30	1282	3727	0	312	79	314	235	52	-13,3	352	0
5	Point	3999875	3229125	57,811429	-8,587707	161,381	8	12,9	4	8,9	1456	491	12,4	26,2	30	1285	3717	0	313	78	314	236	52	-13,2	353	0
6	Point	3999875	3229375	57,812008	-8,583637	241,444	7,6	12,7	3,6	9,1	1468	495	12	25,6	38	1205	3848	0	306	85	310	225	63	-14,2	348	0
7	Point	3999875	3229625	57,812588	-8,579566	148,921	8,1	12,9	4,1	8,8	1450	490	12,5	26,3	28	1307	3683	0	315	76	315	239	49	-13	354	0
8	Point	3999875	3229875	57,813167	-8,575496	93,2513	8,3	13,1	4,4	8,7	1429	483	12,8	27,1	23	1377	3587	0	319	71	318	247	42	-12,4	358	0
9	Point	3999875	3230125	57,813746	-8,571425	74,5423	8,4	13,3	4,5	8,8	1408	477	13,1	27,9	22	1408	3547	0	321	68	319	251	40	-12,1	359	0
10	Point	3999875	3230375	57,814325	-8,567354	72,8297	8,5	13,3	4,5	8,8	1388	470	13,3	28,3	22	1422	3532	0	322	67	320	253	39	-12,1	359	0
11	Point	3999875	3230625	57,814903	-8,563284	91,1057	8,4	13,4	4,4	9	1367	464	13,5	28,9	23	1415	3545	0	321	68	320	252	39	-12,3	357	0

Obr. 9 Náhled na atributovou tabulku s ročními proměnnými vygenerovanými programem ClimateEU.

Pomocí interpolační metody IDW byly z atributů exportovány jednotlivé rastry v rozlišení 250 m na pixel. Použit pro to byl nástroj **IDW** (Inverse Distance Weighing, viz kapitola 3, sekce *Použité metody*) v ArcGIS Pro. Jelikož bylo exportováno velké množství rastrů, byl zvolen dávkový proces operace *IDW* (*Batch IDW*).

Z ročních proměnných byly pro export použity tři atributy – **MAT** (Mean Annual Temperature, Průměrná roční teplota vzduchu = Bio 1), **MAP** (Mean Annual Precipitation, Průměrný roční úhrn srážek = Bio 12) a **TD** (teplotní rozdíl mezi průměrnou teplotou nejteplejšího a nejchladnějšího měsíce = kontinentalita). z měsíčních proměnných byly použity všechny atributy, tedy veškeré průměrné (**Tave01–Tave12**), maximální (**Tmax01–Tmax12**) a minimální (**Tmin01–Tmin12**) teploty vzduchu a měsíční úhrny srážek (**PPT01–PPT12**). Měsíční proměnné byly vstupem pro výpočty většiny bioklimatických charakteristik. Sezónní proměnné sice byly vyexportovány (např. úhrn srážek v zimě **PPT_wt**, teplota vzduchu v létě **TAV_sm** apod.), ale nakonec použity nebyly z důvodů uvedených níže (viz popis Bio 18). Ve výsledku vzniklo 51 rastrů (v rozlišení 250×250 m), z nichž se následně počítaly další proměnné.

Generované rastrové proměnné byly převzaty z předchozích výzkumů a jsou uvedeny v kapitole 4.3.1. z bioklimatických proměnných se jedná o následující:

- **průměrná roční teplota vzduchu (Bio 1)**
- **izotermalita (Bio 3)**
- **roční teplotní rozsah (Bio 7)**
- **úhrn srážek v nejvlhčím měsíci (Bio 13)**
- **srážková sezónnost (Bio 15)**
- **úhrn srážek v nejteplejším čtvrtletí (Bio 18)**
- **úhrn srážek v nejchladnějším čtvrtletí (Bio 19)**

Autor této práce se rozhodl vyexportovat další bioklimatické proměnné (Bio 4, Bio 14), a to pro případ, že by v celoevropském měřítku mohly být některé z nich vlivnější než ty stanovené v lokálních výzkumech. Byly vygenerovány i ty rastry, které byly pouze mezikrokem (např. Bio 2, Bio 5, Bio 10) při výpočtu výše stanovených proměnných.

K počítání většiny rastrů posloužily nástroje **Cell Statistics a Raster Calculator**, ve specifických případech byl použit program RStudio.

Při výpočtech bioklimatických proměnných byly použity vzorce a informace z příručky vydávané americkou USGS s názvem **Bioclimatic Predictors for Supporting Ecological Applications in the Conterminous United States** (O'Donnell et al., 2012):

Bio 1 – Průměrná roční teplota vzduchu byla vygenerována z atributu MAT (Mean Annual Temperature).

Bio 3 – Izotermalita se počítá podle vztahu:

$$Bio\ 3 = \frac{Bio\ 2}{Bio\ 7} \times 100$$

kde **Bio 2** je **průměrný roční rozsah denních teplot** a spočítá se podle vztahu:

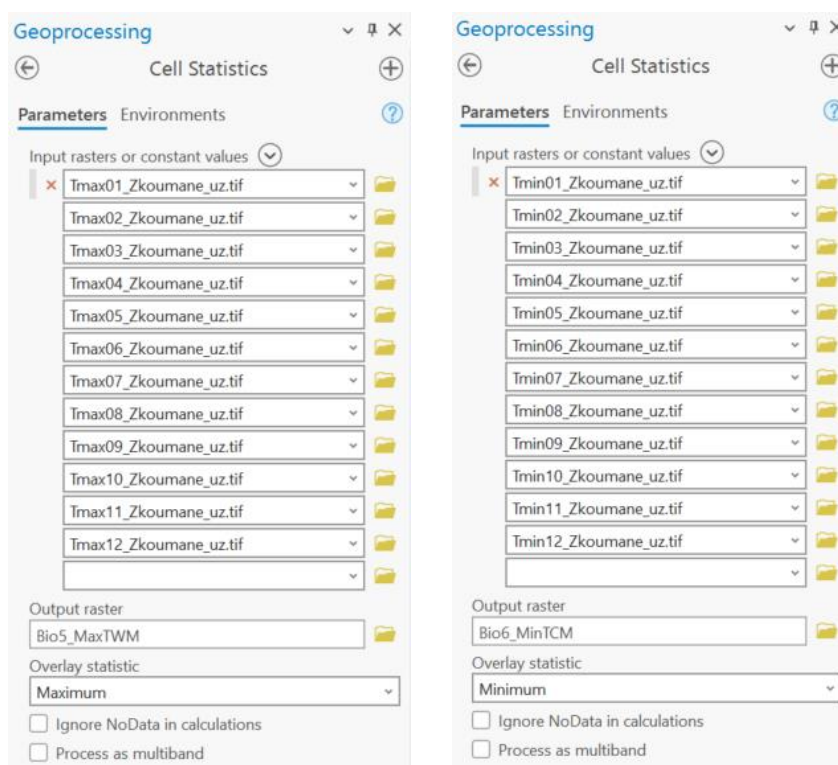
$$Bio\ 2 = \frac{\sum_{i=1}^{12} (Tmax_i - Tmin_i)}{12}$$

Bio 7 – Roční teplotní rozsah vznikne rozdílem:

$$Bio\ 7 = Bio\ 5 - Bio\ 6$$

kde **Bio 5** je **maximální teplota nejteplejšího měsíce** a **Bio 6** je **minimální teplota nejchladnějšího měsíce**.

Proměnná **Bio 5** byla spočítána nástrojem **Cell Statistics**. Do tohoto nástroje byly vloženy rastry maximálních měsíčních teplot vzduchu (Tmax01–Tmax12) a *Overlay statistic* bylo nastaveno na Maximum (Obr. 10). Tím se vytvořil nový rastr, který měl v každé buňce nejvyšší hodnotu ze vstupních rastrů. Je třeba postupovat tímto způsobem, poněvadž v některých oblastech Evropy může být nejvyšší teplota v červenci, jinde v srpnu (případně ještě v jiné měsíce). Nestačí proto uvažovat jako nejteplejší měsíc např. červenec a pro celé území převzít pouze červencovou teplotu a považovat ji za celoevropsky maximální.



Obr. 10 Nastavení nástroje *Cell Statistics* pro vygenerování Bio 5 (vlevo) a Bio 6 (vpravo).

Proměnná **Bio 6** byla vypočítána obdobným způsobem jako Bio 5. Rozdíl byl v tom, že byly vloženy rastry minimálních měsíčních teplot vzduchu (Tmin01–Tmin12) a *Overlay statistic* byla nastavena na Minimum.

Bio 13 – Úhrn srážek v nejvlhčím měsíci se spočítá stejným způsobem jako Bio 5, tedy pomocí nástroje **Cell Statistics**, do něhož se nahrají všechny rastry měsíčních úhrnů srážek (PPT01–PPT12) a *Overlay statistic* se nastaví na Maximum. Navíc byla vygenerována i proměnná **Bio 14**, tedy **úhrn srážek v nejsušším měsíci**, a to stejně jako Bio 13, jen *Overlay statistic* bylo nastaveno na Minimum.

Bio 15 – Srážková sezónnost značí nepravidelné rozložení srážek během běžného roku. Toto nepravidelné rozložení znamená, že většina srážek spadne v určitých měsících (Goméz, 2023). Bio 15 se spočítá podle vztahu:

$$Bio\ 15 = \frac{SD\{PPT01, \dots, PPT12\}}{1 + \left(\frac{Bio\ 12}{12}\right)} \times 100$$

kde **Bio 12** je **roční úhrn srážek** a byl vygenerován z atributu *MAP* (Mean Annual Precipitation).

$SD\{PPT01, \dots, PPT12\}$ je standardizovaná odchylka (SD, Standard Deviation) všech rastrů měsíčních srážek. Ta byla spočítána v programu RStudio z exportovaných rastrů měsíčních srážek pomocí následujícího skriptu:

```
library(raster) #nahraje knihovnu raster
setwd("set_working_directory") #nastaví absolutní cestu k adresáři

#načte z adresáře všechny rastry s příponou .tif/.TIF
ALL_FILES <- list.files(path = dbase, pattern = "*\\.tif$|*.TIF$")

stacked <- stack(ALL_FILES) #navrství rastry na sebe
sdr <- calc(stacked, fun = sd) #z navrstvených rastrů spočítá SD

writeRaster(sdr, filename = "file_name.tif") #exportuje výsledek
```

Standardizovaná odchylka (SD) ze vstupních rastrů se počítá i v případě proměnné **Bio 4** (teplotní sezónnost), na níž jde zmíněný skript též použít. Bio 4 se dá kromě standardizované odchylky vypočítat ještě s použitím koeficientu variance, v této práci však byla použita varianta se standardní odchylkou. Proměnná byla spočítána navíc (podobně jako Bio 14) k ostatním proměnným pro zkoumání případného vlivu na výsledky modelování. Bio 4 se spočítá podle následujícího vztahu:

$$Bio\ 4 = SD\{Tavg01, \dots, Tavg12\}$$

Bio 18 – Úhrn srážek v nejteplejším čtvrtletí se spočítá za pomoci proměnné **Bio 10** (průměrná teplota vzduchu v nejteplejším čtvrtletí). Postup zde však byl složitější, neboť nejprve bylo potřeba identifikovat nejteplejší čtvrtletí, které se ale napříč Evropou může měnit (proto nelze použít sezónní proměnné a obecně stanovit jako nejteplejší čtvrtletí léto, tedy červen–srpen, poněvadž v některých případech může být nejteplejší čtvrtletí o měsíc posunuté (např. červenec–září), což sezónní proměnné nereflektují). Pro zkrácení doby výpočtů bylo uvažováno, že nejteplejší čtvrtletí se bude nacházet v rozsahu květen–září, výpočet Bio 10 tedy proběhlo z atributů *Tavg05–Tavg09*. Pro všechny po sobě jdoucí tříměsíční období v rámci těchto pěti měsíců (květen–červen–červenec (KCC), červen–červenec–srpen (CCS), červenec–srpen–září (CSZ)) byly do atributové tabulky nástrojem **Calculate Field** spočítány tři nové atributy průměrných teplot vzduchu. Na příkladu průměrné teploty vzduchu pro květen–červen–červenec (KCC) vypadal výpočet atributu následovně:

$$Tavg_KCC = \frac{Tavg05 + Tavg06 + Tavg07}{3}$$

Obdobným způsobem byly vypočítány i atributy *Tavg_CCS* a *Tavg_CSZ*. Následně byly pomocí **IDW** z těchto tří atributů vyexportovány rastry (rozlišení 250×250 m) a pomocí nástroje **Cell Statistics** a nastavení *Overlay statistic na Maximum* vybráno nejteplejší čtvrtletí, čímž vznikla rastrová proměnná Bio 10.

Pro spočítání Bio 18 byly použity úhrny měsíčních srážek v období květen–září (atributy PPT05–PPT09). Byl vytvořen samostatný atribut *PPTWQ_bio18* a do něj byly spočítány úhrny srážek na základě atributů *Tavg_KCC*, *Tavg_CCS* a *Tavg_CSZ*. Pokud platilo, že největší hodnotu z těchto tří měl atribut *Tavg_KCC*, *PPTWQ_bio18* bylo spočítáno jako součet atributů *PPT05–PPT07*, pokud byl největší *Tavg_CCS*, *PPTWQ_bio18* se rovnalo

součtu *PPT06–PPT08* apod. Hodnoty z atributu *PPTWQ_bio18* byly nástrojem **IDW** interpolovány a byl vyexportován rastr zobrazující Bio 18.

Bio 19 – Úhrn srážek v nejchladnějším čtvrtletí byl spočítán analogicky k Bio 18. Zároveň bylo potřeba vypočítat **Bio 11**, tedy průměrnou teplotu vzduchu v nejchladnějším čtvrtletí, která byla vypočítána stejným postupem jako Bio 10 (podrobný postup výše), pouze s jinými atributy. Jako nejchladnější čtvrtletí byly vyhodnoceny pouze dvě varianty prosinec-leden-únor a leden-únor-březen, byly proto použity atributy *Tavg01–Tavg03* a atribut *Tave12* a měsíční srážky z atributů *PPT01–PPT03* a *PPT12*. Následnou interpolací **IDW** vznikl rastr Bio 19.

Z celkem 51 vstupních rastrů teplot vzduchu a úhrnů srážek bylo pomocí **interpolační metody IDW** vyexportováno pro **klimatický normál 1991–2020 15 bioklimatických rastrových proměnných**, které budou dále sloužit k modelování vhodnosti habitatu pro výskyt sršně asijské. Každý rastr má **rozlišení 250 metrů na pixel, přes 36 milionů pixelů a pokrývá 11 zemí Evropy**, v nichž byla zaznamenán nálezořvá data sršně *Vespa velutina*.

Pro rekapitulaci jsou zde uvedeny výsledné vygenerované rastry (celkem 15):

- Bio 1 Průměrná roční teplota vzduchu
- Bio 2 Průměrný roční rozsah denních teplot
- Bio 3 Izotermalita
- Bio 4 Teplotní sezónnost
- Bio 5 Maximální teplota nejteplejšího měsíce
- Bio 6 Minimální teplota nejchladnějšího měsíce
- Bio 7 Průměrná teplota nejsuššího čtvrtletí
- Bio 10 Průměrná teplota nejteplejšího čtvrtletí
- Bio 11 Průměrná teplota nejchladnějšího čtvrtletí
- Bio 12 Roční úhrn srážek
- Bio 13 Úhrn srážek v nejvlhčím měsíci
- Bio 14 Úhrn srážek v nejsušším měsíci
- Bio 15 Srážková sezónnost
- Bio 18 Úhrn srážek v nejteplejším čtvrtletí
- Bio 19 Úhrn srážek v nejchladnějším čtvrtletí

S použitím vstupních teplotních rastrů byla vygenerována ještě **kontinentalita** (která však není v seznamu bioklimatických proměnných), a to z atributu *TD* (teplotní rozdíl mezi průměrnou teplotou nejteplejšího a nejchladnějšího měsíce) v ročních datech.

Generování ostatních stanovištních proměnných

Mezi ostatní stanovištní proměnné byly zařazeny **nadmořská výška (DMR EU-DEM)**, jehož tvorba byla popsána dříve v textu), **sklon svahu** (ve stupních, vypočítaný z DMR nástrojem **Slope v** ArcGIS Pro) a **kategorická proměnná krajinného pokryvu Land Cover Map of Europe 2017** (Malinowski, R., et. al., 2020) od Evropské kosmické agentury. Rastr obsahuje 13 kategorií krajinného pokryvu:

- mraky
- umělé plochy a stavby
- zemědělská půda
- vinice
- listnaté lesy
- jehličnaté lesy
- bylinná vegetace
- suhomilná (sklerofylní) vegetace
- bažiny
- rašeliniště
- povrchy z přírodních materiálů (skály, lomy)
- trvale zasněžené plochy
- vodní plochy

Tyto stanovištní proměnné slouží k dokreslení podoby habitatu sršně asijské. Například v Jižní Korei žije až 70 % kolonií sršně asijské *Vespa velutina* v městském prostředí (Diaz et al., 2023). Je tedy zřejmé, že krajinný pokryv a využití území (především urbánní oblasti) budou mít na výskyt sršně asijské velký vliv.

Ostatní stanovištní rastry byly vyexportovány 3. Spolu s 15 bioklimatickými rastry a kontinentalitou bylo celkem vytvořeno **19 rastrů jednotlivých proměnných**.

Úprava zájmového území a export proměnných do WGS

Jelikož pro vstup dat do aplikace Wallace je podmínkou, že rastry musí mít stejné rozlišení, stejný prostorový rozsah dat (extent) a stejné umístění Null hodnot, byly provedeny kroky nezbytné k unifikaci prostorových vlastností všech 19 rastrů.

Rozlišení všech rastrů je totožné, a sice 250×250 m. Vzájemný překryv všech buněk rastrů byl zajištěn nastavením **Snap Raster** při exportu v ArcGIS Pro – všechny rastry byly polohou buněk přichyceny k buňkám zadaného referenčního rastru. **Snap Raster** byl nastaven při použití každého nástroje, jehož výstupem byl rastrový obraz potřebný pro další analýzy.

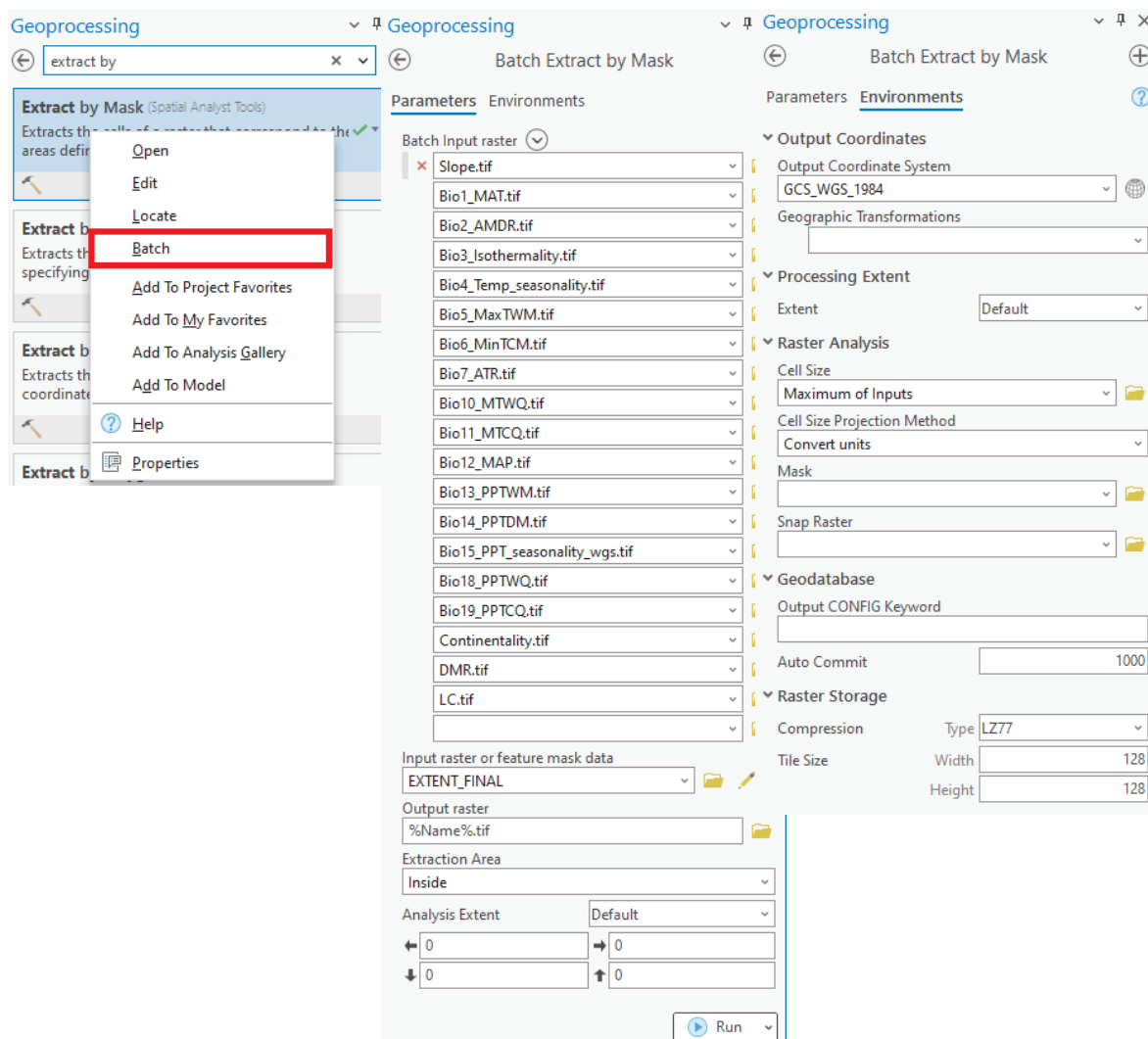
Jednotný prostorový rozsah rastrů a umístění Null hodnot bylo vyřešeno postupně v několika krocích. Nejprve byly všechny rastry nahrány do nástroje **Batch Is Null v** ArcGIS Pro (*batch* = dávkový proces, aby byly všechny rastry zpracovány najednou, spustí se tak, že se na v nabídce Geoprocessingu klikne na nástroj pravým tlačítkem a vybere se možnost *Batch*). Tento nástroj určuje, které buňky ze vstupního rastru mají hodnotu Null. Vrací hodnotu 1 pro buňky, jejichž vstupní hodnota je Null, a 0 pro buňky, které nejsou Null (Esri, 2023b). Vzniklo tak 19 rastrů s hodnotami 0/1. Těchto 19 rastrů poté bylo sečteno pomocí nástroje **Raster Calculator**. Tímto součtem vznikl jeden rastr, v němž byly pixely různé hodnoty od 19 (= všechny rastry mají v dané buňce Null hodnotu) do 0 (= všechny rastry mají v dané buňce hodnotu, která není Null). **Hodnoty 0 značí, že se jedná o společný průnik „not Null“ hodnot všech rastrů, a tedy výsledný prostorový rozsah (extent) společný pro všech 19 proměnných.**

Pomocí nástroje **Raster to Polygon** byl tento rastr převedený do vektorové podoby, hodnoty rastru se tím uložily do atributu *grid_code*. Byly vybrány pouze ty polygony, jejichž hodnota *grid_code* je rovna 0 (= společný extent) a uloženy do nové vrstvy **EXTENT_FINAL**.

Následně bylo všech 19 rastrů oříznuto (dávkovým procesem) nástrojem **Batch Extract By Mask v** ArcGIS Pro. Během exportu byla jako parametr Mask nastavena polygonová vrstva EXTENT_FINAL, nastavena složka, kam se mají výsledné rastry uložit ve formátu GeoTIFF a v Environments byl jako Output Coordinate System zvolený WGS84

(EPSG: 4326), neboť Wallace pracuje právě s tímto souřadnicovým systémem. Přesné nastavení je na Obr. 11.

Takto vyexportované rastrové proměnné, uložené ve formátu GeoTIFF, mají jednotné prostorové vlastnosti (rozdělení, SRS, prostorový rozsah (extent) a stejné umístění Null hodnot) a jsou kompletně připravené pro použití v programu Wallace.



Obr. 11 Spuštění dávkového (Batch) procesu nástrojů v Geoprocessingu (vlevo), nahrání rastrů a zvolení Masky (uprostřed), pojmenování *%Name%.tif* znamená, že se rastr bude jmenovat jako vstupní rastr a bude uložen jako GeoTIFF soubor (mimo databázi). Vpravo je nastavení SRS WGS84 (EPSG: 4326) v Environments.

4.3.3 Finální výběr proměnných statistickými metodami

Ačkoliv nejlivnější proměnné byly stanoveny už v podkapitole 4.3.1, a to na základě rešerše existujících prací, autor této práce se rozhodl vybranými statistickými metodami otestovat i všechny ostatní vygenerované proměnné a zjistit, zda například v celoevropském měřítku nejsou vlivnější jiné bioklimatické charakteristiky. Výzkumy byly totiž pouze lokální – např. pouze pevninské Španělsko, Mallorca, Portugalsko, celoevropský kontext a podmínky jim tedy chybí.

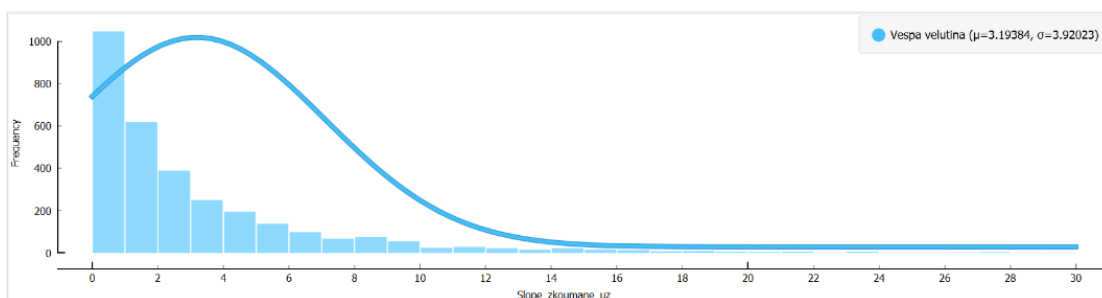
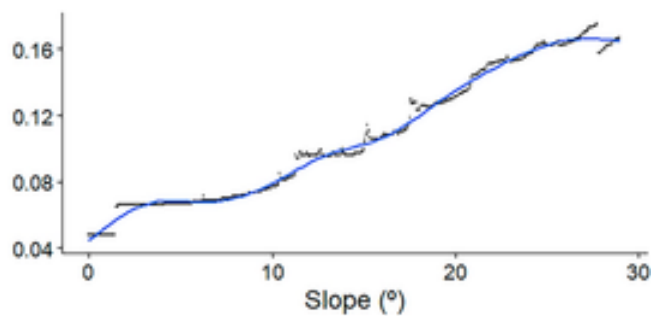
V následující kapitole bude k některým pojmům a provedeným krokům odkazováno do kapitoly 4.4, respektive jejích jednotlivých podkapitol. Výběr výsledného souboru bioklimatických proměnných totiž probíhal mj. na základě znalostí načerpaných při seznamování s modelovacím postupem a na základě datových požadavků aplikace Wallace, v níž bylo modelováno, a jejího přístupu k proměnným. Dle autora je však vhodnější mít informace o tvorbě rastrových proměnných v jedné kapitole, a tak některé zde popsané kroky nebyly v dosavadním textu vysvětleny.

Do modelu by mohly vstupovat všechny vytvořené bioklimatické a stanovištní proměnné, výpočetní doba generování modelu by se tím ale výrazně protáhla. Vstup všech proměnných navíc ani není nutný, neboť některé mají na výskyt druhu zanedbatelný vliv, jiné jsou například navzájem korelované a mohly by naopak predikčnímu výkonu modelu uškodit. Navzájem korelované proměnné by zjednodušeně řečeno do modelu zanesly „několikrát tu samou informaci“, čímž by byla dána této složené proměnné větší váha a model by vykazoval nepřesnosti (Marchi, 2015). Bylo tedy provedeno statistické vyhodnocení proměnných, které vedlo k redukci počtu prediktorů metodou PCA a vybrání pouze těch nejlivnějších, a zároveň vytipování a redukce velmi korelovaných proměnných.

Metoda PCA bývá velmi často používána při modelování ekologických nik a rozšíření druhů (Herrera et al., 2022; Janžekovič, 2012; Marchi, 2015). Metoda se používá na transformaci původních proměnných do menšího počtu proměnných (hlavních komponent) s co nejmenší ztrátou informace. v této práci byla metoda PCA použita podobným způsobem jako v Herrera et al. (2022).

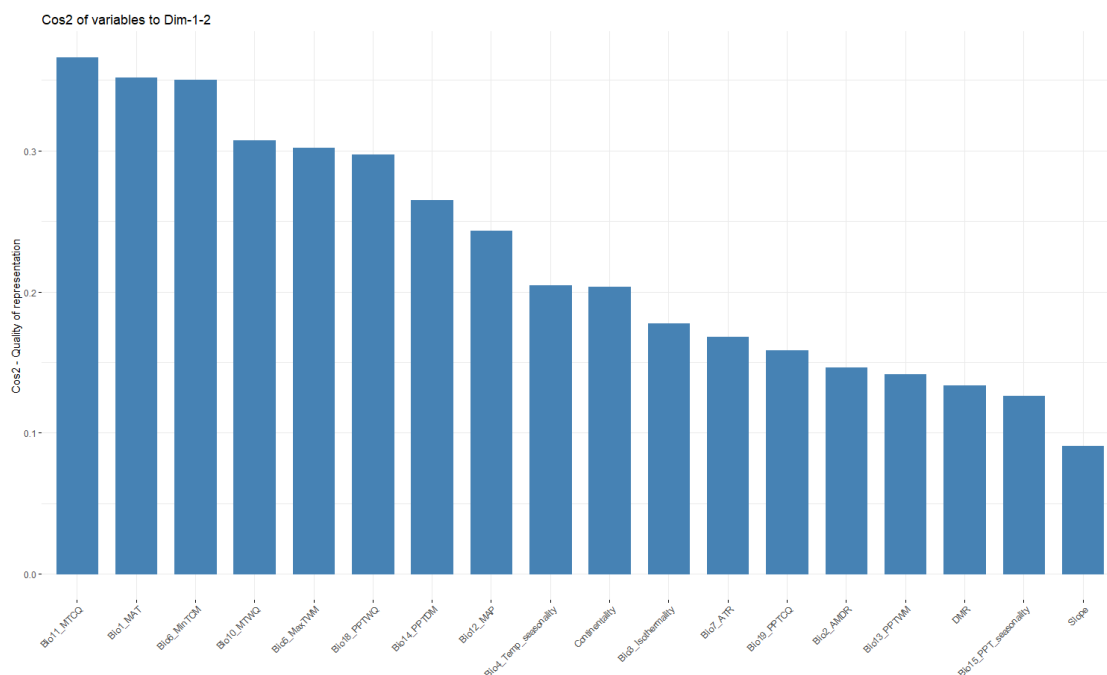
Do analýzy vstupovaly dvě vrstvy – 3 119 bodů výskytu sršní (již osídlené oblasti) a náhodný vzorek 100 000 bodů z **environmentálního pozadí** (tzv. background, dosud neosídlený, viz podkapitola 4.4.4), jenž byl vytvořen nástrojem **Create Random Points** uvnitř polygonu studované oblasti (viz podkapitola 4.4.4). Obě bodové vrstvy byly nahrány do programu ArcGIS Pro a nástrojem **Extract Multi Values To Points** byly do těchto bodů nahrány hodnoty z buněk bioklimatických a stanovištních rastrů (celkem 19 proměnných, viz podkapitola 4.3.2). Bodová vrstva obsahující atribut „scientific_name“, v němž byly hodnoty *Vespa velutina* a *Background*, a dalších 19 atributů odpovídajících jednotlivým proměnným byla vyexportována ve formátu CSV a XLSX. CSV formát byl použit při výpočtu PCA v programu RStudio, XLSX formát byl nahrán do programu Orange pro možnost nahlížení na data.

Některé prediktory byly odebrány z výsledných proměnných rovnou po vizuální inspekci grafů četnosti výskytu sršní v určitých podmínkách. Například zvyšování vhodnosti habitatu se zvyšujícím se sklonem svahu (jenž byl ve výzkumu na Mallorce (Herrera et. al, 2022) určený jako jeden z nejdůležitějších prediktorů a vykazoval stoupavý trend, tedy se zvyšujícím se sklonem rostla vhodnost habitatu) nebylo na základě analýzy celoevropských dat vůbec potvrzeno. Počet nalezených sršní naopak se zvyšujícím se sklonem klesal (Graf). Malý vysvětlující přínos proměnné *Slope* potvrdila i PCA (Graf).



Graf 1 Míra vhodnosti sklonu svahu na Mallorce (Herrera et al., 2022) vůbec neodpovídá pozorovaným datům napříč celou Evropou (dole, výstup z programu Orange). Zatímco dle výzkumu na Mallorce má vhodnost habitatu stoupavý trend se zvyšujícím se sklonem svahu, četnost výskytu podle pozorovaných dat naopak se zvyšujícím se sklonem svahu klesá. Jedná se o výbornou ukázkou toho, že lokální výzkum může potvrdit specifika některé proměnné, které ale v celoevropském měřítku nebudou hrát takovou roli.

V programu RStudio byla následně provedena analýza hlavních komponent (PCA). Nejprve do ní vstupovaly všechny proměnné, přičemž autora práce nejvíce zajímal příspěvek jednotlivých proměnných do prvních dvou komponent (PC1, PC2):



Graf 2 Příspěvek predikčních proměnných do prvních dvou hlavních komponent (PC1, PC2) v případě testování na celé množině proměnných. (výstup z programu RStudio)

Nyní by teoreticky mohlo být vybráno např. 8 proměnných s nejvyšším příspěvkem do PC1 a PC2 a ty vložit do modelu. Mohlo by však dojít k výrazné chybě, jsou-li tyto proměnné výrazně korelované. Pro zjištění vzájemné korelace byla vytvořena korelační matice.

Nad 19 predikčními rastry byla v ArcGIS Pro pomocí nástroje **Band Collection Statistics** spočtena korelační matice. Po jejím nahrání do programu Microsoft Excel byla vizualizována vzájemná korelace všech proměnných, přičemž proměnné s nejvyšší korelací byly vyřazeny z procesu modelování:

	Bio1	Bio2	Bio3	Bio4	Bio5	Bio6	Bio7	Bio10	Bio11	Bio12	Bio13	Bio14	Bio15	Bio18	Bio19	Cont	DMR	Slope
Bio1	1,00	0,47	0,45	0,01	0,88	0,88	0,26	0,95	0,94	-0,52	-0,39	-0,66	0,50	-0,79	-0,20	0,02	-0,48	-0,36
Bio2	0,47	1,00	0,56	0,49	0,76	0,09	0,82	0,61	0,28	-0,42	-0,35	-0,49	0,35	-0,44	-0,40	0,51	0,12	-0,11
Bio3	0,45	0,56	1,00	-0,41	0,36	0,48	0,00	0,30	0,56	-0,18	-0,21	-0,21	0,12	-0,44	0,07	-0,40	-0,19	-0,23
Bio4	0,01	0,49	-0,41	1,00	0,42	-0,44	0,89	0,33	-0,33	-0,27	-0,18	-0,25	0,16	0,05	-0,54	1,00	0,23	0,04
Bio5	0,88	0,76	0,36	0,42	1,00	0,57	0,67	0,97	0,70	-0,59	-0,44	-0,72	0,55	-0,73	-0,41	0,43	-0,23	-0,27
Bio6	0,88	0,09	0,48	-0,44	0,57	1,00	-0,23	0,68	0,98	-0,36	-0,29	-0,47	0,34	-0,72	0,06	-0,43	-0,59	-0,37
Bio7	0,26	0,82	0,00	0,89	0,67	-0,23	1,00	0,54	-0,06	-0,38	-0,26	-0,43	0,34	-0,21	-0,54	0,90	0,26	0,01
Bio10	0,95	0,61	0,30	0,33	0,97	0,68	0,54	1,00	0,78	-0,58	-0,43	-0,71	0,54	-0,74	-0,36	0,34	-0,36	-0,31
Bio11	0,94	0,28	0,56	-0,33	0,70	0,98	-0,06	0,78	1,00	-0,41	-0,31	-0,56	0,44	-0,78	-0,01	-0,32	-0,49	-0,33
Bio12	-0,52	-0,42	-0,18	-0,27	-0,59	-0,36	-0,38	-0,58	-0,41	1,00	0,93	0,81	-0,23	0,74	0,84	-0,28	0,28	0,42
Bio13	-0,39	-0,35	-0,21	-0,18	-0,44	-0,29	-0,26	-0,43	-0,31	0,93	1,00	0,60	0,07	0,63	0,79	-0,19	0,31	0,45
Bio14	-0,66	-0,49	-0,21	-0,25	-0,72	-0,47	-0,43	-0,71	-0,56	0,81	0,60	1,00	-0,69	0,80	0,60	-0,25	0,11	0,23
Bio15	0,50	0,35	0,12	0,16	0,55	0,34	0,34	0,54	0,44	-0,23	0,07	-0,69	1,00	-0,46	-0,08	0,16	0,14	0,10
Bio18	-0,79	-0,44	-0,44	0,05	-0,73	-0,72	-0,21	-0,74	-0,78	0,74	0,63	0,80	-0,46	1,00	0,34	0,04	0,24	0,30
Bio19	-0,20	-0,40	0,07	-0,54	-0,41	0,06	-0,54	-0,36	-0,01	0,84	0,79	0,60	-0,08	0,34	1,00	-0,55	0,05	0,24
Cont	0,02	0,51	-0,40	1,00	0,43	-0,43	0,90	0,34	-0,32	-0,28	-0,19	-0,25	0,16	0,04	-0,55	1,00	0,22	0,03
DMR	-0,48	0,12	-0,19	0,23	-0,23	-0,59	0,26	-0,36	-0,49	0,28	0,31	0,11	0,14	0,24	0,05	0,22	1,00	0,65
Slope	-0,36	-0,11	-0,23	0,04	-0,27	-0,37	0,01	-0,31	-0,33	0,42	0,45	0,23	0,10	0,30	0,24	0,03	0,65	1,00

Obr. 12 Korelační matice proměnných. Tmavé červené odstíny značí silnou pozitivní korelaci, tmavé modré značí silnou negativní korelaci.

Poměrně logicky je **velmi korelovaná Bio 1** (průměrná roční teplota vzduchu) s **Bio 5** (maximální teplota nejteplejšího měsíce), **Bio 6** (minimální teplota nejchladnějšího měsíce), **Bio 10** (průměrná teplota nejteplejšího čtvrtletí) a **Bio 11** (průměrná teplota nejchladnějšího čtvrtletí). Stejně tak je zřejmé, že budou korelovat Bio 5 s Bio 10 a Bio 6 s Bio 11. **Velmi výrazná korelace je i mezi Bio 4** (teplotní sezónnost) a **kontinentalitou**.

Výběr finálních proměnných proběhl na základě porovnání grafu příspěvku jednotlivých proměnných do prvních dvou komponent (Graf 2) a korelační matice (Obr. 12). Bylo dbáno na výběr nejvíce vysvětlujících proměnných a zároveň omezení korelace (např. Bio 11 má nejvyšší příspěvek do PCA, bylo tedy upřednostněno před Bio 1, se kterým koreluje). Snahou bylo omezit proměnné, jejichž korelace je $\text{corr} > 0,8$.

Byly vyzkoušeny jednotlivé varianty proměnných, přičemž nejlepší kombinace z nich (nejvíce vysvětlující s nepříliš velkou korelací) byla nakonec autorem vyhodnocena kombinace následujících proměnných (**jedná se výsledný soubor bioklimatických proměnných, které vstupovaly do modelu**):

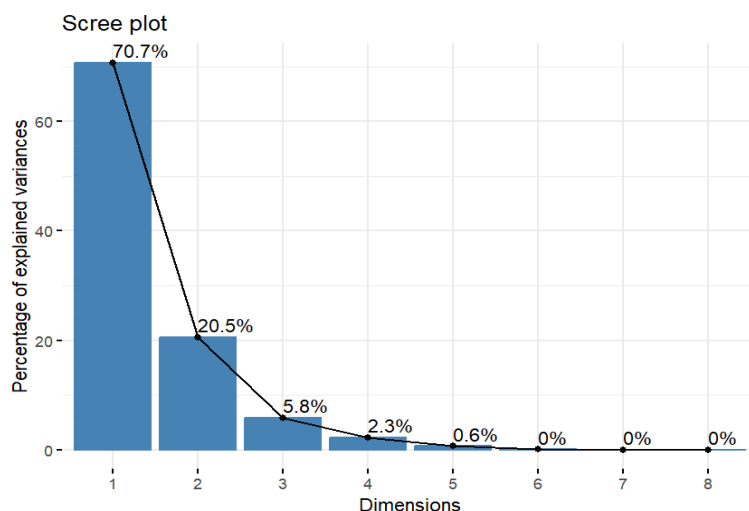
- **Bio 2 Průměrný roční rozsah denních teplot**
- **Bio 3 Izotermalita**
- **Bio 4 Teplotní sezónnost**
- **Bio 10 Průměrná teplota nejteplejšího čtvrtletí**
- **Bio 11 Průměrná teplota nejchladnějšího čtvrtletí**
- **Bio 15 Srážková sezónnost**
- **Bio 18 Úhrn srážek v nejteplejším čtvrtletí**
- **Bio 19 Úhrn srážek v nejchladnějším čtvrtletí**

Výsledná PCA ukazuje, že při použití těchto proměnných, první hlavní komponenta vysvětluje 70,7 % a druhá 20,5 % celkové variability (Graf).

Bioklimatické rastry byly doplněny o kategorickou proměnnou Land Cover (více informací viz sekce *Generování ostatních stanovištních proměnných* v podkapitole 4.3.2. **Celkem tedy bylo zvoleno 9 predikčních proměnných vstupujících do modelu.**

	Bio2	Bio3	Bio4	Bio10	Bio11	Bio15	Bio18	Bio19
Bio2	1,00	0,56	0,49	0,61	0,28	0,35	-0,44	-0,40
Bio3	0,56	1,00	-0,41	0,30	0,56	0,12	-0,44	0,07
Bio4	0,49	-0,41	1,00	0,33	-0,33	0,16	0,05	-0,54
Bio10	0,61	0,30	0,33	1,00	0,78	0,54	-0,74	-0,36
Bio11	0,28	0,56	-0,33	0,78	1,00	0,44	-0,78	-0,01
Bio15	0,35	0,12	0,16	0,54	0,44	1,00	-0,46	-0,08
Bio18	-0,44	-0,44	0,05	-0,74	-0,78	-0,46	1,00	0,34
Bio19	-0,40	0,07	-0,54	-0,36	-0,01	-0,08	0,34	1,00

Obr. 13 Korelační matice výsledného souboru použitých proměnných.



Graf 3 Procento variability vysvětlené jednotlivými hlavními komponentami.

4.3.4 Bioklimatické proměnné pro normál 2041–2060

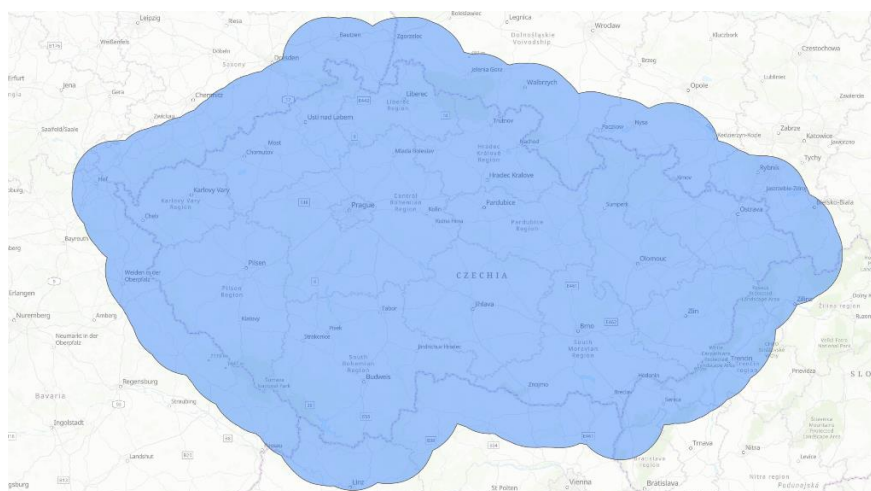
Vzhledem k jednomu z cílů práce, a to predikci výskytu založené na změnách klimatu, byly vybrané proměnné vygenerovány pro **normál 2041–2060**. Tato budoucí data byla vygenerována pouze pro území České republiky, neboť sloužila pouze k aplikaci modelu (natrénovaného na současných datech z celé Evropy) na nové časové období (normál 2041–2060) a vizualizaci vhodnosti habitatu konkrétně v ČR. Nebyly proto exportovány rastry pro celou Evropu.

Pro generování dat pro normál 2041–2060 byl použit dataset **ECLIPS 2.0**. Konkrétně byly použity proměnné z regionálního klimatického modelu (Regional Climate Model, RCM) **MPI-CSC-REMO2009**, který je několika autory (Kumar et al., 2015; Tapiador et al., 2020) považován za jeden z nejpřesnějších, což dokládá i výzkum zaměřený na anomálii rozšiřování ledovců v oblasti pohoří Karákóram navzdory probíhajícímu globálnímu oteplování a ústupu ledovců ve většině části světa. Kumar et al. (2015) právě tuto anomálii úspěšně objasnili a simulovali pomocí modelu REMO. Ostatní RCM tak přesné nebyly.

Jako scénář pro koncentraci skleníkových plynů byl zvolen **RCP 8.5** (Representative Concentration Pathways, hodnota 8,5 odkazuje ke koncentraci uhlíku, která přináší globální oteplování o průměrné hodnotě 8,5 Wattů/m² na celé planetě; Climate Nexus, 2023) z toho důvodu, že je považován za nejpravděpodobnější, pokud lidská společnost nepodnikne potřebné kroky pro redukci emisí skleníkových plynů (Hausfather, 2019).

Dataset ECLIPS 2.0 obsahuje veškeré rastry, které byly vygenerovány v programu ClimateEU (tzn. veškeré průměrné/minimální/maximální, roční/měsíční/sezónní teploty vzduchu a srážky v jednotlivých měsících), jejich rozlišení je ale menší (30" ≈ 1 km na pixel), to však vyřeší následná interpolace.

Nejprve bylo **stanoveno zájmové území** – buffer kolem hranic ČR ve vzdálenosti 30 km (Obr. 14), kterým byly rastry ořezány (nástrojem **Extract By Mask**). Buffer byl vytvořen pro potřeby následné interpolace – kdyby byly rastry oříznuty pouze polygonem ČR, mohlo by na hranici docházet ke generování zkreslených hodnot vlivem extrapolace.



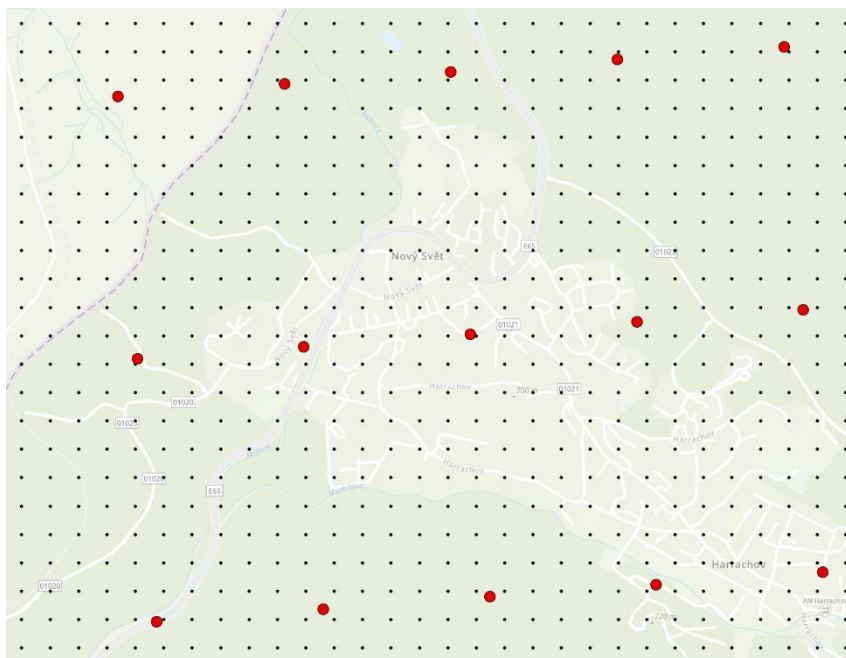
Obr. 14 Zájmové území použité pro ořezání rastrů proměnných pro klimatický normál 2041–2060 a následnou predikci.

Bioklimatické proměnné, které byly v podkapitole 4.3.3 vyhodnoceny jako nejvlivnější na výskyt sršně asijské, byly spočítány nad výchozími rastry z ECLIPS 2.0 (v rozlišení 30"), a to stejnými postupy, které jsou popsány v podkapitole 4.3.2 v sekci *Generování bioklimatických proměnných*.

Následně byla nad rastry vygenerována bodová vrstva („transformací“ jednoho z rastrů na body nástrojem **Raster To Points** v ArcGIS Pro), do níž byly nástrojem **Extract Multi Values To Points** vypsány hodnoty buněk jednotlivých bioklimatických rastrů. Tato vygenerovaná bodová vrstva s názvem **POINT_BIO**, která obsahovala 8 atributů s hodnotami jednotlivých bioklimatických proměnných, byla vstupem do analytického nástroje **Geostatistical Wizard**, kde byly hodnoty metodou **Co-kriging** interpolovány za vzniku finálních rastrových proměnných v požadovaném rozlišení 250×250 metrů.

Interpolace metodou co-kriging byla zvolena z toho důvodu, že umožňuje kromě hlavní predikované proměnné $P1$ nahrát další proměnnou $P2$, na níž je $P1$ závislá a která dokáže proměnnou $P1$ zpřesnit (Ahmed, 2023). Příkladem je použití metody co-kriging při interpolaci teploty vzduchu ($P1$) za použití digitálního modelu reliéfu, resp. nadmořské výšky ($P2$), jelikož teplota vzduchu je na nadmořské výšce závislá.

Data nadmořské výšky byly převzaty z vrstvy EU-DEM převzorkované na rozlišení 100×100 m a oříznuty zájmovým územím (buffer okolo České republiky). Následně byly nástrojem **Raster To Point** převedeny do poměrně podrobné bodové vrstvy.



Obr. 15 Červené body (POINT_BIO) obsahují hodnoty bioklimatických proměnných ($P1$), které jsou závislými proměnnými na nadmořské výšce ($P2$), kterou reprezentují černé body (DMR převedený do bodové vrstvy). Obě tyto vrstvy vstupují do procesu interpolace metodou co-kriging.

V nastavení metody co-kriging nebylo nic měněno, bylo použito např. výchozí počtu sousedních bodů uvažovaných při interpolaci a dalších parametrů. Po vygenerování výstupů interpolace byly tyto rastrové vrstvy uloženy pod stejným názvem jako původní proměnné pro současný klimatický normál. Shoda názvu je podmínkou aplikace Wallace pro přenos modelu do nového prostředí (viz podkapitola 4.4.9). **Pro území České republiky bylo vyexportováno celkem 8 bioklimatických rastrů pro klimatický normál 2041–2060 a svým obsahem jsou shodné s vybranými proměnnými uvedenými v podkapitole 4.3.3. K nim byl přidán ještě rastr krajinného pokryvu Land Cover Map of Europe 2017.**

4.4 Zpracování modelu v aplikaci Wallace

V následující podkapitole bude podrobně popsán postup modelování v aplikaci Wallace, jednotlivé kroky modelování zahrnují následující komponenty:

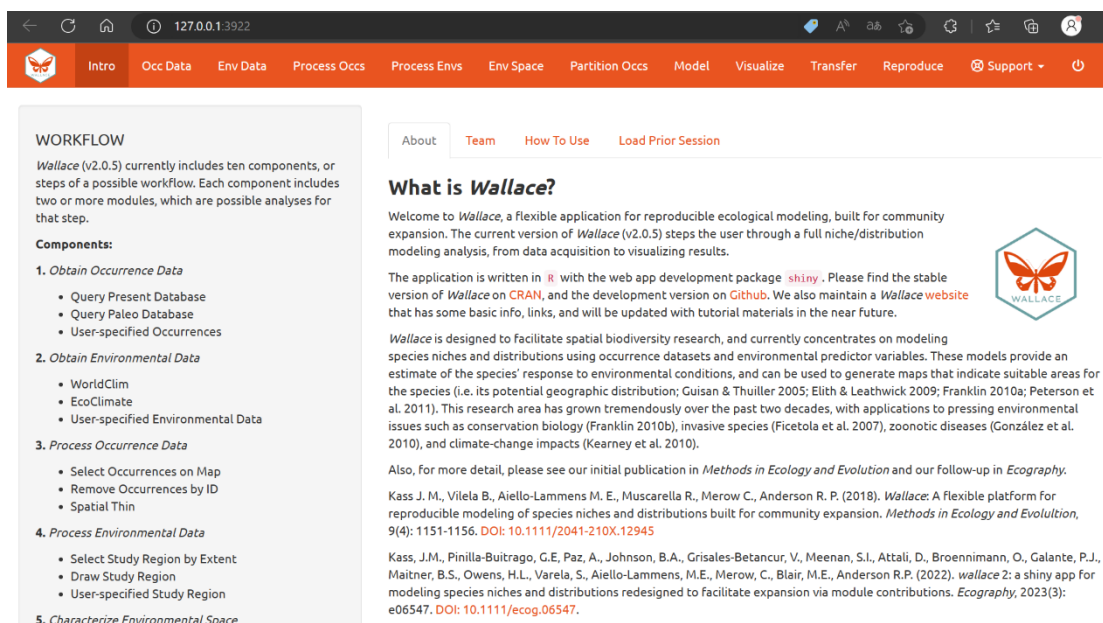
- 1) Obtain Occurrence Data (Získání dat výskytu druhu)
- 2) Obtain Environmental Data (Získání environmentálních dat)
- 3) Process Occurrence Data (Zpracování dat výskytu)
- 4) Process Environmental Data (Zpracování environmentálních dat)
- 5) Characterize Environmental Space (Charakteristika environmentálního prostoru)
- 6) Partition Occurrence Data (Segmentace/Rozdělení dat výskytu)
- 7) Build and Evaluate Niche Model (Vytvoření a vyhodnocení modelu niky)
- 8) Visualize Model Results (Vizualizace výsledků modelu)
- 9) Model Transfer (Přesunutí modelu)
- 10) Reproduce (Reprodukce modelu)

Aplikace Wallace se spouští přes konzoli v programu RStudio, a to tímto skriptem:

```
install.packages("wallace") #instalace knihovny Wallace
library(wallace) #načtení knihovny Wallace
run_wallace() #spuštění aplikace Wallace
```

Aplikace se spustí v okně výchozího prohlížeče. IP adresa 127.0.0.1 udává hostovanou stránku na localhost. Vývojáři pro použití aplikace Wallace vytvořili přehledný online návod s názvem **Wallace Ecological Modeling Application v2.0 Vignette** (Johnson et al., 2023), v němž vysvětlují prováděné kroky a z něhož autor této práce vycházel.

Podrobnější popisy jednotlivých komponentů programu jsou přímo v aplikaci Wallace v záložce *Component guidance*, v záložce *Module Guidance* jsou popsána jednotlivá dílčí nastavení v rámci komponentů. Tento text byl napsán vývojáři s cílem přiblížit uživatelům jednotlivé komponenty a moduly teoreticky (proč by měly být nástroje použity) a metodicky (co nástroje dělají).



The screenshot shows the homepage of the Wallace application. The top navigation bar includes links for Intro, Occ Data, Env Data, Process Occs, Process Envs, Env Space, Partition Occs, Model, Visualize, Transfer, Reproduce, and Support. The main content area is divided into two columns. The left column, titled 'WORKFLOW', lists five components: 1. Obtain Occurrence Data (with sub-points: Query Present Database, Query Paleo Database, User-specified Occurrences), 2. Obtain Environmental Data (with sub-points: WorldClim, EcoClimate, User-specified Environmental Data), 3. Process Occurrence Data (with sub-points: Select Occurrences on Map, Remove Occurrences by ID, Spatial Thin), 4. Process Environmental Data (with sub-points: Select Study Region by Extent, Draw Study Region, User-specified Study Region), and 5. Characterize Environmental Space. The right column, titled 'What is Wallace?', provides a welcome message, describes the application's purpose, and includes a list of references.

WORKFLOW

Wallace (v2.0.5) currently includes ten components, or steps of a possible workflow. Each component includes two or more modules, which are possible analyses for that step.

Components:

1. Obtain Occurrence Data
 - Query Present Database
 - Query Paleo Database
 - User-specified Occurrences
2. Obtain Environmental Data
 - WorldClim
 - EcoClimate
 - User-specified Environmental Data
3. Process Occurrence Data
 - Select Occurrences on Map
 - Remove Occurrences by ID
 - Spatial Thin
4. Process Environmental Data
 - Select Study Region by Extent
 - Draw Study Region
 - User-specified Study Region
5. Characterize Environmental Space

About **Team** **How To Use** **Load Prior Session**

What is Wallace?

Welcome to *Wallace*, a flexible application for reproducible ecological modeling, built for community expansion. The current version of *Wallace* (v2.0.5) steps the user through a full niche/distribution modeling analysis, from data acquisition to visualizing results.

The application is written in **R** with the web app development package **shiny**. Please find the stable version of *Wallace* on **CRAN**, and the development version on **GitHub**. We also maintain a *Wallace* website that has some basic info, links, and will be updated with tutorial materials in the near future.

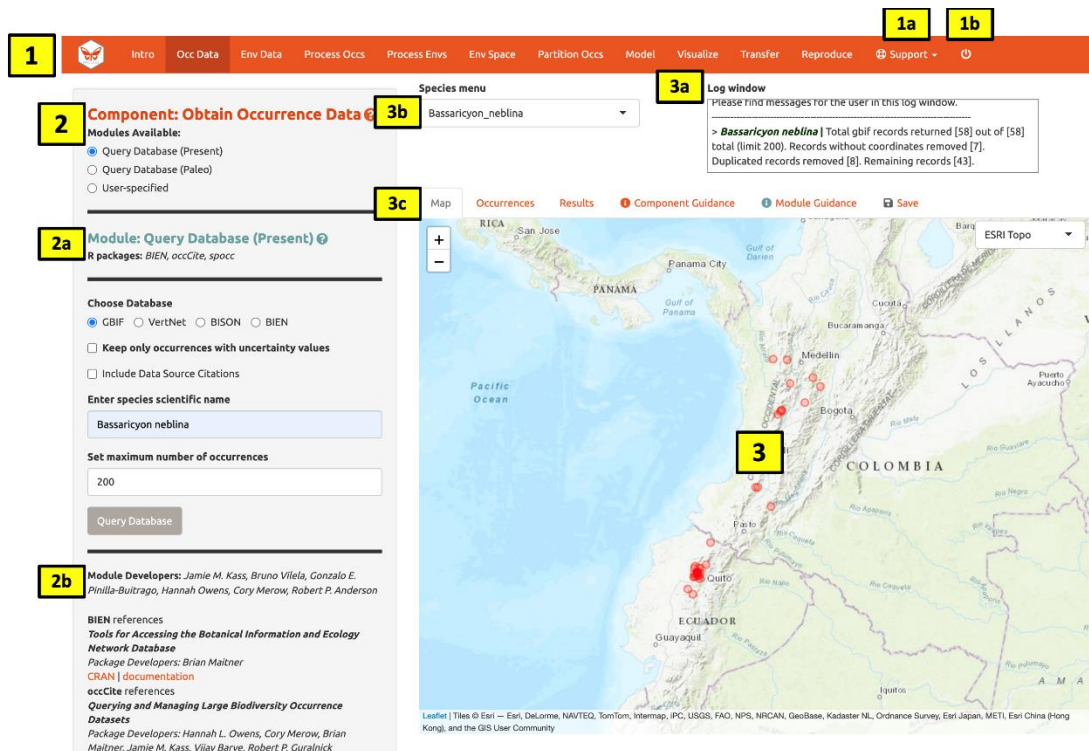
Wallace is designed to facilitate spatial biodiversity research, and currently concentrates on modeling species niches and distributions using occurrence datasets and environmental predictor variables. These models provide an estimate of the species' response to environmental conditions, and can be used to generate maps that indicate suitable areas for the species (i.e. its potential geographic distribution; Guisan & Thuiller 2005; Elith & Leathwick 2009; Franklin 2010a; Peterson et al. 2011). This research area has grown tremendously over the past two decades, with applications to pressing environmental issues such as conservation biology (Franklin 2010b), invasive species (Ficetola et al. 2007), zoonotic diseases (González et al. 2010), and climate-change impacts (Kearney et al. 2010).

Also, for more detail, please see our initial publication in *Methods in Ecology and Evolution* and our follow-up in *Ecography*.

Kass J. M., Vilela B., Aiello-Lammens M. E., Muscarella R., Merow C., Anderson R. P. (2018). *Wallace*: A flexible platform for reproducible modeling of species niches and distributions built for community expansion. *Methods in Ecology and Evolution*, 9(4): 1151-1156. DOI: 10.1111/2041-210X.12945

Kass, J.M., Pinilla-Buitrago, G.E., Paz, A., Johnson, B.A., Crisales-Betancur, V., Meenan, S.I., Attali, D., Broenimann, O., Galante, P.J., Maitner, B.S., Owens, H.L., Varela, S., Aiello-Lammens, M.E., Merow, C., Blair, M.E., Anderson R.P. (2022). *wallace 2*: a shiny app for modeling species niches and distributions redesigned to facilitate expansion via module contributions. *Ecography*, 2023(3): e06547. DOI: 10.1111/ecog.06547.

Obr. 16 Úvodní obrazovka aplikace Wallace.



Obr. 17 Jednotlivé součásti aplikačního rozhraní aplikace Wallace.
(zdroj: <https://wallacecomod.github.io/>)

Následující popis je věnován součástí rozhraní aplikace Wallace (Obr. 17) a vychází z návodu.

Na horní liště jsou za sebou přehledně zobrazeny jednotlivé komponenty (1). Wallace obsahuje tlačítko Podpora (1a), které odkazuje na skupinu Google, e-mail, webové stránky a stránku Github pro nahlášení problémů, a také tlačítko Ukončit (1b), které ukončí relaci.

Vlevo se nachází panel nástrojů (2) se všemi ovládacími prvky uživatelského rozhraní, jako jsou tlačítka, textové vstupy atd. Pro komponent *Obtain Occurrence Data* existují tři moduly: *Query Database (Present)*, *Query Database (Paleo)* a *User-specified*. Komponent i modul mají vedle názvu tlačítko s otazníkem (?). Kliknutím na něj se uživatel dostane na záložku *Component/Module Guidance*, kde jsou příslušné návodné texty. Při klikání na jednotlivé moduly v menu se mění název modulu a jsou zde vypsány balíčky R, které modul používá (2a). Dále se mění i ovládací panel pro vybraný modul (2b). Do modulů mohou přispívat další výzkumní pracovníci a jejich vývojáři; odkazy na síť CRAN a dokumentaci jsou uvedeny v dolní části panelu.

Na pravé straně je prostor pro vizualizace a datové náhledy (3). Jakákoli provedená operace vypíše zprávu v okně protokolu (3a). v tomto okně se také zobrazí případná chybová hlášení. Wallace v2.0.5 umožňuje uživateli pracovat s výskyty více druhů. Pokud je načteno více druhů, přepíná se mezi nimi pomocí rozbalovací nabídky druhů (3b).

Vizualizační prostor obsahuje několik záložek (3c), včetně interaktivní mapy, tabulky záznamů o výskytu (Occurrences), okna s výsledky (Results), textových oken s pokyny pro model a komponenty a záložka pro ukládání výstupů jednotlivých komponent a ukládání aktuální relace.

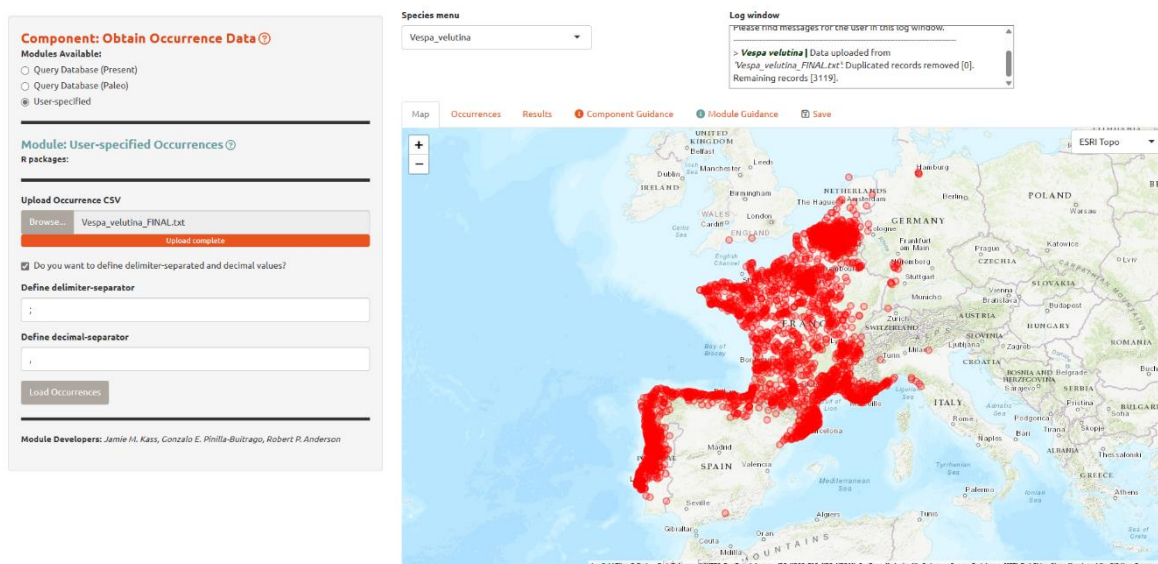
4.4.1 Získání dat výskytu druhu

Následující kapitoly se budou věnovat postupu tvorby modelu v této práci.

Nejprve byla nahrána data výskytu sršně asijské (jejich zpracování je popsáno v podkapitole 4.2). **Byl zvolena možnost *User-specified***, která umožňuje nahrát vlastní data ve formátu CSV nebo TXT. Nahrávaný soubor musí obsahovat tři povinné sloupce, názvy sloupců musí být v tomto pořadí a s anglickými uvozovkami: “scientific_name”, “longitude” a “latitude”. Je možné nastavit vlastní oddělovač položek a desetinných hodnot.

V případě nahrávání souboru CSV však program neustále zobrazoval chybovou hlášku s textem *! ERROR: Please input a file with columns "scientific_name", "longitude", "latitude"*, a to i přesto, že v CSV souboru tyto sloupce obsaženy byly a byly i stejně pojmenovány. Překvapivě, když byl CSV soubor uložen jako TXT, jeho načtení proběhlo bez problémů. Načteno bylo celkem 3 119 bodů výskytu (Obr. 18).

Dále je možné stáhnout zvolené množství bodů např. z GBIF pomocí modulu ***Query Database (Present)***, nebo importovat prehistorická data výskytu pomocí modulu ***Query Database (Paleo)***.



Obr. 18 Ukázka modulu *Obtain Occurrence Data* i s načtenými daty výskytu.

4.4.2 Získání environmentálních dat

Po nahrání dat výskytu byla do programu **nahrána environmentální data, a to opět jako *User-specified data***. Jedná se o 8 vybraných (viz podkapitola 4.3.3) rastrových proměnných ve formátu TIF (resp. GeoTIFF), které musí mít stejné prostorové vlastnosti (stejně rozlišení, prostorový rozsah a umístění Null hodnot, více v podkapitole 4.3.2 v sekci *Úprava zájmového území a export proměnných do WGS*). Pro rychlé načítání rastrů je nejlepší zaškrtnout možnost *Save to memory for faster processing and save/load option*.

V rámci dalších modulů jde pracovat s již existujícími daty ze dvou databází:

V **prvním modulu** jde importovat 19 bioklimatických proměnných ze stránek WorldClim (verze z roku 2005), a to v rozlišení 30“ (≈ 1 km), 2,5‘ (≈ 5 km), 5‘ (≈ 10 km) nebo 10‘ (≈ 20 km). Rozlišení 30“ je sice poměrně podrobné, aplikace však umožní nahrát pouze jednu rastrovou dlaždici

Druhý modul umožňuje pracovat s bioklimatickými proměnnými projektu ecoClimate. v něm jde navíc vybrat konkrétní model obecné cirkulace atmosféry/oceánu a vybírat ze tří časových scénářů – 1) normál 1950–1999, 2) období před 6 000 a 3) před 21 000 lety.

Nejpřesnějších a nejaktuálnějších dat však uživatel dosáhne nahráním vlastních proměnných. Bylo však potřeba, aby byly velmi dobře pohlídány prostorové vlastnosti rastrů, jinak program vrátí chybové hlášky týkající se neodpovídajících si Null hodnot či rozdílných prostorových rozsahů (extents).

Výsledné nahrané rastry odpovídají proměnným vybraným v podkapitole 4.3.3 a jedná se o následující (uvedeny jsou jejich názvy):

- **Bio2_AMDR.tif**
- **Bio3_Isothermality.tif**
- **Bio4_Temp_seasonality.tif**
- **Bio10_MTWQ.tif**
- **Bio11_MTCQ.tif**
- **Bio15_PPT_seasonality.tif**
- **Bio18_PPTWQ.tif**
- **Bio19_PPTCQ.tif**
- **LC.tif** (krajinný pokryv, Land Cover)

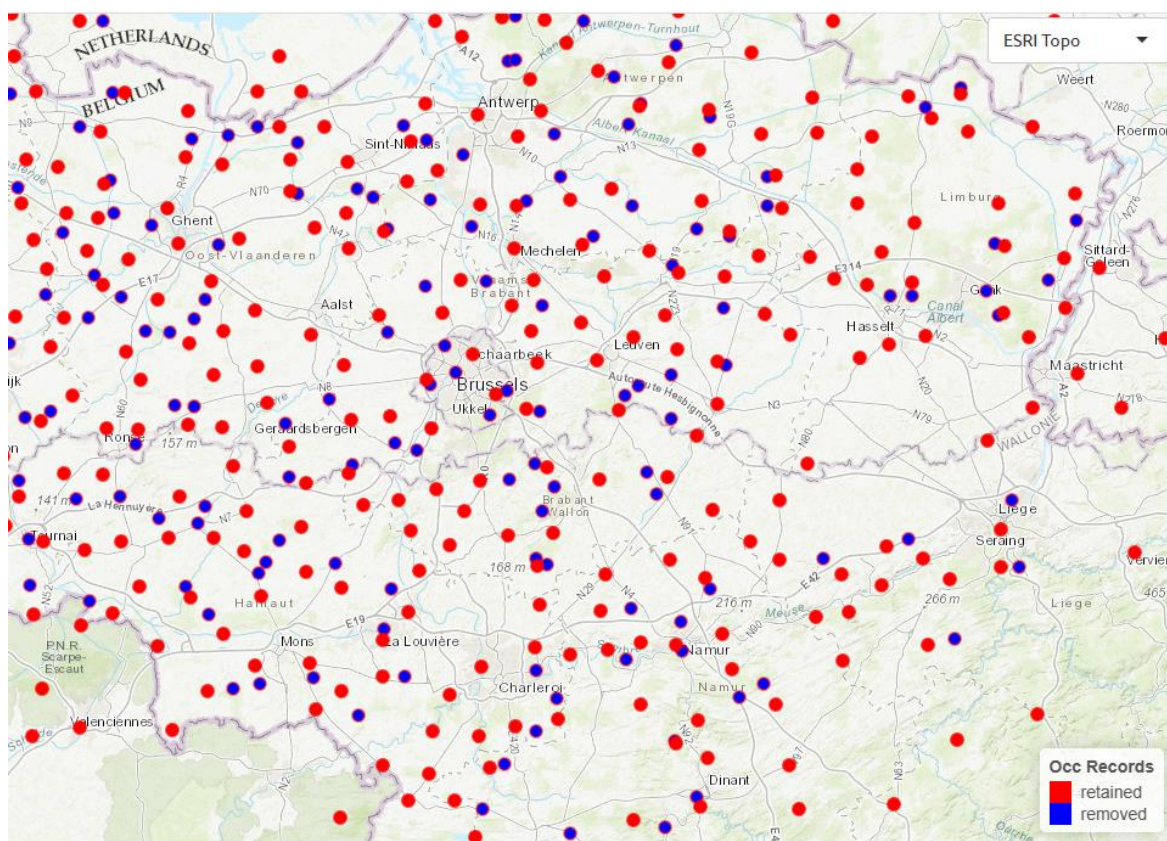
4.4.3 Zpracování dat výskytu

Předtím, než se začnou zpracovávat rastry, byla zpracována data výskytu. Prvotní předzpracování (zředění hustoty bodů) bylo popsáno v kapitole 4.2. Wallace poskytuje několik modulů pro výběr do modelu dále zahrnutých nálezových dat, a to:

- **přímý výběr bodů výskytu nakreslením obalového polygonu v mapě** (*Select Occurrences On Map*),
- **odstranění výběru podle ID** (*Remove Occurrences By ID*) – může se hodit, obsahuje-li vrstva výskytů chyby určené prostou vizuální kontrolou lokalit zakreslených do mapy (Gaiji et al., 2013); například pokud je druh znám pouze z lesa ve vysokých nadmořských výškách a několik lokalit z online databáze se nachází v nížině na pastvinách, mohou tyto podezřelé výskyty vést uživatele k vyloučení těchto záznamů z analýzy,
- **prostorová redukce hustoty bodů** (*Spatial Thin*) – podobná redukce byla použita už v kapitole 4.2, nicméně zde je odlišnost v tom, že algoritmus určuje vzdálenost mezi body a na základě nastavené prahové hodnoty maže body, které jsou blíže sobě, než je prahová hodnota; jelikož údaje o výskytu obvykle trpí negativními efekty (např. velká prostorová autokorelace; Kramer-Schadt, 2013)) souvisejícími s nerovnoměrným (tj. zkresleným) vzorkováním v geografickém prostoru, je dobré body zredukovat tak, aby se zabránilo jakémukoli zkreslení v environmentálním prostoru, které by mohlo zkreslit odhady niky druhu (Kadmon et al., 2004).

V této práci byl použit modul *Spatial Thin*, a to k dodatečným úpravám vložených bodových dat, neboť z velké části již byla data redukována manuálně. Nastavená prahová hodnota pro smazání přebytečných bodů byla nastavena na 5 km (tedy bod, který je od jiného bodu blíže než 5 km, vymaže). z nahraných 3 119 bodů výskytu bylo tímto

nástrojem vygenerována **finální zředitá vrstva bodů čítající 2 339 bodů výskytu** (vrstva tedy byla zredukována o dalších 780 bodů). Ukázka původně nahraných a redukováných dat je na Obr. 19.



Obr. 19 Ukázka zachovaných (červených) a odebraných (modrých) bodů při procesu redukce hustoty bodů pomocí modulu *Spatial Thin*.

4.4.4 Zpracování environmentálních dat

Důležitou součástí celého procesu modelování ve Wallace je komponent 4 – Zpracování environmentálních dat (*Process Environmental Data*). Tento komponent totiž zpracovává definované studované území, resp. environmentální data, a na jeho nastavení je závislá poměrně velká část modelovacího procesu. Komponent vybírá studovanou oblast (definovanou polygonem) pro analýzu, a tímto polygonem poté maskuje (ořezává) jednotlivé vstupní predikční proměnné. Výběr studovaného regionu je rozhodující pro přístupy modelování niky (např. právě Maxent), které porovnávají prostředí, v němž je přítomen určitý druh, s prostředím nějakého srovnávacího souboru dat nazývaného jako pozadí (background). Maxent využívá tzv. *presence-background* přístupu, což znamená, že porovnává hodnoty predikční proměnné v bodech pozadí s hodnotami v bodech výskytu (na rozdíl od technik *presence-absence*, které vyžadují data o nepřítomnosti druhu; Johnson et al., 2023). Rozlišují se tedy druhem osídlené oblasti (presence points) od oblastí, které je potenciálně ještě možné tímto druhem osídlit (background points).

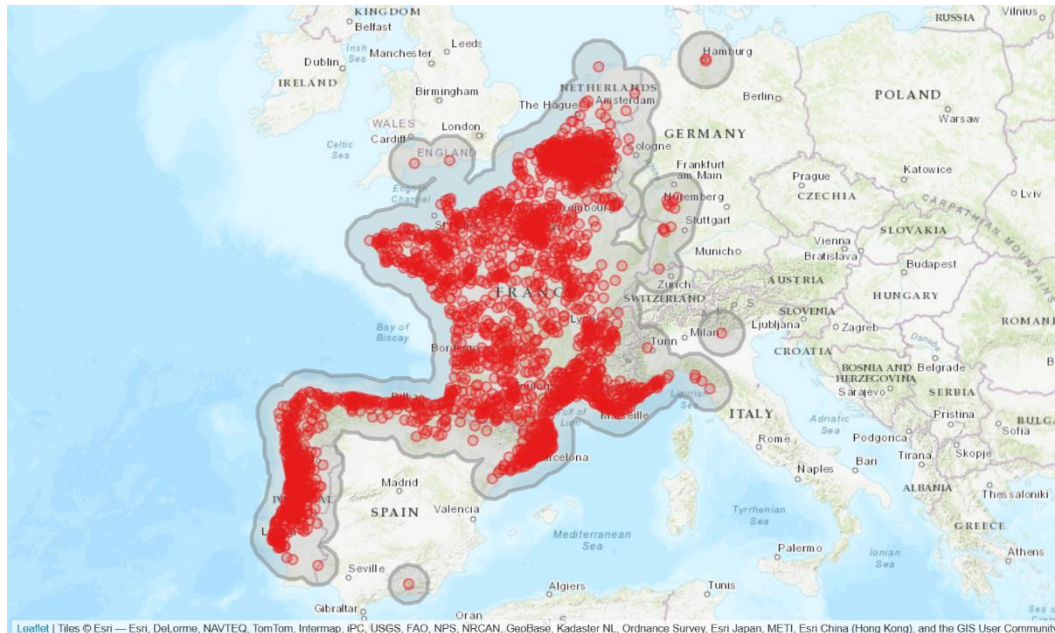
Volba studované oblasti pro generování bodů pozadí byla převzata od Herrera et al. (2022), kteří volili velikost regionu na základě rychlosti šíření sršně asijské. Vzhledem k registrované maximální rychlosti šíření sršně *Vespa velutina* (78 km/rok; Robinet et al., 2017) byl jako studovaný region zvolen polygon, jenž vznikl v programu ArcGIS Pro

nástrojem **Buffer**, který vytvořil obalovou zónu okolo vstupní vrstvy bodů výskytu. Vstupní vrstvou byly všechny výskyty, byla nastavena vzdálenost 78 km, a v *Dissolve Type* byla zvolena možnost *Dissolve all output feature into a single feature*. Vznikl tak polygon (studovaný region), který pokrýval body výskytu a jejich okolí do vzdálenosti 78 km. Jedná se o oblast, které je sršněmi teoreticky dostupná v období jednoho roku, a definuje potenciální pozadí (background), které sršeň může v nejbližší době osídlit.

V komponentu lze použít pro stanovení studované oblasti tři přístupy:

- **výběr studované oblasti** (*Select Study Region*) – třemi různými způsoby lze vybrat studovaný region:
 - **ohraničující rámeček** (*bounding box*), který kolem všech bodů výskytu udělá obdélník,
 - **minimální konvexní polygon** (*minimum convex polygon*), jenž udělá kolem všech bodů konvexní tvar s minimální plochu,
 - **obalové zóny kolem bodů** (*point buffers*) – na základě stanovené vzdálenosti ve *stupních* udělá kolem všech bodů obalovou zónu (buffer),
- **nakreslení studované oblasti** (*Draw Study Region*),
- **uživatelsky specifikovaná studovaná oblast** (*User-specified Study Region*) – lze nahrát polygon ve formátu CSV (obsahující sloupce s definovanými souřadnicemi lomových bodů) nebo v SHP (tři soubory .shp, .shx, .dbf)

V této práci byl použit modul *User-specified Study Region* a byla nahrána studovaná oblast (78km buffer bodů, Obr. 20), jehož vytvoření je popsáno výše.



Obr. 20 Studovaná oblast (šedě) určená jako 78km buffer bodů výskytu (červeně).

Poté byl proveden druhý krok, a to náhodný výběr bodů pozadí (*Sample Background Points*). **Nejsou použita všechna data z proměnných, ale bylo nastaveno vygenerování 50 000 náhodných bodů** v celé studované oblasti (neboť 10 000, které jsou nastaveny jako výchozí, jsou pro takto velké studované území nedostatečný počet). Aplikace Wallace

pak na pozadí procesu tímto regionem oříznula (vymaskovala) jednotlivé rastry a ze všech rastrů vypsalala do každého bodu hodnotu pixelů pod tímto bodem.

4.4.5 Charakteristika environmentálního prostoru

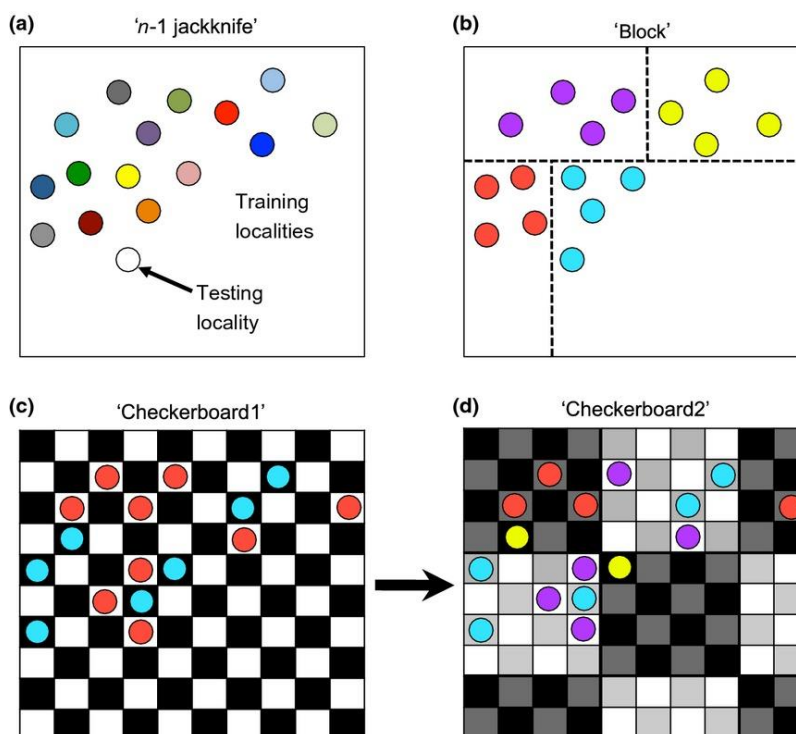
Komponent 5 – Charakteristika environmentálního prostoru slouží pro mezidruhové porovnání stanovištních podmínek dvou druhů:

- metoda PCA (modul *Environmental Ordination*),
- porovnáním grafů ukazujících, které části environmentálního prostoru jsou hustěji obsazeny daným druhem a jaká je dostupnost environmentálních podmínek přítomných na pozadí (modul *Occurrence Density Grid*),
- grafem překryvů nik (modul *Niche Overlap*), který slučuje grafy z předchozího modulu.

Vzhledem k tomu že v práci nebyly porovnávány dva druhy, nebyl tento modul použit.

4.4.6 Segmentace/Rozdělení dat výskytu

V následujícím komponentu došlo k vytvoření trénovacího a testovacího vzorku nálezových dat. Pro hodnocení síly modelu je totiž potřeba vytvořit teoreticky nezávislá data. Pokud žádné nezávislé soubory dat neexistují, je jedním z řešení rozdělit data na podmnožiny, o kterých lze předpokládat, že jsou na sobě nezávislé, pak postupně sestavit model na všech (trénovacích) podmnožinách kromě jedné a vyhodnotit výkonnost tohoto modelu na vynechané (testovací) podmnožině. Tento postup je znám jako k-násobná křížová validace (kde k je celkový počet podmnožin). Po dokončení tohoto postupného sestavování modelu Wallace vypočítá statistiky výkonnosti modelu a poté sestaví model s využitím všech dat.



Obr. 21 Ukázka jednotlivých metod segmentace bodů do skupin. (zdroj: <https://besjournals.onlinelibrary.wiley.com/>)

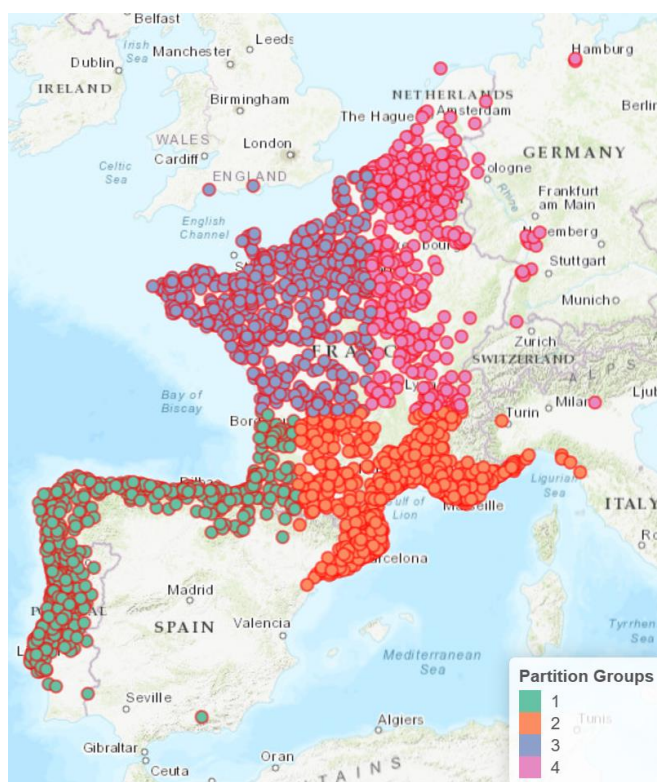
Komponent nabízí dva druhy (moduly) segmentace na trénovací a testovací data: **prostorová** a **neprostorová segmentace**. Příklady některých segmentací jsou na Obr. 21.

Jednou z možností **neprostorové segmentace** je prosté náhodné rozdělení na několik skupin (*Random k-fold*), ale u prostorových dat hrozí, že skupiny na sobě nebudou prostorově nezávislé. Druhá možnost neprostorové segmentace, metoda *Jackknife* ("leave-one-out"), pracuje tak, že každý vzorek (bod výskytu) umístí do samostatné skupiny, což se doporučuje pro druhy s malou velikostí vzorku (Shcheglovitova a Anderson, 2013).

Druhým modulem je **prostorová segmentace**. Prostorové rozdělení s k -násobnou křížovou validací nutí model předpovídat do oblastí, které jsou vzdálené od oblastí použitých k trénování modelu, což se hodí, je-li model přenášen v prostoru nebo čase. Pokud model poskytuje přesné předpovědi na prostorově segmentovaných datech, má pravděpodobně dobrou přenositelnost – může se dobře přenést na nové hodnoty predikčních proměnných (protože vzdálené oblasti jsou obvykle environmentálně odlišnější než blízké).

K dispozici jsou **tři metody prostorové segmentace**:

- **blok** (*Block*, $k = 4$) – metoda nakreslí nad daty vertikální a horizontální čáry, a vytvoří tak čtyři skupiny lokalit ($k = 4$) odpovídající čtyřem prostorovým obdélníkům; tyto obdélníky obsahují podobný počet lokalit,
- **šachovnice 1** (*Checkerboard 1*, $k = 2$) – rozdělí lokality do šachovnicového vzoru a přiřadí lokality do jedné ze dvou skupin ($k = 2$) podle toho, do kterého čtverce šachovnice lokality spadají,
- **šachovnice 2** (*Checkerboard 2*, $k = 4$) – metoda je podobná šachovnici 1, ale přidává další úroveň dělení a výsledkem jsou čtyři skupiny ($k = 4$) lokalit.



Obr. 22 Finální rozdělení bodů výskytu na čtyři skupiny pomocí prostorové segmentační metody *Block*.

V této práci byl použit modul *Spatial Partition* a v rámci něj po pečlivém zvážení použita metoda *Block* (Obr. 22). Jak uvádí Muscarella et al. (2014), metoda *Block* může být žádoucí pro studie zahrnující přenos modelu v prostoru nebo čase, včetně možnosti setkání s neanalogovými podmínkami (tj. nezaznamenanými podmínkami, které nemají žádnou moderní obdobu, buď tu tyto podmínky už nejsou, např. klima v období posledního zalednění, nebo tu ještě nejsou, např. rostoucí teplota vlivem klimatických změn; Williams a Jackson, 2007). Vzhledem k tomu, že model bude přenášen z trénovacích lokalit na Českou republiku, a zároveň i do jiného časového období (normál 2041–2060), byla metoda *Block* vyhodnocena jako nejvhodnější.

4.4.7 Vytvoření a vyhodnocení modelu niky

Komponent 7 – Vytvoření a vyhodnocení modelu niky využívá výstupy z předchozích komponentů k sestavení modelů používajících buď *presence-only* nebo *presence-background* data. *Presence-only* modelovací přístup je nabízený v rámci modulu BIOCLIM, který v této práci proveden nebyl a nebude tedy již více popisován.

Pro modelování byl použit druhý modul *Maxent* (a algoritmus maxnet, nikoli Java verze maxent.jar), jehož implementaci v této práci byla věnována celá příprava a postup zpracování dat.

Vytvoření modelu

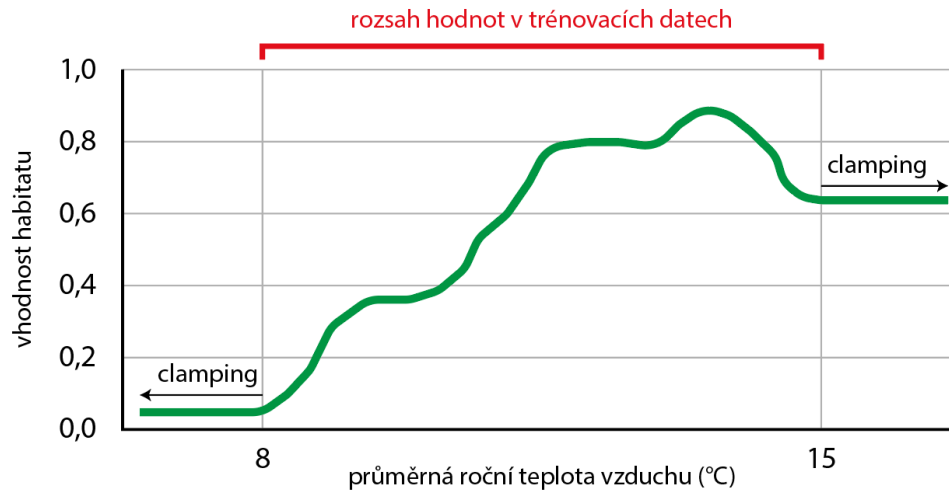
Jakožto algoritmus strojového učení *Maxent* interně rozhoduje o výběru proměnných a vhodnosti modelu (James et al. 2013), nicméně různá nastavení mohou výrazně ovlivnit složitost modelu a výsledné předpovědi. v rámci modulu *Maxent* je několik nastavení, které umožňují **snížit nebo zvýšit složitost modelu** prostřednictvím dvou klíčových faktorů: **funkcí, jimiž se model prokládá** (v algoritmu *Maxent* jsou tyto funkce nazvány jako *Feature Classes*) a **regularizačních multiplikátorů** (RM) neboli penalizací vůči komplexnosti modelu. Čím vyšší je hodnota multiplikátoru, tím větší je penalizace, což vede k tvorbě jednodušších modelů s menším množstvím výsledných proměnných. Multiplikátory lze nastavit v rozsahu 0,5–10, zároveň lze určit hodnotu kroku pro výběr multiplikátoru (*Multiplier Step Value*). Nízké hodnoty multiplikátoru mají za následek větší napasování modelu na trénovací data, což může vést k přetrénování modelu (a tedy tzv. *overfittingu*), tedy přizpůsobení modelu tak moc na trénovací data, že se model dobře nezobecní na nezávislá testovací data (Phillips, 2017).

Modul *Maxent* poté automatizuje dva pracovní postupy, a to (1) sestavení několika modelů s odlišnou mírou složitosti podle určených funkcí a regularizačních multiplikátorů a (2) kvantifikace výkonnosti modelu pomocí několika statistických metod.

Různé funkce (*Feature Classes*) určují tvar dostupných modelovaných vztahů v prostoru prostředí. Standardní funkce jsou *linear (L)*, *quadratic (Q)*, *hinge (H)* a *product (P)*. Více kombinací funkcí (např. LQ, LQHP) vede k potenciálně vyšší složitosti modelu. Podrobnějšímu popisu jednotlivých funkcí se dále věnuje Elith et al. (2011).

Nastavení funkce a RM mohou mít obzvláště silný vliv na výstup modelu *Maxent* (Warren a Seifert, 2011). z těchto důvodů bylo analyzováno více různých nastavení funkcí a multiplikátorů (jemnější hodnoty kroku pro výběr multiplikátoru mohou vést k přesnějšímu určení optimálního nastavení regularizace).

Uživatel musí v nastavení ***Are you using a categorical variable?*** specifikovat nahrané kategorické proměnné (např. kategorie využití území).



Obr. 23 Ukázka upínání (clampingu) hodnot extrémnějších, než jsou hodnoty v trénovacích datech, na smyšlených datech. Vhodnost habitatu je za těmito hranicemi stejná jako pro nejextrémnější hodnoty v rámci trénovaných dat. (zdroj: autor (podle Philips, 2017))

Dalším důležitým nastavením je **Clamping** (neboli upínání hodnot proměnných). Vytváření predikcí pro celé zkoumané území může být komplikováno nutností extrapolace hodnot na podmínky prostředí, které se v souboru trénovacích dat nevyskytují. U algoritmu Maxent lze u rastrových predikcí generovaných modelem nastavit (upnout) hodnoty vhodnosti habitatu pro extrémnější podmínky prostředí, než jsou ty v trénovací sadě dat (tj. neanalogové podmínky), na hodnotu vhodnosti spojenou s minimální, nebo maximální hodnotou proměnných v trénovacích datech. Důvod je následující – to, že se druh vyskytuje v určitých trénovacích podmínkách (např. prům. roční teplota vzduchu nanejvýš 15 °C), neznamená, že by se nemohl adaptovat na extrémnější podmínky (např. až 20 °C), které jsou očekávány s postupnou změnou klimatu. Maxent se s tímto fenoménem vypořádává pomocí tzv. upínání hodnot (*Clamping*), které se k proměnným mimo trénovací rozsah dat chová jako kdyby byly na hraně trénovacích dat (Obr. 23; Phillips, 2017). Tím by se měl omezit výskyt nerealistických vzorů vznikajících v důsledku použití složitých modelů mimo rozsah hodnot, než na kterých byly modely natrénovány. Je pravděpodobně lepší *Clamping* použít, nelze však očekávat o mnoho realističtější výsledky (Smith, 2021).

Nástroj **Parallel** umožňuje aplikaci Wallace využívat paralelně více jader procesoru při počítání modelů. Autor této práce však doporučuje hodnotu nástroje nastavit na *FALSE*, neboť při použití paralelního výpočtu proces vždy skončil chybou upozorňující na nedostatek paměti. v případě menšího vzorku dat, menšího zkoumaného území a méně podrobných rastrů by povolení tohoto nástroje mohlo pomoci zrychlit proces generování modelů, v této práci však bylo výhodnější nechat nastavení *FALSE*, s nímž se sice modely generovaly i více než 20 hodin, vždy ale proces došel s jistým výsledkem do konce.

V rámci tvorby této práce bylo otestováno nad vstupními bioklimatickými proměnnými několik různých nastavení modelů – několik různých kombinací použitých funkcí a hodnot multiplikátorů.

Konečné použité nastavení modelovacích parametrů se opírá o argumenty a doporučení popsané v předcházejícím textu, a je následující:

- algoritmus: maxnet
- vybrané funkce (*Feature Classes*): LQ, H, LQH, LQHP
- vybraný rozsah regularizačních multiplikátorů: 0,5–2,5
- hodnota kroku multiplikátoru: 0,5
- kategorické proměnné: ANO, proměnná LC (Land Cover, krajinný pokryv)
- upínání hodnot (clamping): ANO (TRUE)
- paralelní použití více jader (parallel): NE (FALSE)

Následující nastavení po asi 20 hodinách vygenerovalo celkem 20 různých modelů (4 funkce × 5 multiplikátorů). Ty bylo potřeba vyhodnotit, pro což aplikace Wallace spočítala několik evaluačních statistik.

Přístupy pro hodnocení přesnosti modelu

Wallace v rámci tvorby modelů provádí rozsáhlá hodnocení založená na rozdělení na trénovací a validační data. Wallace uvádí seznam několika běžně používaných hodnotících metrik, včetně **AUC** (Area Under the Curve, plocha pod křivkou), **OR** (Omission Rate, míra vynechání), **CBI** (Continuous Boyce Index, kontinuální Boyceův index) a **ACI** (Akaike Information Criterion, Akaikeho informační kritérium). Neexistuje jediný "nejlepší" způsob hodnocení distribučních modelů, proto se Wallace snaží poskytnout jich několik a nechat uživatele rozhodnout, která kritéria jsou pro něj vyhovující.

AUC (Area Under the Curve) je plocha pod křivkou grafu ROC (Receiver Operating Characteristic). ROC je pravděpodobnostní křivka a AUC představuje stupeň nebo míru separability, posuzuje tedy **diskriminační schopnost modelu** (schopnost určit, zda se na daném místě na základě environmentálních podmínek druh nachází nebo nenachází). **AUC se pohybuje v rozmezí od 0 do 1**, což vyjadřuje, nakolik je model schopen rozlišovat mezi třídami (Narkhede, 2018). Např. AUC = 1 znamená, že model správně zařadil všechna místa výskytu druhu do třídy „zde je výskyt“ a všechna místa pseudo-absence druhu do třídy „zde není výskyt“. Hodnota AUC < 1 (např. 0,8) značí, že model určil některá místa výskytu jako „zde není výskyt“ a některá místa pseudo-absence jako „zde je výskyt“ (model tedy podal falešně negativní/falešně pozitivní informaci). Nejhorší varianta je AUC = 0,5, neboť tato hodnota značí, že model nemá žádnou diskriminační schopnost (schopnost rozlišit jednotlivé třídy). Přehledně princip AUC vysvětluje např. Narkhede (2018).

Velké komplikace pro interpretaci AUC existují v případě, že neexistují skutečná negativní data (nelze říct, že když se v nějakém prostředí druh nevyskytuje, není pro něj prostředí vhodné; prostředí pro něj vhodné být může, pouze se do něj druh ještě nedostal, nejedná se tedy o absenci, ale o pseudo-absenci; Lobo et al. 2008). Model Maxent pracuje na principu „presence-background“ a tedy neobsahuje data o skutečné absenci, z toho důvodu by hodnoty AUC měly být považovány pouze za relativní ukazatele výkonnosti modelu (např. mezi různými nastaveními téhož algoritmu). Další podrobnosti k hodnocení AUC jsou vypsány v záložce **Component Guidance**.

Pět sloupců v tabulce výsledků hodnocení modelů („Results“) se týká AUC:

- **auc.train**: AUC vypočtená s použitím všech lokalit výskytu,
- **auc.val.avg** a **auc.val.sd**: průměr a směrodatná odchylka k počtu validačních AUC (jedna pro každý segment dat),
- **auc.diff.avg** a **auc.diff.sd**: průměr a směrodatná odchylka všech rozdílů mezi k trénovacími a validačními AUC.

Je vhodné vybírat modely s co největší hodnotou auc.val.avg (která určuje diskriminační schopnost modelu) a **naopak s co nejnižší hodnotou auc.diff.avg**, neboť vyšší rozdíl AUC (auc.diff.avg) naznačuje větší přetrénování (overfitting) modelu, protože přetrénované modely fungují na trénovacích datech lépe než na nezávislých testovacích datech (Warren a Seifert, 2011), a tudíž by se jejich použití mělo zabránit.

OR (Omission rate, míra vynechání) je metoda hodnocení schopnosti binárního klasifikátoru (modelu) předpovídat lokality (obvykle neznámé), obvykle po aplikování prahové hodnoty na spojitou předpověď modelu. Aplikací prahu se předpověď stane binární (např. 0 (nevýskyt) a 1 (výskyt)) a **OR se pak rovná procentu testovacích lokalit, které spadly do buněk mřížky s hodnotou 0** (tj. lokalit výskytu druhu, v nichž se druh nachází, ale modelem nejsou určeny jako vhodné pro výskyt druhu). Nízká OR je nezbytnou podmínkou dobrých modelů (Peterson et al. 2011; Phillips et al., 2006). OR 0 znamená, že žádná lokalita nespadá mimo predikci, zatímco OR 1 znamená, že všechny lokality spadají mimo predikci.

Existuje mnoho možných prahových pravidel, ale Wallace nabízí pro vyhodnocení dvě, která se běžně používají: minimální tréninková přítomnost (Minimum Training Presence, MTP) a 10percentilová tréninková přítomnost (10 Percentile Training Presence, 10pct). MTP nastaví prahovou hodnotu na nejnižší míru vhodnosti pro všechny lokality výskytu použité k trénování modelu. Jinými slovy předpokládá, že nejméně vhodný habitat, v němž se trénovací data výskytu druhů ještě vyskytují, je celkově nejmenší hodnota vhodnosti habitatu pro daný druh. 10pct je naopak prahová hodnota, která vynechá (vyloučí) všechny testovací oblasti s vhodností stanoviště nižší, než jsou hodnoty vhodnosti pro nejnižších 10 % trénovacích dat výskytu. Předpokládá, že 10 % trénovacích dat výskytu s nejnižší hodnotou vhodnosti se vyskytuje v regionech, které nejsou reprezentativní pro celkové stanoviště druhu, a proto by měly být vynechány (Morrow, 2019). 10pct vynechává větší oblast než MTP. Hodnota 10pct je tedy přísnější než MTP. Stejně jako u AUC počítá Wallace OR pro každý segment; dělá to tak, že na spojitou předpověď použije práh a zjistí procento validačních dat, která nespádají do výsledné binární předpovědi.

Čtyři sloupce v tabulce výsledků hodnocení modelů („Results“) se týkají OR:

- **or.mtp.avg a or.mtp.sd:** průměrná hodnota a směrodatná odchylka všech testovaných hodnot míry vynechání MTP.
- **or.10p.avg a or.10p.sd:** průměr a směrodatná odchylka všech testovaných 10pct mírek opomenutí

Je vhodné vybírat modely s co nejnižší hodnotou or.10p.avg, která určuje, kolik testovacích lokalit bylo modelem chybně určeno jako nevhodných pro výskyt druhu, ačkoliv se v nich druh vyskytuje (tedy falešně negativní výsledek).

CBI (Continuous Boyce Index, kontinuální Boycův index) hodnotí modely na základě rozdělení rozsahu vhodnosti do b tříd, nikoli pouze do dvou (0 a 1 jako u OR či AUC). Pro každou třídu vypočítává poměr četnosti do této třídy modelem předpovězených lokalit k jejich očekávané četnosti v této třídě – třídy s nízkou vhodností habitatu by měly mít méně predikovaných lokalit než náhodný model a poměr předpovězené/očekávané četnosti by se měl zvyšovat s vyšší vhodností (tedy čím větší vhodnost habitatu, tím větší množství dat výskytu by do ní mělo být zařazeno; Boyce et al., 2002; Hirzel et al., 2006). Hodnoty CBI se pohybují v rozmezí od -1 do 1, přičemž kladné hodnoty označují model, jehož predikce odpovídají reálným datům výskytu, hodnoty kolem 0 značí, že se vytvořený model

příliš neliší od náhodného modelu, a záporné hodnoty ukazují na špatný model (Hirzel et al., 2006).

Dva sloupce v tabulce výsledků hodnocení modelů („Results“) se týkají CBI:

- **cbi.val.avg a cbi.val.sd:** průměr a směrodatná odchylka testovaných CBI.

Na základě výše popsaných vlastností vyplývá, že **je vhodné vybírat modely s co největší hodnotou cbi.val.avg.**

AIC (Akaike Information Criterion, Akaikeho informační kritérium) je metrika hodnocení modelu, která se často používá u technik založených na regresi. Modely s nejnižším AIC jsou mezi kandidátskými modely identifikovány jako optimální. Počítá se tímto způsobem (Burnham a Anderson 2002):

$$AIC = 2k - 2\ln(L)$$

kde k je počet parametrů modelu a $\ln(L)$ je logaritmická pravděpodobnost modelu. Modely s více parametry budou mít větší kladné číslo pro první člen a modely s vyšší pravděpodobností budou mít větší záporné číslo pro druhý člen. Jednodušší modely s menším počtem parametrů a modely s vysokou pravděpodobností proto mají nižší skóre AIC. Platí, že pokud mají dva modely přibližně stejnou pravděpodobnost, ovšem liší se počtem parametrů, bude mít ten s menším počtem parametrů nižší AIC. Tímto způsobem AIC obecně penalizuje složité modely, ale ty složité, které mají navíc vysokou pravděpodobnost, naopak zvyhodňují se záměrem najít nejlepší rovnováhu mezi složitostí modelu a jeho vhodností (Burnham a Anderson, 2002). Wallace počítá pro model Maxent AICc (korigovaný na konečnou velikost vzorku) podle metodiky uvedené v práci Warren a Seifert (2011). AIC se počítá pouze pro celý model, a nezohledňuje tedy segmentaci na trénovací a testovací data.

Čtyři sloupce v tabulce výsledků hodnocení modelů („Results“) se týkají AIC:

- **ncoef:** počet parametrů (koeficientů) v modelu (pro Maxent to zahrnuje i funkce aplikované na proměnné (např. čtyři hinge funkce aplikované na Bio1 znamenají čtyři parametry, kvadratická funkce pro Bio11 znamená dva parametry atd.),
- **AICc,**
- **delta.AICc:** absolutní rozdíl mezi nejnižším AICc a každým AICc,
- **w.AIC:** váha AIC vypočtená jako průměrná relativní pravděpodobnost modelu ($\exp(-0,5 * \text{delta.AICc})$) ve všech modelech.

Je vhodné vybírat modely s co nejnižší hodnotou AICc a delta.AICc, neboť tyto modely mají nejlepší poměr mezi složitostí a silou predikce.

Výběr optimálního modelu

Na základě evaluační tabulky, kterou Maxent poskytl při vygenerování modelů, byly jednotlivé modely zhodnoceny a vybrány ty s neoptimálnějšími hodnotami hodnotících metrik. Autor při hodnocení modelů vycházel z informací z návodu k použití modelovacího komponentu v aplikaci Wallace a z dostupné literatury k jednotlivým metrikám (AUC, OR, CBI, AIC). Pomocí vlastní evaluační metodiky byly vybrány neoptimálnější modely, tedy ty s obecně nejlepším evaluačním skóre pěti sledovaných metrik. Těmito metrikami (podrobnosti k metrikám jsou vypsány v předchozí sekci) jsou:

- **auc.val.avg** – čím větší hodnota, tím lepší,
- **auc.diff.avg** – čím menší hodnota, tím lepší,
- **or.10p.avg** – čím menší hodnota, tím lepší,
- **cbi.val.avg** – čím větší hodnota, tím lepší,
- **AICc** – čím menší hodnota, tím lepší.

Vlastní metodika spočívá ve standardizaci rozpětím, tedy tzv. *min-max normalizaci* podle nejlepší a nejhorší hodnoty dané metriky a určení nejlepšího evaluačního skóre napříč všemi metrikami. Hodnoty musely být standardizovány, neboť jednotlivé proměnné mají různý rozsah a nejsou vzájemně porovnatelné.

Všechny modely byly ohodnoceny pro každou metriku zvlášť, přičemž hodnoty H všech metrik byly standardizovány na hodnoty H_s tak, aby pro nejlepší standardizovanou hodnotu H_{smax} dané metriky platilo, že $H_{smax} = 1$ a pro nejhorší standardizovanou hodnotu H_{smin} platilo $H_{smin} = 0$. v případě metriky, u níž platí, že **nejvyšší hodnota je nejlepší hodnota**, se standardizovaná hodnota H_s vypočítá následovně:

$$H_s = \frac{H - \min(H_m)}{\max(H_m) - \min(H_m)}$$

kde H_m jsou hodnoty dané metriky v rámci všech modelů, $\min(H_m)$ tedy značí nejmenší hodnotu a $\max(H_m)$ značí maximální hodnotu konkrétní metriky napříč všemi modely. Pro metriky, u nichž platí, že **nejnižší hodnota je nejlepší hodnota**, se standardizovaná hodnota H_s vypočítá odečtením předchozího vztahu od 1:

$$H_s = 1 - \frac{H - \min(H_m)}{\max(H_m) - \min(H_m)}$$

Standardizované hodnoty se v rámci jedné metriky pohybují od 0 do 1, přičemž 1 značí nejlepší hodnotu dané metriky napříč všemi modely a 0 nejhorší hodnotu. Standardizované hodnoty jsou poté sečteny za vzniku porovnatelného celkového skóre pro daný model, načež platí, že čím větší skóre, tím obecně ve všech metrikách neoptimálnější model.

Standardizované hodnoty a celkové skóre modelů vyšlo po zpracování následovně:

Tabulka 1 Původní a standardizované hodnoty zkoumaných metrik a výsledné skóre modelů. Barevné označení modelů rozlišuje jednotlivé hodnoty regularizačních multiplikátorů.

Feature Class	RM	auc.diff.avg	Hs(auc.diff.avg)	auc.val.avg	Hs(auc.val.avg)	cbi.val.avg	Hs(cbi.val.avg)	or.10p.avg	Hs(or.10p.avg)	AICc	Hs(AICc)	Skóre
LQHP	0,5	0,083	0,741	0,804	1,000	0,956	0,924	0,125	0,686	74 139,201	1,000	4,350
H	0,5	0,067	0,892	0,798	0,867	0,958	0,945	0,123	0,717	74 171,929	0,965	4,386
LQH	0,5	0,067	0,894	0,798	0,859	0,957	0,940	0,124	0,693	74 174,524	0,962	4,349
LQ	0,5	0,160	0,000	0,761	0,000	0,855	0,000	0,162	0,000	74 838,505	0,251	0,251
LQHP	1	0,068	0,876	0,801	0,930	0,964	1,000	0,116	0,843	74 238,368	0,894	4,543
H	1	0,056	0,993	0,798	0,874	0,947	0,843	0,108	1,000	74 222,534	0,911	4,620
LQH	1	0,056	1,000	0,796	0,827	0,950	0,873	0,110	0,961	74 231,088	0,902	4,563
LQ	1	0,160	0,001	0,765	0,094	0,873	0,159	0,146	0,291	74 891,695	0,194	0,740
H	1,5	0,067	0,890	0,790	0,674	0,934	0,728	0,119	0,787	74 305,201	0,822	3,902
LQHP	1,5	0,068	0,882	0,789	0,659	0,936	0,740	0,125	0,677	74 328,038	0,798	3,755
LQH	1,5	0,062	0,941	0,787	0,598	0,930	0,684	0,119	0,788	74 325,026	0,801	3,812
LQ	1,5	0,159	0,003	0,766	0,110	0,874	0,173	0,142	0,378	74 969,679	0,111	0,774
H	2	0,078	0,780	0,786	0,580	0,929	0,682	0,128	0,630	74 375,371	0,747	3,419
LQHP	2	0,080	0,770	0,785	0,561	0,917	0,569	0,126	0,669	74 392,683	0,729	3,298
LQH	2	0,071	0,850	0,783	0,524	0,919	0,585	0,121	0,756	74 384,221	0,738	3,452
LQ	2	0,158	0,016	0,767	0,139	0,875	0,182	0,134	0,512	75 033,376	0,043	0,892
LQH	2,5	0,083	0,740	0,783	0,519	0,907	0,472	0,127	0,653	74 420,838	0,698	3,083
H	2,5	0,088	0,686	0,783	0,504	0,915	0,548	0,134	0,512	74 414,210	0,706	2,956
LQHP	2,5	0,092	0,654	0,782	0,480	0,901	0,419	0,133	0,543	74 439,188	0,679	2,775
LQ	2,5	0,154	0,056	0,769	0,179	0,878	0,205	0,130	0,583	75 073,153	0,000	1,022

Z tabulky je patrné, že modely používající funkci (*Feature Class*) LQ mají obecně ve všech případech použitého multiplikátoru (*RM*) nejhorší výsledky. Nejlepších výsledků je naopak obecně dosahováno při použití $RM = 1$, přičemž model s nejlepším skóre (a tedy podle metrik neoptimálnější model) je model H1 (použití funkce *hinge* a $RM = 1$).

Jak již bylo uvedeno výše v sekci *Vytvoření modelu*, nízké hodnoty multiplikátoru mohou vést k přetrénování modelu. Proto bylo použito celkem 5 modelů pro aplikaci na data z České republiky, přičemž z každého souboru hodnot multiplikátorů byl vybrán ten nejpřesnější. Jedná se o modely **H0,5**, **H1**, **H1,5**, **LQH2**, **LQH2,5**. Náhled na vygenerovaná data je v následující podkapitole.

4.4.8 Vizualizace výsledků modelu

Pomocí následujícího komponentu lze vizualizovat predikovanou vhodnost habitatu jednotlivými modely. Model vytvoří datový náhled na studovanou oblast, a v ní na barevné stupnici od 0–1 zobrazí predikovanou vhodnost habitatu.

Ještě před začátkem vizualizace byla provedena změna ve zdrojovém kódu modulu *Map prediction*. Obzvláště při práci s velkými daty, jaké byly použity v této diplomové práci, má totiž aplikace Wallace (resp. funkce v jazyku R) problém vyexportovat rastr větší než cca 4 MB. Tuto skutečnost program RStudio zahlásí chybovou hláškou a celá aplikace Wallace spadne. Problém je ve funkci *addRasterImage* v rámci R knihovny *raster*. Tato funkce totiž není uzpůsobena na zpracování velkých rastrů, bohužel je tato funkce – poměrně nepochopitelně – implementována v modulu *Map Prediction*, a tak je ve výchozím stavu velikost exportovaného rastru poměrně omezená.

Aby program bez problému fungoval, byla funkce *AddRasterImage* nahrazena za funkci *addGeoRaster*, která je, už podle názvu, přímo uzpůsobena pro práci s velkými rastry používanými např. v GIS. Aby tato funkce fungovala, byly do RStudia doinstalovány dvě knihovny, a to *leafem* a *stars*, které musí být v R před generováním rastru spuštěny. To se vyřeší přidáním *library(leafem)* a *library(stars)* do R skriptu před generováním rastrů

```

256 # map model prediction raster
257 library(stars)
258 library(leafem)
259 map %>%
260   addCircleMarkers(data = occs(), lat = ~latitude, lng = ~longitude,
261                   radius = 5, color = 'red', fill = TRUE, fillColor = 'red',
262                   fillOpacity = 0.2, weight = 2, popup = ~pop) %>%
263   addGeoRaster(mapPred(), colors = rasPal, opacity = 0.7,
264               group = 'vis', layerId = 'mapPred', method = "ngb") %>%
265   # add background polygon(s)
266   mapBgPolys(bgShpXY())

```

Obr. 24 Úprava funkce *addRasterImage* na *addGeoRaster*.

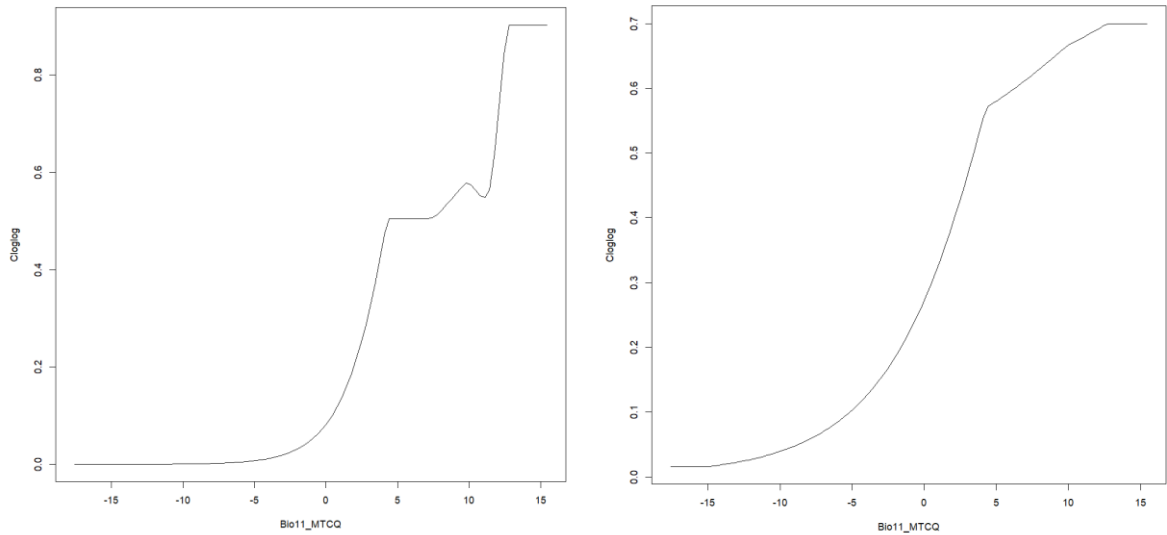
mapových predikcí (... \R\win-library\4.2\wallace\shiny\modules\vis_mapPreds.R):

Jediný problém je, že na funkci *addGeoRaster* nefunguje stejné nastavení barev jako na *addRasterImage*, proto se barvy zobrazí pouze na černobílé stupnici.

Aplikace Wallace umožňuje zobrazit výsledky několika způsoby (4 moduly):

- **Mapa predikcí** (*Map Prediction*) – mapový náhled vhodnosti habitatu ve studované oblasti (Obr. 26), jsou k dispozici tři tzv. *Precision outputs: raw, logistics a cloglog*; bližší informace k těmto nastavením jsou v záložce *Module Guidance*, pro tuto práci bylo ponecháno výchozí nastavení *cloglog*,
- **Evaluační grafy** (*Maxent Evaluation Plots*) – grafy znázorňující přesnosti modelů v jednotlivých metrikách (AUC, OR, AIC) pro všechny hodnoty *RM* a funkce (*Feature Classes*),
- **Response Curves** – grafické znázornění vztahů mezi vhodností habitatu a hodnotami jednotlivých predikčních proměnných; modely využívající přístup Presence-background jsou obvykle určitou variantou aditivního modelu (např. odhadovaná vhodnost v modelu Maxent představuje součet příspěvků několika proměnných; Phillips et al., 2006); vztah mezi vhodností a hodnotami jednotlivých proměnných prostředí lze tedy vyjádřit v podobě grafů označovaných jako *Response Curves* (graf závislosti, Obr. 25); křivky jsou s přibývajícím *RM* hodnotou více vyhlazené,
- **BIOCLIM Envelope Plots** – vyžaduje BIOCLIM model, v této práci neřešeno.

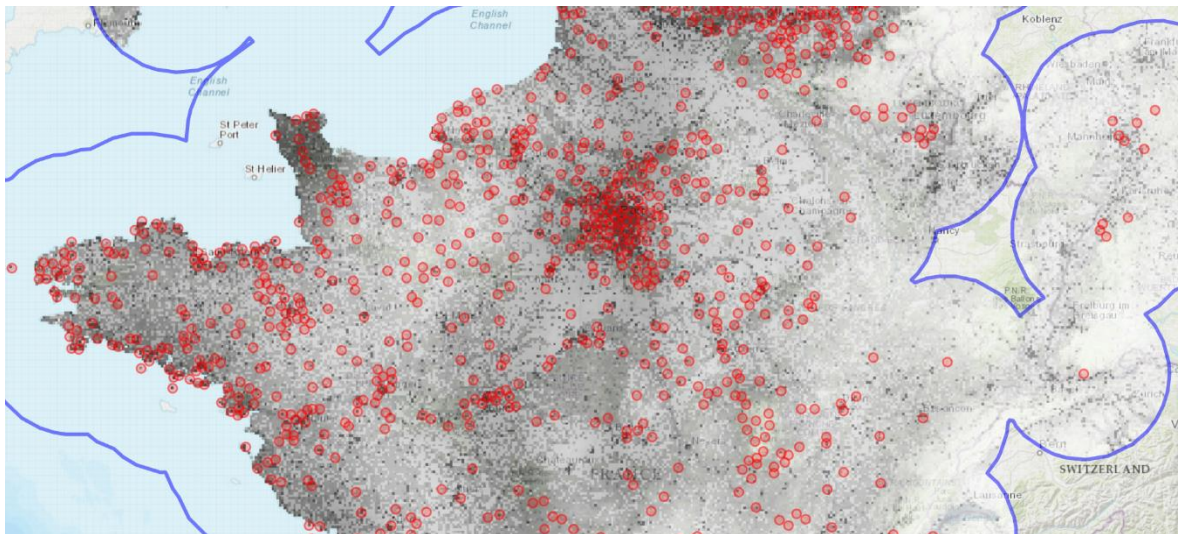
V této práci byl použit modul *Map Prediction*, který se musí povinně spustit před přesouváním modelu na další území. Model, pro nějž se mapová predikce vytvoří (a který následně bude přenesen na další oblasti) se vybere v menu nad mapou.



Obr. 26 Grafy závislosti (*Response Curves*) vhodnosti habitatu na proměnné Bio 11 pro model H0,5 (vlevo) a LQH2,5 (vpravo). v grafu je patrný vliv různých funkcí a různé hodnoty RM na tvar křivky.

Na ose x je hodnota proměnné, na ose y míra vhodnosti.

(zdroj: výstup z aplikace Wallace)



Obr. 25 Náhled na vygenerovanou predikční mapu pro rozsah studovaného území. Čím tmavší barva, tím větší vhodnost habitatu pro šíření sršně *Vespa velutina*.

4.4.9 Přesunutí modelu

Důležitou součástí modelování je možnost přenosu modelu do jiné oblasti nebo časového období, než které byly použity k jeho sestavení (např. pro studie invazivních druhů nebo dopady klimatické změny). Aplikace Wallace v současné době umožňuje přenos do:

- **nové oblasti** (*Transfer to New Extent*) – umožňuje natrénovaný model aplikovat na jiné území, podmínkou je, že sané území musí být v rozsahu predikčních rastrů, které byly nahrány do modelování v komponentu 2 (kapitola 4.4.2),
- **nového časového období** (*Transfer to New Time*) – přenos do podmínek v roce 2050 či 2070, tento modul funguje pouze pokud byly pro modelování v komponentu 2 vybrány bioklimatické proměnné z WorldClim, jinak aplikace hlásí chybu,
- **uživatelé specifikovaného prostředí** (*Transfer to User Environments*) – umožňuje uživateli nahrát vlastní data ve formátu TIF, do nichž bude model přenesen (například vlastní rastry pro budoucí normály apod.), podmínkou však je, že se musí jednat o stejné predikční proměnné, které vstupovaly do modelu (soubory musí mít i stejné jméno), pouze např. pro jiný normál (predikce do budoucnosti),

Navíc model umožňuje spočítat **podobnost prostředí v nové oblasti s prostředím, v němž byl model sestavován** (poslední modul *Calculate Environmental Similarity*).

V této práci byly použity dva moduly – přenos do nové oblasti (pro zjištění aktuální vhodnosti habitatu v ČR) a **přenos do uživatelem specifikovaného prostředí**, kam byly nahrány bioklimatické rastry z normálu 2041–2060 a rastr krajinného pokryvu pro predikci vhodnosti habitatu do budoucnosti. Úskalí použití krajinného pokryvu v predikci do budoucna je okomentováno v kapitole Diskuze.

Pro každý vybraný model (viz sekce *Výběr optimálního modelu* v podkapitole 4.4.7) byla vygenerována vizualizace vhodnosti habitatu pro současný a budoucí (2041–2060) klimatický normál – bylo staženo celkem 10 datových náhledů, resp. rastrů s predikovanou mírou vhodnosti (5 pro současný a 5 pro budoucí normál). Po vizuálním porovnání rastrů v aplikaci ArcGIS Pro bylo nakonec vybráno 6 vizualizací, které reprezentují různé scénáře. Vznikly 3 scénáře pro současnost a 3 scénáře pro predikci do budoucna. Vybrány byly modely podle velikosti regularizačního multiplikátoru, a to **H0,5, H1,5 a LQH2,5**.

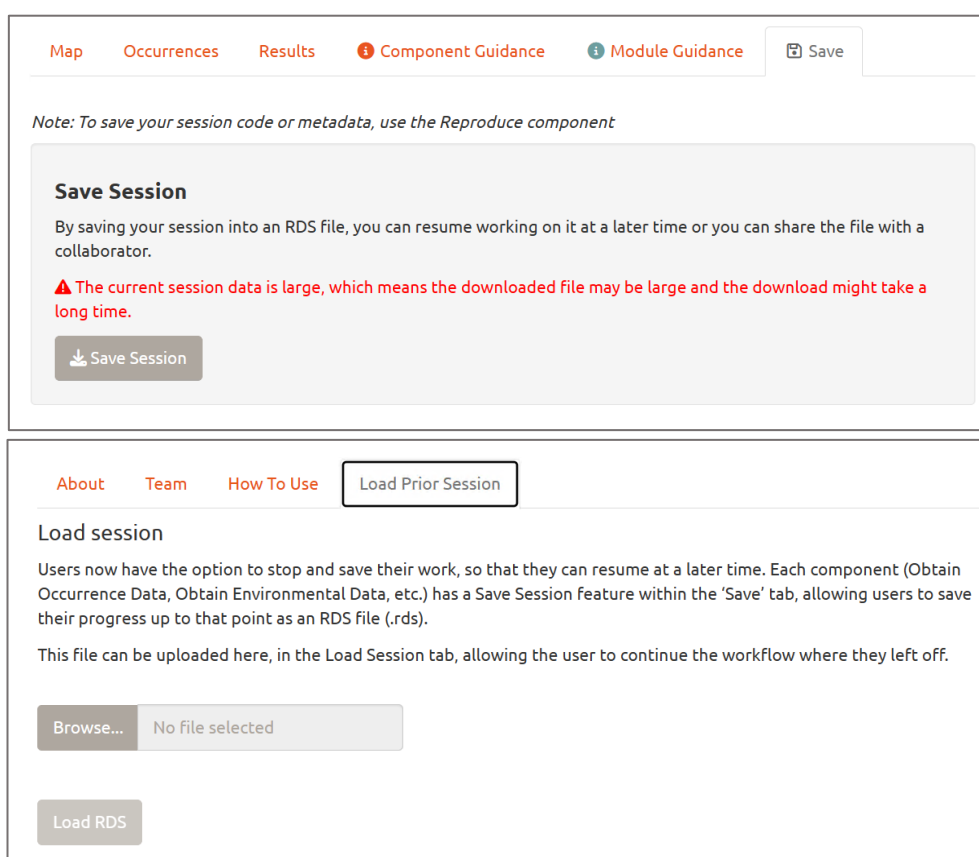
Jak již bylo řečeno v podkapitole 4.4.7 v sekci *Vytvoření modelu*, čím menší je velikost RM, tím větší je napasování modelu na trénovací data. Proto H0,5 představuje nejvíce na data napasovaný model, který dělá nejméně „ústupků“ při posuzování vlivu proměnných na míru vhodnosti, a jedná se tedy o nejpřísnější scénář. Model LQH2,5 je naopak nejvíce shlazený (toleruje hodnoty více odlišné od trénovacích hodnot) a nejméně přísný scénář. Nejpřísnější scénář (H0,5) je ale zároveň pro Českou republiku nejpozitivnější, neboť takový model nachází jen minimum shody mezi trénovacími lokalitami a prostředím ČR, nejméně přísný scénář (LQH2,5) nachází v ČR shody větší, neboť toleruje i hodnoty vzdálenější od trénovacích hodnot. Výsledné vizualizace jsou popsány v kapitole Výsledky.

4.4.10 Reprodukce modelu

Modul reprodukce modelu umožňuje stáhnout kód projektu (celého postupu modelování) ve formátu RMD, který je možné spustit v RStudio, případně ve formátech HTML, PDF, DOC. Dále umožňuje stahovat metadata celého projektu a reference k aplikaci Wallace.

Pomocí tohoto modulu byly staženy kódy projektu ve formátu RMD, tento soubor je součástí volných příloh jako Příloha 4 diplomové práce.

V průběhu celého postupu modelování je možné průběžně ukládat data z aplikace Wallace pomocí záložky *Save* a možnosti *Save session*. Tato možnost uloží poměrně velký soubor RDS, který je možné později znovu nahrát do aplikace Wallace přes záložku *Load Prior Session* na úvodní obrazovce aplikace a pokračovat v práci. Soubor RDS je uložen jako volná Příloha 5 této práce.



Obr. 27 Možnost uložení projektu do formátu RDS a nahrání tohoto souboru do aplikace Wallace.

5 VÝSLEDKY

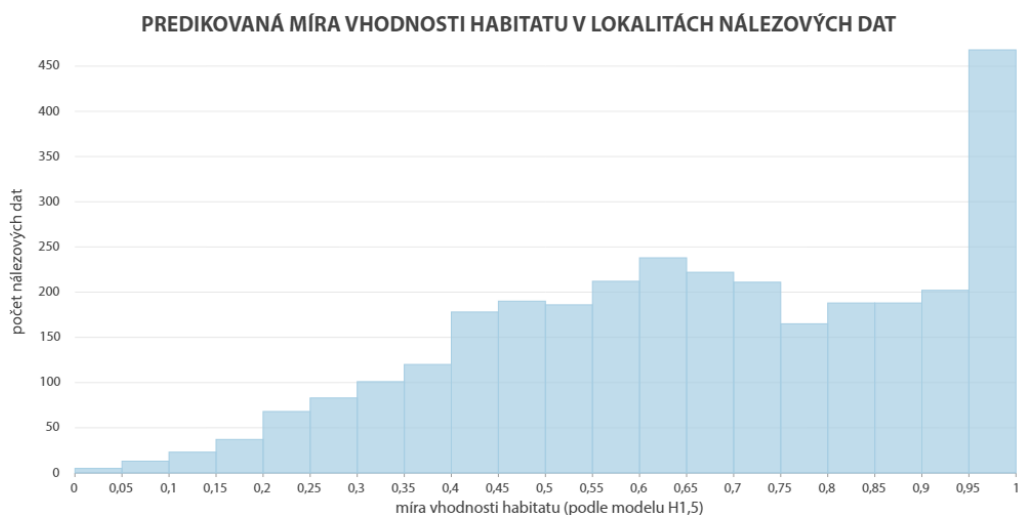
Hlavním výsledkem práce je vytvořený predikční model pro predikci výskytu sršně asijské *Vespa velutina* v České republice v podmínkách současného klimatu a v podmínkách budoucích (normál 2041–2060). Model byl sestavován v aplikaci Wallace na podkladech autorem vytvořených predikčních bioklimatických proměnných a bodových dat výskytu sršně asijské, a je vypočítán algoritmem strojového učení Maxent (celý postup modelování je v kapitole 4.4). Model byl vyexportován ve formátu RMD, který je možné jednoduše nahrát a spustit v programu RStudio, a ve formátu RDS, jenž je možné otevřít v aplikaci Wallace a pokračovat v modelovacím procesu (kapitola 4.4.10).

Z povahy modelovacího přístupu algoritmu Maxent je výsledkem predikční model založený na **určování míry vhodnosti habitatu** pro sršeň asijskou, nikoli model vyjadřující pravděpodobnost výskytu na určitém místě. Predikovaná míra vhodnosti habitatu v České republice se odvíjí od stanovištních a klimatických podmínek, v nichž byla sršeň asijská *Vespa velutina* pozorována jinde v Evropě. **Výběr predikčních proměnných** vychází z rešerše (kapitola 2.2) a z autorem provedeného statistického testování (kapitola 4.3.4). **Výsledný soubor predikčních proměnných zahrnuje následující:**

- **Bio 2 Průměrný roční rozsah denních teplot**
- **Bio 3 Izotermalita**
- **Bio 4 Teplotní sezónnost**
- **Bio 10 Průměrná teplota nejteplejšího čtvrtletí**
- **Bio 11 Průměrná teplota nejchladnějšího čtvrtletí**
- **Bio 15 Srážková sezónnost**
- **Bio 18 Úhrn srážek v nejteplejším čtvrtletí**
- **Bio 19 Úhrn srážek v nejchladnějším čtvrtletí**
- **kategorická proměnná krajinného pokryvu (Land Cover)**

Na těchto datech byl model natrénován a poté aplikován na oblast České republiky. Cílem bylo zjistit, zda jsou už nyní v České republice podmínky pro přežití sršně asijské, a jak moc se zvýší vhodnost těchto podmínek pro přežití sršně asijské v nepříliš vzdálené budoucnosti (v klimatickém normálu 2041–2060).

Výsledné vizualizace vychází ze tří modelů (scénářů; podkapitola 4.4.9), které se liší svými parametry (kapitola 4.4.7, sekce *Výběr optimálního modelu*). **Nejpřísnější model H0,5** (nejvíce upínající se na podmínky v trénovacích datech) **stanovuje maximální vhodnost habitatu na území České republiky na 0,067 pro současný a 0,1 pro budoucí klimatický normál, H1,5 0,091 pro současný a 0,181 pro budoucí klimatický normál a nejméně přísný model LQH2,5 namodeloval maximální vhodnost habitatu v ČR na 0,126 v současném a 0,251 v budoucím normálu.** Minimální hodnota se u všech modelů blíží 0. Ač se výsledné predikované maximální hodnoty poměrně liší na základě použité funkce a regularizačního multiplikátoru, **všechny modely vyhodnotily téměř dvojnásobný nárůst míry vhodnosti habitatu mezi současným a budoucím normálem**, což je informace, která by při interpretaci určitě neměla být opomenuta, ač výsledné hodnoty nejsou nijak extrémně veliké. Mimoto vyhodnocení dat vhodnosti z celé Evropy ukazuje, že se sršně nacházely i v oblastech, které modely vyhodnotily jako téměř nevhodné (Graf 4). Na Mallorce bylo první hnízdo nalezeno v oblasti, jejíž míru vhodnosti vypočetl model sestavený podle Herrera et al. (2022) na hodnotu 0,239. v celoevropském kontextu poskytnutým touto diplomovou prací je však Mallorca územím s mírou vhodnosti habitatu leckdy atakujícím hodnotu 1.

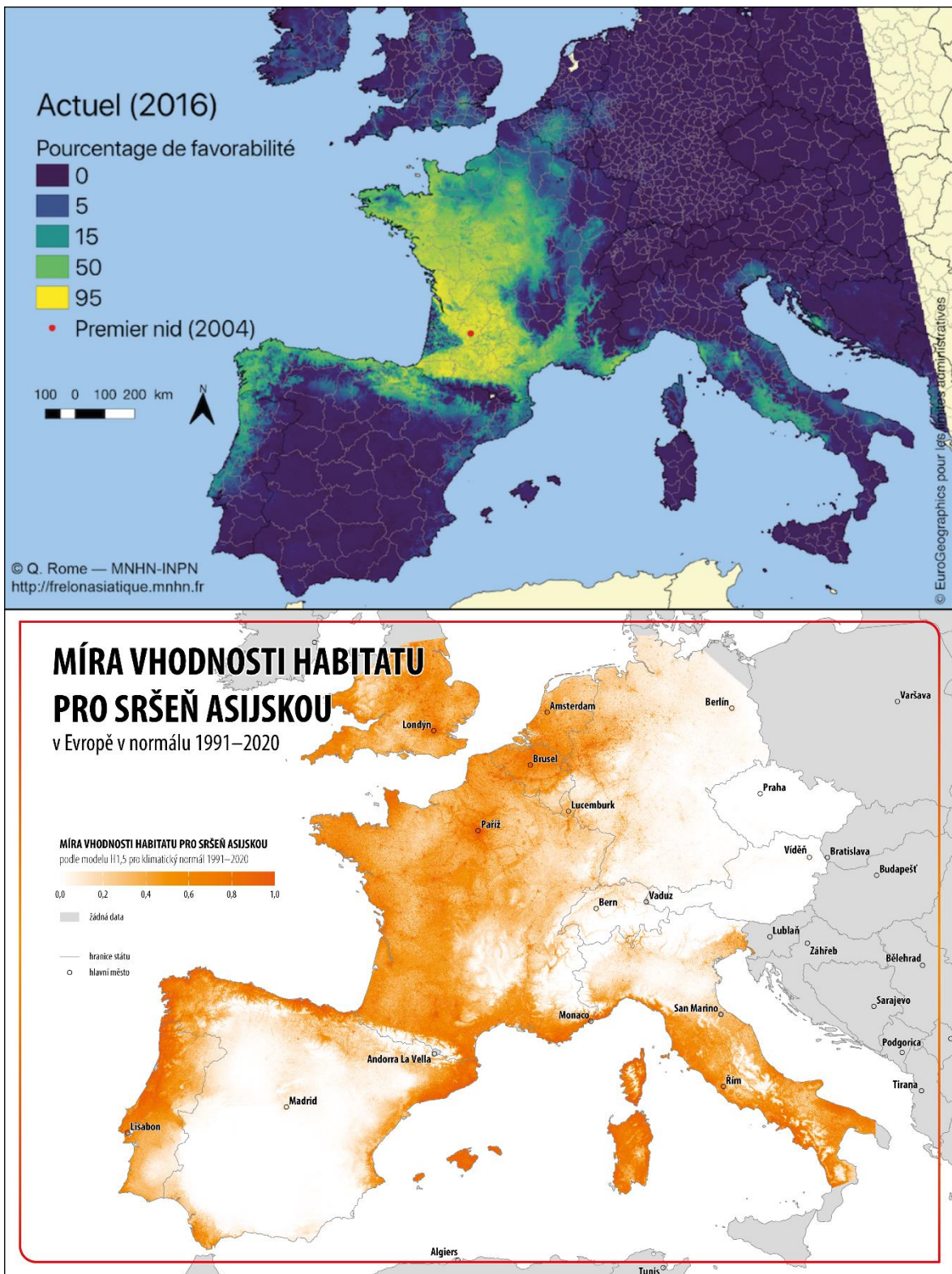


Graf 4 Predikovaná míra vhodnosti habitatu v lokalitách nálezových dat. Asi 90 nálezů bylo zaznamenáno i v lokalitách s vhodností habitatu nižší než 0,2, což naznačuje, že by se sršeň asijská mohla uchytit i v podmínkách České republiky.
(zdroj: výstup z programu ArcGIS Pro)

Výběr výsledného modelu záležel na rozhodnutí autora této práce. Vzhledem k tomu, že sršeň asijská *Vespa velutina* je ve svém šíření Evropou úspěšný druh schopný se dobře adaptovat na podmínky okolí (Diaz et al., 2023), **pro hlavní vizualizace do brožury a letáku byl zvolen střední scénář H1,5**, který není tak přísně upnutý na trénovací data (je benevolentnější v určování míry vhodnosti habitatu v závislosti na jednotlivých proměnných) a zároveň kombinuje dobrý predikční výkon (viz Tabulka 1 v podkapitole 4.4.7).

Pro dokreslení situace byla vytvořena i mapa míry vhodnosti napříč celou Evropou (Obr. 28) a vizuálně porovnána s mapou míry vhodnosti habitatu, která byla vytvořena na základě výzkumu rozšíření sršně ve Francii (Fournier et al., 2017). z výzkumu provedeného v této práci vyplývá větší vhodnost habitatu v Belgii a Německu, a to z toho důvodu, že byl model sestavován i z nálezových dat z těchto lokalit. Výzkum ve Francii s nimi nepočítal. Oba modely se ale shodnou na vysoké míře vhodnosti v přímořských oblastech, a naopak malé míře vhodnosti habitatu ve vnitrozemí.

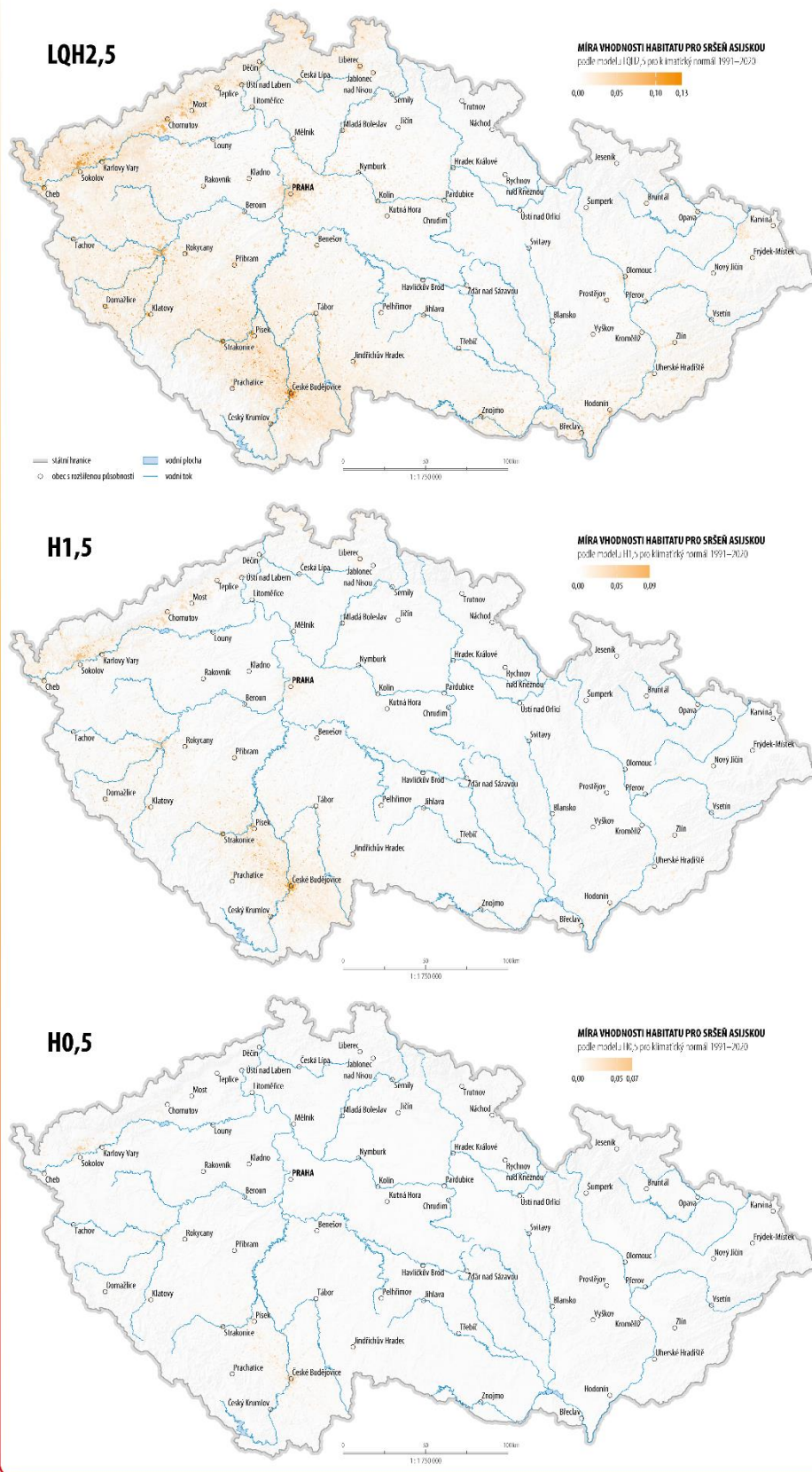
Na mapě míry vhodnosti habitatu pro sršeň asijskou v České republice (Obr. 29, 30) je dobře patrná **zvýšená míra vhodnosti na západě České republiky, směrem k východu vhodnost klesá**. Po vizuálním porovnání vstupních proměnných a grafů závislosti bylo zjištěno, že tento fenomén je nejpravděpodobněji ovlivněn mimo jiné stoupající srážkovou sezónností směrem na východ republiky. Srážková sezónnost je například na území Francie poměrně nízká, tedy nejsou v průběhu roku tak velké rozdíly v úhrnu srážek, což sršním asijským zřejmě vyhovuje. Zároveň je v ČR výraznější teplotní sezónnost (a s ní korelující kontinentalita), která narůstá směrem od moře – sršeň pravděpodobně preferují menší kontinentalitu, a tedy menší rozdíly maximální a minimální teploty v průběhu roku (větší teplotní stálost). Ze stejného důvodu je na předchozí mapě Evropy vidět nevhodnost habitatu např. v centrálním Španělsku – je zde poměrně výrazná kontinentalita, zároveň je střed Španělska chudý na srážky, což sršním nevyhovuje.



Obr. 28 Predikovaná míra vhodnosti habitatu pro sršeň asijskou pro současný klimatický normál 1991–2020 predikovaný pomocí modelu H1,5 (střední scénář, dole). Porovnání s výzkumem z Francie (nahore; Fournier et al., 2017). v mapě jsou viditelné rozdíly v oblasti Beneluxu, neboť model z francouzského výzkumu nepočítal s daty z těchto zemí. v případě území Francie, Itálie a Španělska model z této práce tomu francouzskému odpovídá, pouze je rozšířen o větší množství pozorování, a tedy i větší míru vhodnosti i v jiných lokalitách.

MÍRA VHODNOSTI HABITATU PRO SRŠĚŇ ASIJSKOU

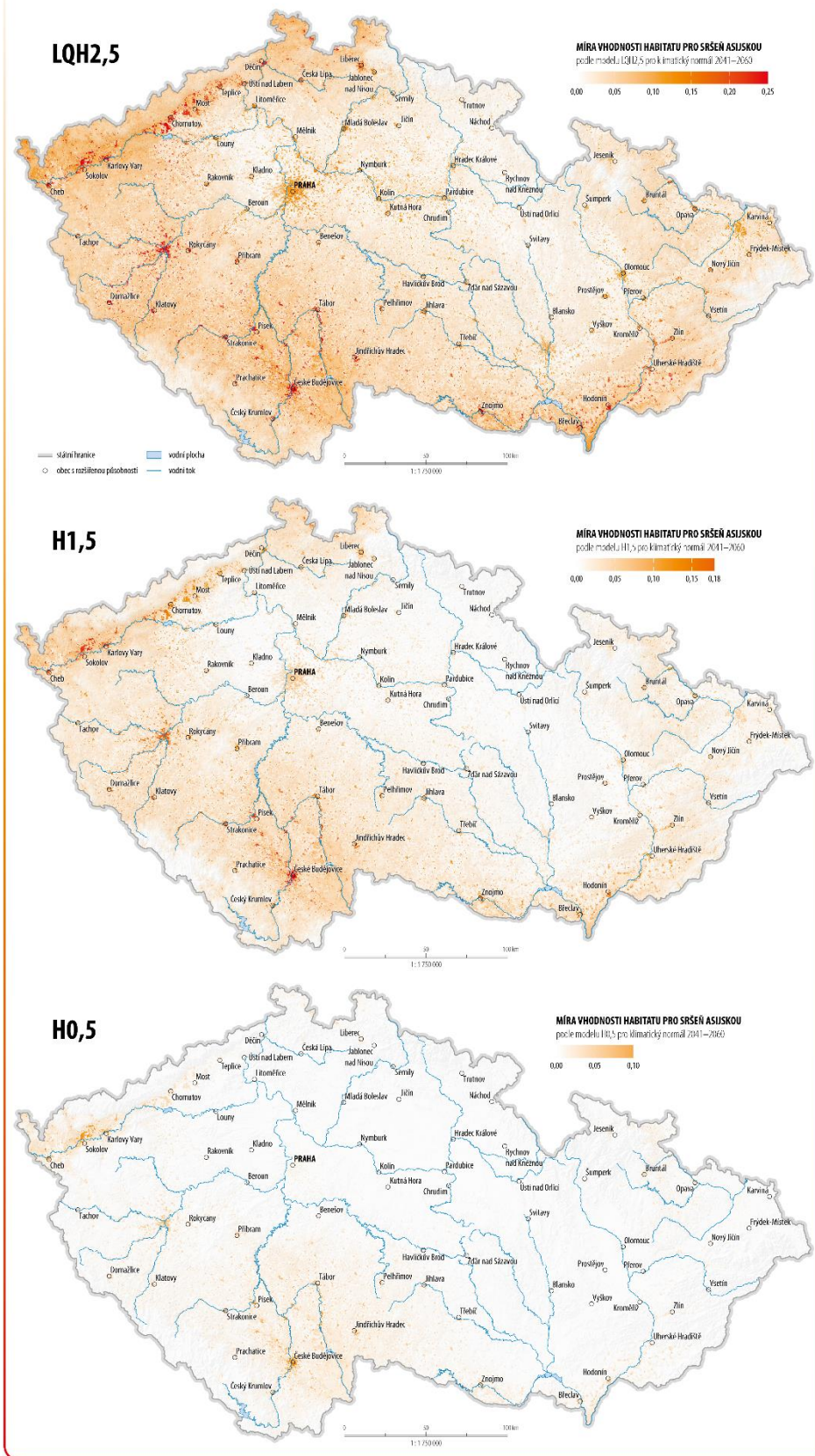
v České republice v normálu 1991–2020



Obr. 29 Srovnání tří různých modelů při predikci míry vhodnosti habitatu pro sršěň asijskou pro současný klimatický normál 1991–2020. Mapy jsou samostatně jako vázané přílohy 6–8. (zdroj: autor)

MÍRA VHODNOSTI HABITATU PRO SRŠĚŇ ASIJSKOU

v České republice v normálu 2041–2060



Obr. 30 Srovnání tří různých modelů při predikci míry vhodnosti habitatu pro sršěň asijskou pro budoucí klimatický normál 2041–2060. Mapy jsou samostatně jako vázané přílohy 9–11.(zdroj: autor)

V České republice je patrný výrazný nárůst vhodnosti v urbánních oblastech, což potvrzují i výzkumy dokládající velkou závislost sršně asijských na urbánním prostředí – například v Jižní Koreji žije až 70 % kolonií sršně asijské *Vespa velutina* v městském prostředí (Diaz et al., 2023), v roce 2010 počet odchytnutých jedinců v urbánních oblastech (městské parky, centra měst) převyšoval odchyt všech ostatních druhů sršní, V. v. tvořila 170 z 245 odchytnutých jedinců (Choi et al., 2012).

V normálu 2041–2060 se v případě použitého modelu H1,5 zvýšila maximální vhodnost míry habitatu oproti současnému normálu téměř dvojnásobně: z maximální vhodnosti 0,091 na 0,181. Může za to nárůst teploty vzduchu, potírání efektu teplotní (kontinentality) i srážkové sezónnosti a další faktory klimatické změny. Například minimální hodnota teplotní sezónnosti (určující roční rozsah teplot) klesla více než dvakrát, což znamená, že se budou postupně snižovat teplotní výkyvy mezi létem a zimou, což sršním vyhovuje. Stejně tak srážková sezónnost mezi dvěma zkoumanými normály klesla místy i o 50 %, což značí mnohem větší vyrovnanost srážkových úhrnů v průběhu roku a tím pádem další faktor, který přidává rozšiřování sršně asijské v ČR.

Podmínky v České republice by mohly v budoucnu postačovat tomu, aby se zde sršeň usadila. Jedná se zejména o oblasti na západě země, v jižních Čechách a na jihu Moravy. Některé oblasti se ale zdají být nevhodné pro šíření sršně asijské i v budoucím normálu, jako například Krkonoše, Šumava, Jeseníky. Jedná se zejména o horské oblasti, kde jsou poměrně nízké sezónní teploty v porovnání se zbytkem České republiky. Obezřetnost je tedy na místě, nicméně podle modelu v této práci není nutné očekávat kalamitu sršně asijských na území České republiky. To vše však ukáže až další šíření sršně *Vespa velutina* a její schopnost adaptace na nové podmínky.

Dalším, dílčím výsledkem práce je **soubor predikčních proměnných** (celkem 19 rastrů; kapitola 4.3). Ty byly vytvořeny pro velkou část západní a střední Evropy a jsou v poměrně podrobném rozlišení 250×250 m a jsou sjednocené z hlediska prostorových vlastností (stejně rozlišení, prostorový rozsah i počet pixelů). Tyto vygenerované rastry tvoří poměrně komplexní soubor 15 bioklimatických a 4 ostatních stanovištních podmínek a mohou být vstupem pro další podobné analýzy rozšiřování druhů.

Výsledkem, který shrnuje informace o sršních a jednotlivé výstupy modelů prezentuje uživatel, je finální **brožura** popisující nejvýraznější charakteristiky sršně a jejího habitatu, popisuje, jak ji rozeznat od v ČR běžné sršně obecné (*Vespa crabro*), poskytuje informace, jak se jí bránit, kde hlásit její pozorování apod. Spolu s **informačním letákem** zobrazujícím nejzákladnější informace o sršních asijských, mapu vhodnosti habitatu v ČR a informace, kde hlásit případné nálezy sršně asijských, jde o hlavní vizualizační výstupy této diplomové práce.

Během práce vznikl jako vedlejší produkt poměrně podrobný a velkou část hlavních informací prezentující český **návod na práci v aplikaci Wallace** (kapitola 4.4), který postupně prochází jednotlivé komponenty a u každého vysvětluje jeho hlavní funkcionalitu včetně popisu jednotlivých modulů, jenž je možné v rámci komponentu použít, a jejich možných nastavení.

6 DISKUZE

V průběhu řešení práce se objevila řada problému a nedokonalostí dat a postupů, které musel autor této práce řešit a které ovlivnily postup zpracování.

Klimatická data

Na samotném začátku stojí za zmínku prvotní generování klimatických dat programem ClimateEU. Do programu vstupovalo CSV obsahující bodovou vrstvu, která měla v atributech souřadnice ve WGS84 a ETRS89 a nadmořskou výšku, na jejichž základě přiřazoval program jednotlivým lokalitám klimatické hodnoty. Při generování hodnot klimatických dat byla vždy vygenerováno výsledné CSV doplněné o roční/sezónní/měsíční proměnné, ale pouze pro prvních cca 24 mil. z celkových 36 mil. bodů, poté program zahlásil konec procesu a dále již nepokračoval. Název tohoto prvního souboru byl doplněný o popisek „cast1“. Zbytek dat (cca 12 mil. bodů) muselo být vygenerováno zvlášť (nazváno jako „cast2“) a poté v programu ArcGIS Pro pomocí nástroje *Append* připojeno k první části. Generování výsledného souboru, z něž byly generovány klimatické rastry, tedy probíhalo složitějším procesem, na kvalitu dat ale toto generování „nadvakrát“ a následné spojení nemá žádný vliv.

Modelovací nástroje a algoritmy

Dalším bodem, s nímž se musel autor vypořádat, byla volba modelovacího nástroje a přístupu. Autor práce si je vědom několika různých modelovacích přístupů a SDM (Species Distribution Modelling) softwarů, které mohly být použity. Finální výběr aplikace Maxent je ovlivněna několika studii uvedenými v rešerši, které vyzdvihují dobrou predikční schopnost tohoto machine-learning nástroje a také poměrná přehlednost při použití aplikace Wallace, která s modelem Maxent pracuje. Knihovna Wallace byla nalezena během rešerše modelovacích softwarů a po nastudování postupu práce v ní byla vybrána jako komplexní, a přesto srozumitelný a přehledný nástroj pro tvorbu modelů niky/distribuce druhů. Vzhledem k tomu, že cílem této práce nebylo srovnávat výstupy různých modelovacích softwarů a algoritmů, nýbrž sestavení komplexního modelu přibližujícího šíření sršně asijské v ČR, nebylo provedeno modelování v jiných softwarech.

Nedostatky statistického testování

Část, u níž si je autor vědom jistých nedostatků, je část statistického vyhodnocení nejvlivnějších proměnných (podkapitola 4.3.3). Vzhledem k tomu, že autor prováděl vyhodnocování proměnných jen základními statistickými metodami, chybí tomuto procesu pokročilé statistické zpracování vstupních predikčních proměnných, a ty proto mohou vykazovat jistá omezení. v rámci ostatních výzkumů (např. Verdasca et al., 2021; Bessa et al., 2016), z nichž autor vycházel, bylo prováděno statistické testování a vyhodnocování nejdůležitějších proměnných na výskyt sršně asijské například regresními analýzami (Generalized Linear Model, GLM; použít by se dala i logistická regrese, která by autorovi zřejmě poskytla přesnější obraz o nejvíce určujících proměnných, než tomu bylo u metody PCA použité v této práci). Modelování ekologických nik a distribuce druhů je složitá vědní disciplína vyžadující dobré statistické znalosti pro efektivní vyhodnocení nejvlivnějších proměnných na výskyt druhů.

I přes tyto nedokonalosti ve statistickém zpracování ovšem výsledný model poskytuje výsledky srovnatelné s ostatními výzkumy prováděnými v Evropě. Vzhledem k tomu, že se jedná o (dle dohledaných informací) první celoevropsky provedené modelování, které využívá data z různých odlišných částí Evropy a nikoli pouze lokální environmentální

podmínky a data (např. pouze z Mallorcy, pouze výskyty z pevninského Španělska apod.), výsledné predikce jsou pro některé oblasti větší, než byly v původních výzkumech.

Dobře to dokládá modelování vhodnosti habitatu pro sršeň asijskou provedené francouzským týmem (Obr. 28 v kapitole Výsledky; Fournier et al., 2017), které bylo v roce 2016 provedeno s daty pouze z území Francie. Vhodnost habitatu např. v Belgii a Nizozemsku je podle této francouzské studie téměř nulová. Tato diplomová práce však ukázala výraznou vhodnost klimatických a stanovištních podmínek Belgie a Nizozemska na šíření sršně asijské, a to zejména z toho důvodu, bylo pracováno i s nálezovými daty z těchto dvou států, a tedy byla přidána další data pro modelování. Je nutné si uvědomit, že ačkoli francouzský výzkum z roku 2016 s použitím nálezových dat pouze z Francie poukázal na téměř nulovou vhodnost podmínek v Belgii a Nizozemsku, o 7 let později jsou tyto země plné nálezových dat referujících o hojném počtu přítomných sršní asijských. Otázkou tak zůstává, na jaké další podmínky prostředí jsou schopny se tyto invazivní sršně dále adaptovat a zda jejich absence v České republice není jen dočasného charakteru.

Nálezová data

K problematice nálezových dat, která již byla zmíněna, je nutno podotknout, že analyzované výzkumy prováděly trénování modelu na základě dat *výskytu hnízd*. Autor v této práci neměl přístup k takovým datům, jelikož se jedná většinou o výstupy terénních prací v konkrétní lokalitě (Herrera et al., 2022) a data pouze o lokalitách hnízd nebyla nalezena. Místo toho byly použity nálezové databáze GBIF.org a iNaturalist.org, z nichž autor získal nálezová data o *pozorovaných jedincích*. Tato data mohou být nepřesná, neboť dle domněnky autora to, že byla sršeň spatřena v některých lokalitách, ještě neznamená, že je pro ni tato lokalita vhodná pro založení hnízda. Pro budoucí zpřesňování modelu by bylo vhodné používat konkrétní lokality hnízd, neboť mají mnohem větší vypovídající hodnotu o podmínkách, v nichž je sršeň asijská schopna založit kolonii, a nikoli o podmínkách, v nichž se, třeba náhodně, vyskytlo několik jedinců.

Data krajinného pokryvu

Poslední nepřesností, již autor zaznamenal, bylo použití dat krajinného pokryvu (Land Cover). Bylo potřeba použít co nejpodrobnější data, a proto se rastrová vrstva *Land Cover Map of Europe 2017 od* Evropské kosmické agentury v rozlišení 10 m na pixel ukázala jako velmi vhodná pro řešenou úlohu. Faktem je, že krajinný pokryv se leckdy mění velmi rychle, a proto již samotné použití dat z roku 2017 pro rok 2023 je poměrně nepřesné a diskutabilní. Navíc je v případě většiny dostupných dat problém najít taková, které by rozlišovala např. více druhů lesů, neboť v jižní Evropě mají lesy jiný charakter než např. ve střední Evropě. Vrstva obsahuje kategorie listnatých a jehličnatých lesů, přičemž toto rozdělení platí pro celou Evropu stejně, nerozlišují se tedy jednotlivé typy lesů. Tím, že jsou data výskytu posbírána v rámci celé Evropy, autor se domnívá, že ve výsledném modelu nehrají lokální anomálie ve struktuře lesů takovou roli. Nesmyslným krokem by bylo například trénovat model jen nad daty např. z Mallorcy a model poté přenést na území České republiky – zde by bylo velmi diskutabilní, do jaké míry by se české lesy podobaly těm na Mallorce. To však platí i pro zbylé stanovištní, zejména bioklimatické charakteristiky, které na Mallorce budou naprosto odlišné od České republiky. Trénování modelu nad daty celé Evropy tyto lokální rozdíly stírá.

Dalším nedostatkem při použití dat krajinného pokryvu je jejich vstup do modelování pro normál 2041–2060. Aplikace Wallace požaduje, aby při modelování do nových uživatelsky definovaných proměnných (např. data z budoucích normálů, podkapitola 4.4.9) byly zachovány všechny proměnné, na nichž byl model trénován. Predikovaná klimatická data pro budoucí normál jsou k dostání poměrně bez problému z několika

databázi (např. ECLIPS 2.0), data o tom, jak bude vypadat krajinný pokryv za cca 30 let neexistují, protože jejich vývoj dle autora není téměř možné predikovat. I tato skutečnost proto může ovlivnit výslednou predikci pro budoucí normály, neboť nezahrnuje konkrétní podobu krajinného pokryvu v budoucnosti.

Možná rozšíření do budoucna

Do budoucna by bylo vhodné práci obohatit o další stanovištní charakteristiky. Vzhledem k tomu, že šíření sršně probíhá i vlivem zvyšující se globalizace a s tím spojené výrazné mobility obyvatel a zboží, bylo by záhodno do modelu použít i informace o tranzitních koridorech a vzdálenostech od silnic, které by mohly ovlivňovat výskyt sršně. Například první sršně královna byla dovezena do Francie v roce 2004 v kontejneru z Číny (Diaz et al., 2023), stejným způsobem by se mohly šířit sršně i v rámci Evropy – s nákladní dopravou. Výzkum v Portugalsku (Verdasca et al., 2022) potvrzuje snižující se počty sršně směrem od dálnic a hlavních tranzitních koridorů. Určitě tedy tato proměnná může hrát roli v dalším šíření sršně asijské do dalších lokalit.

Možným rozšířením by také bylo navázání multioborové spolupráce – tedy přizvat do řešitelského týmu statistika, jenž by byl schopen dodat jednotlivým statistickým přístupům větší odbornost a analyzovat predikční proměnné dalšími způsoby za účelem volby nejlepších prediktorů, a biologa, konkrétně entomologa, který by do procesu modelování vnesl odborné biologické pozadí v posuzování vhodných environmentálních charakteristik nejvíce ovlivňujících výskyt sršně asijské.

Zároveň by bylo zajímavé nepoužívat v analýze průměrná data za celý klimatický normál 1991–2020, nýbrž použít průměrná data např. z posledních 5–10 let a zjistit, zda není v tomto normálu predikovaná míra vhodnosti větší.

Momentálně je dle autora poměrně „velká neznámá“, na jaké podmínky jsou sršně schopny se ještě adaptovat a do jakých lokalit se mohou rozšířit. Další šíření tedy bude pravděpodobně stejně jako doposud monitorováno a vzhledem k vysoké schopnosti a rychlosti šíření sršně *Vespa velutina* do dalších oblastí budou napadené oblasti zřejmě přibývat. Modelování by bylo vhodné periodicky opakovat s cílem sledovat, na jaké další podmínky jsou sršně schopny se adaptovat. Následně by bylo vhodné rozšířit predikci o tato nově nalezená data a zjistit, zda se podobnost habitatu v ČR těmto novým lokalitám více nepřibližuje.

7 ZÁVĚR

Hlavním cílem diplomové práce bylo vytvořit **prostorově založený model a následné vizualizace potenciálního výskytu sršně asijské (*Vespa velutina*) v České republice**. Modelování potenciálních lokalit výskytu v ČR bylo provedeno na základě současných přístupů a dat o výskytu v Evropě. Modelování bylo založeno na environmentálních charakteristikách a stanovištních podmínkách. Těmito podmínkami byly klimatické charakteristiky a typ krajinného pokryvu, v nichž se sršním asijským nejvíce daří. **Porovnávány byly podmínky v současném (1991–2020) a budoucím (2041–2060) klimatickém normálu.**

Rešerše byla věnována sršni asijské *Vespa velutina*, jejím charakteristikám, způsobu života a nastíněním problému rozšiřování sršně asijské v Evropě. Jsou uvedeny důvody, **proč je sršeň asijská v Evropě výrazným problémem** a jak se jí mohou lidé bránit. Bylo porovnáno několik výzkumů, které se věnovaly distribuci sršně asijské v Evropě a na základě těchto výzkumů byly stanoveny **základní bioklimatické a stanovištní charakteristiky, které nejvíce vysvětlují a podporují její výskyt**. Na základě těchto informací byly analyzovány dostupné datové sady potřebné pro modelování. v rámci rešerše byly také popsány vybrané modelovací přístupy pro modelování distribuce druhů a na základě této rešerše **byl vybrán výsledný modelovací algoritmus Maxent**, jenž byl použit v rámci aplikace Wallace postavené na jazyku R.

Nejprve byla stažena ze dvou náleзовých databází **nálezová data** výskytu sršně asijské v Evropě, byla očištěna a zpracována tak, aby byla použitelná v procesu modelování. Na rozsahu těchto dat bylo **stanoveno zájmové území** pro trénování modelu, které pokrývá 11 zemí Evropy včetně České republiky. Následně byly pomocí programů RStudio a ClimateEU pro celé zájmové území vygenerovány **klimatické charakteristiky** (teploty vzduchu, srážky) pro roční, sezónní a měsíční období. K těmto klimatickým datům byla **přidána data krajinného pokryvu a digitální model reliéfu**. z rastrových dat bylo v programech ArcGIS Pro a RStudio zpracováno celkem **19 predikčních proměnných**. Proměnné byly **statisticky zpracovány a vybrány ty, které mají největší vliv na vhodnost habitatu** pro sršeň asijskou.

Veškerá další práce probíhala v aplikaci **Wallace**, která pracuje s algoritmem Maxent a slouží pro **modelování distribuce druhů**. v rámci modelování v aplikaci Wallace bylo postupně řešeno: (1) **nahrání environmentálních a nálezových dat**, (2) **jejich zpracování**, (3) **výběr vhodného, sršněmi doposud neobývaného environmentálního pozadí** pro porovnání podmínek v místech výskytu sršně a jejich pseudo-absence, (4) **sestavení modelu a jeho evaluace** a (5) **přenesení modelu do nové oblasti a časového období a následná vizualizace výsledků**. **Výsledný model** je vyexportován do dvou formátů, které je možné pomocí programu RStudio, resp. aplikace Wallace otevřít, modifikovat (např. doplnit data výskytu) a spustit výpočet nad novými daty.

Z výsledných vygenerovaných dat míry vhodnosti byly v programu ArcGIS Pro vytvořeny **mapy**. Ty byly následně programy desktop publishing (Adobe Illustrator a Adobe InDesign) graficky upraveny, obohaceny o informace o sršních a zpracovány do formy **brožury**, která je spolu s **informačním letákem** hlavními vizualizačními materiály vzniklými v této práci a popisuje problematiku šíření sršně na území ČR laické veřejnosti. Výsledné vizualizace byly v širším geografickém kontextu okomentovány. Tato práce dokázala, že **sršeň asijská v České republice momentálně nemá vhodné podmínky pro šíření**, a tedy výrazný strach z tohoto invazivního druhu není na místě. To se ovšem může v budoucnu změnit, a tak **obezřetnost je na místě**. Vše nyní záleží na vývoji klimatu v následujících letech a schopnosti sršně asijských adaptovat se na nové podmínky.

POUŽITÁ LITERATURA A INFORMAČNÍ ZDROJE

AHMED, Zia. Co-Kriging. *Geospatial Data Science in R* [online]. 2023 [cit. 2023-04-29]. Dostupné z: <https://zia207.github.io/geospatial-r-github.io/cokriging.html>

Asian Hornet (*Vespa velutina*), 2023. *Vita Bee Health* [online]. [cit. 2023-05-01]. Dostupné z: <https://www.vita-europe.com/beehealth/diseases/asian-hornet/>

BEAUMONT, Linda J., Lesley HUGHES a Michael POULSEN, 2005. Predicting species distributions: use of climatic parameters in BIOCLIM and its impact on predictions of species' current and future distributions. *Ecological Modelling* [online]. **186**(2), 251-270 [cit. 2023-05-03]. ISSN 03043800. Dostupné z: doi:10.1016/j.ecolmodel.2005.01.030

BESSA, Ana S., João CARVALHO, Alberto GOMES, Frederico SANTARÉM, Raphael DIDHAM a Seán BRADY, 2016. Climate and land-use drivers of invasion: predicting the expansion of *Vespa velutina nigrithorax* into the Iberian Peninsula. *Insect Conservation and Diversity* [online]. **9**(1), 27-37 [cit. 2021-12-16]. ISSN 1752458X. Dostupné z: doi:10.1111/icad.12140

BOYCE, Mark S, Pierre R VERNIER, Scott E NIELSEN a Fiona K.A SCHMIEGELOW, 2002. Evaluating resource selection functions. *Ecological Modelling* [online]. **157**(2-3), 281-300 [cit. 2023-04-29]. ISSN 03043800. Dostupné z: doi:10.1016/S0304-3800(02)00200-4

BURNHAM, Kenneth P. a David R. ANDERSON, 2022. *Model selection and multimodel inference: a practical information-theoretic approach* [online]. Springer New York, NY [cit. 2023-04-30]. Dostupné z: doi:10.1007/b97636

CARNAVAL, Ana Carolina a Craig MORITZ, 2008. Historical climate modelling predicts patterns of current biodiversity in the Brazilian Atlantic forest. *Journal of Biogeography* [online]. **35**(7), 1187-1201 [cit. 2023-05-03]. ISSN 03050270. Dostupné z: doi:10.1111/j.1365-2699.2007.01870.x

CLIMATENEXUS. RCP 8.5: Business-as-usual or a worst-case scenario?. *Climate Nexus* [online]. [cit. 2023-04-29]. Dostupné z: <https://climatenexus.org/climate-change-news/rcp-8-5-business-as-usual-or-a-worst-case-scenario/>

ČESKÁ METEOROLOGICKÁ SPOLEČNOST, 2023. Kontinentalita klimatu. *Elektronický meteorologický slovník výkladový a terminologický* [online]. [cit. 2023-05-02]. Dostupné z: <http://slovník.cmes.cz/heslo/1676>

DE SOUZA MUÑOZ, Mauro Enrique, Renato DE GIOVANNI, Marinez Ferreira DE SIQUEIRA, Tim SUTTON, Peter BREWER, Ricardo Scachetti PEREIRA, Dora Ann Lange CANHOS a Vanderlei Perez CANHOS, 2011. OpenModeller: a generic approach to species' potential distribution modelling. *GeoInformatica* [online]. **15**(1), 111-135 [cit. 2023-05-03]. ISSN 1384-6175. Dostupné z: doi:10.1007/s10707-009-0090-7

DIAZ, Adriana, Sophie GRÜNEWALD, Helena PROKOVÁ a Wolfgang WIMMER, 2023. *Sršeň asijská (Vespa velutina), její biologie, monitoring, kontrola a prevence šíření* [online]. Opatovice: Pracovní společnost nástavkových včelařů CZ [cit. 2023-05-03]. ISBN 978-80-907079-7-9. Dostupné z: <https://www.psnv.cz/attachment/s-49-d469a72334c2/1024-srsen-asijska-prirucka.pdf>

DORMANN, Carsten F., Stanislaus J. SCHYMANSKI, Juliano CABRAL, et al., 2012. Correlation and process in species distribution models: bridging a dichotomy. *Journal of Biogeography* [online]. **39**(12), 2119-2131 [cit. 2023-05-03]. ISSN 03050270. Dostupné z: doi:10.1111/j.1365-2699.2011.02659.x

DRAKE, John M., 2014. Ensemble algorithms for ecological niche modeling from presence-background and presence-only data. *Ecosphere* [online]. **5**(6) [cit. 2023-05-03]. ISSN 2150-8925. Dostupné z: doi:10.1890/ES13-00202.1

DRBOHLAVOVÁ, Lenka, 2013. Sršeň obecná – Vespa crabro. *Příroda.cz* [online]. [cit. 2023-05-01]. Dostupné z: <https://www.priroda.cz/clanky.php?detail=2453>

ELITH, Jane, Catherine H. GRAHAM, Robert P. ANDERSON, et al., 2006. Novel methods improve prediction of species' distributions from occurrence data. *Ecography* [online]. **29**(2), 129-151 [cit. 2023-04-29]. ISSN 09067590. Dostupné z: doi:10.1111/j.2006.0906-7590.04596.x

ELITH, Jane, Steven J. PHILLIPS, Trevor HASTIE, Miroslav DUDÍK, Yung En CHEE a Colin J. YATES, 2011. a statistical explanation of MaxEnt for ecologists. *Diversity and Distributions* [online]. **17**(1), 43-57 [cit. 2023-04-29]. ISSN 13669516. Dostupné z: doi:10.1111/j.1472-4642.2010.00725.x

ESRI, 2023a. How inverse distance weighted interpolation works. *Esri* [online]. [cit. 2023-05-02]. Dostupné z: <https://pro.arcgis.com/en/pro-app/latest/help/analysis/geostatistical-analyst/how-inverse-distance-weighted-interpolation-works.htm>

ESRI, 2023b. Is Null (Spatial Analyst). *Esri* [online]. [cit. 2023-04-27]. Dostupné z: https://pro.arcgis.com/en/pro-app/latest/tool-reference/spatial-analyst/is-null.htm#L_

EU-DEM v1.1, 2017. *European Environment Agency* [online]. [cit. 2023-05-01]. Dostupné z: <https://www.eea.europa.eu/data-and-maps/data/copernicus-land-monitoring-service-eu-dem>

EUROPEAN ENVIRONMENT AGENCY, 2018. EEA coastline for analysis. *European Environment Agency* [online]. [cit. 2023-05-01]. Dostupné z: <https://www.eea.europa.eu/data-and-maps/data/eea-coastline-for-analysis-2>

EUROPEAN ENVIRONMENT AGENCY, 2017. Copernicus Land Monitoring Service - EU-DEM: EU-DEM v1.1. *European Environment Agency* [online]. [cit. 2023-05-02]. Dostupné z: <https://www.eea.europa.eu/data-and-maps/data/copernicus-land-monitoring-service-eu-dem>

FOURNIER, Alice, Morgane BARBET-MASSIN, Quentin ROME a Franck COURCHAMP, 2017. Predicting species distribution combining multi-scale drivers. *Global Ecology and Conservation* [online]. **12**, 215-226 [cit. 2023-05-02]. ISSN 23519894. Dostupné z: doi:10.1016/j.gecco.2017.11.002

FRANKLIN, Daniel N., Mike A. BROWN, Samik DATTA, Andrew G. S. CUTHBERTSON, Giles E. BUDGE a Matt J. KEELING, 2017. Invasion dynamics of Asian hornet, *Vespa velutina* (Hymenoptera: Vespidae). *Applied Entomology and Zoology* [online]. **52**(2), 221-229 [cit. 2023-05-02]. ISSN 0003-6862. Dostupné z: doi:10.1007/s13355-016-0470-z

GAIJI, Samy, Vishwas CHAVAN, Arturo H. ARIÑO, Javier OTEGUI, Donald HOBERN, Rajesh SOOD a Estrella ROBLES, 2013. Content assessment of the primary biodiversity data published through GBIF network: Status, challenges and potentials. *Biodiversity Informatics* [online]. **8**(2) [cit. 2023-04-28]. ISSN 15469735. Dostupné z: doi:10.17161/bi.v8i2.4124

Global Biodiversity Information Facility [online], 2023. [cit. 2023-05-02]. Dostupné z: <https://www.gbif.org/>

GOMÉZ, Jorge García, 2023. Rainfall seasonality. *Desertification Indicator System for Mediterranean Europe* [online]. [cit. 2023-05-01]. Dostupné z: https://esdac.jrc.ec.europa.eu/public_path/shared_folder/projects/DIS4ME/indicator_descriptions/rainfall_seasonality.htm

GUTIÉRREZ-HERNÁNDEZ, O. a L.V. GARCÍA, 2021. Relationship between precipitation and species distribution. *Precipitation* [online]. Elsevier, 2021, 239-259 [cit. 2023-05-03]. ISBN 9780128226995. Dostupné z: doi:10.1016/B978-0-12-822699-5.00010-0

HAUSFATHER, Zeke. Explainer: The high-emissions 'RCP8.5' global warming scenario. *CarbonBrief: Clear On Climate* [online]. 2019 [cit. 2023-04-29]. Dostupné z: <https://www.carbonbrief.org/explainer-the-high-emissions-rcp8-5-global-warming-scenario/>

HERRERA, Cayetano, José JURADO-RIVERA a Mar LEZA, 2022. Ensemble of Small Models As a Tool For Alien Invasive Species Management Planning: Evaluation of *Vespa Velutina* (Hymenoptera: Vespidae) Under Mediterranean Island Conditions. *Research Square* [online]. Universitat de les Illes Balears [cit. 2021-12-16]. Dostupné z: doi:10.21203/rs.3.rs-809089/v1

HIRZEL, Alexandre H., Gwenaëlle LE LAY, Véronique HELFER, Christophe RANDIN a Antoine GUIBAN, 2006. Evaluating the ability of habitat suitability models to predict species presences. *Ecological Modelling* [online]. **199**(2), 142-152 [cit. 2023-04-29]. ISSN 03043800. Dostupné z: doi:10.1016/j.ecolmodel.2006.05.017

CHAKRABORTY, Debojyoti, Laura DOBOR, Anita ZOLLES, Tomáš HLÁSNÝ a Silvio SCHUELER, 2020. High-resolution gridded climate data for Europe based on bias-corrected EURO-CORDEX: The ECLIPS dataset. *Geoscience Data Journal* [online]. **8**(2), 121-131 [cit. 2023-05-02]. ISSN 2049-6060. Dostupné z: doi:10.1002/gdj3.110

CHENG, Ya-Nan, Ping WEN, Ken TAN a Eric DARROUZET, 2022. Designing a sex pheromone blend for attracting the yellow-legged hornet (*Vespa velutina*), a pest in its native and invasive ranges worldwide. *Entomologia Generalis* [online]. **42**(4), 523-530 [cit. 2023-05-03]. ISSN 0171-8177. Dostupné z: doi:10.1127/entomologia/2022/1395

CHOI, Moon Bo, Stephen J. MARTIN a Jong Wook LEE, 2012. Distribution, spread, and impact of the invasive hornet *Vespa velutina* in South Korea. *Journal of Asia-Pacific Entomology* [online]. **15**(3), 473-477 [cit. 2023-05-02]. ISSN 12268615. Dostupné z: doi:10.1016/j.aspen.2011.11.004

IBM, 2023. What is a neural network?. *IBM* [online]. [cit. 2023-05-03]. Dostupné z: <https://www.ibm.com/topics/neural-networks>

iNaturalist [online], 2023a. [cit. 2023-05-02]. Dostupné z: <https://www.inaturalist.org/>

iNaturalist, Frequently Asked Questions: How should I cite iNaturalist?, 2023b. *iNaturalist* [online]. [cit. 2023-05-02]. Dostupné z: <https://www.inaturalist.org/pages/help#cite>

JANŽEKOVIČ, Franc a Tone NOVAK, 2012. PCA – a Powerful Method for Analyze Ecological Niches. *Principal Component Analysis - Multidisciplinary Applications* [online]. InTech, 2012-02-29 [cit. 2023-04-30]. ISBN 978-953-51-0129-1. Dostupné z: doi:10.5772/38538

JOHNSON, Bethany A., , Andrea PAZ, Jamie M. KASS, Sarah I. MEENAN a Robert P. ANDERSON, 2023. Wallace Ecological Modeling Application v2.0 Vignette. *Wallace* [online]. [cit. 2023-05-03]. Dostupné z: <https://wallaceecomod.github.io/wallace/articles/tutorial-v2.html#references>

KRAMER-SCHADT, Stephanie, Jürgen NIEDBALLA, John D. PILGRIM, et al., 2013. The importance of correcting for sampling bias in MaxEnt species distribution models. *Diversity and Distributions* [online]. **19**(11), 1366-1379 [cit. 2023-04-28]. ISSN 13669516. Dostupné z: doi:10.1111/ddi.12096

KUMAR, Pankaj, Sven KOTLARSKI, Christopher MOSELEY, Kevin SIECK, Holger FREY, Markus STOFFEL a Daniela JACOB, 2015. Response of Karakoram-Himalayan glaciers to climate variability and climatic change: a regional climate model assessment. *Geophysical*

Research Letters [online]. **42**(6), 1818-1825 [cit. 2023-04-29]. ISSN 0094-8276. Dostupné z: doi:10.1002/2015GL063392

LOBO, Jorge M., Alberto JIMÉNEZ-VALVERDE a Raimundo REAL, 2008. AUC: a misleading measure of the performance of predictive distribution models. *Global Ecology and Biogeography* [online]. **17**(2), 145-151 [cit. 2023-05-03]. ISSN 1466-822X. Dostupné z: doi:10.1111/j.1466-8238.2007.00358.x

MALINOWSKI, Radek, Stanisław LEWIŃSKI, Marcin RYBICKI, et al., 2020. Automated Production of a Land Cover/Use Map of Europe Based on Sentinel-2 Imagery. *Remote Sensing* [online]. **12**(21) [cit. 2023-05-02]. ISSN 2072-4292. Dostupné z: doi:10.3390/rs12213523

MARCHI, Maurizio. Why we should refine MaxEnt model by removing highly correlated variables?. *ResearchGate* [online]. 2015 [cit. 2023-04-30].

MARCHI, Maurizio, Dante CASTELLANOS-ACUÑA, Andreas HAMANN, Tongli WANG, Duncan RAY a Annette MENZEL, 2020. ClimateEU, scale-free climate normals, historical time series, and future projections for Europe. *Scientific Data* [online]. **7**(1) [cit. 2023-05-03]. ISSN 2052-4463. Dostupné z: doi:10.1038/s41597-020-00763-0

MONCEAU, Karine, Mariangela ARCA, Lisa LEPRÉTRE, et al., 2013. Native Prey and Invasive Predator Patterns of Foraging Activity: The Case of the Yellow-Legged Hornet Predation at European Honeybee Hives. *PLoS ONE* [online]. **8**(6) [cit. 2023-05-03]. ISSN 1932-6203. Dostupné z: doi:10.1371/journal.pone.0066492

MORROW, Cecina B. *Thresholding species distribution models* [online]. 2019 [cit. 2023-04-29]. Dostupné z: <https://babichmorrowc.github.io/post/2019-04-12-sdm-threshold/>

NARKHEDE, Sarang. Understanding AUC - ROC Curve. *Towards Data Science* [online]. [cit. 2023-04-29]. Dostupné z: <https://towardsdatascience.com/understanding-auc-roc-curve-68b2303cc9c5>

O'DONNELL, M. S. a D. A. IGNIZIO, 2012. Bioclimatic predictors for supporting ecological applications in the conterminous United States: U.S. Geological Survey Data Series 691. USGS [online]. [cit. 2023-05-02]. Dostupné z: <https://pubs.usgs.gov/ds/691/ds691.pdf>

OPENMODELLER, 2015. Documentation for Users. *OpenModeller* [online]. [cit. 2023-05-03]. Dostupné z: <https://openmodeller.sourceforge.net/documentation.html>

PETERSON, A. Townsend, Monica PAPEŞ a Jorge SOBERÓN, 2008. Rethinking receiver operating characteristic analysis applications in ecological niche modeling. *Ecological Modelling* [online]. **213**(1), 63-72 [cit. 2023-04-29]. ISSN 03043800. Dostupné z: doi:10.1016/j.ecolmodel.2007.11.008

PHILLIPS, Steven J. *a Brief Tutorial on Maxent* [online]. 2017 [cit. 2023-04-29]. Dostupné z:
https://biodiversityinformatics.amnh.org/open_source/maxent/Maxent_tutorial_2021.pdf

PHILLIPS, Steven J., Robert P. ANDERSON a Robert E. SCHAPIRE, 2006. Maximum entropy modeling of species geographic distributions. *Ecological Modelling* [online]. **190**(3-4), 231-259 [cit. 2023-04-29]. ISSN 03043800. Dostupné z:
doi:10.1016/j.ecolmodel.2005.03.026

POLECHOVÁ, Jitka a David STORCH, 2019. Ecological Niche. *Encyclopedia of Ecology* [online]. Elsevier, 2019, 72-80 [cit. 2023-05-03]. ISBN 9780444641304. Dostupné z:
doi:10.1016/B978-0-12-409548-9.11113-3

PROKOVÁ, Helena, 2023. Dřív, než bude pozdě. *Moderní včelař*. Pracovní společnost nástavkových včelařů CZ, **2023**(02), 8–14.

ROBINET, Christelle, Christelle SUPPO, Eric DARROUZET a Tim DIEKÖTTER, 2017. Rapid spread of the invasive yellow-legged hornet in France: the role of human-mediated dispersal and the effects of control measures. *Journal of Applied Ecology* [online]. **54**(1), 205-215 [cit. 2023-04-28]. ISSN 00218901. Dostupné z: doi:10.1111/1365-2664.12724

SKUHROVEC, Jiří. Biologie sršně asijské a monitoring jejího šíření. Dřív, než bude pozdě. 23. 2. 2023, Praha. Pracovní společnost nástavkových včelařů CZ, domácí konference.

SMITH, Tyler, 2021. *Maxent Best Practices* [online]. [cit. 2023-04-29]. Dostupné z:
<https://plantarum.ca/notebooks/maxent/#auc>

ŠTĚRBA, Roman Richard. *Bioklimatické modelování* [online]. Brno, 2010 [cit. 2023-05-02]. Dostupné z: https://is.muni.cz/th/ui8j8/bakalarska_prace.pdf. Bakalářská práce. Masarykova univerzita, Přírodovědecká fakulta, Centrum pro výzkum toxických látek v prostředí. Vedoucí práce Mgr. Klára Kubošová, PhD.

TAPIADOR, Francisco J., Andrés NAVARRO, Raúl MORENO, José Luis SÁNCHEZ a Eduardo GARCÍA-ORTEGA, 2020. Regional climate models: 30 years of dynamical downscaling. *Atmospheric Research* [online]. **235** [cit. 2023-04-29]. ISSN 01698095. Dostupné z: doi:10.1016/j.atmosres.2019.104785

UNIVERSITY OF EAST ANGLIA CLIMATIC RESEARCH UNIT, I. C. HARRIS, P. D. JONES a T. OSBORN, 2021. CRU TS4.05: Climatic Research Unit (CRU) Time-Series (TS) version 4.05 of high-resolution gridded data of month-by-month variation in climate (Jan. 1901-Dec. 2020). *NERC EDS Centre for Environmental Data Analysis* [online]. [cit. 2023-05-02]. Dostupné z: <https://catalogue.ceda.ac.uk/uuid/c26a65020a5e4b80b20018f148556681>

VERDASCA, Maria João, Hugo REBELO, Luísa G. CARVALHEIRO a Rui REBELO, 2021. Invasive hornets on the road: motorway-driven dispersal must be considered in management plans of *Vespa velutina*. *NeoBiota* [online]. **69**, 177-198 [cit. 2021-12-16]. ISSN 1314-2488. Dostupné z: doi:10.3897/neobiota.69.71352

VLČKOVÁ, Eliška. Srovnání výpočetních algoritmů pro prostorovou distribuci druhu [online]. Olomouc, 2016 [cit. 2023-05-02]. Dostupné z: <http://www.geoinformatics.upol.cz/dprace/bakalarske/vlckova16/>. Bakalářská práce. Univerzita Palackého v Olomouci, Přírodovědecká fakulta, Katedra geoinformatiky. Vedoucí práce Mgr. Jitka Doležalová

WARREN, Dan L. a Stephanie N. SEIFERT, 2011. Ecological niche modeling in Maxent: the importance of model complexity and the performance of model selection criteria. *Ecological Applications* [online]. **21**(2), 335-342 [cit. 2023-04-30]. ISSN 1051-0761. Dostupné z: doi:10.1890/10-1171.1

WILLIAMS, John W. a Stephen T. JACKSON, 2007. Novel climates, no-analog communities, and ecological surprises. *Frontiers in Ecology and the Environment* [online]. **5**(9), 475-482 [cit. 2023-04-28]. ISSN 1540-9295. Dostupné z: doi:10.1890/070037

PŘÍLOHY

SEZNAM PŘÍLOH

Kapitola se nečísluje. Uvést kompletní číslovaný seznam vázaných a volných příloh.

Příloha 1 (elektronická volná) **Brožura – Nebezpečná sršeň asijská**

Příloha 2 (elektronická volná) **Informační leták – Nebezpečná sršeň asijská**

Příloha 3 (volná) **Poster**

Příloha 4 (elektronická volná) **Kód modelu ve formátu RMD**

Příloha 5 (elektronická volná) **Kód (soubor) relace aplikace Wallace ve formátu RDS**

Příloha 6 (vázaná) Mapa Míra vhodnosti habitatu pro sršeň asijskou v České republice pro normál 1991–2020, model H0,5

Příloha 7 (vázaná) Mapa Míra vhodnosti habitatu pro sršeň asijskou v České republice pro normál 1991–2020, model H1,5

Příloha 8 (vázaná) Mapa Míra vhodnosti habitatu pro sršeň asijskou v České republice pro normál 1991–2020, model LQH2,5

Příloha 9 (vázaná) Mapa Míra vhodnosti habitatu pro sršeň asijskou v České republice pro normál 2041–2060, model H0,5

Příloha 10 (vázaná) Mapa Míra vhodnosti habitatu pro sršeň asijskou v České republice pro normál 2041–2060, model H1,5

Příloha 11 (vázaná) Mapa Míra vhodnosti habitatu pro sršeň asijskou v České republice pro normál 2041–2060, model LQH2,5

Struktura odevzdávaných digitálních dat:

Poster

Text_Prace

Vstupni_Data

Vstupni_CSV_klima

Vstupni_CSV_srsne

Zkoumane_uzemi_polygon

Vystupni_Data

Elektronicke_Prilohy (obsahuje volné elektronické přílohy viz výše)

Envri_promenne_CR_204160

Envir_promenne_Evropa

Srazky_mesicni

Teplota_mesicni

Vystupy_Modelu

WEB

assets

images

text

MÍRA VHODNOSTI HABITATU PRO SRŠEŇ ASIJSKOU

v České republice v normálu 1991–2020, model H0,5



MÍRA VHODNOSTI HABITATU PRO SRŠEŇ ASIJSKOU

v České republice v normálu 1991–2020, model H1,5



MÍRA VHODNOSTI HABITATU PRO SRŠEŇ ASIJSKOU

v České republice v normálu 1991–2020, model LQH2,5



MÍRA VHODNOSTI HABITATU PRO SRŠEŇ ASIJSKOU

v České republice v normálu 2041–2060, model H0,5



MÍRA VHODNOSTI HABITATU PRO SRŠEŇ ASIJSKOU

v České republice v normálu 2041–2060, model H1,5



MÍRA VHODNOSTI HABITATU PRO SRŠEŇ ASIJSKOU

v České republice v normálu 2041–2060, model LQH2,5

