

BOMHUM_salidzinajums

Betija Rubene

2026-01-22

Brūnjoslas kamenes (*Bombus humilis*) modeļu salīdzinājums

Šī ir tāda suga, kura pēc literatūras ir malu suga (mežmalas) bet arī zālājos un virsājos, it kā īpaši bieži kalķainos zālājos un piekrastes kāpās. Manā pieredzē suga ir gana plaši sastopama dažādās vidēs, tomēr fragmentēti, nav tā, ka ir visur. Līdz ar to ekspektācijas mēreni neskaidras, izņemot, ka varētu būt saistība ar malām.

Modeļu varianti:

v1 - modelis bez bias kontroles;

v1_1 - modelis bez bias kontroles, ar labotu novērojumu filtrēšanu (to laboju arī turpmākajos modeļos, jo ar Rūtu un Aināru nospriedām, ka tā ir kļūda);

v2 - modelis ar telpisko filtrēšanu 1 km šūnās (bet bez bias slāņa) - turpmākajos modeļos 1km filtrēšana saglabāta;

Modeļi ar bias slāni, kas veidojies no visiem novērojumiem:

v3 - bez apakšgala limitācijas;

v4 - ar limitāciju 10% no vidējā;

v5 - ar limitāciju kā vidējo vērtību;

Modeļi ar sezonāli svaroto bias slāni:

v6 - bez apakšgala limitācijas;

v7 - ar limitāciju 10% no vidējā;

v8 - ar limitāciju kā vidējo vērtību;

Augstākminētā un zemāk aprakstīto pārdomu sakarā - **papildu modeļi**, kuros intereses pēc ieviesu kādas izmaiņas:

v4_1 - ar bias slāni, kas veidojies no visiem novērojumiem, **bez 1km filtrēšanas**, ar limitāciju 10% no vidējā;

v7_1 - ar sezonāli svaroto bias slāni, **bez 1 km filtrēšanas**, ar limitāciju 10% no vidējā;

v7_1a - izmetot meža tipu un vecumgrupu egv, vietā liekot kokus kopumā

v7_2 - ar sezonāli svaroto bias slāni, ar 1 km filtrēšanu, ar limitāciju 10% no vidējā, **r = 3000**;

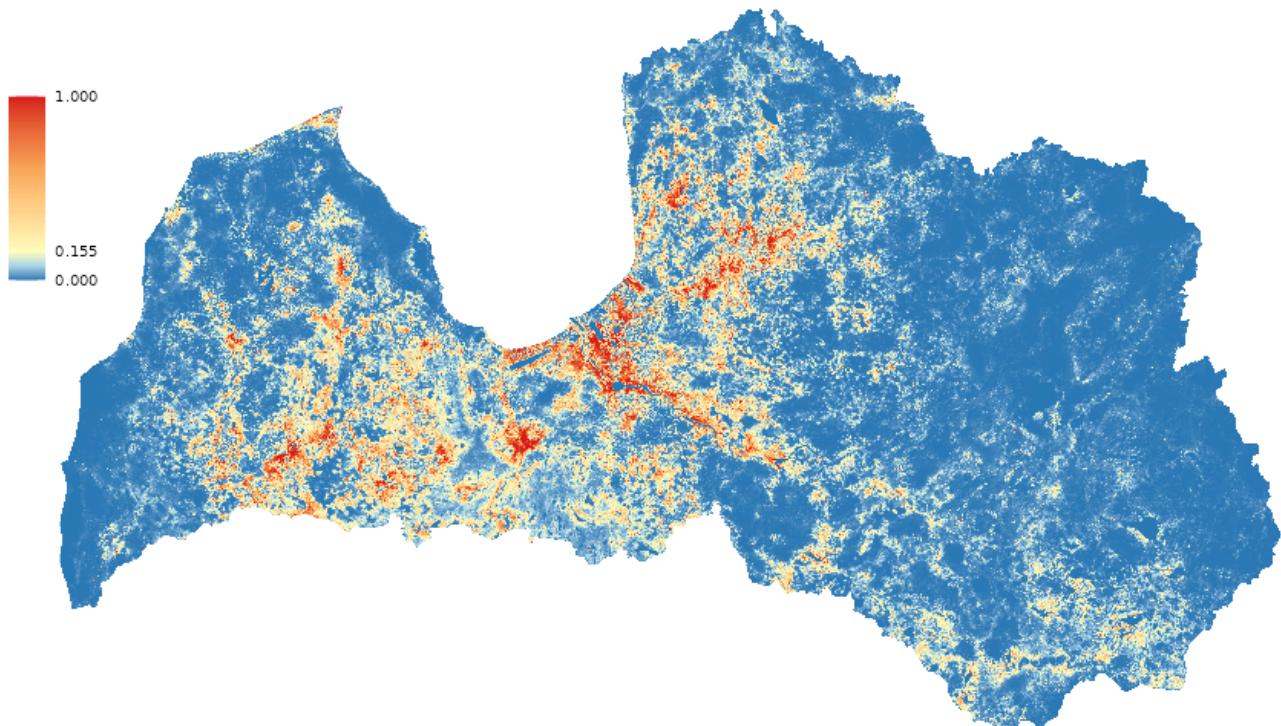
v7_3 - ar sezonāli svaroto bias slāni, bez 1 km filtrēšanas, ar limitāciju 10% no vidējā, **r = 3000**;

Arī citām sugām, kurām kaut ko mēģināju, pieturējos pie šādas numerācijas.

Jau pirms modeļu rezultātu apskatīšanas - ir zināms, ka datu ievākšanā pavisam noteikti ir novirzes no nejaušības, un ar to kaut kas ir jāiesāk, tādēļ v1 un v1_1 neizvēlēšos kā labāko modeli. Teorētiski sezonāli svarotais variants ir pamatootāks. Turklat apakšgala limitācijai 10% no vidējā vajadzētu palīdzēt - vienkārši vidējais šķiet par daudz, bez limitācijas paraugotā vide šķiet varbūt pārāk ierobežota, tāpēc v7 varētu būt labākais variants. Bet tad ļeros pie pašu rezultātu pētīšanas.

v1_1 karte diezgan izteikti velk uz pilsētām:

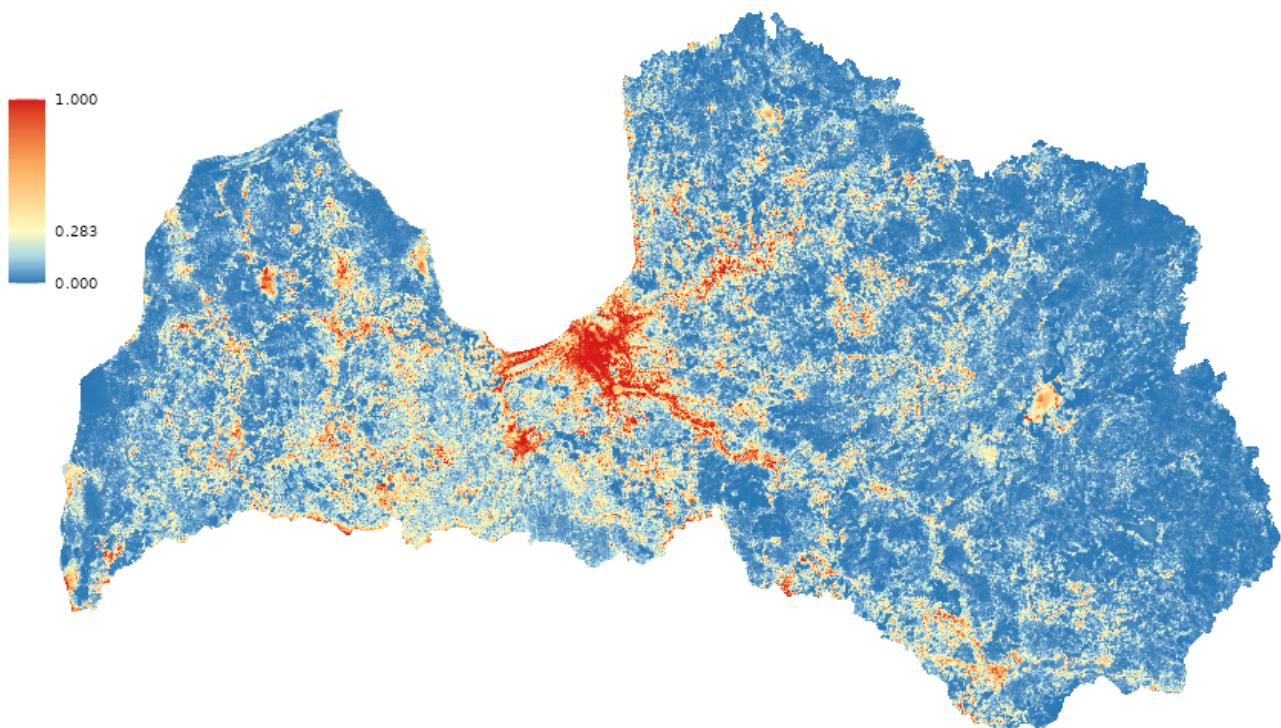
Hymenoptera: BOMHUM



BOMHUM v1_1 HSmap

Ar telpisko filtrēšanu situācija šoreiz kļūst vēl sliktāka:

Hymenoptera: BOMHUM



BOMHUM v2 HSmap

v1 - v2 tad atļaujos vairs īpaši tālāk neapskatīt.

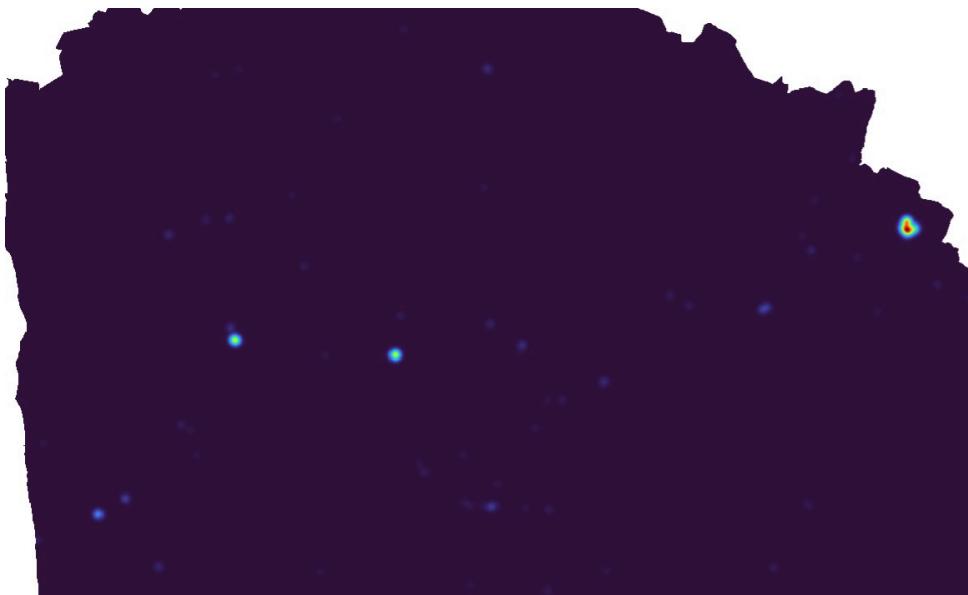
Ieviesu vēl v4_1 un v7_1 - pamatnosacījumi no mājasdarba modeļu versijām, bet bez filtrēšanas 1 km tīklā. Andra skripti, kuros telpiskā filtrēšana 1km tīklā nebija veikta, liek man domāt, ka tā tas bija paredzēts jau sākotnēji. Nu, būs, ar ko salīdzināt.

AUC vērtības no ROC līknes:

Modelis	AUC (train)	AUC (test)
v1	0.964	0.646
v1_1	0.956	0.881
v2	0.934	0.619
v3	0.912	0.69
v4	0.951	0.508
v4_1	0.908	0.565
v5	0.895	0.663
v6	0.913	0.712
v7	0.927	0.503
v7_1	0.884	0.651
v7_1a	0.866	0.476
v7_2	0.901	0.502
v7_3	0.871	0.84
v8	0.92	0.521

No pamatmodeļiem v4 un v7 ir viszemākās vērtības, lai gan teorijā man šie divi patika vislabāk, bet te aiziet visa tā diskusija par to, kādēļ vērtību pazemināšanās ir diezgan loģiska. Intereses pēc salīdzināšu, cik ļoti atšķiras bias slāņi (vienkārši novērojumi / sezonālā svarošana). Liekas jocīgi, ka v5 un v8 ir labāki rādītāji - fona punktu vide tajos ir krieti mazāk ierobežota.

(Update no vēlāka, kad domāju par bias rūpīgāk) Skatījos QGIS - grūti parādīt kopainu šajā pārskatā, jo tie ir mazi punktiņi lielā kartē, bet atšķiras diezgan ievērojami - šo jāpatur prātā. Piemēram - daļa vidzemes:



Sezonāli nesvarotais bias slānis



Sezonāli svarots bias slānis

Zinot lietas par datu iegūšanu, vairāk sliecos piekrist variantam, kurā nav sezonālā svarošana - ir "iztīrījušās" Agrihorta pētījumu vietas, kuros pētījums noticis vairākus gadus, lielāku uzsvaru liekot vietās, kur tas veikts tikai vienu gadu, kā arī kaut kādās citās vietās - pieņemu, ka kaut kas no DD vai personīgajiem novērojumiem.

Bišu gadījumā sezonālā svarošana veikta, nesmot vērā novērojumu dienu - tāda pieejama mazliet vairāk ka pusei novērojumu, tomēr jāpatur prātā, ka diena nav pieejama tieši Agrihorta datiem. Tam gan īsti nevajadzētu radīt problēmas, manuprāt.

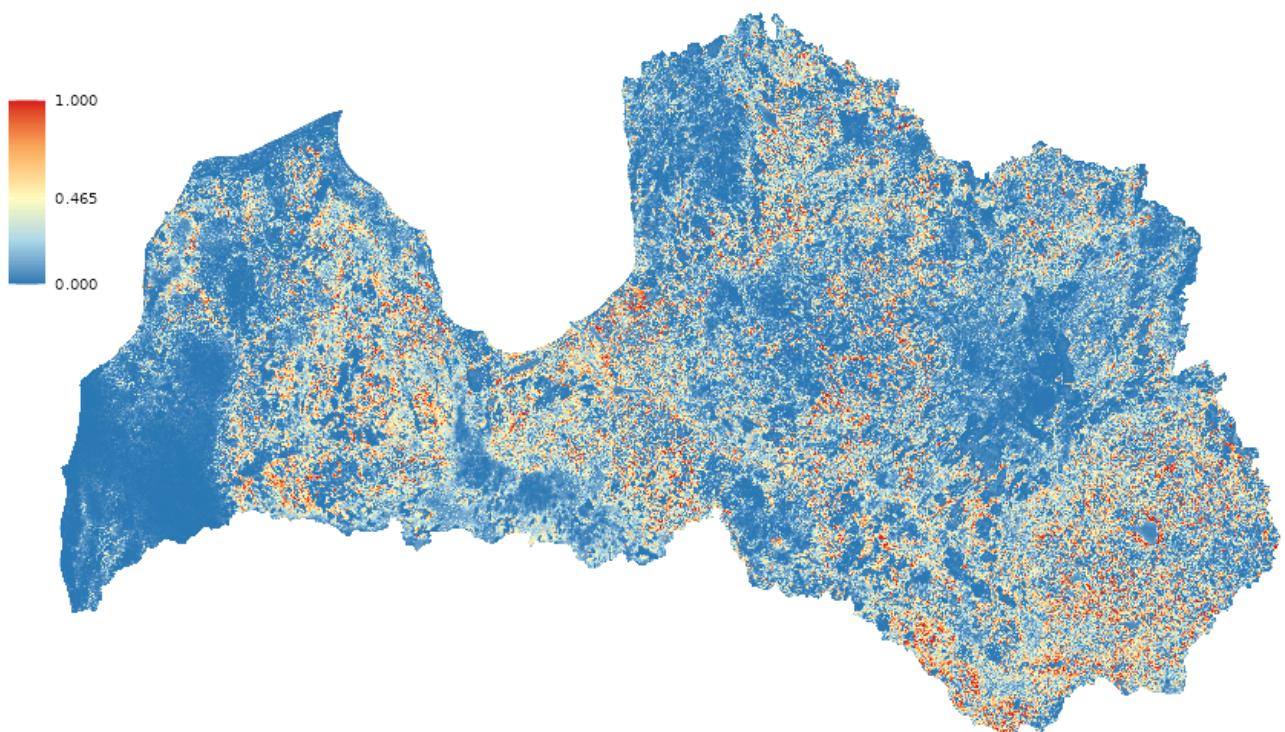
Aizdomājos par to, ka bitēm šajā projektā sugu loks ir stipri ierobežots (modelēšanā 18 sugas no ~300) un bias slāņa izveidei šobrīd tiek izmantoti tikai šī projekta ietvaros apkopotie dati - tas nozīmē, ka bias raksturošanai tiek izmantota tikai izlase no DD novērojumiem. Ir ļoti daudz DD novērojumu, kuros suga nav noteikta/nosakāma, bet tas jau nenozīmē, ka piepūle šajās vietās nav bijusi - attiecīgi šobrīd izmantotie dati "neredz", ka šajās vietās kāds ir vērojis bites. Šo visu nesmot vērā, man ir pamats domāt, ka izmantotie dati DD novērotāju piepūli īpaši labi neraksturo. Bet, ja novērojumam nav piesaistāma konkrētā suga, tad nav iespējams veikt sezonālo svarošanu. Attiecībā uz Agrihorta datiem arī sanāk jocīgi - šādā aprēķina gadījumā svars tiek dots nevis izmantojot faktiski zināmo par izmantoto piepūli (un tas ir zināms - lamatdienas konkrētā vietā), bet gan izmantojot mērķsugu novērojumu skaitu, attiecīgi aprēķināto piepūli nepamatoti ietekmē tas, cik daudz izlases sugu novērojumu ir bijis pētījuma vietā (pat ja vietas paraugotas ar vienādu piepūli, lielāks svars būs vietai, kur vairāk izlases sugu novērojumu). To var labi redzēt, piemēram, Ziemeļgaujā (Attēlā "Sezonāli nesvarotais bias slānis" košais punkts labajā pusē, kas svarotajā variantā principā nav redzams, bet skaidri zināms, ka tur piepūle ir bijusi liela, jo lamatas esam likuši vairākās vietās jau gadiem no aprīļa līdz jūlijām, +/- kas noklāj visu sezonu). Mans skaidrojums tam, ka šī vieta iztīrījusies, ir tāds, ka tur ir sugu novērojumi, kuri nesakrīt ar modelētās sugas aktivitātes laiku (kas pēc fakta liekas nepareizi, jo kamenes ir aktīvas vairāk vai mazāk visu sezonu - tātad vēl viens pamats domāt, ka šajā gadījumā sezonālā svarošana nav labs variants), tādēļ šai vietai dots ļoti mazs svars.

Īsti nezinu, kā šim labāk pieiet. Šobrīd nenāk prātā domas par to, kā visu informēti un pamatoti sakombinēt (izņemot tā, kā darījām maģistra darbā, kombinējot blīvumu DD novērojumiem ar Agrihorta lamatdienām, svarojot atbilstoši mērķsugas novērojumu ipatsvaram katrā kopā), bet šī noteikti ir viela diskusijai un arī citām organismu grupām gan jau vērts izpētīt sezonāli svaroto vs nesvaroto bias slāni un apdomāt, cik labi modelēšanai izvēlēto sugu novērojumu izlase raksturo piepūli. Zinu, ka, piemēram, Rūtai attiecībā uz taisnspārniem tas diži daudz neko nemainīja.

Pagaidām šo atstāju un salīdzināšu modeļu rezultātus.

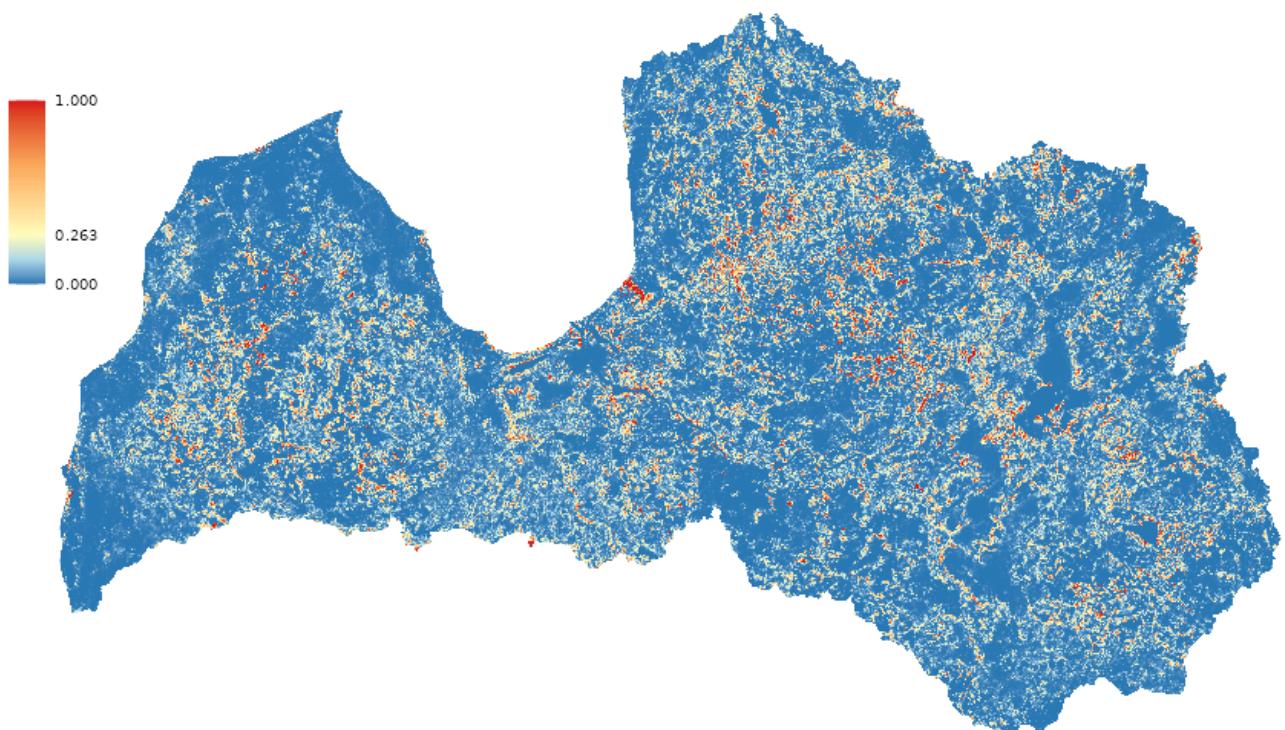
Sāku ar v3 - v5 HSmaps:

Hymenoptera: BOMHUM



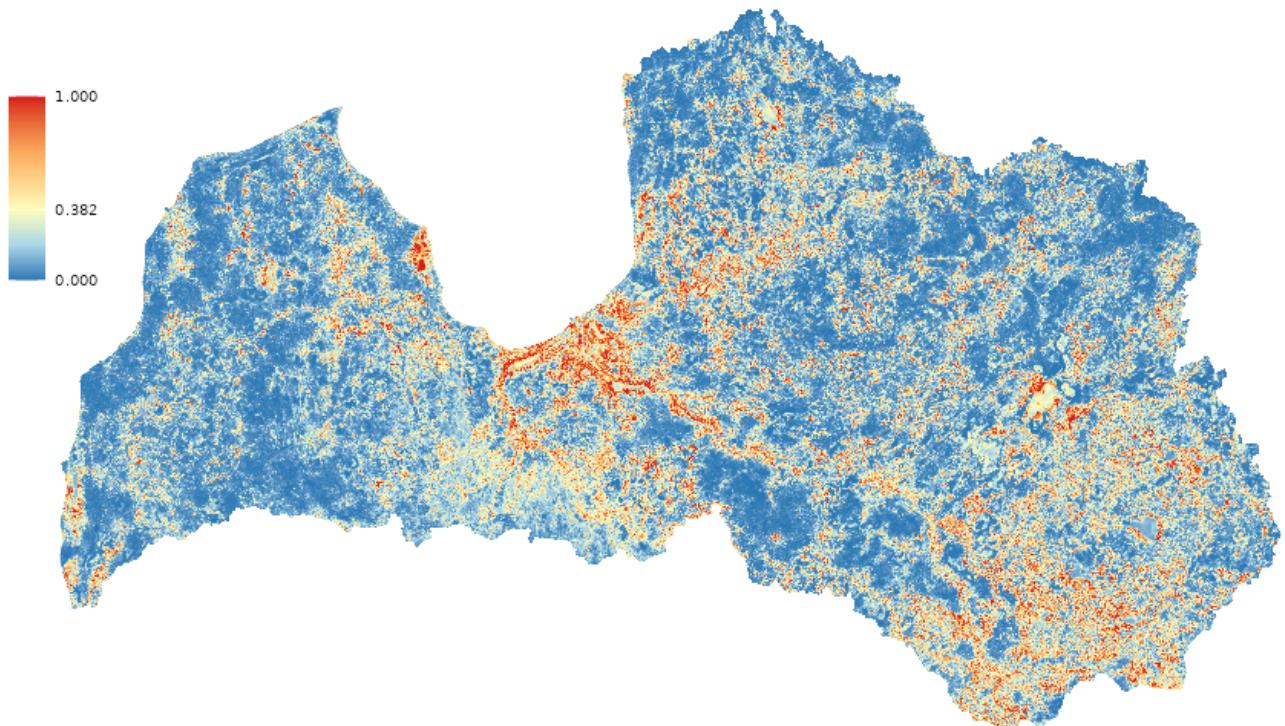
BOMHUMv3 HMap

Hymenoptera: BOMHUM



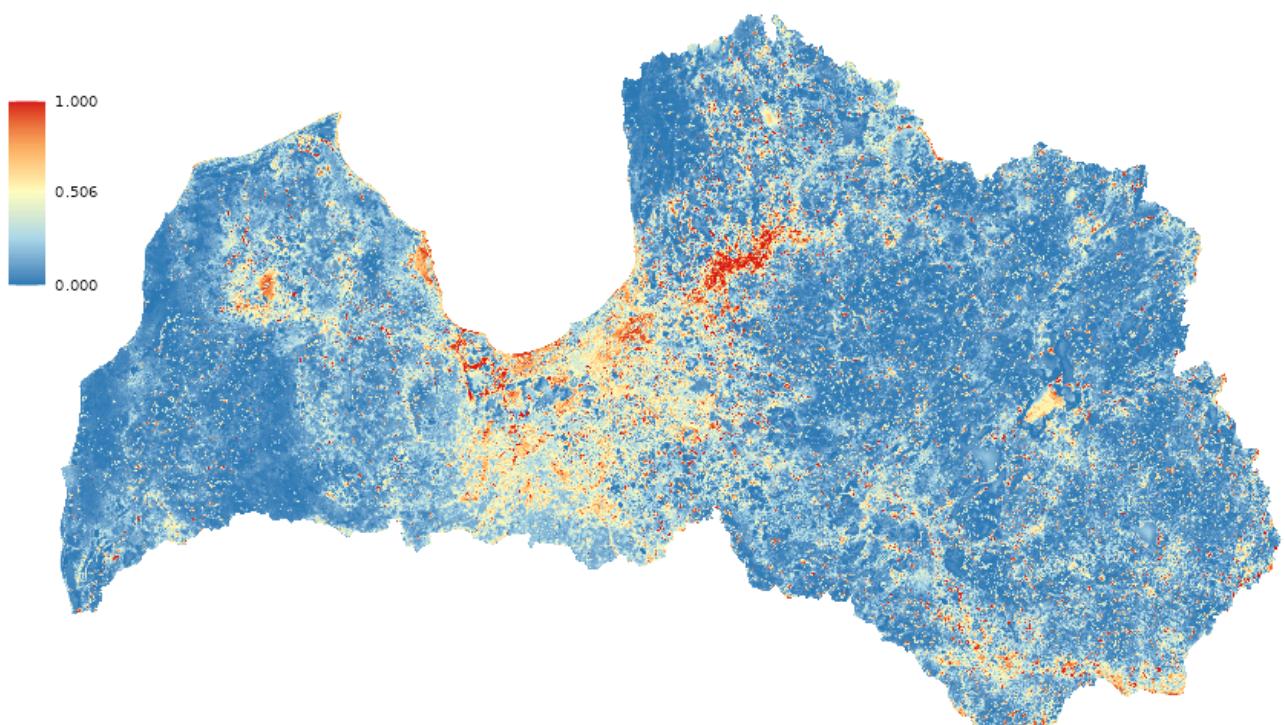
BOMHUM v4 HMap

Hymenoptera: BOMHUM

**BOMHUM v5 HSmap**

v5 atkal ir "aizvilcis" uz Rīgu. Šķiet diezgan loģiski - ja bias slānim ieviestā apakšgala limitācija ir vidējais (būtībā tas nozīmē to, ka visas vērtības rastrā, kas ir zem vidējā, tiek aizvietotas ar vidējo vērtību), un novērojumi ir ap pilsētām un ceļiem (kas sagaidāms DD novērojumu dēļ), tā ir tā sakarība, ko modelis atrod. Sanāk, ka šeit fona punktu izloze ir pietuvināta tam variantam, kas ir nejauša izloze (bez bias slāņa). Šī iemesla pēc modeļus ar šo apakšgala limitāciju (v5 un v8) neizvēlēšos kā labākos un vairs īpaši neapskatīšu. Bet intereses pēc v8 karte:

Hymenoptera: BOMHUM

**BOMHUM v8 HSmap**

Nujā.

Kas attiecas uz v3 un v4 - no kartes grūti tā uzreiz pateikt, kas varētu noteikt izplatību. Apskatīšu mainīgo nozīmību, vairāk uzmanības pievēršot v4, jo teorijā tas varētu būt labāks variants:



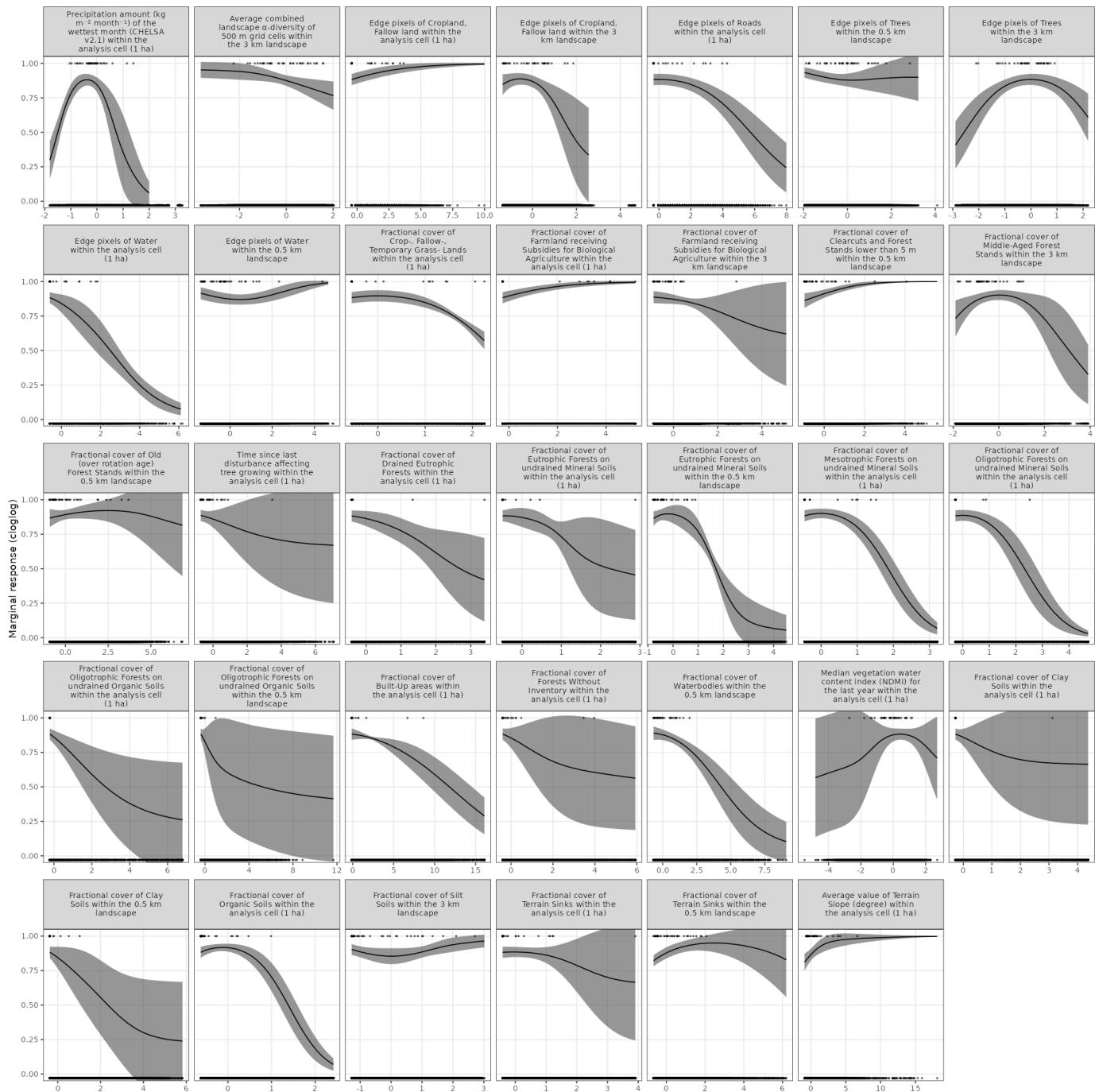
v3

BOMHUM_salidzinajums

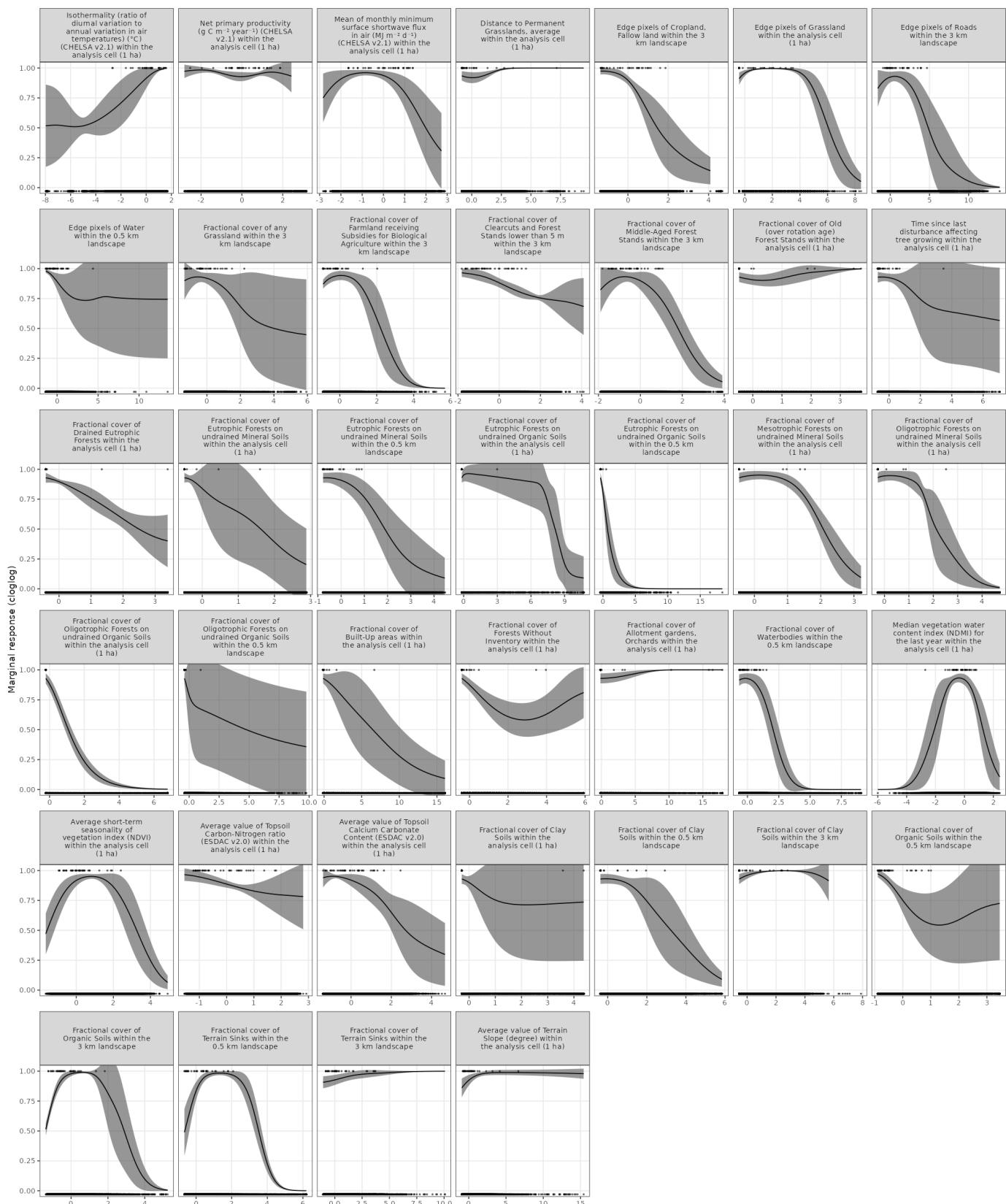


v4

Abos nozīmīgākais ir konkrētu tipu mežu īpatsvars analīzes šūnā, lai gan starp variantiem atšķiras, par kuriem tipiem ir runa. Bet šo jāapskata kopā ar atbildes līknēm:



v3

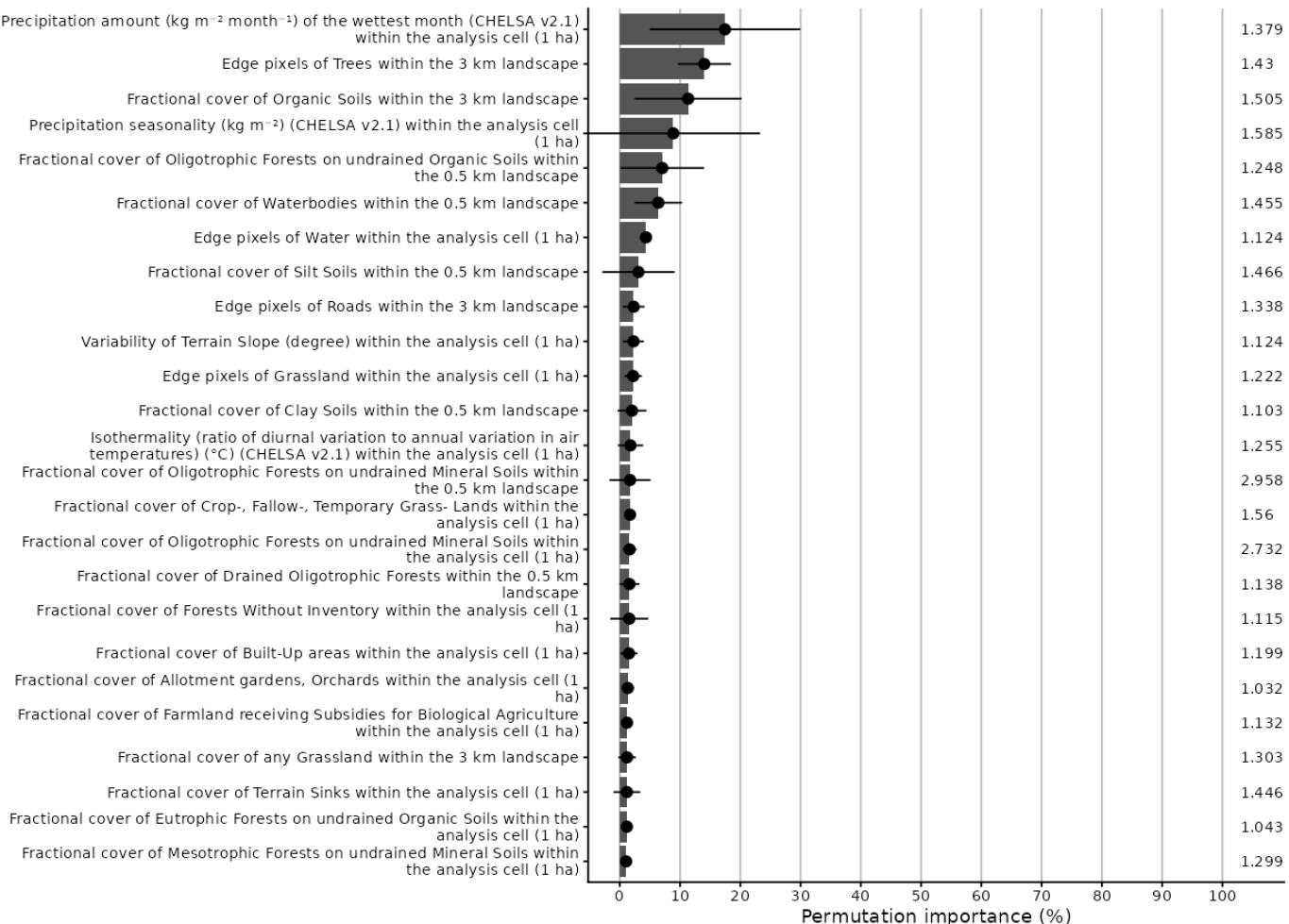


v4

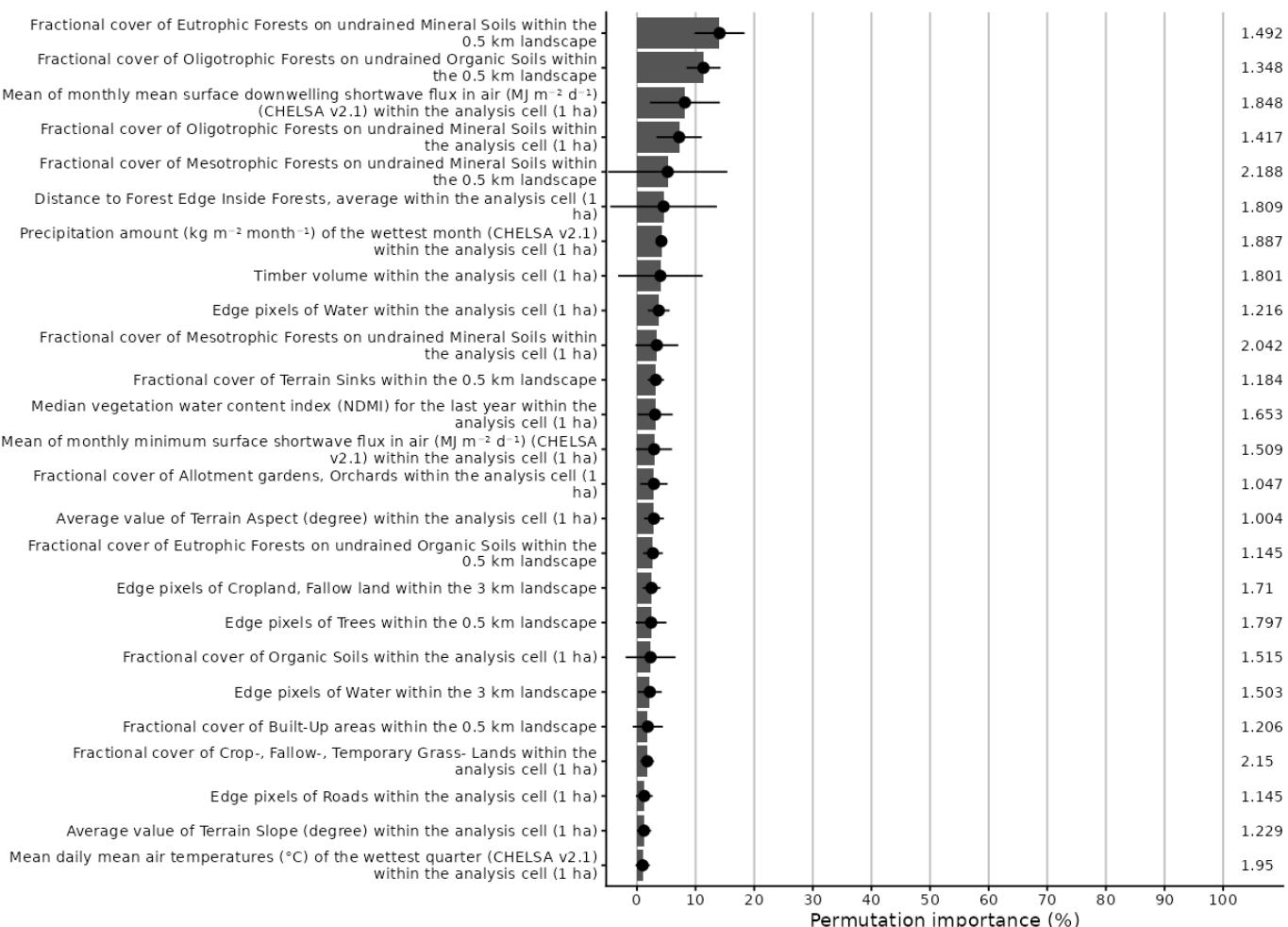
Kopumā atbildes līknes, straujāk vai lēnāk dilstošas, tomēr norāda uz to, ka jebkāda veida lielāks meža īpatsvars ir slikti. Diezgan loģiski, nemot vērā, ka malu suga - lai gan tad es līknes sagaidītu vairāk kā v4 versijas "Fractional cover of Mesotrophic forests on undrained mineral soils within the analysis cells" - mazliet mežs ir labi, jo, ja mežs ir, tad ir arī meža malas. Katrā ziņā šobrīd drīzāk sanāk, ka šī ir atklāto vietu un arī malu suga, nevis tikai malu suga. Tomēr v3 koku malu pikseļu skaita 3 km ainavā atbildes līkne rāda, ka vidējs malu daudzums ir labi, maz vai daudz - slikti. Tā ka tomēr ar malām saistīs, lai gan šī konkrētā EGV nozīmība neuzrādās īpaši liela. No v4 šis mainīgais ir izmests....

v4 trešais svarīgākais EGV ir malu pikseļu skaits "Cropland, Fallow" 3 km ainavā. Atbildes līkne tomēr rāda, ka daudz malu ir slikti tomēr mazliet malu ir labi (līkne sākumā nekrīt strauji), kas +/- sakrīt ar sagaidāmo.

Pie šī ilgi varētu sēdēt un pētīt, bet tagad apskatīšu v6 un v7. Sagaidu atšķirīgu situāciju no v3 un v4.

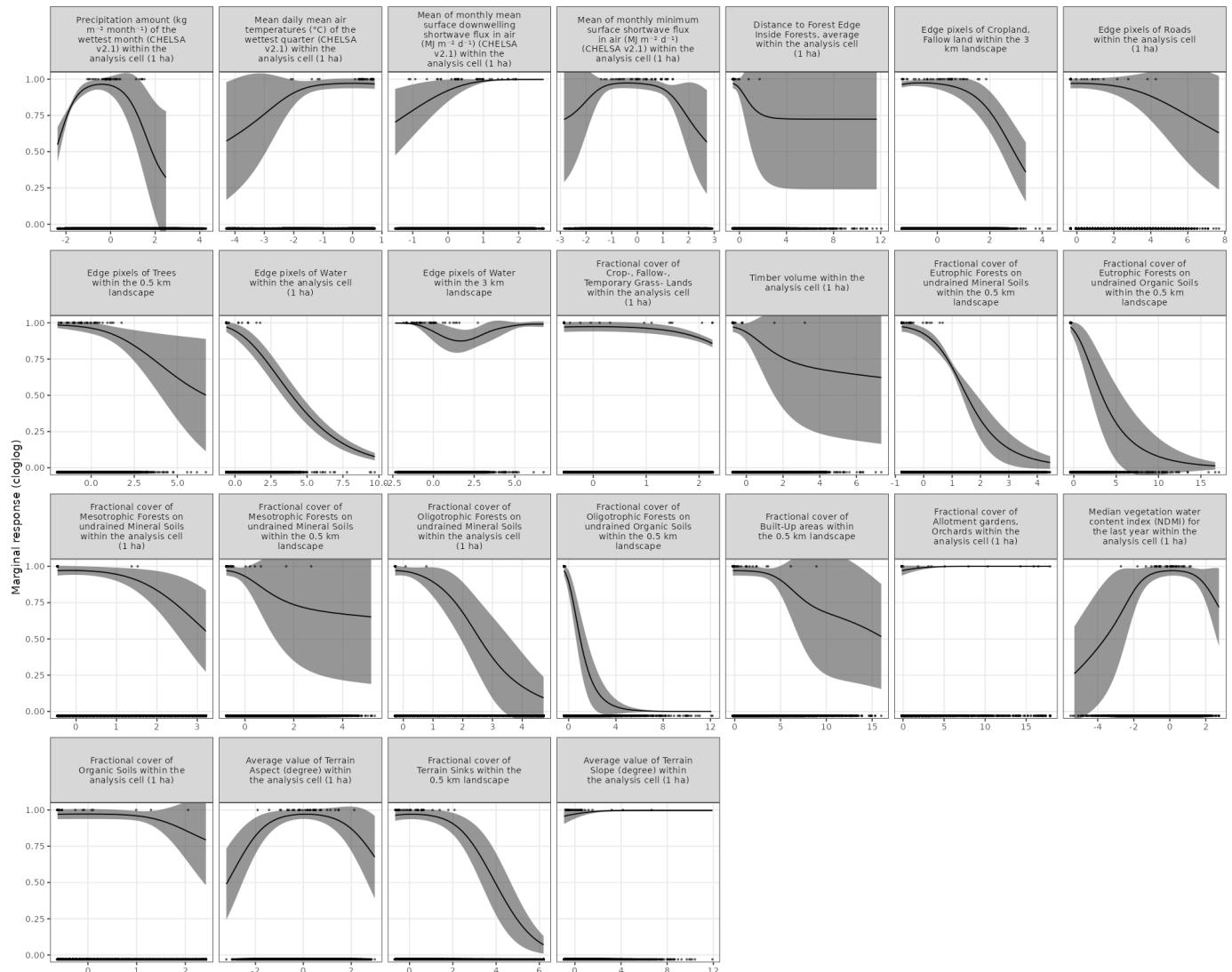


v6



v7

Vispār v7 rāda kaut ko līdzīgu v4. Vienīgi tas, ka svarīgākie nav vairs šūnu līmeņa, bet 0,5 km rādiusa meži. Un kā ar atbildes līknēm?



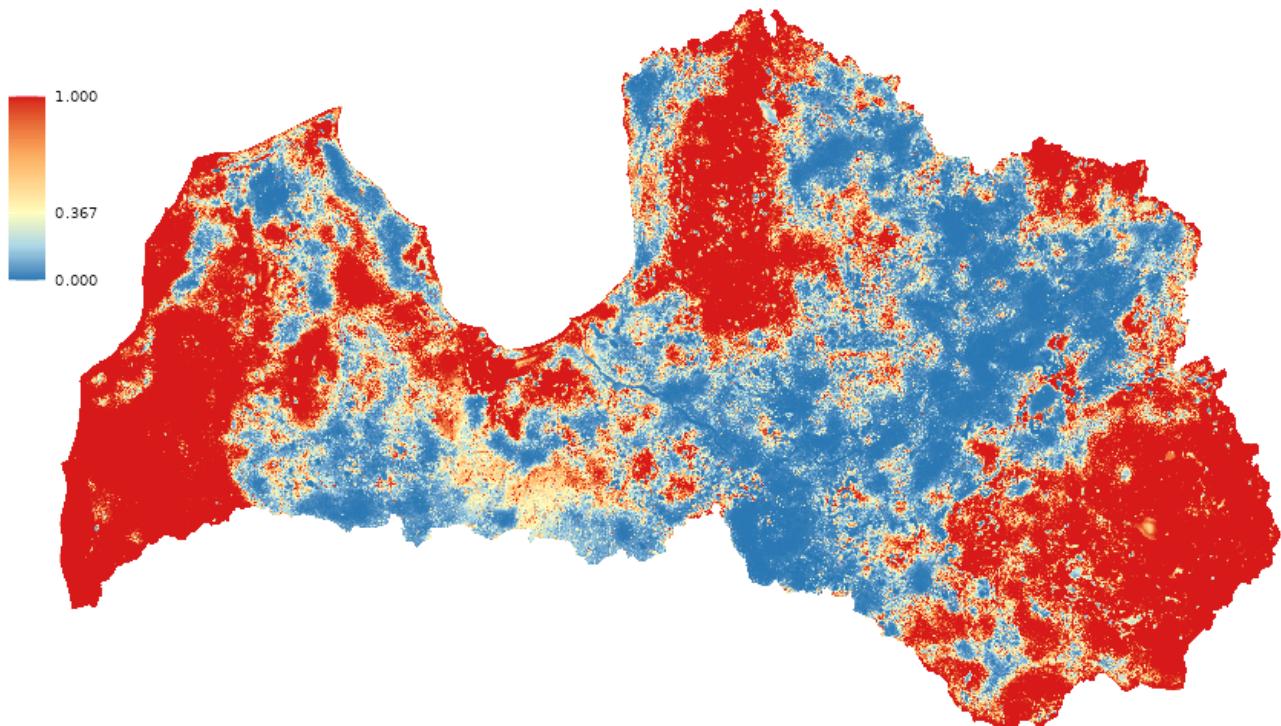
v7

Vispār līdzīgas tendences. Vienīgi mezotrofajiem mežiem līkne šūnas un 0,5 km ainavā nekrīt tik strauji kā citiem mežiem - arī vidējs un liels mežu īpatsvars ir ok, lai gan mazāk ir piemērotāks.

v7 trešais svarīgākais mainīgais ir vidējā dienas vidējā surface downwelling shortwave flux in air - vidējā Zemes virsmu sasniedzošā saules radiācija šūnā. Vispār interesanti un loģiski, ja skatās uz atbildes līknī - vairāk ir labi, lai gan arī maz ir ok (tomēr plats ticamības intervāls). Attiecināms gan uz ziedaugiem un to augšanu (barības resursi), tāpat agri sezonā, gan pašām kamenēm - mazāk jāpatērē enerģija siltuma uzturēšanai (kamenes spēj lidot tikai tad, ja ķermeņa temperatūra ir virs aptuveni 30 grādiem, pirms lidošanas vēsākā laikā sevi uzsilda, izmantojot krūšu daļas muskuļus un arī ligzdu viņas uztur siltu, ap 30 grādiem).

v6 izteikti svarīgākais uzrādās mitrākā mēneša nokrišņu daudzums - nepatīk, jo klimata mainīgais, ko arī nevar tik labi izskaidrot. Un arī karte (zemāk)... nu ārprāts :D

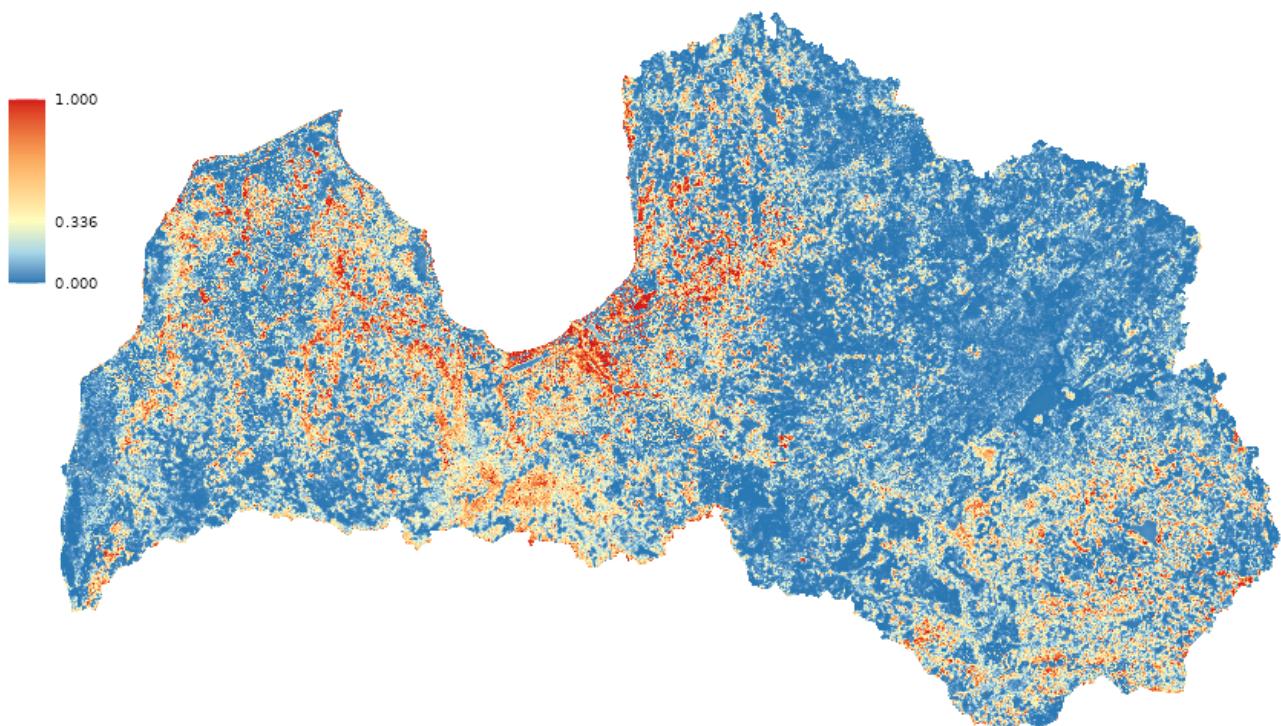
Hymenoptera: BOMHUM



v6

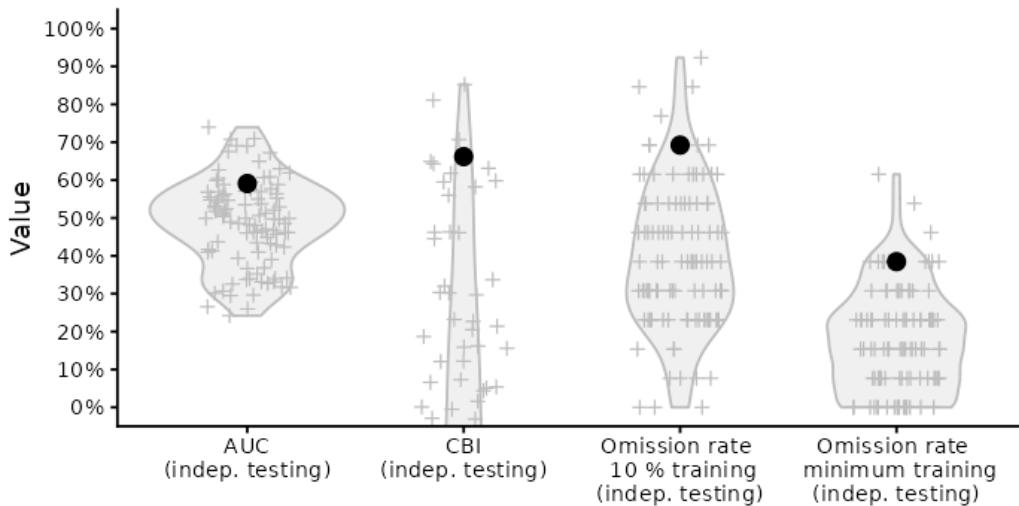
Nu, katrā ziņā šķiet, ka v7 varētu būt labākais variants. Bet karte rāda, ka atkal diezgan aizvilkts uz Rīgu :(

Hymenoptera: BOMHUM

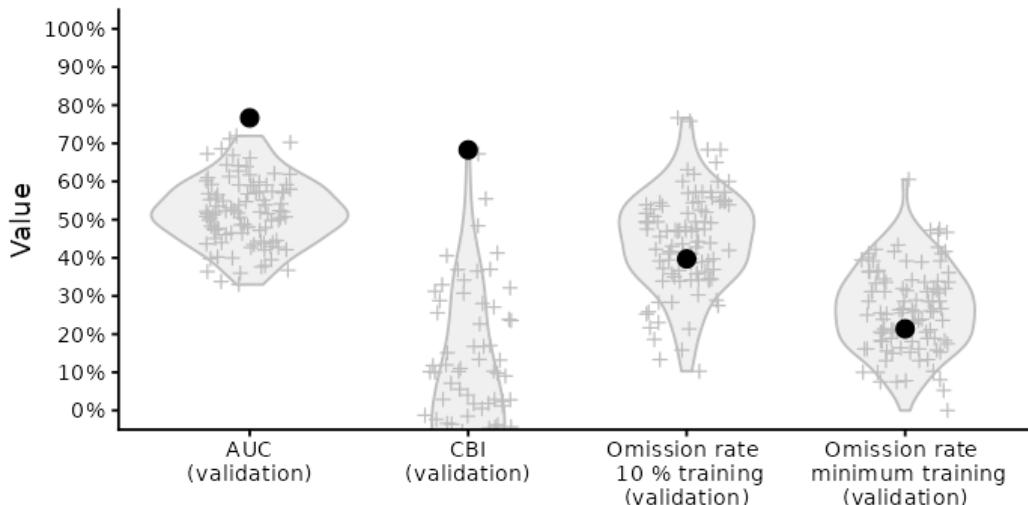


v7

Vispār, līdz šim gandrīz nemaz neesmu skatījusies uz modeļu metrikām. Es labprāt apspriestu, neatkarīgā testa vs validācijas vērtības - kam vairāk pievērst uzmanību, sagaidāmās atšķirības, u. tml.



Nulls test v7



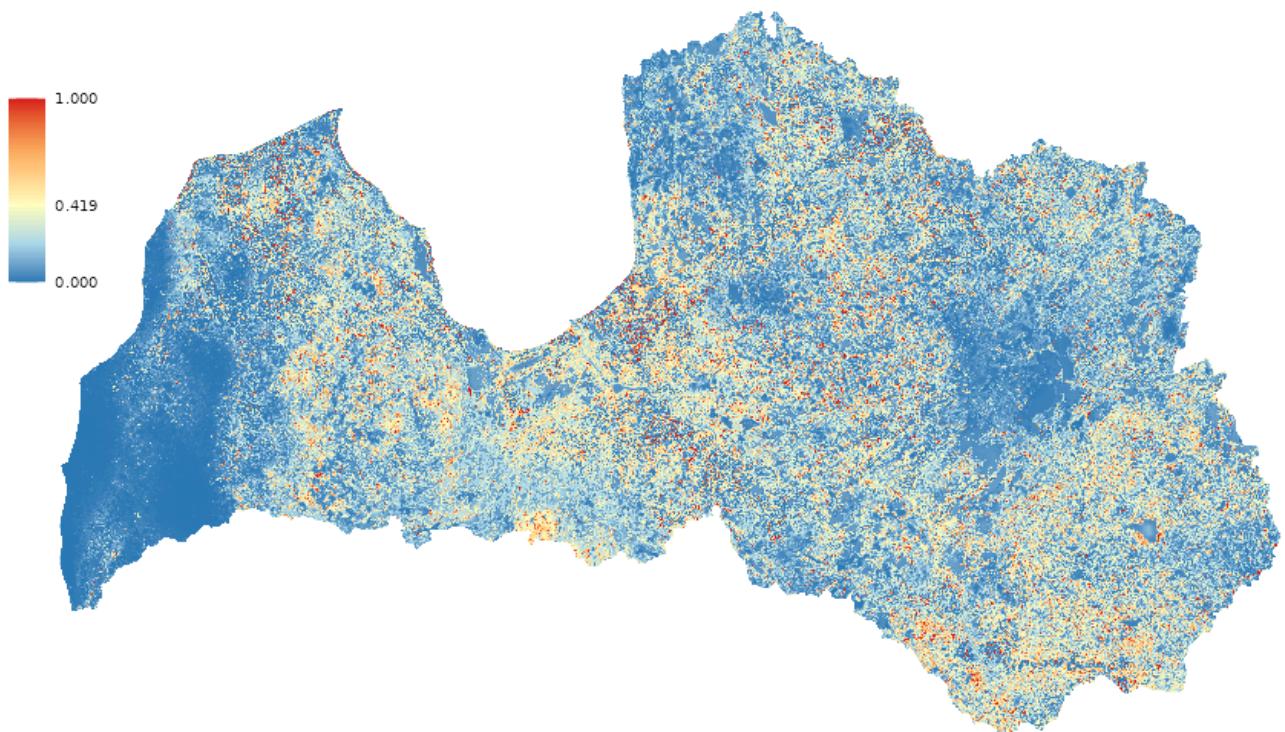
Nulls val v7

AUC un CBI ir ok, jo nulles modeļu augšdaļā un vērtības šķiet diezgan augstas (varbūt izņemot testa AUC). Omission rate gan neiepriecina, īpaši neatk. testa datos - diezgan augstas vērtības, acīmredzot atsevišķiem novērojumiem ir liela ietekme uz rezultātu, modelis varētu būt pārpielāgots. Par to gan neliecina atbildes līknes (nav jocīgu, neizskaidrojamu "zobu").

Kopējais iespaids no visiem variantiem - ir diezgan lielas atšķirības. Bet jāpieņem lēmumi.

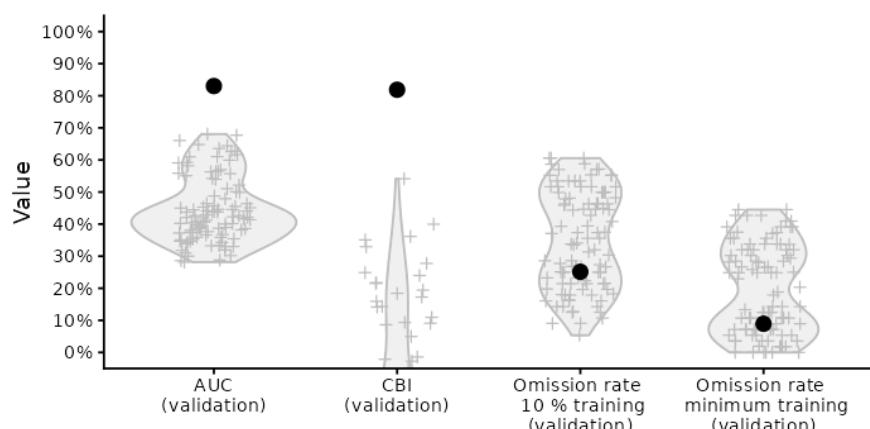
Uztaisīju v7_1, kas ir tas pats, kas v7, tikai bez telpiskās filtrēšanas, karte izskatās šādi:

Hymenoptera: BOMHUM



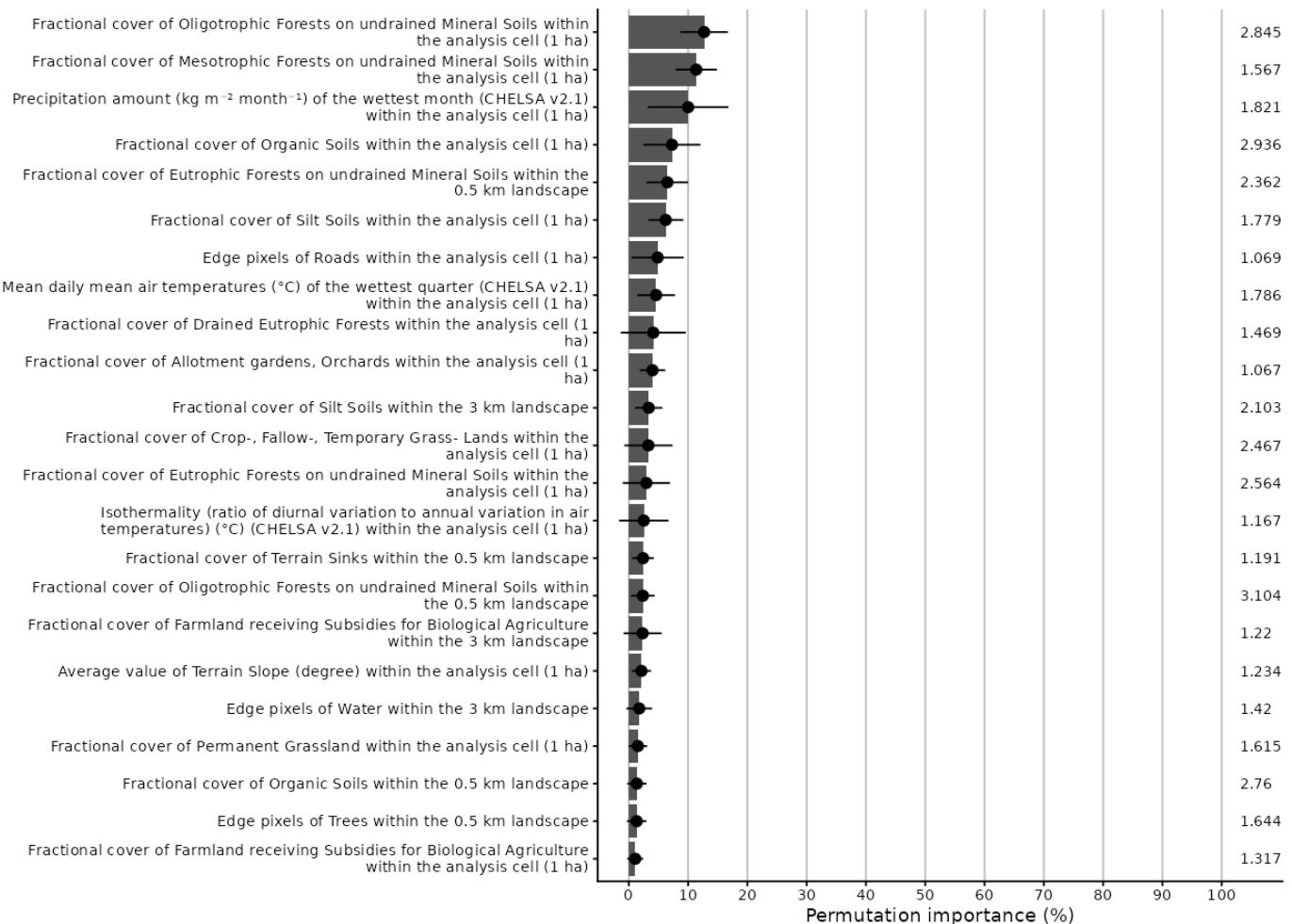
v7_1

Karte ievērojami atšķiras no v7. Vairs nav pilsētas, bet ir jocīga Kurzeme. Spekulēju, ka tas vismaz daļēji ir dēļ nokrišņu daudzuma mitrākajā mēnesī. Katrā ziņā man ne pārāk patīk. Krieti pacēlies AUC (no 0,503 uz 0,651). Atbildes līknēm smukāki ticamības intervāli, līknes līdzēnas. Omission rate vērtības un novietojums uzlabojies:

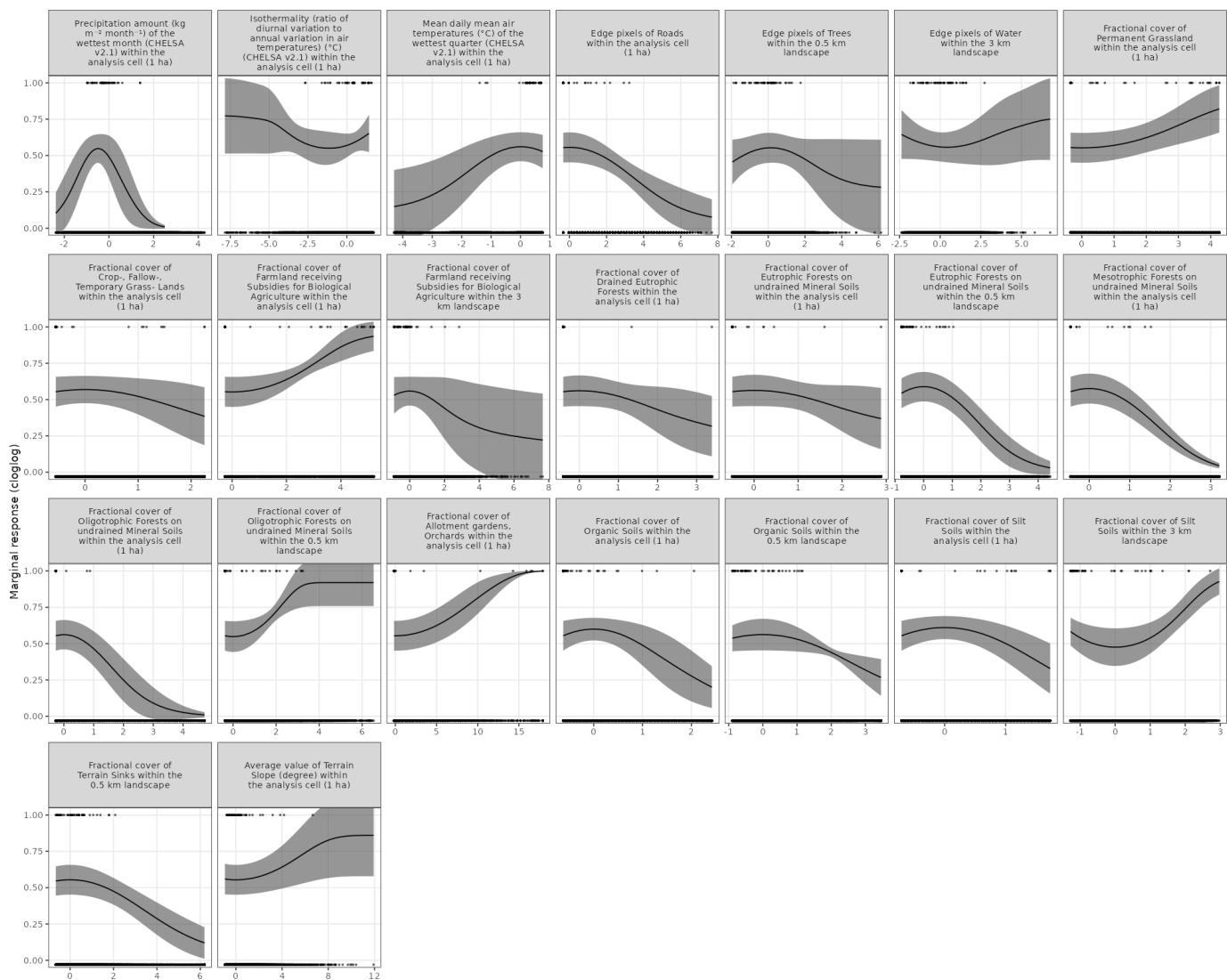


v7_1

Un mainīgo nozīmība:



v7_1

**v7_1**

Svarīgs kļuvis organisko augšņu platības īpatsvars analīzes šūnā - līkne lēzena, tomēr rāda, ka mazāk ir labāk. Ainavā (3 km) pozitīva ietekme ir lielākam augsnēs granulometriskās klases "smilšmāls un mālsmilts" platības īpatsvaram. Tomēr analīzes šūnā ir otrādi.

Tomēr jebkura tipa mežu īpatsvara kopējā tendence (stāvāka vai lēzenāka, tomēr vairāk = slikti, lai gan mazliet ir labāk nekā necik) liek domāt, ka mežus varētu apvienot zem viena EGV, neizdalot tipus, tomēr paturot augsnēs granulometriskā sastāva EGV. Lai gan tagad skatos, ka oligotrofajiem mežiem homorange līmenī ir pozitīva ietekme - nevienā citā mērogā šie meži gan nav saglabāti.

Pagaidām tad strādāšu tālāk ar versiju v7_1. Varētu izvērtēt arī rādiusa ietekmi bias veidošānā (v7_2 un v7_3, tos arī uztaisīju) tomēr pagaidām to atlieku un pieturos pie 1,5 km.

Tātad versija v7_1a - apvienošu mežus zem viena EGV. Skatos, ka esmu dalījusi arī vecumgrupas, bet nekas no tā nav saglabājies - to arī izmetīšu, saglabājot pēdējā traucējuma laiku un kopējo krāju, kam vajadzētu būt pietiekami šajā gadījumā. Izsekojamībai - ko nēmu nost un ko liku klāt - zemāk skripts, ko pievienoju v7_1a:

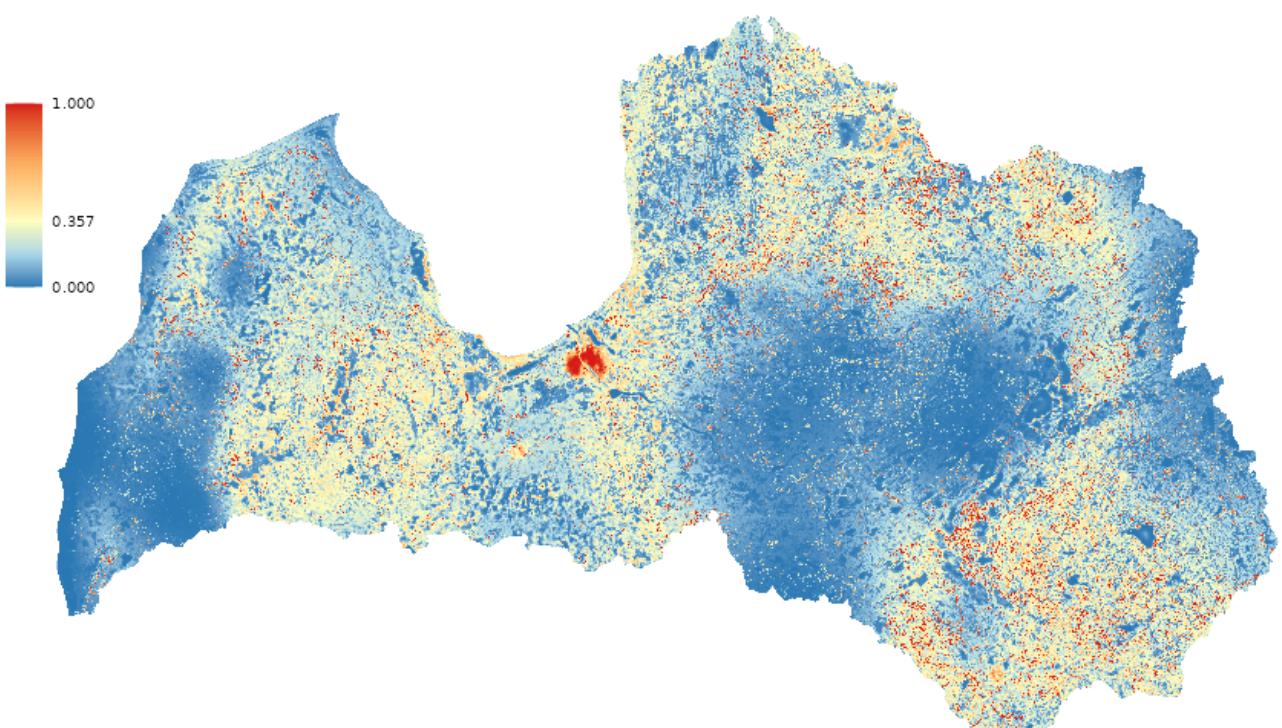
```

izsledzamie_egv=c("egv_270",
                 "egv_271",
                 "egv_273", # izcirtumi un Līdz 5m meži visos mērogos
                 "egv_275",
                 "egv_276",
                 "egv_278", # vidēja vecuma un briesaudzes visos mērogos
                 "egv_280",
                 "egv_281",
                 "egv_283", # vecas audzes visos mērogos
                 "egv_285",
                 "egv_286",
                 "egv_288", # krūmāji un jaunaudzes visos mērogos
                 "egv_308", # un visi meža tipi secīgi kā tabulā 0 un 500 m
                 "egv_309",
                 "egv_313",
                 "egv_314",
                 "egv_318",
                 "egv_319",
                 "egv_323",
                 "egv_324",
                 "egv_328",
                 "egv_329",
                 "egv_333",
                 "egv_334",
                 "egv_338",
                 "egv_339",
                 "egv_423", # un vēl netaksētie
                 "egv_424")
pievienojamie_egv=c("egv_449",
                     "egv_450") # visi koki 0 un 500 m rādiusā, jo ainavā (3 km) jau bija

```

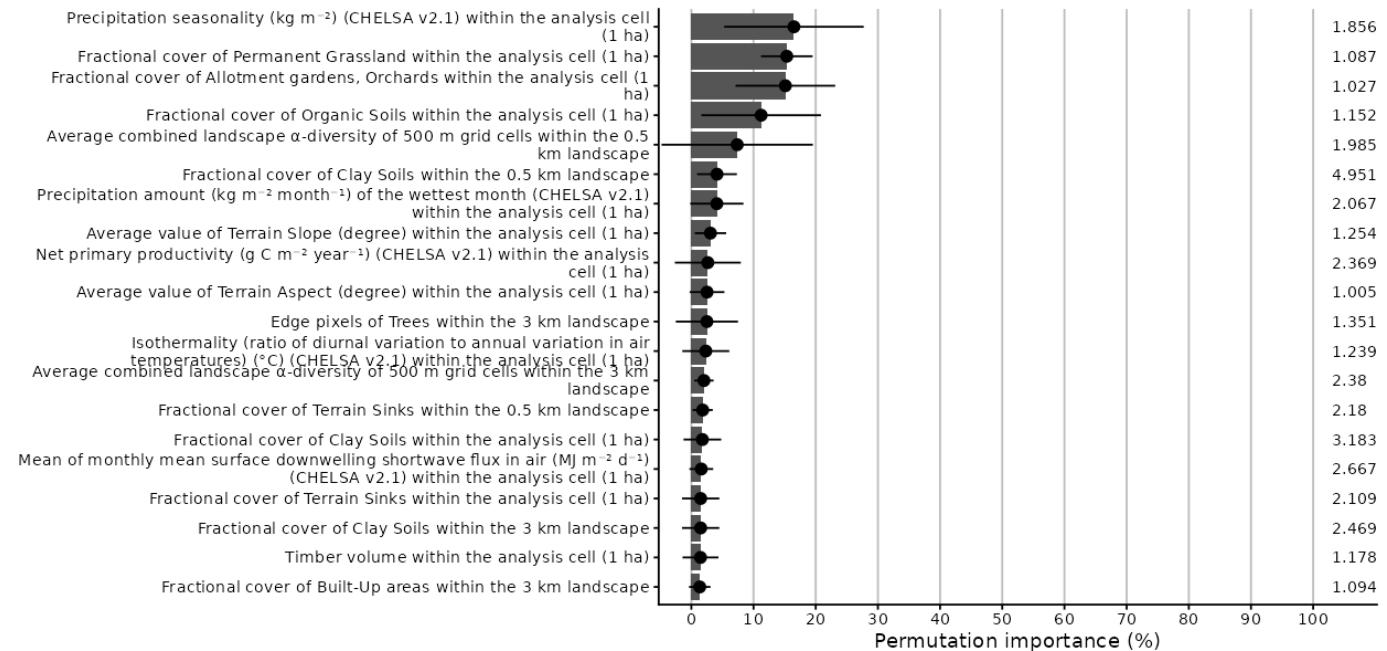
Rezultāti - ahem... nē, paldies.

Hymenoptera: BOMHUM

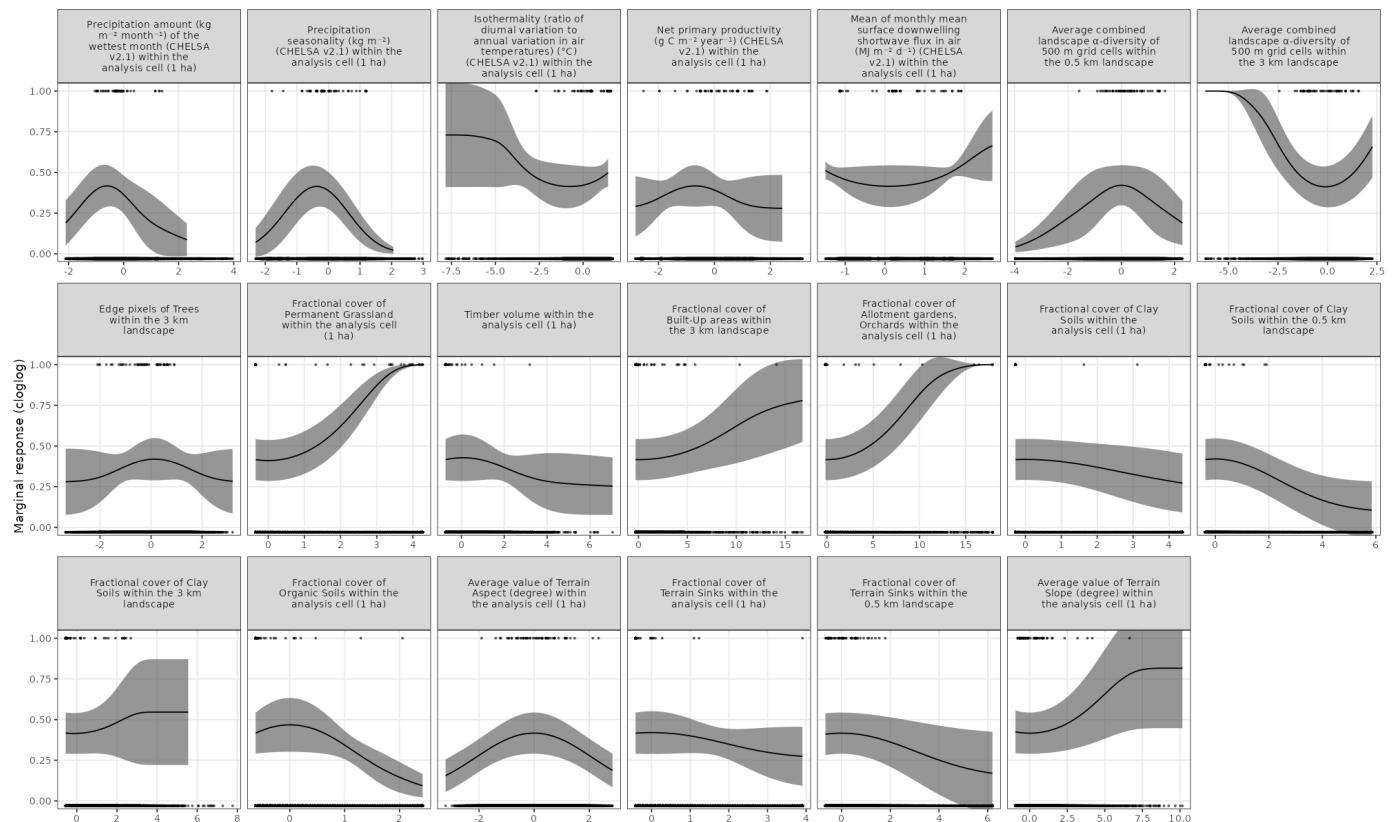


v7_1a

Meži kā tādi vispār izsviesti ārā jebkurā mērogā:



v7_1a

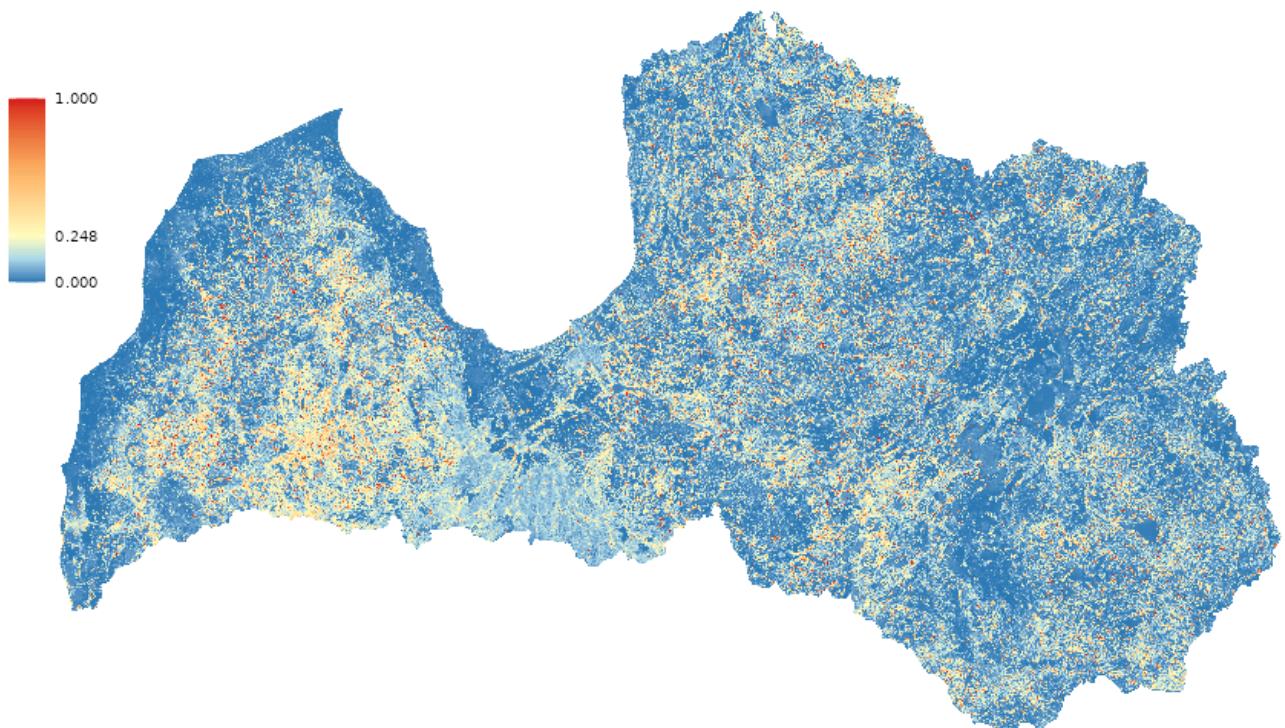


v7_1a

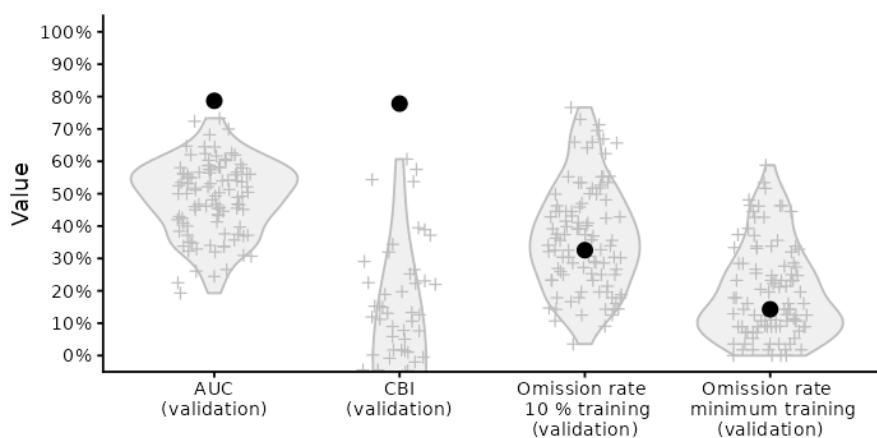
Taisnību sakot, es baigi negribu uz šito vairs skatīties.

Virzoties uz izvēli par labāko modeli - sākumā domāju par v7, bet ņemot vērā pārdomas šīs izdrukas sākumā par bias raksturošanu (kuras man radās pēc tam, kad darbojos ar v7 un v7_1), tomēr sliecos uz v4. Esmu uztaisījusi arī v4_1, kas ir tas pats v4, bet bez telpiskās filtrēšanas 1 km šūnā:

Hymenoptera: BOMHUM



v4_1



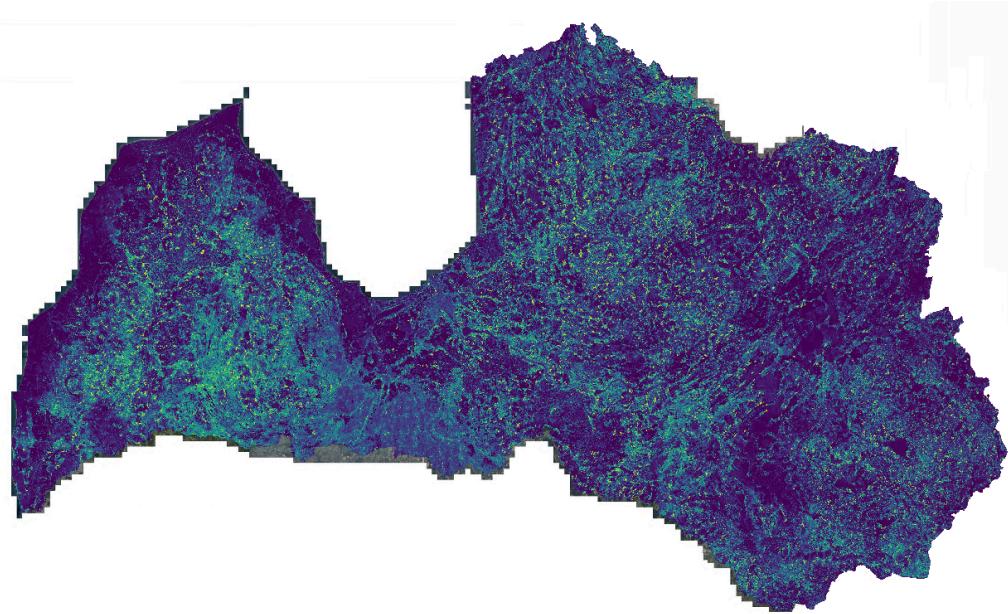
v4_1

Papētīšu dzīvotņu piemērotības kartes GIS, lai pavirzītu uz priekšu domas. Salīdzinu v4 un v4_1. Bildēm visur secība v4 - v4_1. Atziņas:

- Lai gan sākotnēji tā neizskatījās, v4 un v4_1 nav ļoti konfliktējoši un izzīmē līdzīgas vietas, tikai v4_1 krieti vairāk vietas ir ar vidēji augstu piemērotību, kamēr v4 ir kontrastaināks. (bildes zemāk) Abos augstāka piemērotība ir atklātām teritorijām, ne mežiem. Bet nevienā no modeļiem es īsti nevaru saprast, kas padara augstās piemērotības vietas par tik augstām. Ir šur un tur kaut kur tās vietas izmētātas;

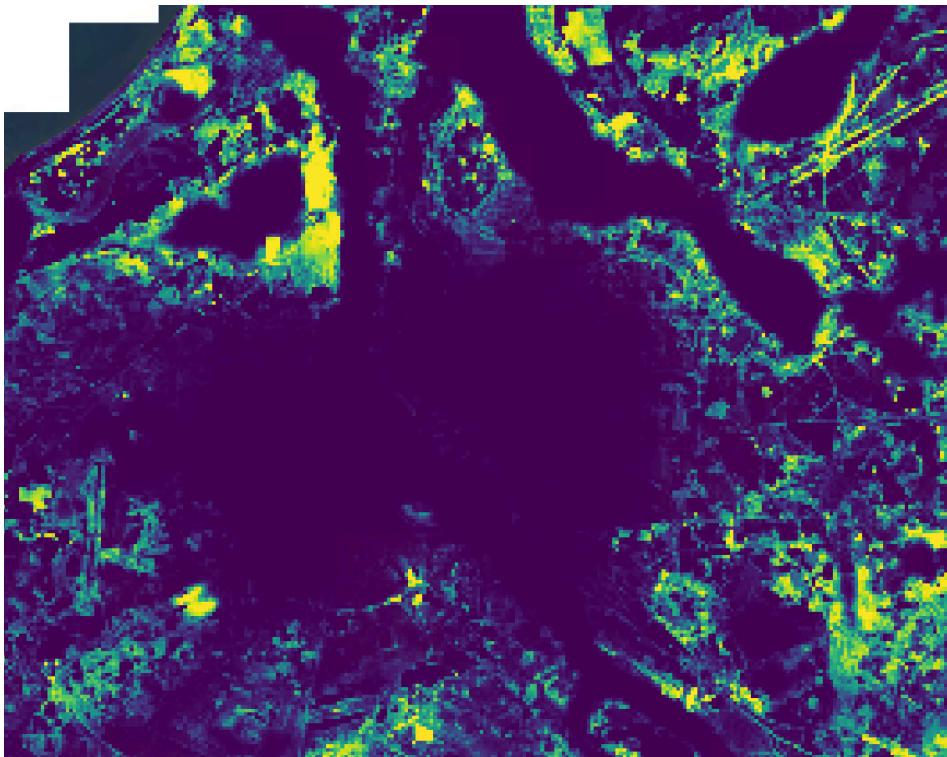


v4

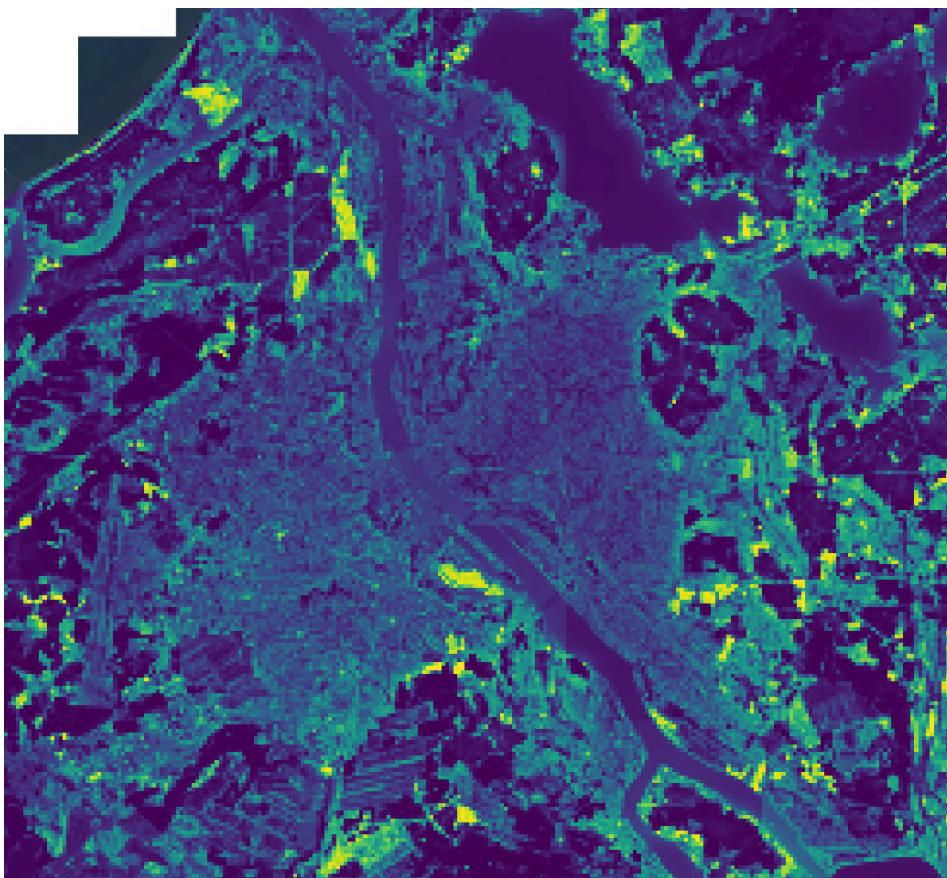


v4_1

- Priecē, ka abos modeļos ūdeņi ir nepiemēroti un skatot Rīgu apbūvē v4 ir ļoti zemas vērtības, v4_1 minimāli augstākas tajās vietās, kur apbūve drusku mazāk blīva - sliecos vairāk piekrist **v4_1**, jo pilsētās šī kamene varētu būt sastopama un ir redzēta, lai gan pareti, un es nesagaidu, ka Rīga būs tāds melnais caurums, kā zīmējas v4;



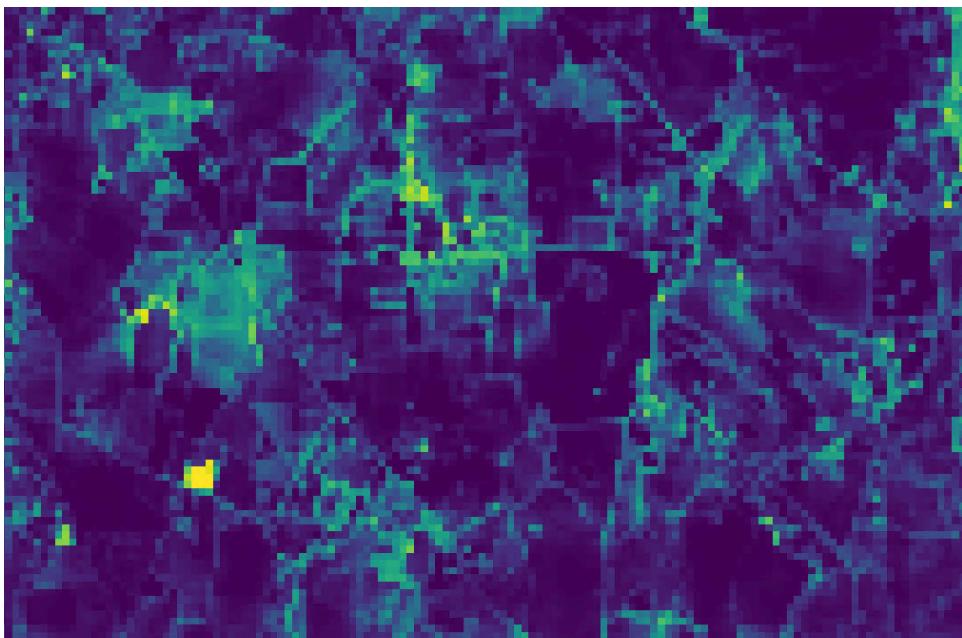
v4



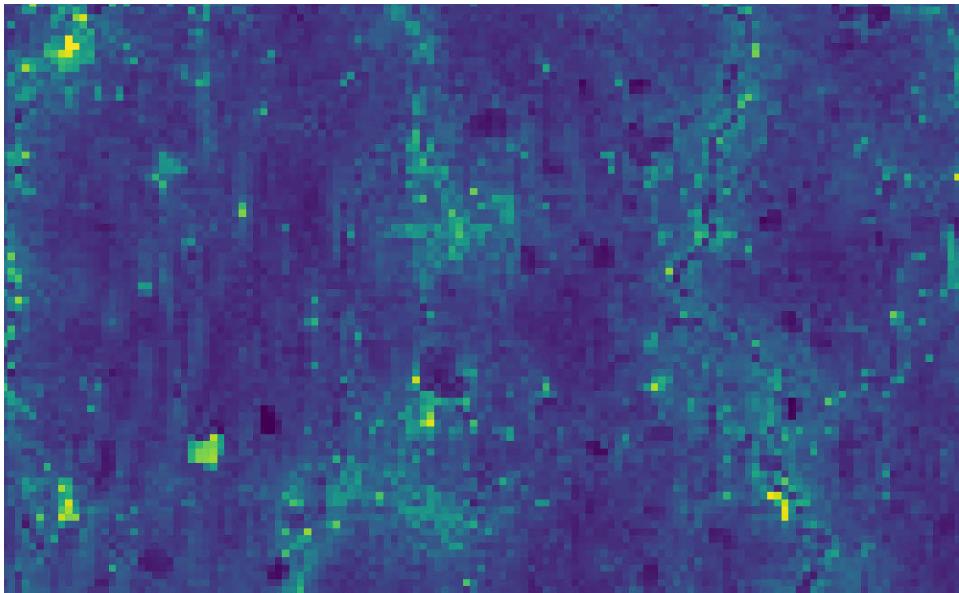
v4_1

- Abos variantos purvi ir diezgan nepiemēroti - jauki;
- Mani neiepriecina tas, ka abiem modeļiem Kurzemē jūras tuvumā vērtības kopumā ir ievērojami mazākas nekā citur, tāpēc izteikti tas ir v4_1 - man nav, kā to ekoloģiski pamatoti izskaidrot. Kaut ko dara klimats, tikai es vēl nesaprotu, kurš mainīgais;
- v4 ar ļoti augstu piemērotību izzīmējas Ādažu poligons. It kā virsāji un tā varētu būt piemērota vieta, lai gan grūti skaidrot, kāpēc tieši tur ir tik augsta piemērotība. Arī blakus esošie sausie meži iezīmējas kā ļoti piemēroti;

- Ja skatās intensīvi apsaimniekoto Zemgali, v4 ir vairāk redzami ceļi, koku un krūmu puduri, kā arī mājas starp laukiem (kā augstāk piemēroti), sliecos tam piekrist, kamēr v4_1 ir vienmērīgāks ar kopumā vidēji zemu piemērotību. Spilgtās vietas +/- sakrīt (tie ir ābeļdārzi, mājas, koku puduri un tā). Sliecos vairāk piekrist **v4**, jo šādā ainavā ceļmalas ir svarīgas;

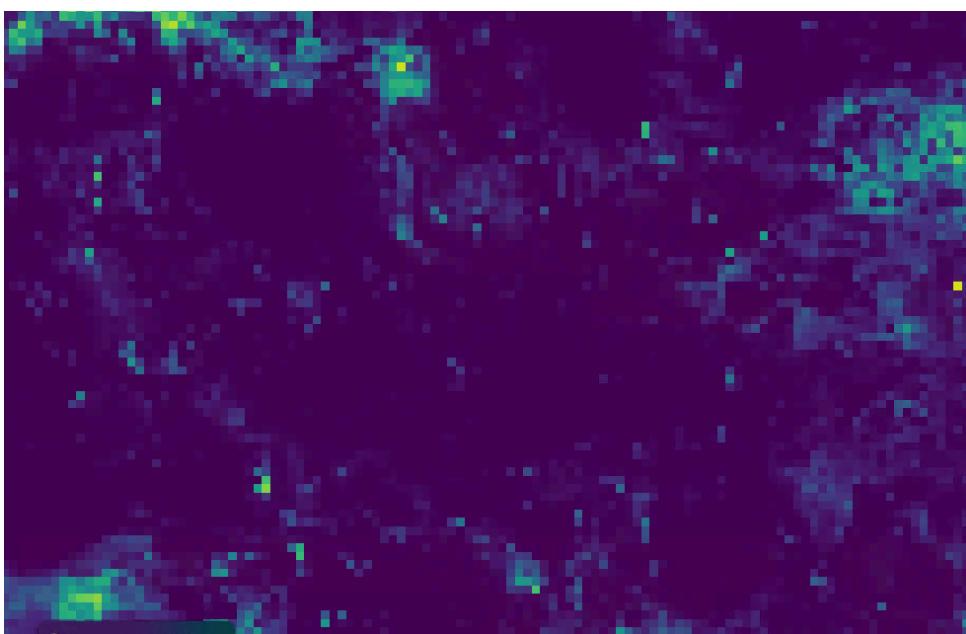
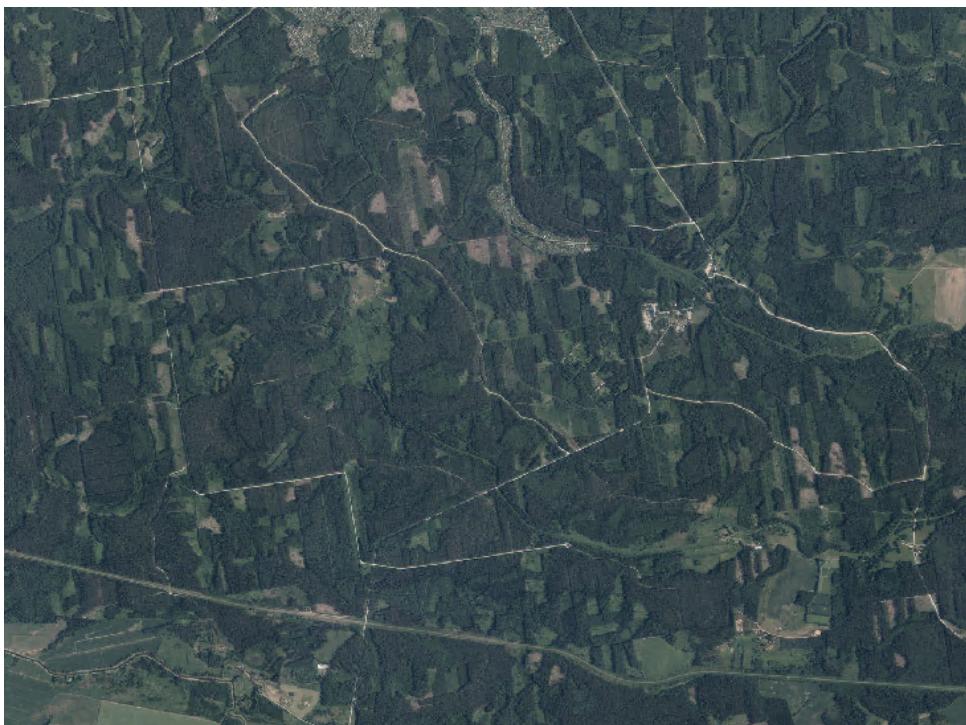


v4

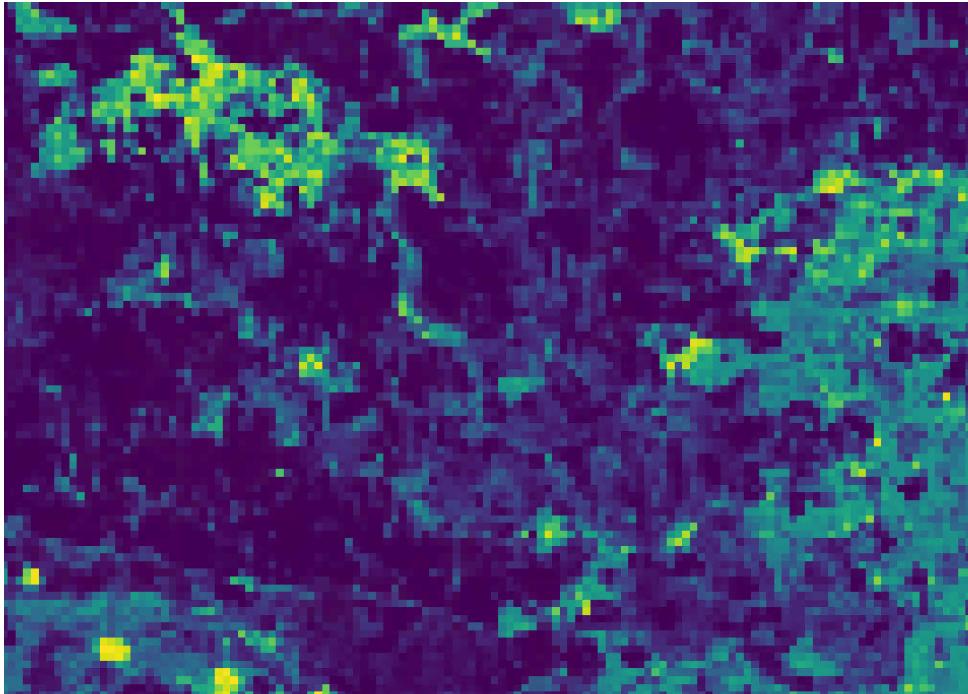


v4_1

- Meži kopumā ir diezgan nepiemēroti, bet v4 tajos ir vispārīgi zemākas vērtības, kamēr v4_1 "redz" visādus izcirtumus, skrajākus kokus, ceļus, atklātās vietas starp kokiem, kas teorijā būtu sagaidāms kā piemērotāks; tādā ziņā **v4_1** šķiet it kā labāks;

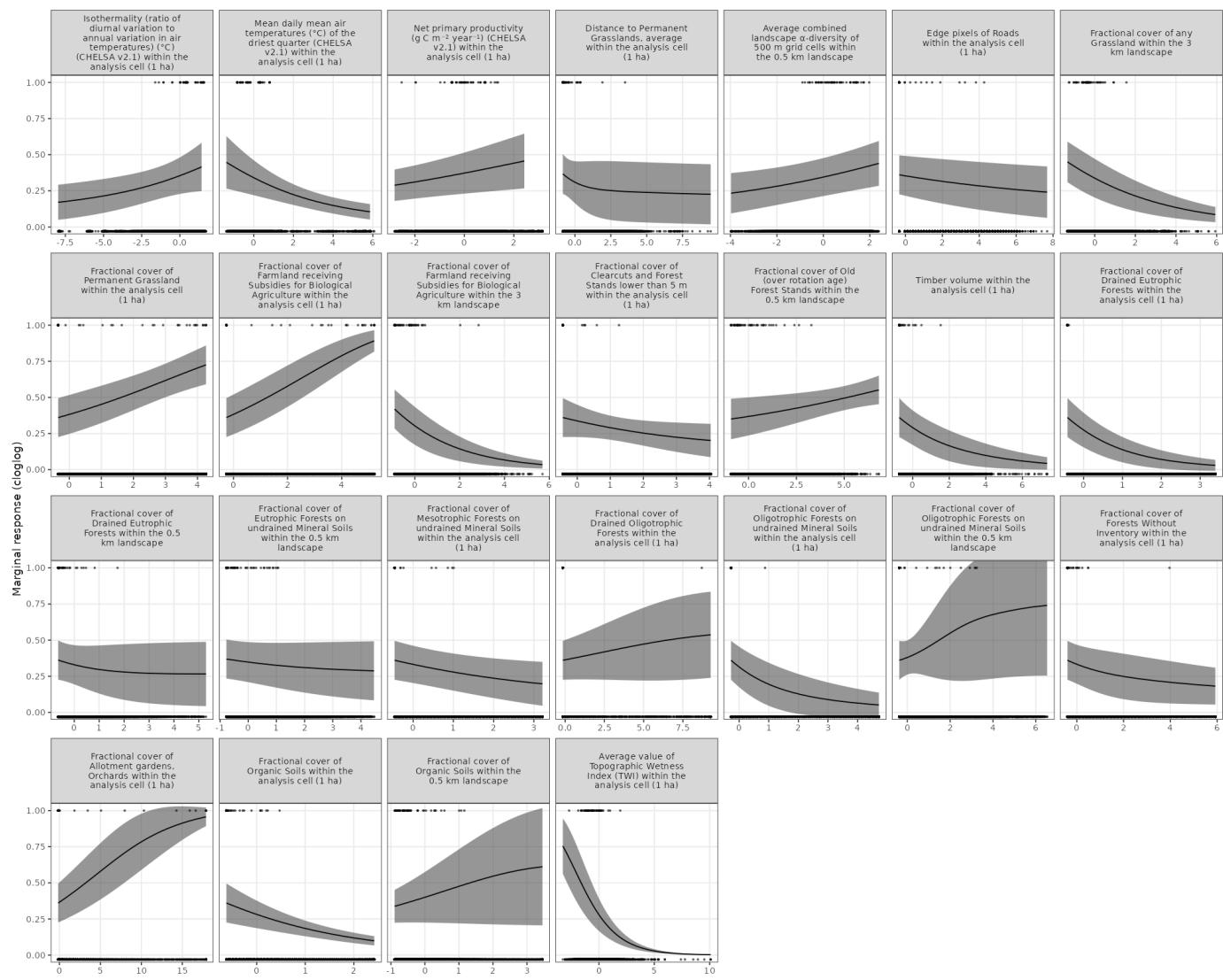


v4



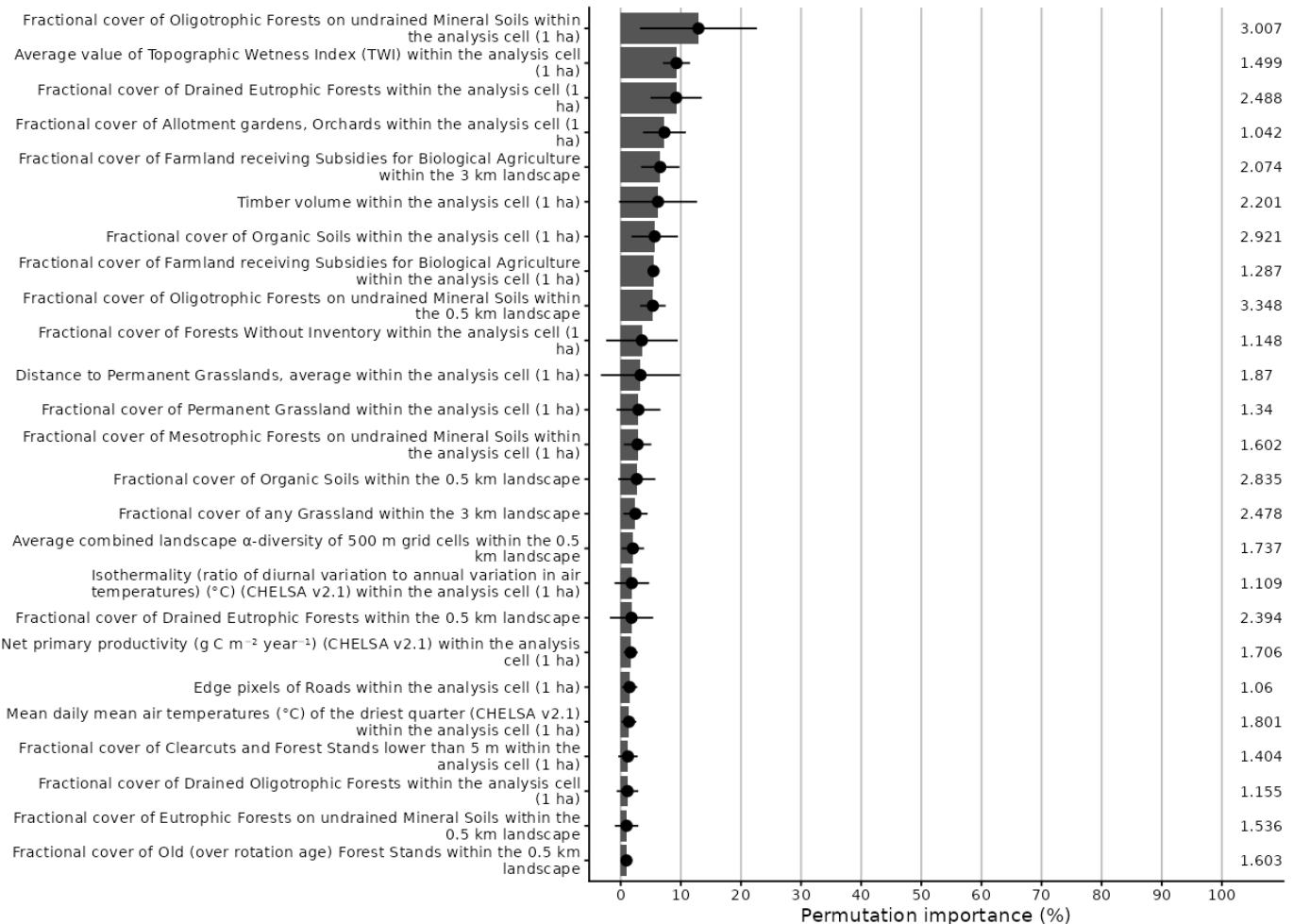
v4_1

Nemot vērā visu pārskatīšanu, vairāk gribētos piekrist v4_1, kur nav veikta telpiskā filtrēšana 1 km tīklā. Arī AUC ir mazliet labāks (v4 - 0.508, v4_1 - 0.565). Papētīšu atbildes līknes un nozīmību:



v4_1

BOMHUM_salidzinajums



Ar tiem mežiem nevar īsti saprast. Tagad līknes kļuvušas krietiņi lēzenākas, turklāt oligotrofo mežu īpatsvaram dažos mērogos ir kļuvusi pozitīva ietekme.

Vēl svarīgs ir TWI - ļoti izteikti piemērotākas ir sausās vietas. Tas varētu būt diezgan loģiski, kamene ligzdas taisa blīvā veģetācijā un kopumā saistāma ar sausākām vietām.

Organiskajai augsnei šūnā negatīva ietekme, bet homerange pozitīva - sākumā šķita nelogiski, bet, ja ligzdvjeta sausākā vietā, tad plašākā mērogā gan vajadzīgs būt barības resursiem. Andris Rūtai ieteica, ka varētu izmest augsnī ainavā - varētu pamēģināt.

Bioloģiskajai lauksaimniecībai šūnā izteikti pozitīva ietekme, bet 3 km ainavā negatīva. Moš loģiski, jo bio lauksaimniecība ir laba, bet, ja daudz ainavā, tad arī kopumā daudz lauksaimniecības ainavā, kas vairs nav tik labi.

Vairāk ilggadīgo zālāju šūnā ir labi. Mazāk koku krājas ir labi.

Mean daily mean temperatures of the driest quarter varētu būt tas, kurš taisa to Kurzemes malu. Varbūt jāpamēģina izsviest ārā.

Nu, taisnību sakot, jo vairāk skatos, jo vairāk saprotu, ka nav nemaz tik slikti. Tā ka vismaz pagaidām kā labāko izvēlos v4_1. Šis fails tad tiks kaut kad papildināts.