



Statistilise metsainventuuri (SMI) ja Maakasutuse, maakasutuse muutuse ja metsanduse (LULUCF) andmehõive metoodika arendamine

Tartu Ülikooli
matemaatika ja statistika instituudi (MSI)
lõpparuanne

Tartu Ülikool

Reg. kood 74001073

Ülikooli 18, Tartu 50090

E-post: info@ut.ee

Kodulehekülg: www.ut.ee



Versiooni nr.	Kuupäev	Muudatus	Autor
1	15.11.2025	Lõpparuanne	Kalev Pärna, TÜ MSI
2	08.12.2025	Lõpparuande täiendatud versioon	Kalev Pärna, TÜ MSI
3	12.12.2025	Lõpparuande parandatud versioon	Kalev Pärna, TÜ MSI
4	19.12.2025	Lõpparuande parandatud versioon	Kalev Pärna, TÜ MSI

Sisukord

1. Sissejuhatus ja tulemuste kokkuvõte	4
2. Uue mudeli väljatöötamine puu kõrguse jaoks	6
2.1. Sissejuhatus.....	6
2.2. Uus puu kõrguse mudel	7
2.3. Kõrguse prognoosimise meetodid	11
2.4. Kõrgusmodelite võrdlus kõrgusprognooside täpsuse alusel.....	15
2.5. Kõrgusmodelite võrdlus mahuprognooside täpsuse alusel	19
3. Puistu juurdekasvu ja suremuse mudelite loomine	22
3.1. Üksikpuu juurdekasvu hindamise meetodid	22
3.1.1. Puu juurdekasvu hindamise meetod 1.....	22
3.1.2. Puu juurdekasvu hindamise meetod 2.....	23
3.2. Meetodite 1 ja 2 võrdlus puu juurdekasvu hindamisel.....	25
3.3. Kogujuurdekasvu arvutamine mudeli põhiselt	29
3.3.1. Andmete ja abifunktsioonide sisselugemine	30
3.3.2. Juurdekasvude leidmine klupitud puude jaoks.....	31
3.3.3. Klupitud puude juurdekasvu leidmine osaproovitükkide jaoks	33
3.3.4. Peenpuude juurdekasv.....	34
3.3.5. Raiutud puude juurdekasv	34
3.3.6. Hinnangu arvutamine kogu Eesti metsamaa kohta	38
3.3.7. Keskmise juurdekasvu hindamine kogu metsamaal	39
3.4. Surnud puude mahu leidmine	41
3.5. Kogujuurdekasvu hindamine mitmese imputeerimise meetodil	42
3.5.1. Abifunktsioonid ja mudelpuude andmete organiseerimine mahuarvutuseks	43
3.5.2. Imputeerimine kluppimise andmetel.....	45
3.5.3. Hinnangu arvutamine kogu Eesti metsamaa kohta	46
3.5.4. Keskmise juurdekasvu hindamine kogu metsamaal	48
4. Kirjanduse loetelu.....	49
Lisa 1. Uue kõrgusmudeli detailne põhjendus.....	50
Lisa 2. Diameetri, kõrguse ja inventeeritud kõrguse seoste graafikud.....	58
Lisa 3. Optimaalne ajaaken uue kõrgusmudeli hindamiseks.....	63
Lisa 4. Uue kõrgusmudeli täiendusvõimaluse uurimine	69
Lisa 5. Uue kõrgusmudeli hindamise ja mahu veaarvutuse R-kood.....	73

1. Sissejuhatus ja tulemuste kokkuvõte

Aruanne käsitleb TÜ MSI projektimeeskonna poolt saadud tulemusi Keskkonnaagentuuri tellitud projekti *Statistilise metsainventuuri (SMI) ja Maakasutuse, maakasutuse muutuse ja metsanduse (LULUCF) andmehõive metoodika arendamine* täitmisel ajaperioodil 15.07.2024 - 15.11.2025.

Projekti lähteülesanded olid järgmised (väljavõte):

- 1) Uue metoodika väljatöötamine **puu kõrguse** arvutamiseks puu mahu ning selle kaudu proovitükil ja kogu metsamaal kasvava metsa tagavara (mahu) täpsemaks hindamiseks, mille tulemused on kasutatavad tagavara arvutamiseks nii SMI kui kaugseire meetoditega;
- 2) Puistu **juurdekasvu** ja **suremuse** mudelite loomine, et täpsemalt modelleerida proovitüki ja kogu metsamaal kasvava metsa tagavara muutust ning süsiniku sidumist tekke-kaos meetodiga.

Laiemas plaanis on antud projekt suunatud SMI-s kasutatava statistilise metoodika edasisele täiendamisele, olles mitmes mõttes jätkuks projektidele [3] ja [4]. Projektimeeskonna käsutuses olid seni kogutud SMI andmed (kuni 2023 k.a.) ja ametlikud tööjuhendid (Arvutusmetoodiline juhend, Välitööde juhend). Töö käigus konsulteeriti SMI spetsialistidega Keskkonnaagentuurist ning EMÜ metsandusteadlastega. Teadusliku kirjanduse kaudu tutvuti ka teiste riikide teadlaste tööga samalaadsete probleemide lahendamisel.

Tulemuste lühikokkuvõte

1. **Uue mudeli väljatöötamine puu kõrguse jaoks.** On koostatud uus metoodika puu kõrguse prognoosimiseks rinnasdiameetri ja vastavat proovitükki kirjeldavate tunnuste põhjal, mis on tunduvalt täpsem kui seni SMI-s kasutatav kõrguse mudel.
 - Läbiviidud analüüs näitas, et uue kõrgusmudeli (vt valem (5)) kasutamine kõrguse prognoosimisel vähendab prognoosi viga võrreldes SMI praeguse mudeliga tervikuna 30% (vastavad RMSE väärtused 2,54 m ja 1,77 m). Detailsem analüüs näitas, et uus mudel on paegusest kõrguse mudelist täpsem ka kõikide suuremate puuliikide lõikes: prognoosivea vähenemine okaspuudel MA ja KU on vastavalt 44% ja 20%, lehtpuudel KS 25%, LM 37%, LV 30% ja HB 27%. Uue kõrgusmudeli täpsust saab veelgi tõsta selle parameetrite rekaliibreerimisega lühemate (nt 5a või 10a) andmeperioodide pealt.

- Analüüs näitas ka, et uue kõrgusmudeli (5) kasutamine mahuarvutuses vähendab mahuproгноosi viga võrreldes SMI praeguse mudeliga tervikuna 24% (vastavad RMSE väärtused 0,068 ja 0,052 m³). Detailsem analüüs näitas, et uus kõrguse mudel on praegusest mudelist täpsem ka kõikide suuremate puuliikide lõikes: mahuhinnangu vea vähenemine okaspuudel MA ja KU on vastavalt 38% ja 22%, lehtpuudel LM 32% ja LV 21%, tagasihoidlikum on vea vähenemine HB ja KS korral (vastavalt 13% ja 7%).

2. Puistu juurdekasvu ja suremuse mudelite loomine. On välja pakutud kaks meetodit puistu juurdekasvu hindamiseks:

- 1) mudelipõhine meetod, mis kasutab uut kõrguse mudelit (5) ja sellega seotud kõrguse muudu mudelit (9);
- 2) mitmene imputeerimine, mis otseselt ei kasuta statistilisi mudeleid.

Mõlemad meetodid lähtuvad mudelpuude andmetest (kus puude tegelik juurdekasv loetakse teadaolevaks), liiguvad edasi alalistele proovitükkidele ja sealt omakorda juhumetsa meetodi abil ajutistele proovitükkidele. Mudelipõhise meetodi juures on üksikpuude mahumuudu hindamiseks välja pakutud uudne kõrguse muudu mudel (9), mille eeliseid tavameetodi ees on põhjendatud nii teoreetiliselt kui ka numbriliselt.

Mitmese imputeerimise korral arvutatakse samale proovitükile üksteise järel välja terve seeria erinevaid prognoose (meie uurimuses 20 prognoosi), mistõttu prognoositud andmetesse tekib ka vajalik juhuslikkus ning see võimaldab veaarvutust läbi viia nõutud viisil.

2. Uue mudeli väljatöötamine puu kõrguse jaoks

2.1. Sissejuhatus

Puu kõrguse võimalikult täpne hindamine on metsastatistikas üks keskseid küsimusi, mille lahendamisest sõltub metsa kogutagavara hinnang. SMI-s kasutatav puu tagavara (mahu) arvutamise valem, mis on fikseeritud Keskkonnaministri 16.jaan. 2009 määruse nr. 2 “Metsa korraldamise juhend“ lisas 11, lähtub puu kõrgusest ja rinnasdiameetrist:

$$V_{puu} = 0,0000785 \cdot d^2 \cdot h \cdot \left(c_1 + \frac{c_2}{d} + \frac{c_3}{h} + \frac{c_4}{d \cdot h} \right), \quad (1)$$

kus d on puu rinnasdiameeter,

h on puu kõrgus,

c_1, c_2, c_3, c_4 on puuliigist sõltuvad parameetrid.

Antud projekti raames eeldatakse, et mudel (1) on usaldusväärne ning kui seal kasutada d ja h mõõdetud (mitte mudeldatud) väärtusi, siis saadavat puu mahtu võib käsitleda kui puu 'tegelikku' mahtu, mille vastu saab testida teiste mudelitega saadud mahtusid.

Puistu ja kogu Eesti metsade tagavara hinnang leitakse üksikpuude tagavarade summana. SMI tagavaraproovitükkidel mõõdetakse iga puu diameeter, aga kõrgus (mille mõõtmine on suhteliselt keeruline ja aeganõudev) mõõdetakse vaid 2-4 mudelpuul. Ülejäänud puude kõrgused modelleeritakse, s.t. arvutatakse mudeli abil, mis luuakse mudelpuude andmete põhjal. Hetkel **SMI-s kasutatav kõrguse mudel** on järgmine:

$$h = 1,3 + (a_1 + a_2 \cdot H100 + a_3 \cdot OnSaar) \left(\frac{d}{d + b_1} \right)^{c_1 + c_2 \cdot H100}, \quad (2)$$

kus

h on puu kõrguse prognoos (m),

d on puu rinnasdiameeter (cm),

$H100$ on puistu kõrgusindeks (proovitüki boniteet),

$OnSaar$ on Saare-, Hiiu- või Läänemaa indikaator (1-0 tunnus),

$a_1, a_2, a_3, b_1, c_1, c_2$ on puuliigist lähtuvad mudelite parameetrite hinnangud (hinnatakse mudelpuude põhjal).

Võrdlusbaasina pakub huvi ka 2023. a. välja pakutud alternatiivne kõrgusmudel (vt. valemid (1.3) & (1.4) projekti aruandes [5]):

$$a_1 = \beta_0 + \beta_1 \cdot \left(1 - \exp\left(-\beta_2 \cdot \frac{A}{100}\right)\right), \quad (3)$$

$$h = (a_0 + a_1 \cdot H100) \left(\frac{d}{d+5}\right)^{c_0+c_1 \cdot H100}, \quad (4)$$

kus

d on puu rinnasdiameeter,

A ja $H100$ on puistu vanus ja kõrgusindeks,

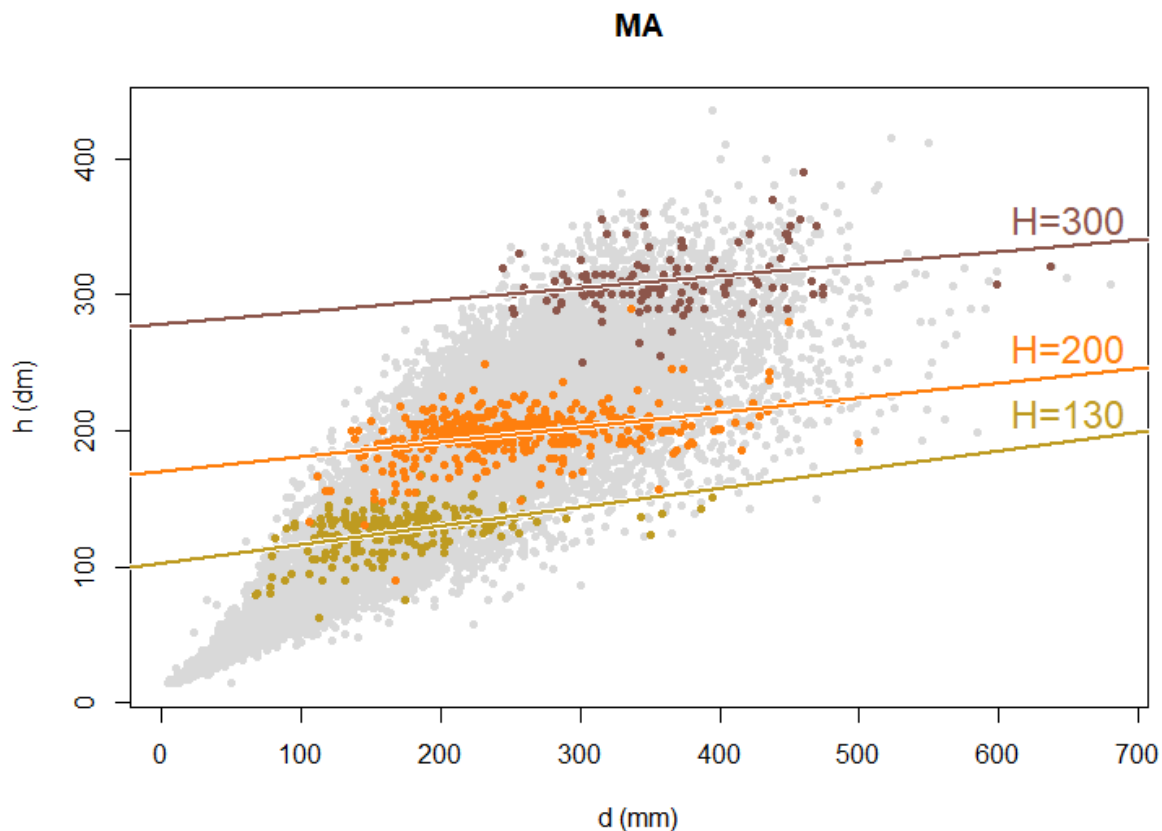
$a_0, \beta_0, \beta_1, c_0, c_1$ on puuliigist sõltuvad parameetrid.

Erinevalt SMI praegusest kõrgusmudelist (2) on alternatiivses mudelis tunnuse $H100$ kordaja seatud sõltuvusse puistu vanusest A ning samas on mudelist välja jäetud tunnus *OnSaar*.

Vead puu kõrguse modelleerimisel kanduvad edasi metsa kogutagavara hinnangutesse. Meie eesmärk on pakkuda välja uus puu kõrguse mudel, mille statistiline viga oleks väiksem kui seni kirjeldatud mudelitel (2) – (4) ja seda nii kõrguse enese prognoosi mõttes kui ka järgneva mahu hinnangu mõttes. Erinevate kõrgusmudelite täpsust saab hinnata mudelpuude andmestikul, kus on olemas instrumentaalsed mõõteandmed puu diameetri ja kõrguse kohta ning valemi (1) kaudu on teada ka puu 'tegelik' maht.

2.2. Uus puu kõrguse mudel

Uue mudeli aluseks on asjaolu, mida peegeldab hästi järgmine joonis, mis kujutab mändide rinnasdiameetri d ja kõrguse h vahelist seost sõltuvalt proovitüki keskmisest kõrgusest H . On näha, et kõrguse hajuvus antud d korral on üldiselt väga suur, aga kui proovitüki keskmine (inventeeritud) kõrgus H fikseerida, siis kõrguse hajuvus väheneb järsult. Vaatame näiteks mände, mis kasvavad proovitükkidel, kus $H=200$ (joonisel oranžid punktid). Selliste puude diameetrid võivad muutuda üsna suures vahemikus (ca 100...500 dm), kuid peaaegu kõigi nende puude kõrgus on ligilähedaselt 200 dm. Sõltuvus puu enda diameetrist on suhteliselt väike ning see seos on modelleeritav lihtsa sirge abil.



Joonis 1. Männi diameetri d ja kõrguse h vaheline seos inventeeritud kõrguse H fikseeritud tasemete korral.

Kirjeldatud asjaolust lähtuvalt on alljärgnevas välja pakutud uus puu kõrguse mudel (5), mis on segu kolmest komponendist, millest esimene annab kõrguse hinnangu seemnepuudele, teine noorpuudele ja kolmas nõ tavapuudele. Seejuures järgib tavapuude kõrguse prognoos peamiselt vastava proovitükiosa inventeeritud kõrgust H ning sõltuvus puu rinnasdiameetrist on suhteliselt tagasihoidlik. Põhjenduseks on äsjane tõdemus, et täiskasvanud metsas on kõik puud (v.a. järelkasv) suhteliselt ühetaolise kõrgusega, mis on ühtlasi puistu keskmine kõrgus H . Seemnepuude ja noorpuude korral on olukord teistsugune ning siin on kõrguse prognoosimiseks on kasutatud Hossfeldi kõrguskõverat [18] (vt alltoodud valem (8)). Mudeli igal komponendil on oma kaal, milleks on tõenäosus, et puu on kas seemnepuu, noorpuu (järelkasv) või tavapuu. Nimetatud kaalud (tähistatud vastavalt P_S , P_{NP} ja $1 - P_S - P_{NP}$) arvutatakse puu rinnasdiameetri d ja proovitükiosa keskmise diameetri D erinevuse kaudu (valemid (6) ja (7)). Mudeli detailne põhjendus ja täiendavad selgitused on toodud aruande Lisas 1.

$$h = P_s \cdot f(d, a_s, b_s) + P_{NP} \cdot f(d, a_{NP}, b_{NP}) \cdot c_2^{I(\text{puuliik} \neq \text{peapuu})} + (1 - P_s - P_{NP}) c_2^{I(\text{puuliik} \neq \text{peapuu})} \left(c_3 H - \left[c_0 + c_1 \left(\frac{1}{R} \right) \right] (d - D) \right), \quad (5)$$

$$P_s = \frac{1 - \exp \left(-10 \left(\frac{d - D}{d} \cdot I(d > D) \right) \right)}{1 + \exp \left(-10 \left(\frac{d - D}{d} \cdot I(d > D) - c_s \right) \right)} \quad (6)$$

$$P_{NP} = \frac{1 - \exp \left(-10 \left(\frac{D - d}{D} \cdot I(d < D) \right) \right)}{1 + \exp \left(-10 \left(\frac{D - d}{D} \cdot I(d < D) - c_{NP} \right) \right)} \quad (7)$$

$$f(d, a, b) = 1,3 + \frac{d^{1,3}}{a + b \cdot d^{1,3}} \quad (8)$$

kus

h, d – puu kõrgus ja rinnasdiameeter;

H, D – proovitüki osa inventeeritud kõrgus ja diameeter;

R – relaskoobi mõõtmistulemus (metsa tihedus);

$a_s; b_s; c_s; a_{NP}; b_{NP}; c_{NP}; c_0; c_1; c_2; c_3$ on iga puuliigigrupi jaoks andmete põhjal hinnatavad parameetrid, kus

$a_s; b_s; c_s$ iseloomustavad seemnepuid;

$a_{NP}; b_{NP}; c_{NP}$ kirjeldavad noorpuude käitumist;

c_0 ja c_1 kirjeldavad kuidas teatud kõrgusega metsas üksikpuu kõrgus sõltuvub selle puu diameetrist;

c_2 parameeter korrigeerib hinnanguid siis, kui peapuu puuliigigrupp erineb mudeldava puu puuliigigrupist (sel juhul indikaatortunnus $I(\text{puuliik} \neq \text{peapuu}) = 1$), vastasel juhul indikaatortunnus on 0 ning $c_2^0 = 1$);

c_3 on kordaja, mis vajadusel korrigeerib inventeeritud kõrguste väärtuseid).

Antud valem võib teoreetiliselt prognoosida üksikpuu kõrguse, mis on väiksem kui 1,3 m. Soovitavalt võiks sellised prognoosid asendada väärtusega 1,3.

Tabel. Kõrgusmodeli (5) parameetrid, hinnatud puuliigigruppide jaoks 2015-2024 mudelpuude andmete põhjal

	HB	KS	KU	LM	LV	MA	RE	XX
c_s	0.621124	0.469963	0.486609	0.778848	0.403154	0.805878	2.532236	0.715073
c_NP	0.449462	0.456187	0.104016	0.425854	0.433777	0.469280	0.300265	0.263205
a_s	2.404942	2.348334	3.514819	1.839891	2.161592	6.080749	1.456700	1.943751
a_NP	1.131174	2.542609	3.790082	1.269522	3.201142	5.940132	0.000000	0.333156
b_s	0.003601	0.003417	0.002044	0.004861	0.003574	0.000000	0.005110	0.004046
b_NP	0.002953	0.003290	0.003114	0.003163	0.003631	0.000830	0.000000	0.003210
c0	0.165860	0.233164	0.309831	0.101021	0.228985	0.119803	0.308700	-0.017603
c1	-0.674324	1.090901	-0.677039	1.587901	1.301196	0.581938	-2.999049	1.905059
c2	1.028058	1.029750	0.939448	0.957757	0.941489	0.967937	0.888586	0.914324
c3	1.010450	0.985544	0.975348	0.989155	0.979220	0.989213	0.957874	0.988127

Ülaltoodud tabelis kasutatud puuliigigruppide tähistused on lahti seletatud järgnevalt:

Puuliigigruppide tähistus ja koosseis kõrgusmodeli (5) jaoks

Tähistus	Koosseis
HB	HB , PP
KS	KS , PN , TL , EL
KU	KU , NU , TO , TS
LM	LM
LV	LV
MA	MA , LH , SD
RE	RE , PI , TM , SP , OP , TY
XX	TA , SA , VA , JA , KP

Kõrgusmodeli (5) parameetrite hindamiseks sobiva ajaakna pikkuse detailne analüüs ja sellest tulenevad praktilised soovitused on toodud Lisas 3.

Kõrgusmodeli (5) hindamise ja kasutamise kood on toodud Lisas 5.

Järgnevas asume uut kõrgusmodelit võrdlema teiste kõrguse prognoosimise meetoditega.

2.3. Kõrguse prognoosimise meetodid

Eesmärgiks on võrrelda järgnevate kõrguse prognoosimise meetodite täpsust:

- 1) SMI-s **praegu** kasutatav puu kõrguse mudel,
- 2) **Alternatiivne** kõrgusmudel¹,
- 3) Juhumetsa (**rf**) meetod,
- 4) Üldistatud aditiivne mudel (**gam**),
- 5) Antud projekti raames välja pakutud **uus** kõrguse mudel.

(Rasvaselt on trükitud vastavate meetodite tähistused alljärgnevas koodis).

Selleks, et võrdlus oleks võimalikult aus, jaotame andmestiku masinõppe meetodite jaoks kaheks - mudeli sobitame treeningandmestikul ja prognooside headust võrdleme testandmestikul, mida treenimisel ei kasutatud. Osadeks jaotamise printsiip on järgmine: jätame välja proovitükiosasid, säilitades peapuuliikide proportsioone. Treeninguks kasutatakse 70% proovitükiosasid, testimiseks jääb 30%.

Kood koos selgitustega

Kõigepealt tingimustele vastavate osade nimekiri koos peapuuliigiga:

```
prtosad=mudelpuud|>left_join(prtosa,by="prto_aid")|>
  filter(maakategooria %in% c("M","MM"))|>
  group_by(prto_aid,peapuuliik)|>summarize(mitu=n(),.groups="drop")
```

Seejärel jaotame need osadeks ja tekitame vastavad mudelpuude andmestikud

```
set.seed(20250905)
osad=initial_split(prtosad,prop=0.7,strata=peapuuliik)
mudelpuud_treening=mudelpuud|>inner_join(training(osad),by="prto_aid")
mudelpuud_testing=mudelpuud|>inner_join(testing(osad),by="prto_aid")
save(mudelpuud_treening,mudelpuud_testing,file="mudelpuud_valimid_lihtne.RData")

load("mudelpuud_valimid_lihtne.RData")
```

Toome sisse puuliikide grupid, mida on ka seni mudelites kasutatud

```
plgrp <- list(
  "MA" = c("MA","LH","SD"),
  "KU" = c("KU","NU","TO"),
  "KS" = c("KS","PN","TL"),
  "HB" = c("HB","PP","EL"),
  "LM" = c("LM"),
  "LV" = c("LV"),
  "XX" = c("TA","SA","VA","JA","KP")
)
pl_grupp=tibble()
for(nimi in names(plgrp)){
```

¹ Töövõtulepingu nr 2-1/27/2021 aruanne (8.veebr. 2023) „KHG inventuuri LULCF sektori heite arvutamiseks vajalike metsakasvumudelite väljatöötamine“

```
tmp=tibble(pliik=plgrp[[nimi]],grupp=nimi)
pl_grupp=rbind(pl_grupp,tmp)
}
```

Järgnevas leiame prognoosid erinevate meetoditega ning seejärel võrdleme tulemusi. Selleks aga tekitame eelnevalt vajalikku infot sisaldava testandmestiku.

```
testandmed=modelpuud_testing|>
  inner_join(prtosa,by=c("prto_aid","aasta","peapuuliik"))|>
  dplyr::select(amudelpuud_id,puuliik,aasta,d_13,h_puu,peapuuliik,vanus,inv_korgus,
    arenguklass,inv_bon,onsaar,kasvukohatyp,inv_diameeter,arv_h100,inv_g1)|>
  inner_join(pl_grupp,by=c("puuliik"="pliik"))|>
  mutate(inv_korgus=coalesce(inv_korgus,0),inv_diameeter=coalesce(inv_diameeter,0))|>
  filter(!is.na(h_puu),!is.na(d_13))
```

1) Praegune mudel

Defineerime funktsiooni, mille abil arvutada kõrguseid praegu SMI-s kasutatava mudeli (vt eespool olev valem (2)) kohaselt puuliigiti ning tekitame vastava prognooside andmestiku.

```
#kordajate                                tabel
kordajad_korgus=read_csv("pkorguskover.csv",show_col_types = FALSE)
korgus_praegu_puuliik=function(arv_h100,onsaar,d_13,puu_L,kordajad=kordajad_korgus){
  nimed=c("a1","a2","a3","b1","c1","c2")
  rida=kordajad|>filter(grepl(puu_L,kaasliigid))
  if(nrow(rida)==0){return(NA)}
  for(nimi in nimed){
    assign(nimi,unlist(rida[nimi]))
  }
  vastus=13+10*(a1+a2*arv_h100+a3*onsaar)*(d_13/(d_13+10*b1))^(c1+arv_h100*c2)
  return(vastus)
}
}
prognoosid=NULL
for(pl in unique(testandmed$puuliik)){
  tmp=testandmed|>filter(puuliik==pl)|>mutate(prognoos_praegu=korgus_praegu_puuliik(arv_h100,
    onsaar,d_13,pl))|>select(amudelpuud_id,puuliik,h_puu,prognoos_praegu)
  if(is.null(prognoosid)){
    prognoosid=tmp
  }else{
    prognoosid=rbind(prognoosid,tmp)
  }
}
progn_test=prognoosid
```

2) Alternatiivne mudel

Alternatiivne mudel on esitatav valemitena (3) ja (4), mis on pärit lepingu aruandest [1]. Sama allikast on võetud ka allolev mudeli kordajate tabel. Põhimõtteliselt saaks mudeli uuesti andmestikule sobitada, aga erinevused on väikesed, mistõttu neid arvutusi pole siia lisatud.

Mudeli kordajate tabel:

```
tabel="Puuliik Kood b0 b1 b2 a0 c0 c1 Arv St_viga
Mänd MA -0.010 1.408 1.272 3.446 1.609 -0.033 16603 1.50
Kuusk KU 0.083 1.277 1.232 7.722 2.649 -0.031 7059 2.15
Kask KS 0.351 0.838 2.699 1.335 1.408 -0.007 11161 1.74
Haab HB 0.360 0.834 2.145 1.473 2.389 -0.048 1119 1.75
Sanglepp LM 0.319 1.004 2.137 -1.367 1.189 -0.008 1272 1.66
Hall-lepp LV 0.279 1.002 2.271 -0.483 2.292 -0.047 3054 1.55
Muud_liigid XX 0.148 1.143 1.681 1.222 1.102 -0.005 311 1.70"
kordajadAlt=read.table(header=TRUE,text=tabel,as.is=TRUE)
```

```
kordajadAltPl=pl_grupp|>inner_join(kordajadAlt,by=c("grupp"="Kood"))
```

Defineerime funktsiooni, mis arvutab kõrgused alternatiivse mudeli järgi puuliigiti

```
korgus_Alt=function(arv_h100, vanus, d_13, puu_l, kordajad=kordajadAltPl){
  nimed=c("a0", "b0", "b1", "b2", "c0", "c1")
  rida=kordajad|>filter(pliik==puu_l)
  if(nrow(rida)==0){return(NA)} else {
    for(nimi in nimed){
      assign(nimi, unlist(rida[nimi]))
    }
    a1=b0+b1*(1-exp(-b2*vanus/100))
    vastus=(a0+a1*arv_h100)*(d_13/(d_13+5))^(c0+c1*arv_h100)
    return(vastus)
  }
}
```

Lisame prognoosid tabelisse:

```
}

puuliigid=testandmed$puuliik|>unique()
progn_test=progn_test|>mutate(prognnoos_Alt=NA)
for(puu_l1 in puuliigid){
  #print(puu_l1)
  tmp=progn_test|>filter(puuliik==puu_l1)|>inner_join(testandmed,by="amudelpuud_id")
  progn_test$prognnoos_Alt[progn_test$puuliik==puu_l1]=10*korgus_Alt(tmp$arv_h100,tmp$vanus,
tmp$d_13/10, puu_l1)
}
```

3) Juhumets

Baastaseme leidmises sobitame juhumetsa mudeli treeningandmetele ja leiame selle prognooside mõõdikud testandmete peal puuliigiti. Lisaks leiame tunnuste tähtsuse mõõdikud, et teiste mudelite sobitamisel/väljatöötamisel oskaksime arvestada, millised tunnused tulemust mõjutavad. Juhumetsa puhul on hea see, et mõju mõõdikud võtavad arvesse ka mittelineaarseid mõjusid. Kuna hetkel kasutatav mudel ja ka alternatiivne mudel kasutavad väga vähe tunnuseid,

siis anname juhumetsale ette samuti üsna piiratud tunnuste valiku. Nimekirja võib suurendada, kui on kahtlus, et midagi olulist on välja jäetud.

```
treening=modelpuud_treening|>
  inner_join(prtosa,by=c("prto_aid","aasta","peapuuliik"))|>
  dplyr::select(amudelpuud_id,puuliik,aasta,d_13,h_puu,peapuuliik,vanus,inv_korgus,
    arenguklass,inv_bon,onsaar,kasvukohatyp,inv_diameeter,inv_g1)|>
  inner_join(pl_grupp,by=c("puuliik","plik"))|>
  mutate(inv_korgus=coalesce(inv_korgus,0),inv_diameeter=coalesce(inv_diameeter,0))|>
  filter(!is.na(h_puu),!is.na(d_13))
rec1=recipe(h_puu~.,data=treening)|>update_role(amudelpuud_id,new_role="modelpuu id")
spec1=rand_forest(mode="regression")
wf1=workflow()|>add_recipe(rec1)|>add_model(spec1)|>set_engine("ranger",importance="impurity",
  respect.unordered.factors=TRUE))
m1=wf1|>fit(treening)
```

Leiame tunnuste tähtsuse

```
tahtsus_m1=m1|>extract_fit_engine()
tahtsus_m1$variable.importance|>sort()

##      onsaar      peapuuliik      aasta      grupp      puuliik
##      1064871      1921564      2581617      3435821      4063653
## kasvukohatyp      inv_bon      vanus inv_diameeter      arenguklass
##      7499535      10952505      20679115      22777504      31677107
##      inv_korgus      d_13      inv_g1
##      79305494      95974120      72756925569
```

Lisame prognooside andmetele juhumetsa prognoosid

```
progn_test=progn_test|>left_join(augment(m1,new_data=testandmed)|>select(amudelpuud_id,.pred)|>
  rename(prognoos_rf=.pred),by="amudelpuud_id")
```

4) GAM model

Juhumets võib nõuda väga palju andmeid täpsuse saavutamiseks, kui argumenttunnuste hulgas on pidevaid tunnuseid ja on alust arvata, et nendest sõltuvus on ka mingi mitme muutuja pidev funktsioon.

Seega sobitame juhumetsa alternatiivina küllaltki vähe tunnuseid kasutavad GAM modelid.

Sobitame ja prognoosime rühmade kaupa

```
prognoosid_gam=NULL
for(gr in unique(testandmed$grupp)){
  m_gam=mgcv::gam(h_puu~s(inv_korgus,d_13,vanus,inv_diameeter)+areng*bon,data=treening|>filter(grupp==gr)|>
    mutate(bon=as.numeric(inv_bon),areng=factor(coalesce(arenguklass,"0"))))
  tmp=testandmed|>filter(grupp==gr)|>mutate(bon=as.numeric(inv_bon),areng=factor(coalesce(arenguklass,"0")))
  tmp1=augment(m_gam,newdata=tmp)|>select(amudelpuud_id,.fitted)|>
    rename(prognoos_gam=.fitted)
  if(is.null(prognoosid_gam)){
    prognoosid_gam=tmp1
  } else {
```

```
    prognoosid_gam=rbind(prognoosid_gam,tmp1)
  }
}
progn_test=progn_test|>left_join(prognoosid_gam,by="amudelpuud_id")
```

5) Uus model

Uus kõrguse model (5) on siinkohal kalibreeritud mudelpuude peal aastaist 2005-2024.

Lisame prognoosid testandmestikule

```
load("hmodel3.RData",verbose=TRUE)

## Loading objects:
##   prognoositavad

progn_test=progn_test|>left_join(prognoositavad|>select(amudelpuud_id,hprognoos),by="amudelpuud_id")|>rename(prognoos_uus=hprognoos)
```

2.4. Kõrgusmodelite võrdlus kõrgusprognooside täpsuse alusel

1) Summaarne viga üle kogu testandmestiku

Kõigepealt vaatame summarset ruutkeskmist viga üle kogu testandmestiku

```
progn_test|>inner_join(mudelpuud|>select(amudelpuud_id,aasta))|>filter(aasta>2004)|>summarize(RMSE_praegu=rmse_vec(truth=h_puu,estimate=prognoos_praegu)/10,RMSE_Alt=rmse_vec(truth=h_puu,estimate=prognoos_Alt)/10,RMSE_rf=rmse_vec(truth=h_puu,estimate=prognoos_rf)/10,RMSE_gam=rmse_vec(truth=h_puu,estimate=prognoos_gam)/10,RMSE_uus=rmse_vec(truth=h_puu,estimate=prognoos_uus)/10,puude_arv=n())

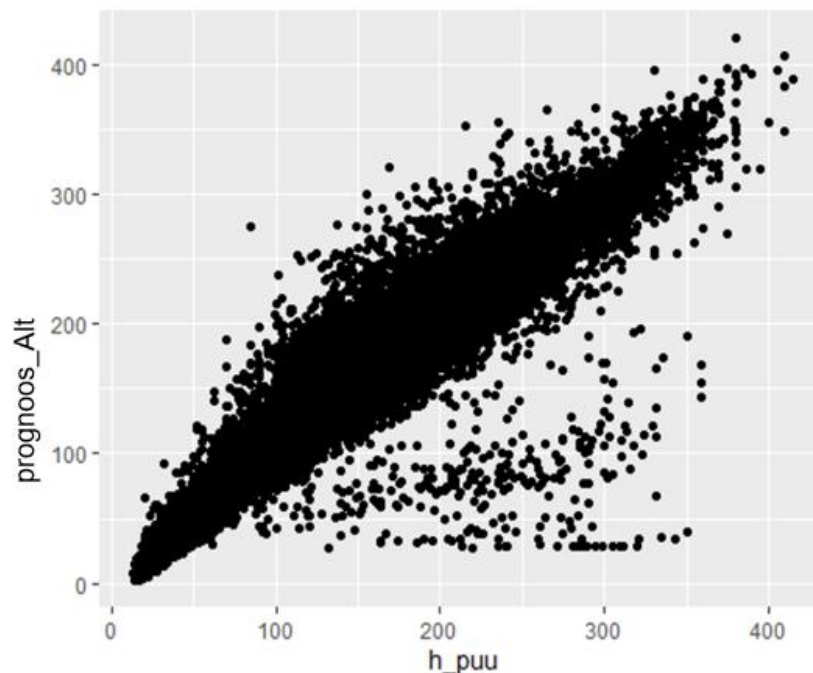
## Joining with `by = join_by(amudelpuud_id)`

##      RMSE_praegu  RMSE_Alt  RMSE_rf  RMSE_gam  RMSE_uus  puude_arv  uus-praegu%
## 1      2.539799    2.945588  1.692535  1.705824  1.770406    18248      -30%
```

Siin on üllatav, et alternatiivse mudeli Alt prognoosiviga on kõige suurem. Põhjus võib olla selles, et model sobib ainult täiendavatel tingimustel ehk nn ‘headel’ andmetel (mis nt välistab olukorra, kus suur mudelpuu asub proovitükil, mille inventeeritud kõrgus ja vanus on väiksed).

Selliseid võimalikke efekte näitab ka järgnev joonis:

```
progn_test|>ggplot(aes(x=h_puu,y=prognoos_Alt))+geom_point()
```



Jooniselt on näha, et peaaegu kõigi puukõrguste `h_puu` korral leidub arvukas grupp mudelpuid, mille korral `prognoos_Alt` on kordades väiksem tegelikust kõrgusest. Tagajärjeks on suur RMSE väärtus.

GAM ja juhumets käituvad üsna sarnaselt ja nende erinevus võib olla tingitud lihtsalt testandmestiku valikust, st juhuslik. Kindel on aga see, et oluliselt saab parandada praegust SMI kõrgusmudelit. Hiljem võrdleme mudelite käitumist ka nn 'headel' andmetel, kus on eemaldatud mõned raskesti prognoositavad juhud tingimustega, mida kasutati alternatiivse mudeli Alt sobitamisel.

2) Vead puuliigiti üle kogu testandmestiku

```
progn_test|>inner_join(mudelpuud|>select(amudelpuud_id,aasta))|>filter(aasta>2004)|>group_by(puuliik)|>
  summarize(RMSE_praegu=rmse_vec(truth=h_puu,estimate=prognoos_praegu)/10,
            RMSE_Alt=rmse_vec(truth=h_puu,estimate=prognoos_Alt)/10,
            RMSE_rf=rmse_vec(truth=h_puu,estimate=prognoos_rf)/10,
            RMSE_gam=rmse_vec(truth=h_puu,estimate=prognoos_gam)/10,
            RMSE_uus=rmse_vec(truth=h_puu,estimate=prognoos_uus)/10,
            n=n())

## Joining with `by` = join_by(amudelpuud_id)`

## # A tibble: 14 × 7
##   puuliik  RMSE_praegu RMSE_Alt RMSE_rf RMSE_gam RMSE_uus    n  uus-praegu
```

##	<chr>	<dbl>	<dbl>	<dbl>	<dbl>	<dbl>	<int>	%
##	1 HB	2.45	2.96	1.98	1.78	1.80	858	-27%
##	2 JA	2.61	4.68	3.64	3.07	3.50	16	34%
##	3 KP	3.45	3.74	2.13	3.18	2.99	2	-13%
##	4 KS	2.25	2.68	1.63	1.61	1.68	5277	-25%
##	5 KU	2.63	3.37	1.98	2.02	2.11	3908	-20%
##	6 LH	3.48	2.29	1.61	1.61	1.63	8	-53%
##	7 LM	2.75	2.66	1.69	1.74	1.74	866	-37%
##	8 LV	2.31	2.86	1.52	1.55	1.62	1638	-30%
##	9 MA	2.75	2.81	1.45	1.49	1.54	5338	-44%
##	10 PN	3.87	5.91	2.65	4.02	4.10	41	6%
##	11 PP	2.08	0.690	1.17	0.791	0.785	2	-62%
##	12 SA	2.60	3.29	1.90	2.17	1.99	144	-23%
##	13 TA	3.08	4.52	2.55	2.64	2.55	102	-17%
##	14 VA	2.65	4.98	2.62	2.30	2.20	48	-17%

3) Summarne viga üle 'heade' andmete

Tekitame nn 'head' andmed, mis ühtivad aruandes [5] kõrgusmudeli hindamiseks ja testimiseks kasutatud andmetega.

```
head_test_andmed=progn_test|>
  left_join(mudelpuud|>select(amudelpuud_id,kahj_aste,prto_aid,d_13,aasta))|>
  left_join(prtos|>select(prto_aid,maakategooria,vanus,inv_korgus))|>
  filter(vanus>15,d_13>50,is.na(kahj_aste),maakategooria %in% c("M","MM"),!is.na(inv_korgus),aasta>2004)

## Joining with `by` = join_by(amudelpuud_id)`
## Joining with `by` = join_by(prto_aid)`
```

Arvutame vead:

```
head_test_andmed|> summarize(RMSE_praegu=rmse_vec(truth=h_puu,estimate=prognos_praegu)/10,
  RMSE_Alt=rmse_vec(truth=h_puu,estimate=prognos_Alt)/10,RMSE_rf=rmse_vec(truth=h_puu,estimate=prognos_rf)/10,
  RMSE_gam=rmse_vec(truth=h_puu,estimate=prognos_gam)/10,
  RMSE_uus=rmse_vec(truth=h_puu,estimate=prognos_uus)/10,
  n=n())

##      RMSE_praegu RMSE_Alt  RMSE_rf  RMSE_gam  RMSE_uus      n  uus-praegu%
## 1      2.592357  2.230725  1.647337  1.662136  1.703885 13478      -34%
```

4) Vead puuliigiti 'headel' andmetel

```
head_test_andmed|> group_by(puuliik)|> summarize(RMSE_praegu=rmse_vec(truth=h_puu,estimate=prognos_praegu)/10,
  RMSE_Alt=rmse_vec(truth=h_puu,estimate=prognos_Alt)/10,RMSE_rf=rmse_vec(truth=h_puu,estimate=prognos_rf)/10,
  RMSE_gam=rmse_vec(truth=h_puu,estimate=prognos_gam)/10,
  RMSE_uus=rmse_vec(truth=h_puu,estimate=prognos_uus)/10,n=n())
```

```
## # A tibble: 14 × 7
##   puuliik  RMSE_praegu RMSE_Alt RMSE_rf RMSE_gam RMSE_uus  n  uus-praegu(%)
##   <chr>      <dbl>      <dbl>  <dbl>  <dbl>  <dbl> <int>
## 1 HB          2.67        2.49    2.22    2.03    2.06   418   -23%
## 2 JA          1.76        2.92    2.35    1.67    2.25    12    28%
## 3 KP          3.45        3.74    2.13    3.18    2.99     2   -13%
## 4 KS          2.29        2.16    1.60    1.59    1.65  3729  -28%
## 5 KU          2.65        2.82    1.91    1.95    2.01  2883  -24%
## 6 LH          3.48        2.29    1.61    1.61    1.63     8  -53%
## 7 LM          2.76        2.17    1.64    1.69    1.67   595  -39%
## 8 LV          2.34        2.44    1.48    1.53    1.57   934  -33%
## 9 MA          2.78        1.64    1.42    1.42    1.46  4652  -47%
## 10 PN         3.92        4.97    2.83    4.47    4.55    30    16%
## 11 PP         2.91        0.972    1.38    0.757    0.493     1  -83%
## 12 SA         2.84        2.68    1.98    2.29    2.15    93   -24%
## 13 TA         2.82        3.30    2.39    2.42    2.29    82   -19%
## 14 VA         2.59        4.64    2.65    2.30    2.23    39   -14%
```

Kokkuvõte (kõrguse uus mudel kõrguse prognoosimise kontekstis):

Toodud analüüs näitas, et uue kõrguse mudeli kasutamine kõrguse prognoosimisel vähendab prognoosi viga tervikuna 30% võrra (võetuna üle kõigi andmete). Detailsem analüüs näitas, et uus mudel on paegusest kõrguse mudelist täpsem ka kõikide suuremate puuliikide lõikes: prognoosivea vähenemine okaspuudel MA ja KU on vastavalt 44% ja 20%, lehtpuudel KS (25%), LM (37%), LV (30%), HB (27%). Kahel väikesearvulisel lehtpuul (PN, JA) võib täheldada prognoosivea teatavat kasvu. 'Headel' andmetel on uue mudeliga saadud kõrgusprognoosid võrreldes praeguse mudeliga veelgi täpsemad – tervikuna on täpsuse kasv sel juhul 34%. Arvestades asjaolu, et antud analüüsis kasutatud kõrguse uus mudel oli kalibreeritud pikima võimaliku andmestiku pealt (2005-2024), võib lühema perioodi (5 või 10 a) pealt hinnatud mudel ilmselt töötada veelgi täpsemalt.

2.5. Kõrgusmodelite võrdlus mahuprognoside täpsuse alusel

Järgnevas eeldame, et puu 'tegelik' maht on arvutatav valemiga (1), mis on fikseeritud 'Metsa korraldamise juhendis' 2009.a. Mudelpuude tegelike mahtude vastu võrdleme samade puude prognoositud mahtusid, mis on saadud endiselt mahuvalemi (1) abil, kus aga puu tegeliku kõrguse asemel kasutatakse kõrguse erinevaid prognoose.

Defineerime funktsiooni mahtude arvutamiseks

```
olemas=c('MA','LH','SD','KU','NU','TS','TO','KS','PN','HB','LM','LV',  
         'PP','RE','TL','OP','TA','VA','SA','TM','JA','KP','SP','PI','EL','TY')  
kordajad_maht=data.frame(puuliigid=c("MA","LH","SD","","KU","NU","TS","TO","","KS","PN","","H  
B","LM","LV",  
         'PP','RE","","TL","OP","TA","VA","SA","TM","JA","KP","SP","PI","EL","TY"),a=c(0.3571,0.4216  
,0.4080,0.4723,0.4033),b=c(0.660,0.181,0.757,-0.608,0),c=c(2.156,1.190,0.801,0,1.586),d=c(-  
8.312,-1.309,-10.707,12.724,1.440))  
kordajad_maht_maatriks=matrix(NA,nrow=length(olemas),ncol=4)  
row.names(kordajad_maht_maatriks)=olemas  
colnames(kordajad_maht_maatriks)=c('a','b','c','d')  
for(i in 1:length(olemas)){  
  kordajad_maht_maatriks[i,]=kordajad_maht|>filter(grepl(olemas[i], puuliigid))|>select(a:d  
)|>as.matrix()  
}  
  
V_puu=function(d13,h,a=0.3571,b=0.66,c=2.156,d=-8.312){  
  #eeldame, et d on millimeetrites ja h on detsimeetrites  
  #puu mahu valem metsakorralduse eeskirja kohaselt, vaikeparameetrid  
  #männi, lehise ja seedermänni jaoks  
  D=d13/10  
  H=h/10  
  return(0.0000785*D^2*H*(a+b/D+c/H+d/(D*H)))  
}  
  
V_maht_koik=function(puuliigid,d13,h,kordajadM=kordajad_maht_maatriks){  
  #eeldus on, et puuduvaid andmeid pole  
  D=d13/10  
  H=h/10  
  return(0.0000785*D^2*H*(rowSums(kordajadM[puuliigid,]*cbind(1,1/D,1/H,1/(D*H)))))  
}
```

Lisame mahud prognooside andmetele

```
progn_test_maht=progn_test|>  
  inner_join(mudelpuud|>select(amudelpuud_id,d_13))|>  
  mutate(V=V_maht_koik(puuliik,d_13,h_puu),V_praegu=V_maht_koik(puuliik,d_13,prognoos_praegu),  
         V_Alt=V_maht_koik(puuliik,d_13,prognoos_Alt),  
         V_rf=V_maht_koik(puuliik,d_13,prognoos_rf),  
         V_gam=V_maht_koik(puuliik,d_13,prognoos_gam),  
         V_uus=V_maht_koik(puuliik,d_13,prognoos_uus))|>  
  select(amudelpuud_id,h_puu,d_13,starts_with("V"))  
  
## Joining with `by = join_by(amudelpuud_id)`
```

1) Mahuproгноoside vead kogu testandmestikul alates 2005

```
progn_test_maht|>inner_join(mudelpuud|>select(amudelpuud_id,aasta))|>filter(aasta>2004)|>summarize(RMSE_praegu=rmse_vec(truth=V,estimate=V_praegu),RMSE_Alt=rmse_vec(truth=V,estimate=V_Alt),RMSE_rf=rmse_vec(truth=V,estimate=V_rf),RMSE_gam=rmse_vec(truth=V,estimate=V_gam),RMSE_uus=rmse_vec(truth=V,estimate=V_uus),n=n())
```

```
## Joining with `by = join_by(amudelpuud_id)`
```

	RMSE_praegu	RMSE_Alt	RMSE_rf	RMSE_gam	RMSE_uus	n	uus-praegu%
## 1	0.068365	0.099866	0.055359	0.049498	0.051948	18248	-24%

2) Mahuproгноoside vead puuliigiti, kogu testandmestik al. 2005

```
progn_test_maht|>inner_join(mudelpuud|>select(amudelpuud_id,aasta,puuliik))|>filter(aasta>2004)|>group_by(puuliik)|>summarize(RMSE_praegu=rmse_vec(truth=V,estimate=V_praegu),RMSE_Alt=rmse_vec(truth=V,estimate=V_Alt),RMSE_rf=rmse_vec(truth=V,estimate=V_rf),RMSE_gam=rmse_vec(truth=V,estimate=V_gam),RMSE_uus=rmse_vec(truth=V,estimate=V_uus),n=n())
```

```
## Joining with `by = join_by(amudelpuud_id)`
```

```
## # A tibble: 14 × 7
```

	puuliik	RMSE_praegu	RMSE_Alt	RMSE_rf	RMSE_gam	RMSE_uus	n	uus-praegu%
	<chr>	<dbl>	<dbl>	<dbl>	<dbl>	<dbl>	<int>	%
## 1	HB	0.117	0.236	0.141	0.102	0.102	858	-13%
## 2	JA	0.0788	0.120	0.114	0.0948	0.108	16	37%
## 3	KP	0.177	0.0798	0.0525	0.105	0.0199	2	-89%
## 4	KS	0.0444	0.0737	0.0450	0.0390	0.0414	5277	-7%
## 5	KU	0.0767	0.105	0.0614	0.0572	0.0596	3908	-22%
## 6	LH	0.0854	0.0655	0.0581	0.0550	0.0587	8	-31%
## 7	LM	0.0633	0.0648	0.0424	0.0425	0.0432	866	-32%
## 8	LV	0.0369	0.0436	0.0249	0.0251	0.0293	1638	-21%
## 9	MA	0.0728	0.0940	0.0404	0.0421	0.0448	5338	-38%
## 10	PN	0.107	0.141	0.0743	0.0976	0.111	41	4%
## 11	PP	0.0481	0.0159	0.0276	0.0191	0.0193	2	-60%
## 12	SA	0.0477	0.0640	0.0415	0.0427	0.0332	144	-30%
## 13	TA	0.225	0.255	0.114	0.136	0.162	102	-28%
## 14	VA	0.0396	0.0669	0.0584	0.0644	0.0355	48	-10%

3) Mahuproгноoside vead 'headel' andmetel

```
progn_test_maht|>inner_join(head_test_andmed|>select(amudelpuud_id))|>summarize(RMSE_praegu=rmse_vec(truth=V,estimate=V_praegu),RMSE_Alt=rmse_vec(truth=V,estimate=V_Alt),RMSE_rf=rmse_vec(truth=V,estimate=V_rf),RMSE_gam=rmse_vec(truth=V,estimate=V_gam),RMSE_uus=rmse_vec(truth=V,estimate=V_uus),n=n())
```

```
## Joining with `by = join_by(amudelpuud_id)`
```

	RMSE_praegu	RMSE_Alt	RMSE_rf	RMSE_gam	RMSE_uus	n	uus-praegu%
## 1	0.0659	0.0524	0.0471	0.0437	0.0452	13478	-31%



Kokkuvõte (kõrguse uus mudel mahu hindamise kontekstis):

Ülaltoodud analüüs näitas, et uue kõrguse mudeli kasutamine mahuarvutuses vähendab mahuproгноosi viga 24% võrra (võetuna üle kõigi andmete al. 2005). Detailsem analüüs näitas, et uus kõrguse mudel on paegusest mudelist täpsem ka kõikide suuremate puuliikide lõikes. Eriti märgatav on mahuhinnangu vea vähenemine okaspuudel MA ja KU (vastavalt 38% ja 22%), samuti lehtpuudel LM (32%) ja LV (21%), tagasihoidlikum aga HB ja KS korral (vastavalt 13% ja 7%). Kahel väikese arvulisel lehtpuul (PN, JA) võib täheldada mahuhinnangu RMSE mõningast kasvu. 'Headel' andmetel on uue mudeli täpsuse kasv võrreldes senise kõrgusmudeliga veelgi suurem, olles 31% võetuna üle kõigi puuliikide.

3. Puistu juurdekasvu ja suremuse mudelite loomine

Sissejuhatus

Juurdekasvu teemat käsitledes lähtume järgmistest määratlustest.

Kogujuurdekasv (*gross annual increment*) on kõigi kasvavate puude tagavara muutus aasta jooksul ehk puistu poolt aastas kasvatatav puidukogus. Väljendatakse tavaliselt tihumeetrites hektari kohta aastas (tm/ha/a).

Netojuurdekasvu (*net annual increment*) saame, kui kogujuurdekasvust maha arvata aasta jooksul surnud s.t. looduslikult väljalangenud puude maht.

Tagavara muut on kahe järjestikuse aasta tagavara vahe, kus on arvesse võetud nii kasvavate puude juurdekasv, suremus kui ka raied.

Meie eesmärk on hinnata SMI andmete põhjal Eesti metsade kogujuurdekasvu ja suremust (seega ka netojuurdekasvu).

Juurdekasvu hindamiseks on järgnevas välja pakutud kaks erinevat meetodit: 1) mudelipõhine meetod ja 2) mitmene imputeerimine. Mõlemad lähtuvad mudelpuude andmetest (kus 'tõde' on teada), liikudes edasi alalistele ja sealt omakorda juhumetsa meetodi abil ajutistele proovitükkidele. Tulemuseks saab kogujuurdekasvu hinnangu hektari kohta aastas.

Kogujuurdekasvu hindamiseks on esmalt vaja osata hinnata üksikpuu juurdekasvu.

3.1. Üksikpuu juurdekasvu hindamise meetodid

3.1.1. Puu juurdekasvu hindamise meetod 1

Vaatleme üksikpuu juurdekasvu ehk puu mahumuudu arvutamise tavameetodit, mis on ka kirjanduses laialt levinud ning mida me edaspidi nimetame paralleelselt ka '**lihtmeetodiks**'.

Meetod 1

Olgu d_1 ja d_2 puu rinnasdiameetri mõõtetulemused perioodi alg- ja lõpphetkel. Olgu antud mingi kõrgusmudel $h(d)$ (mis võib näiteks olla ka mudel (5)) rinnasdiameetri d järgi puu kõrguse h prognoosimiseks. Meetodi 1 korral

- 1) prognoositakse kõrgusmudeli abil kõrgused $h_1 = h(d_1)$ ja $h_2 = h(d_2)$,
- 2) kehtivat mahuvalemit (1) järgides arvutatakse puu alg- ja lõppmaht M_1 ja M_2 ,
- 3) leitakse puu juurdekasv ehk mahu muut $M_2 - M_1$ vaadeldaval ajaperioodil.

Osutub, et selline kahe lihtsa kõrgusprognoosi h_1 ja h_2 abil mahu juurdekasvu leidmine ei ole optimaalne lähenemine, kuna ei kasuta kogu olemasolevat infot. Nimelt on alalisel proovitükil

viimase mõõtmise hetkel olemas info ka 5 aasta taguse mõõtmise kohta ning juurdekasvu leidmisel on otstarbekas kasutada kogu olemasolev info ära hindamisprotsessi igas etapis.

Matemaatiliselt on asi selge: kõrguse tinglik keskväärus $E(h | \text{Info_t1}, \text{Info_t2})$ ei ole üldjuhul sama, mis $E(h | \text{Info_t1})$ - kui teame kahe ajamomendi infot, saab prognoose täpsustada ja seda tuleks ka võimalusel teha. Isegi juhul, kui valem $h(d)$ on kõrguse prognoosimiseks väga täpne (mis eeldatavalt annab ka täpse mahuhinnangu), ei pruugi meetod 1 samal ajal olla hea variant kõrguse ja mahu **juurdekasvu** arvutamiseks. Seda demonstreerime allpool ka numbriliselt.

3.1.2. Puu juurdekasvu hindamise meetod 2

Järgnevas kirjeldame teistsugust juurdekasvu hindamise meetodit 2, mis põhineb kõrguse juurdekasvu mudeldamisel alalistel proovitükkidel. Meetod 2 erineb tavapärasest puu juurdekasvu arvutamise viisist (meetodist 1 ehk lihtmeetodist) selle poolest, et pärast kõrgusmudeli abil h_2 prognoosimist (d_2 põhjal) prognoositakse järgmisena mitte h_1 vaid kõrguse juurdekasv Δh (lähtudes h_2 , d_1 ja d_2 väärtustest). Algekõrgus saadakse aga juba vahena $h_1 = h_2 - \Delta h$.

Täpsemalt, kordusklupitud puu mahumuudu prognoosimiseks pakume välja järgmise skeemi.

Meetod 2

- 1) Prognoosime kordusklupitud puu lõppkõrguse h_2 , kasutades d_2 ja uut kõrgusvalemit (5).
- 2) Prognoosime puu kõrguse muutuse Δh eelnenud 5 aasta jooksul, kasutades allolevat mudelit (9), mis on hinnatud mudelpuude andmetel.
- 3) Hindame puu algkõrgust vahega $h_1 = h_2 - \Delta h$.
- 4) Arvutame puu algmahu M_1 ja lõppmahu M_2 , kasutades kehtivat mahuvalemit (1).
- 5) Leiame puu mahumuudu $\Delta M = M_2 - M_1$ vaadeldaval ajaperioodil.

Mudel üksikpuu kõrguse muudu Δh prognoosimiseks on meie poolt pakutuna järgmine:

$$\Delta h = c_0 + \alpha_{aasta} + c_1 \cdot (h_2/d_2) \cdot \Delta d_+ + c_2 \cdot h_2 \cdot \Delta d_+, \quad (9)$$

kus

h_2 on puu prognoositud lõppkõrgus (kõrgusmudeli (5) abil);

d_2 on puu diameeter viimase mõõtmise ajal;

Δd_+ on puu diameetri muutus viimase 5 aasta jooksul (kui kasvas suuremaks; diameetri kahanemise korral loetakse tunnuse väärtuseks 0);

c_0 , c_1 , c_2 , α_{aasta} – on mudelpuude andmete põhjal hinnatavad parameetrid. Parameeter α_{aasta} hinnatakse iga lõppaasta jaoks eraldi.

Tabel.

Kõrguse muudu mudeli (9) parameetrid, hinnatud 2015-2024.a. andmete põhjal

	HB	KS	KU	LM	LV	MA	RE	XX
c_0	2.046290	6.077654	3.515979	12.092251	6.961233	7.401516	2.297512	6.896253
a_2016	7.924727	-1.816518	0.665491	-4.197583	-1.239273	-2.652187	16.101933	-3.046339
a_2017	-5.580465	-11.143787	-8.828188	-14.241974	-13.537465	-12.148698	-11.351970	-9.259179
a_2018	-4.239059	-14.055343	-10.883271	-19.722100	-11.080320	-12.340747	0.852223	-8.818883
a_2019	-2.580029	-13.714222	-10.732431	-18.363029	-12.036814	-9.874836	-0.213274	-6.931408
a_2020	-4.747453	-15.467464	-13.493807	-16.872140	-13.864637	-15.856894	-2.858215	-17.509667
a_2021	-6.335817	-14.822451	-11.180744	-14.322267	-13.640038	-13.298403	-3.593672	-8.430774
a_2022	6.254119	-4.599290	-2.626751	-9.927725	-3.912810	-6.903019	17.358022	-3.649344
a_2023	7.112609	-6.295737	-5.208066	-13.921469	-5.538007	-8.124841	-0.601819	-7.101276
a_2024	3.514699	-6.731310	-5.613751	-10.914466	-3.041607	-7.745524	18.480458	-1.472832
c_1	0.888106	0.991681	0.896172	0.965762	1.019773	0.953142	0.467936	0.857786
c_2	-0.001704	-0.001922	-0.000728	-0.002325	-0.003158	-0.001742	-0.000193	-0.001971

Märkus: Tabelis kasutatud puuliigigruppide tähitused on toodud punktis 2.2

Mudeli (9) selgitus. Suhe h_2/d_2 kirjeldab noore kasvava puu sihvakust. Arvutusvalem $(h_2/d_2) \cdot \Delta d$ kirjeldaks suhteliselt hästi puu kõrguse muutust nooremate puude jaoks, kus me ei pea arvestama asjaoluga, et puude vananedes see suhe muutub – vanuse lisandumisel kõrguse kasv taltub, aga diameeter jätkuvalt suureneb. Seega vanematel (suurema diameetriga puude) puhul juurdekasvu tempo aeglustumise tõttu lisandub kõrgust veidi vähem, kui suhte (h_2/d_2) järgi võiks oletada. Seda ajas (st puu diameetri kasvades) järjest väiksemat panust kõrgusesse kasvamisele saame kirjeldada liikme $(h_2/d_2)d_2 \Delta d = h_2 \Delta d$ abil, mille ees olev andmete põhjal hinnatav kordaja c_2 kirjeldab puu vananedes (diameetri kasvades) kahanevat kõrgusesse kasvamise tähtsust.

Kuna puu diameetri muutust on mõõdetud suhteliselt suure veaga, siis ei tasu kõrguse muutuse arvutamisel tugineda vaid diameetri mõõdetud muudule vaid kombineerida kahte hinnangut – võtta mingi kaaluga viie aasta „keskmist“ kõrguse juurdekasvu ja kombineerida seda siis üksikpuu mõõtmiste tulemusi kasutava hinnanguga $(h_2/d_2)\Delta d$ (mis kirjeldab konkreetset puud, aga on tänu Δd suurele mõõtmisveale ja ka puu lõppkõrguse prognoosi h_2 ebatäpsuse tõttu üsnagi ebatäpne). Kaalumise jaoks vajame liikme $(h_2/d_2)\Delta d$ ette ka hinnatavat parameetrit c_1 . Juurdekasvud on väga tugevalt sõltuvad aastast, seega tuleks iga viisaastaku jaoks hinnata eraldi parameeter, mis näitab kas tegemist oli heade kasvuaastatega või mitte (α_{aasta}). Aasta mõju on antud mudelis ülimalt oluline. Näiteks mändide korral suudab mudel ilma aasta mõjuta ära kirjeldada 17% kõrguse muudu varieeruvusest, koos viie aasta eripära arvestades aga 49,7% kõrguse muudu varieeruvusest.

Antud mudelit võib vajadusel mõnevõrra parandada. Näiteks võime konstandi c_0 – mis kirjeldab keskmist juurdekasvu – asendada millegagi, mis paremini võtaks arvesse

proovitüki valitsevaid tingimusi; näiteks suurusega $c_0' + c_3 (H/D)$, kus H ja D on proovitüki inventeeritud kõrgus ja diameeter. Saadud mudeli prognoosivõime võiks sel moel veidi paraneda, kuid arvutustest selgus, et täiendav võit täpsuses on vägagi minimaalne (mändide puhul suudaksime nüüd kirjeldada 50,0% kõrguse muudu varieeruvusest).

Mudeli (9) võimet prognoosida kõrguse muutust Δh iseloomustab järgmine tabel:

<u>plgrp</u>	<u>RMSEA</u>	<u>R²</u>
HB	1.519	0.421
KS	1.287	0.526
KU	1.128	0.540
LM	1.263	0.462
LV	1.264	0.473
MA	1.038	0.497
RE	1.205	0.510

Olles nüüd varustatud mudelitega, mis võimaldavad prognoosida puu diameetri järgi puu kõrgust (uus kõrguskõvera mudel (5)) ja diameetri muutuse järgi prognoosida kõrguse muutust (mudel (9)), võime kehtivat mahuvalemit kasutades leida üksikpuude mahu muutuse. Kui täpsed on sel viisil saadud hinnangud üksikpuude mahu muutustele? Järgnev tabel kirjeldab mudeli võimet prognoosida mudelpuude mahtude muutust viie aasta jooksul:

<u>plgrp</u>	<u>RMSEA</u>	<u>R²</u>
HB	0.1102	0.924
KS	0.0320	0.828
KU	0.0350	0.889
LM	0.0211	0.871
LV	0.0160	0.848
MA	0.0284	0.852
RE	0.0140	0.947
XX	0.0390	0.867

3.2. Meetodite 1 ja 2 võrdlus puu juurdekasvu hindamisel

Meie eesmärgiks on siin demonstreerida numbriliselt, et puu juurdekasvu hindamise meetod 2 töötab paremini kui meetod 1 (lihtmeetod). Huvi pakub eeskätt hinnangu süstemaatiline viga (hinnangu nihe). Meetodite omavaheliseks võrdlemiseks on kasutatud mudelpuude andmeid, kus tegelik puude juurdekasv on teada. Kuna mõlemal juhul kasutatakse uut kõrgusmudelit (mis

on kalibreeritud pika andmeakna (20a) pealt), siis märgime Lisas 3 toodud analüüsile toetudes, et suurte puuliikide korral oli aastail 2017-2022 näha selged kõrguse ülepõrguosiid.

On selge, et kohati ülehinnatud kõrgus tekitab nihkeid ka puude mahuhinnangutes. Alltoodud arvutused näitavad, et lihtmeetodit kasutades on juurdekasvu ülehinnang kordades suurem kui meetodi 2 korral.

```
library(tidyverse)

## — Attaching core tidyverse packages — tidyverse 2.0.0 —
## ✓ dplyr      1.1.4      ✓ readr      2.1.5
## ✓ forcats    1.0.1      ✓ stringr    1.5.2
## ✓ ggplot2    4.0.0      ✓ tibble     3.3.0
## ✓ lubridate  1.9.4      ✓ tidyr      1.3.1
## ✓ purrr      1.1.0
## — Conflicts — tidyverse_conflicts() —
## X dplyr::filter() masks stats::filter()
## X dplyr::lag()     masks stats::lag()
## i Use the conflicted package (<http://conflicted.r-lib.org/>) to force all conflicts to
become errors

setwd('xxxx')
load("uued_smi_andmed.RData")
load("k6rgusmudel.RData")
source("prognoosifunktsioon_h_ver3.R") #selle nimega fail on toodud aruande Lisas 5
```

1) Abifunktsioon mahtude arvutamiseks:

```
olemas=c('MA','LH','SD','KU','NU','TS','TO','KS','PN','HB','LM','LV',
         'PP','RE','TL','OP','TA','VA','SA','TM','JA','KP','SP','PI','EL','TY')
kordajad_maht=data.frame(puuliigid=c("MA","LH","SD","KU","NU","TS","TO","KS","PN","HB",
B","LM","LV",
'PP','RE',"","TL","OP","TA","VA","SA","TM","JA","KP","SP","PI","EL","TY"),a=c(0.3571,0.4216
,0.4080,0.4723,0.4033),b=c(0.660,0.181,0.757,-0.608,0),c=c(2.156,1.190,0.801,0,1.586),d=c(-
8.312,-1.309,-10.707,12.724,1.440))
kordajad_maht_maatriks=matrix(NA,nrow=length(olemas),ncol=4)
row.names(kordajad_maht_maatriks)=olemas
colnames(kordajad_maht_maatriks)=c('a','b','c','d')
for(i in 1:length(olemas)){
  kordajad_maht_maatriks[i,]=kordajad_maht|>filter(grepl(olemas[i], puuliigid))|>select(a:d
)|>as.matrix()
}

V_puu=function(d13,h,a=0.3571,b=0.66,c=2.156,d=-8.312){
  #eeldame, et d on millimeetrites ja h on detsimeetrites
  #puu mahu valem metsakorralduse eeskirja kohaselt, vaikeparameetrid
  #männi, lehise ja seederänni jaoks
  D=d13/10
  H=h/10
  return(0.0000785*D^2*H*(a+b/D+c/H+d/(D*H)))
}

V_maht_koik=function(puuliigid,d13,h,kordajadM=kordajad_maht_maatriks){
  #eeldus on, et puuduvaid andmeid pole
```

```
D=d13/10
H=h/10
return(0.0000785*D*H*(rowSums(kordajadM[puuliigid,]*cbind(D,1,D/H,1/H))))
}
puugrupp=function(puuliik){
return( as.character( factor(puuliik,
levels=
c("MA","LH","SD", "KU","NU","TO","TS", "KS","PN","TL","EL", "HB","PP", "LM", "LV",
"TA","SA","VA","JA","KP", "RE","PI","TM","SP","OP","TY"),
labels=
c("MA","MA","MA", "KU","KU","KU","KU", "KS","KS","KS","KS", "HB","HB", "LM", "LV",
"XX","XX","XX","XX","XX", "RE","RE","RE","RE","RE","RE"))))
}
```

2) Puu juurdekasvu hinnangu nihe 'lihtmeetodil'

Lihtmeetodile vastavalt prognoositakse (lähtudes alg- ja lõppdiameetrist) mudeli (5) abil puu algkõrgus ja lõppkõrgus, siis ametliku mahualemi (1) järgi alg- ja lõppmaht ning leitakse nende vahe ehk juurdekasv. Saadud juurdekasve võrreldakse aastate 2019-2023 tegelike mahumuutudega mudelpuudel ning leitakse keskmine prognoosiviga (hinnangu nihe) üle kõigi puuliikide.

Tekitame andmestiku kaks korda klupitud mudelpuudest

```
mudelpuud2=mudelpuud|>
  filter(aasta>2004,!is.na(d_13))|>
  inner_join(prtosaa|>select(prto_aid,prto_aid,peapuuliik,inv_diameeter,inv_g1,inv_g2,inv_korgus,onsaar,inv_bon,vanus,
                           maakategooria, majandkategooria))|>
  inner_join(prt|>select(prt_aid,koord_n,koord_e))|>
  filter(maakategooria %in% c('M','MM'),puuliik %in% row.names(kordajad_maht_maatriks))|>
  mutate(valepuu=1*(puuliik!=peapuuliik),
         plgrp=puugrupp(puuliik),
         V=V_maht_koik(puuliik,d_13,h_puu),
         hprognoos=prognoos_h(k6rgusmudel, d_13, puuliik, peapuuliik, inv_diameeter, inv_korgus, inv_g1, inv_g2),
         relaskoop=pmax((coalesce(inv_g1,inv_g2,5)+coalesce(inv_g2,inv_g1,5))/2,5)
         )|>
  mutate(V_progn=V_maht_koik(puuliik,d_13,hprognoos))

## Joining with `by` = join_by(prto_aid)`
## Joining with `by` = join_by(prt_aid)`

## [1] "Prognoosime puuliiki: HB"
## [1] "Prognoosime puuliiki: KS"
## [1] "Prognoosime puuliiki: KU"
## [1] "Prognoosime puuliiki: LM"
## [1] "Prognoosime puuliiki: LV"
## [1] "Prognoosime puuliiki: MA"
## [1] "Prognoosime puuliiki: RE"
## [1] "Prognoosime puuliiki: XX"

mudelpuud_teada=mudelpuud2|>
  select(amudelpuud_id,h_puu,d_13,puuliik,V,em_id,aasta,
        puuliik,peapuuliik,plgrp,inv_diameeter,inv_korgus,onsaar,inv_bon,vanus,majandkategooria,koord_n,koord_e,valepuu,V,hprognoos,relaskoop,maakategooria,V_progn)
mudelpuud_teada=mudelpuud_teada|>
  inner_join(mudelpuud_teada,by=c("em_id"="amudelpuud_id"),suffix=c("", ".enne"))|>
  filter(aasta>2004,!is.na(d_13),!is.na(d_13.enne),puuliik==puuliik.enne,aasta-aasta.enne==
```

```
5)|>
  select(-puuliik.enne,-plgrp.enne,-em_id.enne,-onsaar.enne,-aasta.enne,-vanus.enne,-valepuu.enne,-koord_e.enne,-koord_n.enne,-maakategooria.enne)|>
  mutate(V_muut=V-V.enne,V_muut_progn=V_progn-V_progn.enne)
puugrupid=unique(mudelpuud_tead$a$plgrp)
aasta_tasemed=unique(mudelpuud_tead$a$aasta)

tabel=mudelpuud_tead|>filter(aasta>2018,d_13>=80)|>#group_by(plgrp)|>
  summarize(kesk_maht=mean(V),kesk_mahu_viga=mean(V_progn-V),sd_mahu_viga=sd(V_progn-V)/sqrt(n()),kesk_juurde=mean(V_muut),kesk_juurde_viga=mean(V_muut_progn-V_muut),sd_kesk_juurde_viga=sd(V_muut_progn-V_muut)/sqrt(n()))
tabel

##   kesk_maht kesk_mahu_viga sd_mahu_viga kesk_juurde kesk_juurde_viga
## 1 0.5188201    0.007325239    0.0007400904    0.0375859    0.006463315
##   sd_kesk_juurde_viga
## 1          0.0005457623
```

Tulemus: Lihtmeetodit kasutades on puude juurdekasv ülehinnatud keskmiselt 0.00646 võrra (hinnangu nihe), mis moodustab umbes 17 protsenti tegelikust keskmisest juurdekasvust (0.03758).

3) Puu juurdekasvu hinnangu nihe meetodil 2

```
mudelpuud_tead$a$V_muut_progn2=NA
for(pg in puugrupid){
  prognoosimiseks=mudelpuud_tead|>filter(plgrp==pg)|>mutate(aasta_f=factor(aasta,levels=aasta_tasemed),dmuut_p=pmax(d_13-d_13.enne,0),suhe=hprognoos/d_13,h_muut=h_puu-h_puu.enne)
  hmuut=lm(h_muut~aasta_f+suhe*dmuut_p+hprognoos*dmuut_p,data=prognoosimiseks)
  indeks=(mudelpuud_tead$a$plgrp==pg) & (!is.na(mudelpuud_tead$a$d_13.enne))
  V_muut_progn=prognoosimiseks|>mutate(hprognoos.enne=hprognoos-predict(hmuut,prognoosimiseks))|>
  mutate(V_muut=V_maht_koik(puuliik,d_13,hprognoos)-V_maht_koik(puuliik,d_13.enne,hprognoos.enne))|>
  pull(V_muut)
  mudelpuud_tead$a$V_muut_progn2[indeks]=V_muut_progn
}

tabel=mudelpuud_tead|>filter(aasta>2018,d_13>=80)|>#group_by(plgrp)|>
  summarize(kesk_maht=mean(V),kesk_mahu_viga=mean(V_progn-V),sd_mahu_viga=sd(V_progn-V)/sqrt(n()),kesk_juurde=mean(V_muut),kesk_juurde_viga=mean(-V_muut+V_muut_progn2),sd_kesk_juurde_viga=sd(V_muut-V_muut_progn2)/sqrt(n()))
tabel

##   kesk_maht kesk_mahu_viga sd_mahu_viga kesk_juurde kesk_juurde_viga
## 1 0.5188201    0.007325239    0.0007400904    0.0375859    0.001729916
##   sd_kesk_juurde_viga
## 1          0.0003644336
```

Tulemus: Meetodit 2 kasutades on puude juurdekasv ülehinnatud keskmiselt 0.0017 võrra (hinnangu nihe), mis moodustab umbes 4% tegelikust keskmisest juurdekasvust (0.03758), olles seega märgatavalt väiksem kui lihtmeetodi korral tekkinud hinnangu nihe.

Samas on mõlemad veahinnangud küllalt väikese standardhälbega, nii et nende erinevus ei ole käsitletav juhuslikuna.



Kokkuvõttes ilmnes, et lihtmeetodil (meetod 1) juurdekasvu arvutamine pole parim tee - juba mudelpuude andmetel tekib juurdekasvu hinnangutes tõsine ülehindamine. Samas seda ei tohiks lugeda lihtsalt uue kõrgusmodeli (5) veaks – meetod 1 poleks optimaalne ka siis, kui kõrguse mudeliga saadavad prognoosid ise oleksid nihketa. Läbiviidud võrdluse tulemusena võtame edaspidi **kogujuurdekasvu arvutuse aluseks meetodi 2**.

3.3. Kogujuurdekasvu arvutamine mudeli põhiselt

Järgnevas on kirjeldatud kogujuurdekasvu arvutamise detailne algoritm. Üksikpuu juurdekasvu arvutamisel (millele tugineb kõik ülejäänu) kasutatakse meetodit 2, mille paremus 'lihtmeetodi' ees sai äsja tõestatud.

Mõned üldised märkused mudelite kasutamise kohta:

- Oluline eeldus on, et mudelpuude mõõtmised (diameeter ja kõrgus) on võimalikult täpsed ja ilma süstemaatiliste vigadeta. Sel juhul mudelpuud annavad parimat infot teiste puude kõrguse ja juurdekasvu kohta ajas, võimaldades kohandada mahu ja juurdekasvu hinnanguid vastavalt looduslike tingimuste muutusele ajas.
- Kui mingi kõrguse või juurdekasvu mudel annab vaadeldaval perioodil nihkega hinnangud mudelpuude andmetel (kus tegelikud väärtused on teada), siis kanduvad need nihked üle ka koguhinnangutesse ja seda isegi siis, kui muud vaatlusandmed on kõrge kvaliteediga. Seega on esmalt vaja kontrollida mudeli töökindlust mudelpuudel.
- Kui näiteks mudelpuude kõrguse (ja/või proovitükkide inventeeritud kõrguste) mõõtmises toimub sisuline muutus (nt hakatakse täpsemalt mõõtma ja siis selgub, et varem oli hinnangutes väike nihe), siis tekib arvutatud juurdekasvudes ja seega ka hinnangutesse pikaajaline süstemaatiline nihe.
- Kuna andmeid mõjutavad üldised tingimused muutuvad ajas ja need kõik ei ole mõõdetavate tunnustega igal proovitükil otseselt kirjeldatavad, siis on eriti ainult osalist infot omavate prognooside tegemisel hea kasutada ajaliselt lähedasi andmeid (nt sama perioodi tagavara proovitükkide andmed).
- Juurdekasvu prognooside laiendamine kasvukoha proovitükkidele, kus suur osa täpsemaid prognoose võimaldavaid tunnuseid on puudu, võib olla kasutatav üleeestiliste hinnangute mõningase täpsustamise jaoks, kuid nende prognooside kasutamine osapiirkondades arvutuste tegemisel ei pruugi olla põhjendatud.
- Käesolevas projektis pakume kombinatsiooni statistiliste mudelitega prognoosimisest (et mudelpuude abil prognoosida võimalikult hästi andmeid alalistel proovitükkidel), ja prognoosimisest masinõppe meetoditega (juhumets), mida kasutame üleminekul alalistelt proovitükkidelt ajutistele proovitükkidele. Juhumetsa mudelid on võimalised

kohanduma keeruliste seostega tunnuste vahel ning on küllalt kindlad ülesobitamise vältimisel.

Juurdekasvu hinnangu leidmise sammud on järgmised:

1. Leiame juurdekasvud perioodil 2020-2024 kaks korda klupitud puudele, kasutades puu juurdekasvu hindamise meetodit 2 (kus kõrgusmudel ja kõrguse juurdekasvu mudel hinnatakse 2015-2024 mudelpuudel).
2. Laiendame juurdekasvud kõigile klupitud puudele (sh ka ajutistel proovitükkidel klupitud puud), kasutades sammul 1. leitud kaks korda klupitud puude juurdekasvu andmeid. Sel eesmärgil kasutame juhumetsa meetodit, mis 5a taguse diameetri puudumist kompenseerib rea taustatunnuste infoga.
3. Leiame proovitükiosadele klupitud puudele vastavad mahu juurdekasvud, korrutades väikseid klupitud puid alalistel proovitükkidel 16-ga ja ajutistel 4-ga.
4. Leiame peenpuude juurdekasvud kordusmõõtmistega alalistele proovitükkidele.
5. Laiendame peenpuude juurdekasvud ajutistele tagavara proovitükkidele.
6. Lisame raiutud puude juurdekasvud.
7. Arvutame hinnangu kogu Eesti kohta.

3.3.1. Andmete ja abifunktsioonide sisselugemine

```
library(tidyverse)

## — Attaching core tidyverse packages — tidyverse 2.0.0 —
## ✓ dplyr      1.1.4      ✓ readr      2.1.5
## ✓ forcats    1.0.1      ✓ stringr    1.5.2
## ✓ ggplot2    4.0.0      ✓ tibble     3.3.0
## ✓ lubridate  1.9.4      ✓ tidyr      1.3.1
## ✓ purrr      1.1.0
## — Conflicts — tidyverse_conflicts() —
## X dplyr::filter() masks stats::filter()
## X dplyr::lag()    masks stats::lag()
## I Use the conflicted package (<http://conflicted.r-lib.org/>) to force all conflicts to become errors

load("uued_smi_andmed2.RData")
load("kõrgusmudel_al2015.RData") #sobitatud kõrgusmudelid
source("prognoosifunktsioon_h_ver3.R") #Uue kõrgusmudeli kohaselt prognooside arvutamise funktsioon
#Loeme sisse abifunktsioonid
# V_maht_koik() mahu arvutamiseks riikliku valemiga
# puugrupp() - grupi leidmiseks liigi järgi
# arvuta_tingimustel - hinnangu ja vahhinnangu leidmine varasema projekti meetodiga.
source("abifunktsioonid_maht_viga.R")
```

3.3.2. Juurdekasvude leidmine klupitud puude jaoks

Mudeli sobitamise alusandmed: kaks korda klupitud mudelpuud

```
modelpuud2=modelpuud|>
  filter(is.na(seisukord)|grepl("U",seisukord))|> #surnud ja raiutud puude eemaldamine
  filter(aasta>2004,!is.na(d_13))|> #enne 2004 oli olulisi tunnuseid puudu
  inner_join(prtosa|>select(prto_aid,prt_aid,peapuuliik,inv_diameeter,inv_g1,inv_g2,inv_kor
gus,onsaar,inv_bon,vanus,maakategooria, majandkategooria,arenguklass),by="prto_aid")|>
  inner_join(prt|>select(prt_aid,koord_n,koord_e),by="prt_aid")|>
  filter(maakategooria %in% c('M','MM'),puuliik %in% row.names(kordajad_maht_maatriks))|>
  mutate(valepuu=1*(puuliik!=peapuuliik),
         plgrp=puugrupp(puuliik),
         V=V_maht_koik(puuliik,d_13,h_puu),
         hprognos=prognos_h(k6rgusmodel, d_13, puuliik, peapuuliik, inv_diameeter, inv_ko
rgus, inv_g1, inv_g2),
         relaskoop=pmax((coalesce(inv_g1,inv_g2,5)+coalesce(inv_g2,inv_g1,5))/2,5)
  )

modelpuud_teada=modelpuud2|>
  select(amudelpuud_id,h_puu,d_13,puuliik,V,em_id,aasta,
        puuliik,peapuuliik,plgrp,inv_diameeter,inv_korgus,onsaar,inv_bon,vanus,majandkate
goria,koord_n,koord_e,valepuu,V,hprognos,relaskoop,maakategooria,arenguklass)
modelpuud_teada=modelpuud_teada|>
  inner_join(modelpuud_teada,by=c("em_id"="amudelpuud_id"),suffix=c("", ".enne"))|>
  filter(aasta>2004,!is.na(d_13),!is.na(d_13.enne),puuliik==puuliik.enne,aasta-aasta.enne==
5)|>
  mutate(V_muut=V-V.enne)|>
  select(-puuliik.enne,-plgrp.enne,-onsaar.enne,-aasta.enne,-vanus.enne,-valepuu.enne,-koor
d_e.enne,-koord_n.enne,-maakategooria.enne)#|>
modelpuud_teada_orig=modelpuud_teada
```

Fikseerime ajaperioodi, mille jaoks mahumuute arvutada ja tekitame vastava andmestiku.

```
aasta_alg=2020
aasta_lopp=2024
veerud=c('puu_id','h_puu','d_13','puuliik','aasta','peapuuliik','plgrp','inv_diameeter','re
laskoop','inv_korgus','onsaar','inv_bon','hprognos','valepuu','vanus','koord_n','koord_e',
'maakategooria','majandkategooria','h_puu.enne','d_13.enne','inv_diameeter.enne','relaskoop
.enne','inv_korgus.enne','inv_bon.enne','hprognos.enne','arenguklass','arenguklass.enne')
klupp2=klupp|>filter(arv_grp=='ES',aasta>2004,!is.na(puuliik),!is.na(d_13uus),!(inf_vp %in%
c("V","W")),puuliik %in% row.names(kordajad_maht_maatriks) )|>
  select(klupp_id,atrakt_id,d_13uus,aasta,prto_aid,puuliik,em_id)|>
  inner_join(prtosa|>filter(maakategooria %in% c("M","MM"))|>select(prto_aid,peapuuliik,van
us,inv_diameeter,inv_bon,inv_g1,inv_g2,inv_korgus,onsaar,prt_aid,maakategooria,majandkate
goria,arenguklass),by="prto_aid")|>
  mutate(plgrp=puugrupp(puuliik),
         valepuu=1*(puuliik!=peapuuliik),
         hprognos=prognos_h(k6rgusmodel, d_13uus, puuliik, peapuuliik, inv_diameeter, inv
_korgus, inv_g1, inv_g2),
         relaskoop=pmax((coalesce(inv_g1,inv_g2,5)+coalesce(inv_g2,inv_g1,5))/2,5),
         V=V_maht_koik(puuliik,d_13uus,hprognos)
  )|>
  select(-inv_g1,-inv_g2)|>
  inner_join(prt|>select(prt_aid,koord_n,koord_e),by="prt_aid")
klupp_imputeerimiseks=klupp2|>filter(aasta=aasta_alg,aasta<=aasta_lopp)|>
  left_join(klupp2,by=c("em_id"="klupp_id"),suffix=c("", ".enne"))|>
```

```
mutate(puu_id=paste0("k",klupp_id),h_puu=NA,h_puu.enne=NA,d_13=d_13uus,d_13.enne=d_13uus.enne,V_muut=NA)|>
  select(all_of(veerud))
puugrupid=unique(mudelpuud_teada$plgrp)
tulemused=NULL
#leiam proovitükiosad, mida on ka varem klupitud
on_varem_klupitud=klupp2|>filter(!is.na(em_id))|>select(prto_aid,aasta)|>distinct_all()
```

Sobitame iga puuliigi korral mudeli kõrguse muudu prognoosimiseks korduvkluppimisel mõõdetud andmeid kasutades (st uue ja varasema diameetri ja muude proovitüki andmete põhjal), leiam lõppkõrguse prognoosi, seejärel algkõrguse prognoosi (= lõppkõrgus – muut), ja täidame mahu valemi abil mahumuudud kirjetel, kus on kaks mõõtmist.

Väikeste klupitud puude puhul (alla 8cm), millele vastavat proovitükki on 5 aastat tagasi klupitud, kuid mis on uued puud, loeme juurdekasvuks kogu mahu.

Seejärel kasutame vaadeldaval perioodil kaks korda klupitud puude andmeid selleks, et prognoosida mahumuudud ainult üks kord klupitud puudele.

```
set.seed(20251201) #juhuslikkuse fikseerimiseks juhumetsa mudelis
#veerud=c('V_muut','d_13','puuliik','aasta','peapuuliik','plgrp','inv_diameeter','relaskoop',
', 'inv_korgus','onsaar','inv_bon','hprognoos','valepuu','vanus','koord_n','koord_e','maakat',
'egooria','majandkategoria','d_13.enne','inv_diameeter.enne','relaskoop.enne','inv_korgus.e',
'nne','inv_bon.enne','hprognoos.enne','arenguklass','arenguklass.enne')
aasta_tasemed=unique(mudelpuud_teada$aasta)
klupp_imputeerimiseks$V_muut=NA
for(pg in puugrupid){
  print(pg)
  puud_modelisse=mudelpuud_teada|>filter(plgrp==pg)
  hmuut=lm(h_muut~aasta_f+suhe*dmuut_p+hprognoos*dmuut_p,
    data=puud_modelisse|>
      mutate(
        aasta_f=factor(aasta,levels=aasta_tasemed),
        dmuut_p=pmax(d_13-d_13.enne,0),
        suhe=hprognoos/d_13,h_muut=h_puu-h_puu.enne
      )
  )
  indeks=(klupp_imputeerimiseks$plgrp==pg) & (!is.na(klupp_imputeerimiseks$d_13.enne))
  prognoosimiseks=klupp_imputeerimiseks[indeks,]|>mutate(aasta_f=factor(aasta,levels=aast
a_tasemed),dmuut_p=pmax(d_13-d_13.enne,0),suhe=hprognoos/d_13)
  V_muut_progn=prognoosimiseks|>
    mutate(hprognoos.enne=hprognoos-predict(hmuut,prognoosimiseks))|>
    mutate(V_muut=V_maht_koik(puuliik,d_13,hprognoos)-V_maht_koik(puuliik,d_13.enne,hprog
noos.enne))|>
    pull(V_muut)
  klupp_imputeerimiseks$V_muut[indeks]=V_muut_progn
  #loeme 4-8 cm puude, mis on korduvalt klupitud proovitükiosal, aga eelmisel korral olid
puudu, juurdekasvu võrdseks mahuhinnanguga
  vp_uus=klupp_imputeerimiseks|>filter(plgrp==pg,is.na(V_muut),d_13<80)|>
  mutate(klupp_id=as.numeric(substring(puu_id,2))|>inner_join(klupp|>select(klupp_id,prto_ai
d))|>inner_join(on_varem_klupitud)|>select(puu_id)
  indeks2= klupp_imputeerimiseks$puu_id %in% vp_uus$puu_id
  klupp_imputeerimiseks$V_muut[indeks2]=V_maht_koik(klupp_imputeerimiseks$puuliik[indeks2
],klupp_imputeerimiseks$d_13[indeks2],klupp_imputeerimiseks$hprognoos[indeks2])

  #kanname andmed üle üks kord klupitud puudele juhumetsa mudeliga
  mudel_esmakordne=ranger::ranger(V_muut~.,data=klupp_imputeerimiseks[indeks | indeks2,]|
>select(-ends_with(".enne")))
  print(mudel_esmakordne$r.squared)
```

```

    indeks3=(klupp_imputeerimiseks$plgrp==pg) & (is.na(klupp_imputeerimiseks$V_muut))
    klupp_imputeerimiseks$V_muut[indeks3]=predict(mudel_esmakordne,klupp_imputeerimiseks[in
deks3,])$predictions
  }

## [1] "KS"

## Joining with `by = join_by(klupp_id)`
## Joining with `by = join_by(aasta, prto_aid)`

# [1] "KU"

## Joining with `by = join_by(klupp_id)`
## Joining with `by = join_by(aasta, prto_aid)`

## [1] "MA"

## Joining with `by = join_by(klupp_id)`
## Joining with `by = join_by(aasta, prto_aid)`

## [1] "HB"

## Joining with `by = join_by(klupp_id)`
## Joining with `by = join_by(aasta, prto_aid)`

## [1] "XX"

## Joining with `by = join_by(klupp_id)`
## Joining with `by = join_by(aasta, prto_aid)`

## [1] "LM"

## Joining with `by = join_by(klupp_id)`
## Joining with `by = join_by(aasta, prto_aid)`

## [1] "LV"

## Joining with `by = join_by(klupp_id)`
## Joining with `by = join_by(aasta, prto_aid)`

## [1] "RE"

## Joining with `by = join_by(klupp_id)`
## Joining with `by = join_by(aasta, prto_aid)`

```

3.3.3. Klupitüd puude juurdekasvu leidmine osaproovitükkide jaoks

Kõigepealt nn suured puud (üle 8 cm diameetriga):

```

prtosuured=klupp_imputeerimiseks|>
  mutate(on_suur=(d_13 >= 80)*1)|>
  mutate(klupp_id=as.numeric(substring(puu_id,2))|>
  inner_join(klupp|>select(klupp_id,prto_aid))|>
  group_by(prto_aid)|>
  summarize(muudu_maht=sum(on_suur*V_muut))|>
  ungroup()|>

```

```
inner_join(prtosa|>select(prto_aid,pindosak,pindala))|>
mutate(muudu_maht_ha=muudu_maht/pindala*10000/5)

## Joining with `by = join_by(klupp_id)`
## Joining with `by = join_by(prto_aid)`
```

Peenemate klupid puude puhul lähtume sellest, et jaotamata alalistel proovitükkidel toimub selliste puude kluppimine sageli sektoris, st iga puu esindab tegelikult 16 puud proovitükil ja jaotamata ajutistel proovitükkidel toimub kluppimine 3,5m raadiusega ringis, kus iga puu esindab 4 puud kogu proovitükil.

```
prtosa_prog_vaiksed=klupp_imputeerimiseks|>
mutate(on_vaike=(d_13 < 80)*1)|>
mutate(klupp_id=as.numeric(substring(puu_id,2))|>
inner_join(klupp|>select(klupp_id,prto_aid),by="klupp_id")|>
inner_join(prtosa|>select(prto_aid,atrakt_id,pindosak,on_sektor),by="prto_aid")|> # võtame
e külge info, kas on sektorit kasutatud
inner_join(trakt|>select(atrakt_id,tr_tyyp),by="atrakt_id")|> #trakti tüübi info
mutate(kordaja=if_else(on_sektor=="J",if_else(tr_tyyp==2,4*pindosak,16*pindosak),1))|> #
Leiame väikestele puudele kordaja mahu arvestamiseks, arvestades sektori olemasolu
mutate(V_muut=on_vaike*kordaja*V_muut)|> #arvestame iga puu muutu vastavalt kordajale
group_by(prto_aid)|>
summarize(muudu_maht=sum(V_muut))|>
ungroup()|>
inner_join(prtosa|>select(prto_aid,pindosak,pindala),by="prto_aid")|>
mutate(muudu_maht_ha=muudu_maht/pindala*10000/5)
```

3.3.4. Peenpuude juurdekasv

Peenpuude (diameter alla 4 cm) jaoks kasutame andmestikus olemasolevaid mahtude arvutusi. Ekspertide hinnangul on mõistlik eeldada, et kõik peenpuud on kasvanud viimase viie aasta jooksul, kusjuures peenpuude suremus on mahu mõttes marginaalne. Seetõttu arvestame peenpuid ainult juurdekasvude leidmisel. Alalistel tagavara proovitükkidel on seega andmed peenpuude juurdekasvu arvutamiseks olemas ning me laiendame need ajutistele tagavara proovitükkidele.

Peenpuude mahud üle kõigi puuliikide proovitükiosade kaupa:

```
prtosa_prog_peenpuud=peenpuud|>
mutate(prto_id=substring(prto_aid,6))|>
group_by(prto_aid,prto_id,aasta)|>
summarize(maht=sum(arv_tagavara*kordne),.groups="drop")|>
inner_join(prtosa|>select(prto_aid,pindosak,pindala),by="prto_aid")|>
mutate(muudu_maht_ha=coalesce(pindosak*maht/5,0)/pindala*10000)
```

3.3.5. Raiutud puude juurdekasv

Leiame alaliste proovitükkide korral kõikide raiutuks märgitud puude jaoks juurdekasvud, lugedes selliste puude juurdekasvud, mis eelmisel vaatlusel ei olnud grupis ES, nulliks ning hinnates ülejäänud rajatud puude juurdekasvuks 0.8 eelmise perioodi jaoks prognoositud juurdekasvust hooajalise raide puhul ning 0.3 varemraiutud puude viie aasta juurdekasvust. Kordajad on hinnangulised ja tuginevad sellel, et mahtude mõttes on hooajaliseks märgitud umbes 40% kõigist ER ja VR klassi puudest, mistõttu võib eeldada, et tegelikult on hooajalise

raide all kirjas umbes kahe vaatluseelse aasta raiutud puud ning VR all on siis esimese kolme aasta puud.

Selle abil kanname andmed üle ajutistele tagavara proovitükkidele eraldi ER ja VR klassides.

Kõigepealt leiame juurdekasvu prognoosid eelmise perioodi andmetele. See sisuliselt kordab varasemat koodi, aga andmed tuleb valida vastavalt perioodile.

```
aasta_alg2=aasta_alg-5
aasta_lopp2=aasta_lopp-5
veerud=c('puu_id','h_puu','d_13','puuliik','aasta','peapuuliik','plgrp','inv_diameeter','re
laskoop','inv_korgus','onsaar','inv_bon','hprognoos','valepuu','vanus','koord_n','koord_e',
'maakategooria','majandkategooria','h_puu.enne','d_13.enne','inv_diameeter.enne','relaskoop
.enne','inv_korgus.enne','inv_bon.enne','hprognoos.enne','arenguklass','arenguklass.enne')
klupp_imputeerimiseks_eelmine=klupp2|>filter(aasta>=aasta_alg2,aasta<=aasta_lopp2)|>
  left_join(klupp2,by=c("em_id"="klupp_id"),suffix=c("", ".enne"))|>
  mutate(puu_id=paste0("k",klupp_id),h_puu=NA,h_puu.enne=NA,d_13=d_13uus,d_13.enne=d_13uu
s.enne,V_muut=NA)|>
  select(all_of(veerud))
set.seed(20251201) #juhuslikkuse fikseerimiseks juhumetsa mudelis
aasta_tasemed=unique(mudelpuud_teada$aasta)
for(pg in puugrupid){
  print(pg)
  puud_mudelis=muudpuud_teada|>filter(plgrp==pg)
  hmuut=lm(h_muut~aasta_f+suhe*dmuut_p+hprognoos*dmuut_p,
    data=puud_mudelis|>
      mutate(
        aasta_f=factor(aasta,levels=aasta_tasemed),
        dmuut_p=pmax(d_13-d_13.enne,0),
        suhe=hprognoos/d_13,h_muut=h_puu-h_puu.enne
      )
  )
  indeks=(klupp_imputeerimiseks_eelmine$plgrp==pg) & (!is.na(klupp_imputeerimiseks_eelmin
e$d_13.enne))
  prognoosimiseks=klupp_imputeerimiseks_eelmine[indeks,]|>mutate(aasta_f=factor(aasta,lev
els=aasta_tasemed),dmuut_p=pmax(d_13-d_13.enne,0),suhe=hprognoos/d_13)
  V_muut_progn=prognoosimiseks|>
    mutate(hprognoos.enne=hprognoos-predict(hmuut,prognoosimiseks))|>
    mutate(V_muut=V_maht_koik(puuliik,d_13,hprognoos)-V_maht_koik(puuliik,d_13.enne,hprog
noos.enne))|>
    pull(V_muut)
  klupp_imputeerimiseks_eelmine$V_muut[indeks]=V_muut_progn
  #Loeme 4-8 cm puude, mis on korduvalt klupitud proovitükiosal, aga eelmisel korral
  olid puudu, juurdekasvu võrdseks mahuhinnangu
  vp_uus=klupp_imputeerimiseks_eelmine|>filter(plgrp==pg,is.na(V_muut),d_13<80)|>
  mutate(klupp_id=as.numeric(substring(puu_id,2)))|>inner_join(klupp|>select(klupp_id,prto_ai
d))|>inner_join(on_varem_klupitud)|>select(puu_id)
  indeks2= klupp_imputeerimiseks_eelmine$puu_id %in% vp_uus$puu_id
  klupp_imputeerimiseks_eelmine$V_muut[indeks2]=V_maht_koik(klupp_imputeerimiseks_eelmine
$puuliik[indeks2],klupp_imputeerimiseks_eelmine$d_13[indeks2],klupp_imputeerimiseks_eelmine
$hprognoos[indeks2])

  #kanname andmed üle üks kord klupitud puudele juhumetsa mudeliga
  model_esmakordne=ranger::ranger(V_muut~.,data=klupp_imputeerimiseks_eelmine[indeks | in
deks2,]|>select(-ends_with(".enne")))
  print(model_esmakordne$r.squared)
  indeks3=(klupp_imputeerimiseks_eelmine$plgrp==pg) & (is.na(klupp_imputeerimiseks_eelmin
e$V_muut))
  klupp_imputeerimiseks_eelmine$V_muut[indeks3]=predict(model_esmakordne,klupp_imputeerim
```

```
iseks_eelmine[indeks3,])$predictions
}

## [1] "KS"

## Joining with `by = join_by(klupp_id)`
## Joining with `by = join_by(aasta, prto_aid)`

## [1] "KU"

## Joining with `by = join_by(klupp_id)`
## Joining with `by = join_by(aasta, prto_aid)`

## [1] "MA"

## Joining with `by = join_by(klupp_id)`
## Joining with `by = join_by(aasta, prto_aid)`

## [1] "HB"

## Joining with `by = join_by(klupp_id)`
## Joining with `by = join_by(aasta, prto_aid)`

## [1] "XX"

## Joining with `by = join_by(klupp_id)`
## Joining with `by = join_by(aasta, prto_aid)`

## [1] "LM"

## Joining with `by = join_by(klupp_id)`
## Joining with `by = join_by(aasta, prto_aid)`

## [1] "LV"

## Joining with `by = join_by(klupp_id)`
## Joining with `by = join_by(aasta, prto_aid)`

## [1] "RE"

## Joining with `by = join_by(klupp_id)`
## Joining with `by = join_by(aasta, prto_aid)`
```

Tekitame andmestiku, kus on juurdekasv (vähemalt prognooside põhjal) teada.

```
set.seed(20251201)
raie_juurde=klupp|>filter(aasta>=aasta_alg,aasta<=aasta_lopp,arv_grp %in% c('ER','VR','SR'))
|>
  inner_join(prtos|>filter(maakategooria %in% c("M","MM"))|>select(prto_aid),by="prto_aid"
)|>
  select(prto_aid,klupp_id,em_id,arv_grp)|>
  left_join(klupp_imputeerimiseks_eelmine|>mutate(klupp_id=as.numeric(substring(puu_id,2)))
|>select(klupp_id,V_muut),by=c("em_id","klupp_id"))|>
  mutate(V_muut=if_else(is.na(em_id),NA,coalesce(V_muut,0)*if_else(arv_grp=='ER',0.8,if_els
e(arv_grp=='VR',0.3,0.5))))
raie_juurde_tead=raie_juurde|>filter(!is.na(em_id))
treening_raie=raie_juurde_tead|>
  filter(arv_grp!='VR')|>
  inner_join(klupp|>select(klupp_id,puuliik,d_0),by="klupp_id")|>
  select(prto_aid,V_muut,puuliik,d_0,arv_grp)|>
  inner_join(prtos|>select(prto_aid,prto_aid,peapuuliik,inv_diameeter,inv_g1,inv_g2, inv_ko
```

```
rgus,onsaar,inv_bon,vanus,maakategooria, majandkategooria,arenguklass,pindosak), by="prto_a  
id")|>  
  inner_join(prt|>select(prt_aid,koord_n,koord_e),by="prt_aid")|>  
  select(-prto_aid,-prt_aid)  
model_raie=ranger::ranger(V_muut~.,data=treening_raie)
```

Tekitame andmestiku raiutud puude juurdekasvude prognoosidega hooajalise ja surnud
puude raide osas

```
indeks=is.na(raie_juurde$em_id)  
prognoosimiseks=raie_juurde[indeks,]>inner_join(klupp|>select(klupp_id,puuliik,d_0),by="kl  
upp_id")|>  
  select(prto_aid,V_muut,puuliik,d_0,arv_grp)|>  
  inner_join(prtosa|>select(prto_aid,prt_aid,peapuuliik,inv_diameeter,inv_g1,inv_g2, inv_ko  
rgus,onsaar,inv_bon,vanus,maakategooria, majandkategooria,arenguklass,pindosak), by="prto_a  
id")|>  
  inner_join(prt|>select(prt_aid,koord_n,koord_e),by="prt_aid")|>  
  select(-prto_aid,-prt_aid)  
raie_juurde$V_muut[indeks]=predict(model_raie,prognoosimiseks)$predictions  
  
prtosa_prog_raie_hooaeg=raie_juurde|>  
  filter(arv_grp!='VR')|>  
  group_by(prto_aid)|>  
  summarize(muudu_maht=sum(V_muut))|>  
  ungroup()|>  
  inner_join(prtosa|>select(prto_aid,pindosak,pindala),by="prto_aid")|>  
  mutate(muudu_maht_ha=muudu_maht/pindala*10000/5)
```

Leiame varemraaiutud puude andmed proovitükiosade kaupa

```
prtosa_prog_raie_VR_alaline=raie_juurde|>  
  filter(arv_grp=='VR',!is.na(em_id))|>  
  group_by(prto_aid)|>  
  summarize(muudu_maht=sum(V_muut))|>  
  ungroup()|>  
  inner_join(prtosa|>select(prto_aid,pindosak,pindala),by="prto_aid")|>  
  mutate(muudu_maht_ha=muudu_maht/pindala*10000/5)
```

Kasutame neid andmeid koos proovitükiosade jaoks leitud muude raiete summaarsete
andmetega selleks, et prognoosida varasemat raiet ajutistele tagavara proovitükkidele. Kuna
ajutistel proovitükkidel ei ole seda tüüpi raie kohta mingeid mõõdetud andmeid, siis on
sellised prognoosid paratamatult üsna ebatäpsed ja ei pruugi olla kasutatavad lokaalsemate
(st mitte kogu Eesti metsamaa kohta käivate) hinnangute arvutamisel.

```
treening_VR=klupp|>filter(aasta>=aasta_alg,aasta<=aasta_lopp)|>  
  filter(!is.na(em_id))|>  
  select(prto_aid)|>distinct_all()|>  
  inner_join(prtosa|>filter(maakategooria %in% c("M","MM"))|>select(prto_aid,peapuuliik,van  
us,inv_diameeter,inv_bon,inv_g1,inv_g2,inv_korgus,onsaar,prt_aid,maakategooria,majandkatego  
ria,arenguklass),by="prto_aid")|>  
  left_join(prtosa_prog_raie_VR_alaline|>select(prto_aid,muudu_maht_ha),by="prto_aid")|>  
  left_join(prtosa_prog_raie_hooaeg|>mutate(maht_hooaeg=muudu_maht_ha)|>select(prto_aid,mah  
t_hooaeg))|>  
  left_join(prtosa_prog_suured|>mutate(maht_suured=muudu_maht_ha)|>select(prto_aid,mah  
t_suured))|>  
  mutate(  
    muudu_maht_ha=coalesce(muudu_maht_ha,0),  
    maht_hooaeg=coalesce(maht_hooaeg,0),
```

```

    maht_suured=coalesce(maht_suured,0)
  )

## Joining with `by = join_by(prto_aid)`
## Joining with `by = join_by(prto_aid)`

  mudel_vr=ranger::ranger(muudu_maht_ha~.,data=treening_VR)

  #prognoosimise andmestik sama loogikaga
  prognoosimiseks=klupp|>filter(aasta>=aasta_alg,aasta<=aasta_lopp)|>
    select(prto_aid)|>distinct_all()|>
    anti_join(treening_VR|>select(prto_aid))|>
    inner_join(prtosa|>filter(maakategooria %in% c("M","MM"))|>select(prto_aid,peapuulik,v
anus,inv_diameeter,inv_bon,inv_g1,inv_g2,inv_korgus,onsaar,prt_aid,maakategooria,majandkate
gooria,arenguklass),by="prto_aid")|>
    left_join(prtosa_prog_raie_hooaeg|>mutate(maht_hooaeg=muudu_maht_ha)|>select(prto_aid,m
aht_hooaeg))|>
    left_join(prtosa_prog_suured|>mutate(maht_suured=muudu_maht_ha)|>select(prto_aid,maht_s
uured))|>
    mutate(
      maht_hooaeg=coalesce(maht_hooaeg,0),
      maht_suured=coalesce(maht_suured,0)
    )

## Joining with `by = join_by(prto_aid)`
## Joining with `by = join_by(prto_aid)`
## Joining with `by = join_by(prto_aid)`

  prognoosid=predict(mudel_vr,prognoosimiseks)$predictions
  prtosa_prog_raie_VR_ajutine=prognoosimiseks|>select(prto_aid)|>
    mutate(muudu_maht=NA,pindosak=1,pindala=157,muudu_maht_ha=prognoosid)
  prtosa_prog_raie_VR=rbind(prtosa_prog_raie_VR_alaline,prtosa_prog_raie_VR_ajutine)

```

3.3.6. Hinnangu arvutamine kogu Eesti metsamaa kohta

Paneme kõigepealt kõik prognoosid kokku

```

prtosa_prog=prtosa|>filter(aasta>=aasta_alg,aasta<=aasta_lopp)|>select(prto_aid)|>
  left_join(prtosa_prog_peenpuud|>select(prto_aid,muudu_maht_ha)|>rename(dVpeen=muudu_maht_
ha))|>
  left_join(prtosa_prog_vaiksed|>select(prto_aid,muudu_maht_ha)|>rename(dVvaike=muudu_maht_
ha))|>
  left_join(prtosa_prog_suured|>select(prto_aid,muudu_maht_ha)|>rename(dVsuur=muudu_maht_ha
))|>
  left_join(prtosa_prog_raie_hooaeg|>select(prto_aid,muudu_maht_ha)|>rename(dVraie_h=muudu_
maht_ha))|>
  left_join(prtosa_prog_raie_VR|>select(prto_aid,muudu_maht_ha)|>rename(dVraie_vr=muudu_mah
t_ha))|>
  mutate(pindosak=1,muudu_maht_ha=coalesce(dVpeen,0)+coalesce(dVvaike,0)+
    coalesce(dVsuur,0)+coalesce(dVraie_h,0)+coalesce(dVraie_vr,0))

## Joining with `by = join_by(prto_aid)`
## Joining with `by = join_by(prto_aid)`
## Joining with `by = join_by(prto_aid)`
## Joining with `by = join_by(prto_aid)`
## Joining with `by = join_by(prto_aid)`

```

3.3.7. Keskmise juurdekasvu hindamine kogu metsamaal

Arvutame hinnangu koos veahinnanguga imputeeritud andmete baasil, kasutades mitmese imputeerimise metoodikat.

Kõigepealt tekitame hinnangute leidmiseks vajalikud abiandmestikud (kood varasemast projektist)

```
tmp_prt=prt%>%dplyr::filter(aasta>=aasta_alg,aasta<=aasta_lopp,psees=="J")
tmp_prtos_orig=tmp_prt%>%dplyr::select(aproovitykk_id)%>%inner_join(prtos,by="aproovitykk_id")%>%filter(on_sees=="J")
#tmp_klupp=tmp_prt%>%dplyr::select(aproovitykk_id)%>%inner_join(klupp)%>%filter(kaugus<=100,arv_d13>=8)
tmp_trakt=trakt%>%filter(aasta>=aasta_alg,aasta<=aasta_lopp)

#teisendame regulaarsele võrgule

theta=0 #Kui palju on regulaarne võrk pööratud horisontaaltelje suhtes
kyljepikkus=5000 #meetrites, regulaarse võrgustiku ühe ruudu küljepikkus
vec=c(cos(theta),sin(theta))
vec2=c(-sin(theta),cos(theta))
M=matrix(c(vec,vec2),byrow=TRUE,nrow=2)
abi_nihe=tmp_prt%>%pivot_wider(names_from = c(pkylg,pkaugus),values_from=c(koord_e,koord_n),id_cols = atrakt_id)%>%
  mutate(v1x=(koord_e_N_08-koord_e_W_08)/2,v1y=(koord_n_N_08-koord_n_W_08)/2,v2x=(koord_e_N_08-koord_e_E_08)/2,v2y=(koord_n_N_08-koord_n_E_08)/2)
vecs=abi_nihe %>%summarize(across(v1x:v2y,~mean(.x,na.rm=TRUE)))
v1=c(vecs$v1x[1],vecs$v1y[1])
v2=c(vecs$v2x[1],vecs$v2y[1])
teisendused=tibble(pkylg=rep(c("W","N","E","S"),each=4),pkaugus=rep(c("02","04","06","08"),times=4),
  muut_x=rep(c(v1[1],-v2[1],-v1[1],v2[1]),each=4)+rep(c(0.5,0,-0.5,-1),times=4)*rep(c(v2[1],v1[1],-v2[1],-v1[1]),each=4),
  muut_y=rep(c(v1[2],-v2[2],-v1[2],v2[2]),each=4)+rep(c(0.5,0,-0.5,-1),times=4)*rep(c(v2[2],v1[2],-v2[2],-v1[2]),each=4))
trakt_keskpunktiga=tmp_prt%>%inner_join(teisendused,by=c("pkylg","pkaugus"))%>%group_by(atrakt_id)%>%filter(row_number()==1)%>%
  ungroup()%>%mutate(koord_e=koord_e+muut_x,koord_n=koord_n+muut_y)%>%mutate(id=row_number())%>%dplyr::select(atrakt_id,koord_e,koord_n,id)
teisendatud=as.matrix(trakt_keskpunktiga%>%dplyr::select(koord_e,koord_n))%*%t(M)
colnames(teisendatud)=c("x","y")
trakt_teisendatud=trakt_keskpunktiga%>%dplyr::bind_cols(as_tibble(teisendatud))

xlimits=range(trakt_teisendatud$x)
ylimits=range(trakt_teisendatud$y)

xp=seq(xlimits[1]-2.5*kyljepikkus,xlimits[2]+2.5*kyljepikkus,by=kyljepikkus)
yp=seq(ylimits[1]-2.5*kyljepikkus,ylimits[2]+2.5*kyljepikkus,by=kyljepikkus)
i_x=as.numeric(cut(trakt_teisendatud$x,breaks=xp))
i_y=as.numeric(cut(trakt_teisendatud$y,breaks=yp))
trakt_piirkond=mutate(trakt_teisendatud,i=i_x,j=i_y)%>%dplyr::select(atrakt_id,i,j)
koik_indeksid=tibble(i=rep(1:max(i_x),each=max(i_y)),j=rep(1:max(i_y),times=max(i_x)))

#tekitame andmestiku, kus kasvukoha proovitükkide mahud on imputeeritud
tmp_prtos=tmp_prtos_orig|>filter(substr(prt_aid,11,12) %in% c("04","08")) |>
  left_join(prtos_prog|>dplyr::select(-pindosak),by = "prto_aid")|>
  mutate(maht=coalesce(muudu_maht_ha*pindosak,0))
cond_A=tmp_prtos$maakategooria %in% c("M","MM")
cond_y=cond_A
```

```
tmp=arvuta_tingimustel(cond_A,cond_y,"maht",1,8)

kv_imp=tmp$mu[1]
#arvutame standardhälbe hinnangu
sd_imp=tmp$sd[1]

viga_imp=sd_imp#sqrt(mean(sd_imp^2)+(m+1)/m*var(kv_imp))
hinnang_imp=mean(kv_imp)
paste0("Imputeerimise abil leitud hinnang: ",hinnang_imp,"", hinnangu standardhälve: ",viga_imp)

if(is.null(tulemused)){
  tulemused=data.frame(hinnang_imp=hinnang_imp,viga_imp=viga_imp)
} else {
  tulemused=rbind(tulemused,c(hinnang_imp,viga_imp))
}
tulemused
```

Arvutame hinnangu ainult alaliste proovitükkide pealt

```
tmp_prtosa_orig|>filter(substr(prt_aid,11,12) %in% c("04","08"))|>
  inner_join(trakt|>select(atrakt_id,tr_tyyp)|>filter(tr_tyyp%in% c(1,3)))|>
  left_join(prtosa_prog|>select(prto_aid,muudu_maht_ha),by="prto_aid")|>
  mutate(muut=coalesce(muudu_maht_ha*pindosak,0),pindosak=if_else(maakategooria%in% c('M','MM'),pindosak,0))|>summarize(hinnang=sum(muut)/sum(pindosak))

## Joining with `by` = join_by(atrakt_id)`
```

Ainult ajutiste proovitükkide pealt:

```
tmp_prtosa_orig|>filter(substr(prt_aid,11,12) %in% c("04","08"))|>
  inner_join(trakt|>select(atrakt_id,tr_tyyp)|>filter(tr_tyyp%in% c(2)))|>
  left_join(prtosa_prog|>select(prto_aid,muudu_maht_ha),by="prto_aid")|>
  mutate(muut=coalesce(muudu_maht_ha*pindosak,0),pindosak=if_else(maakategooria%in% c('M','MM'),pindosak,0))|>summarize(hinnang=sum(muut)/sum(pindosak))

## Joining with `by` = join_by(atrakt_id)`
```

Vaatame osade kaupa:

```
tmp_prtosa=tmp_prtosa_orig|>filter(substr(prt_aid,11,12) %in% c("04","08")) |>
  left_join(prtosa_prog_suured|>dplyr::select(prto_aid,muudu_maht_ha),by = "prto_aid")|>
  mutate(maht=coalesce(muudu_maht_ha*pindosak,0))
cond_A=tmp_prtosa$maakategooria %in% c("M","MM")
cond_y=cond_A
tmp_suured=arvuta_tingimustel(cond_A,cond_y,"maht",1,8)
tmp_prtosa=tmp_prtosa_orig|>filter(substr(prt_aid,11,12) %in% c("04","08")) |>
  left_join(prtosa_prog_vaiksed|>dplyr::select(prto_aid,muudu_maht_ha),by = "prto_aid")|>
  mutate(maht=coalesce(muudu_maht_ha*pindosak,0))
tmp_vaiksed=arvuta_tingimustel(cond_A,cond_y,"maht",1,8)
tmp_prtosa=tmp_prtosa_orig|>filter(substr(prt_aid,11,12) %in% c("04","08")) |>
  left_join(prtosa_prog_peenpuud|>dplyr::select(prto_aid,muudu_maht_ha),by = "prto_aid")|>
  mutate(maht=coalesce(muudu_maht_ha*pindosak,0))
tmp_peenpuud=arvuta_tingimustel(cond_A,cond_y,"maht",1,8)
tmp_prtosa=tmp_prtosa_orig|>filter(substr(prt_aid,11,12) %in% c("04","08")) |>
  left_join(prtosa_prog_raie_hooaeg|>dplyr::select(prto_aid,muudu_maht_ha),by = "prto_aid")|>
  mutate(maht=coalesce(muudu_maht_ha*pindosak,0))
tmp_raie_hooaeg=arvuta_tingimustel(cond_A,cond_y,"maht",1,8)
tmp_prtosa=tmp_prtosa_orig|>filter(substr(prt_aid,11,12) %in% c("04","08")) |>
  left_join(prtosa_prog_raie_VR|>dplyr::select(prto_aid,muudu_maht_ha),by = "prto_aid")|>
```

```
mutate(maht=coalesce(muudu_maht_ha*pindosak,0))
tmp_raie_VR=arvuta_tingimustel(cond_A,cond_y,"maht",1,8)
tulem=rbind(tmp_suured,tmp_vaikesed,tmp_peenpuud,tmp_raie_hooaeg,tmp_raie_VR)
cbind(tibble(nimed=c("suured","vaikesed","peenpuud","hooaja raie","varasem raie")),tulem)
```

3.4. Surnud puude mahu leidmine

Alalistel proovitükkidel saame leida surnuks märgitud puu korral, kas see oli eelmisel korral elus ning kui oli, siis loeme eelmise kluppimise mahu viie aasta suremise mahu hulka. Samuti loeme suremise hulka sellised puud, mis eelmisel korral olid klupitud elusana, aga on praegu kadunud. Sellised surnud puud, millel ei ole viie aasta tagust mõõtmist, ei loe alalisel proovitükil suremise hulka, kui seda tükki oli 5 aastat tagasi klupitud.

```
set.seed(20251201)
surnud=klupp|>filter(aasta>=aasta_alg,aasta<=aasta_lopp,arv_grp %in% c('SL','SS','XX','SR'))
|>
  inner_join(prtos|>filter(maakategooria %in% c("M","MM"))|>select(prto_aid),by="prto_aid")
|>
  select(prto_aid,klupp_id,em_id)|>
  left_join(klupp_imputeerimiseks_eelmine|>mutate(klupp_id=as.numeric(substring(puu_id,2)),
V=V_maht_koik(puuliik,d_13,hprognoos))|>select(klupp_id,V),by=c("em_id"="klupp_id"))
# kui oli em_id, aga polnud elus, siis loeme mahu nulliks
surnud=surnud|>mutate(V=if_else(!is.na(em_id)&is.na(V),0,abs(V)))
#sobitame mudeli ülejäänute prognoosimiseks
treenimiseks=surnud|>filter(!is.na(em_id))|>
  inner_join(klupp|>select(klupp_id,puuliik,d_13uus),by="klupp_id")|>
  select(prto_aid,V,puuliik,d_13uus)|>
  inner_join(prtos|>select(prto_aid,prto_aid,peapuuliik,inv_diameeter,inv_g1,inv_g2, inv_ko
rgus,onsaar,inv_bon,vanus,maakategooria, majandkategooria,arenguklass,pindosak), by="prto_a
id")|>
  inner_join(prt|>select(prt_aid,koord_n,koord_e),by="prt_aid")|>
  select(-prto_aid,-prt_aid)
model_surnud=ranger::ranger(V~,data=treenimiseks,importance="impurity")

#lisame prognoosid
indeks=is.na(surnud$em_id)
prognoosimiseks=surnud[indeks,]|>inner_join(klupp|>select(klupp_id,puuliik,d_13uus),by="klu
pp_id")|>
  select(prto_aid,V,puuliik,d_13uus)|>
  inner_join(prtos|>select(prto_aid,prto_aid,peapuuliik,inv_diameeter,inv_g1,inv_g2, inv_ko
rgus,onsaar,inv_bon,vanus,maakategooria, majandkategooria,arenguklass,pindosak), by="prto_a
id")|>
  inner_join(prt|>select(prt_aid,koord_n,koord_e),by="prt_aid")|>
  select(-prto_aid,-prt_aid)
surnud$V[indeks]=predict(model_surnud,prognoosimiseks)$predictions
```

Lisame tabelisse sellised eelmisel ajamomendil olnud puud, mida ei ole 5 aastat hiljem

```
kadunud_puudega=klupp2|>filter(aasta>=aasta_alg2,aasta<=aasta_lopp2)|>left_join(klupp,by=c(
"klupp_id"="em_id"))|>
  group_by(prto_aid,x)|>mutate(puudu=sum(is.na(aasta.y)),olemas=sum(!is.na(aasta.y))|>
  ungroup())|>filter(olemas>0,puudu>0)|>
  group_by(prto_aid.x,prto_aid.y)|>summarize(V_kokku=sum(V),.groups="drop")
kadunud_puudega
```

Vähemalt viimasel perioodil on kadunud puid nii vähe, et neid võib mahuarvestuses ignoreerida

```
#tekitame proovitükiosade andmed
prtosa_prog_surnud=surnud|>
  group_by(prto_aid)|>
  summarize(muudu_maht=sum(V))|>
  ungroup()|>
  inner_join(prtosa|>select(prto_aid,pindosak,pindala),by="prto_aid")|>
  mutate(muudu_maht_ha=muudu_maht/pindala*10000/5)
```

Arvutame surnud puude mahuhinnangud

```
if(!exists("tulemused_surnud")){tulemused_surnud=NULL}
#tekitame andmestiku, kus kasvukoha proovitükkide mahud on imputeeritud
tmp_prtosa=tmp_prtosa_orig|>filter(substr(prt_aid,11,12) %in% c("04","08")) |>
  left_join(prtosa_prog_surnud|>dplyr::select(-pindosak),by = "prto_aid")|>
  mutate(maht=coalesce(muudu_maht_ha*pindosak,0))
cond_A=tmp_prtosa$maakategooria %in% c("M","MM")
cond_y=cond_A
tmp=arvuta_tingimustel(cond_A,cond_y,"maht",1,8)

kv_imp=tmp$mu[1]
#arvutame standardhälbe hinnangu
sd_imp=tmp$sd[1]

viga_imp=sd_imp
hinnang_imp=mean(kv_imp)
paste0("Imputeerimise abil leitud hinnang: ",hinnang_imp,"", hinnangu standardhälve: ",viga_imp)

if(is.null(tulemused_surnud)){
  tulemused_surnud=data.frame(hinnang_imp=hinnang_imp,viga_imp=viga_imp)
} else {
  tulemused_surnud=rbind(tulemused_surnud,c(hinnang_imp,viga_imp))
}
tulemused_surnud
```

3.5. Kogujuurdekasvu hindamine mitmese imputeerimise meetodil

Täiendava kontrollina kogujuurdekasvu hindamisel pakume suhteliselt riskivaba alternatiivi näol välja mitmese imputeerimise meetodi, mis ei kasuta konkreetseid statistilisi mudeleid. Mitmese imputeerimise korral arvutatakse samale proovitükile üksteise järel välja terve seeria erinevaid prognoose (meie uurimuses 20 prognoosi), mistõttu prognoositud andmetesse tekib ka vajalik juhuslikkus ning see võimaldab veaarvutust läbi viia nõutud viisil (Zhang (2003)). Mitmese imputeerimise korral on vaja otsustada, millist algoritmi kasutada üksikprognoosi tekitamiseks. Antud juhul on sel eesmärgil kasutatud juhumetsa meetodit (Random Forests), mis on piisavalt paindlik suutmaks kasutada suurt hulka argumenttunnuseid, arvestades nii

nende omavahelisi seoseid kui ka võimalikke mittelineaarseid seoseid sihttunnusega ehk puistu juurdekasvuga (Breiman (2001)).

Mitmese imputeerimise rakendamisel

- 1) tekitatakse 20 täielikku juurdekasvu andmestikku,
- 2) iga sellise täisandmestiku põhjal leitakse juurdekasvu punkthinnang ning selle statistiline viga (standardhälve),
- 3) kõigi punkthinnangute ja nende standardhälvete põhjal leitakse lõpuks mitmese imputeerimise meetodi koondviga.

Mitmese imputeerimise meetodi koondviga (*viga_imp*) koosneb kahest komponendist, millest üks on 20 üksikandmestiku pealt saadud veadispersioonide aritmeetiline keskmine ja teine on kogujuurdekasvu 20 erineva punkthinnangu pealt arvutatud dispersioon. Lühidalt on see esitatav alljärgneva programmeerimise näitega

```
viga_imp=sqrt(mean(sd_imp^2)+(m+1)/m*var(kv_imp)),
```

kus m on korduste arv ($m=20$).

NB! Mitmese imputeerimise abil on alljärgnevas hinnatud üksnes see juurdekasvu komponent, mis vastab elusatele klupitud **suurtele puudele** (diameeter > 8cm).

Mitmese imputeerimise R-kood

```
library(tidyverse)

## — Attaching core tidyverse packages — tidyverse 2.0.0 —
## ✓ dplyr      1.1.4      ✓ readr      2.1.5
## ✓ forcats    1.0.1      ✓ stringr    1.5.2
## ✓ ggplot2    4.0.0      ✓ tibble     3.3.0
## ✓ lubridate  1.9.4      ✓ tidyr      1.3.1
## ✓ purrr      1.1.0
## — Conflicts — tidyverse_conflicts() —
## ✖ dplyr::filter() masks stats::filter()
## ✖ dplyr::lag()     masks stats::lag()
## i Use the conflicted package (<http://conflicted.r-lib.org/>) to force all conflicts to become errors

load("uued_smi_andmed2.RData")
```

3.5.1. Abifunktsioonid ja mudelpuude andmete organiseerimine mahuarvutuseks

Abifunktsioon mahtude arvutamiseks:

```
olemas=c('MA','LH','SD','KU','NU','TS','TO','KS','PN','HB','LM','LV',
         'PP','RE','TL','OP','TA','VA','SA','TM','JA','KP','SP','PI','EL','TY')
kordajad_maht=data.frame(puuliigid=c("MA","LH","SD","","","KU","NU","TS","TO","","","KS","PN","","","H
B","LM","LV",
'PP','RE','','','TL','OP','TA','VA','SA','TM','JA','KP','SP','PI','EL','TY'),a=c(0.3571,0.4216
,0.4080,0.4723,0.4033),b=c(0.660,0.181,0.757,-0.608,0),c=c(2.156,1.190,0.801,0,1.586),d=c(-
8.312,-1.309,-10.707,12.724,1.440))
```

```
kordajad_maht_maatriks=matrix(NA,nrow=(length(olemas)),ncol=4)
row.names(kordajad_maht_maatriks)=olemas
colnames(kordajad_maht_maatriks)=c('a','b','c','d')
for(i in 1:length(olemas)){
  kordajad_maht_maatriks[i,]=kordajad_maht|>filter(grepl(olemas[i], puuliigid))|>select(a:d)
}|>as.matrix()
}

V_puu=function(d13,h,a=0.3571,b=0.66,c=2.156,d=-8.312){
  #eeldame, et d on millimeetrites ja h on detsimeetrites
  #puu mahu valem metsakorralduse eeskirja kohaselt, vaikeparameetrid
  #männi, lehise ja seederänni jaoks
  D=d13/10
  H=h/10
  return(0.0000785*D^2*H*(a+b/D+c/H+d/(D*H)))
}

V_maht_koik=function(puuliigid,d13,h,kordajadM=kordajad_maht_maatriks){
  #eeldus on, et puuduvaid andmeid pole
  D=d13/10
  H=h/10
  return(0.0000785*D*H*(rowSums(kordajadM[puuliigid,]*cbind(D,1,D/H,1/H))))
}
puugrupp=function(puuliik){
  return( as.character( factor(puuliik,
    levels=
      c("MA","LH","SD", "KU","NU","TO","TS", "KS","PN","TL","EL", "HB","PP", "LM", "LV",
        "TA","SA","VA","JA","KP", "RE","PI","TM","SP","OP","TY"),
    labels=
      c("MA","MA","MA", "KU","KU","KU","KU", "KS","KS","KS","KS", "HB","HB", "LM", "LV",
        "XX","XX","XX","XX","XX", "RE","RE","RE","RE","RE","RE"))))
}

#kõrgusmudeli andmed ja funktsioonid
load("kõrgusmudel_al2015.RData") #juba sobitatud kõrgusmudel
source("prognoosifunktsioon_h_ver3.R")
```

Tekitame kõigepealt andmestiku, kus “tõde” on teada.

```
modelpuud2=modelpuud|>
  filter(is.na(seisukord)|grepl("U",seisukord))|> #surnud ja raiutud puude eemaldamine
  filter(aasta>2004,!is.na(d_13))|>
  inner_join(prtos|>select(prto_aid,prt_aid,peapuuliik,inv_diameeter,inv_g1,inv_g2,inv_korgus,onsaar,inv_bon,vanus,
    maakategooria, majandkategooria,arenguklass))|>
  inner_join(prt|>select(prt_aid,koord_n,koord_e))|>
  filter(maakategooria %in% c('M','MM'),puuliik %in% row.names(kordajad_maht_maatriks))|>
  mutate(valepuu=1*(puuliik!=peapuuliik),
    plgrp=puugrupp(puuliik),
    V=V_maht_koik(puuliik,d_13,h_puu),
    hprognoos=prognoos_h(kõrgusmudel, d_13, puuliik, peapuuliik, inv_diameeter, inv_korgus, inv_g1, inv_g2),
    relaskoop=pmax((coalesce(inv_g1,inv_g2,5)+coalesce(inv_g2,inv_g1,5))/2,5)
  )

## Joining with `by` = join_by(prto_aid)`
## Joining with `by` = join_by(prt_aid)`

modelpuud_tead=modelpuud2|>
  select(amodelpuud_id,h_puu,d_13,puuliik,V,em_id,aasta,
    puuliik,peapuuliik,plgrp,inv_diameeter,inv_korgus,onsaar,inv_bon,vanus,majandkate
```

```
ooria,koord_n,koord_e,valepuu,V,hprognoos,relaskoop,maakategooria,arenguklass)
mudelpuud_teada=mudelpuud_teada|>
  inner_join(mudelpuud_teada,by=c("em_id","amudelpuud_id"),suffix=c("", ".enne"))|>
  filter(aasta>2004,!is.na(d_13),!is.na(d_13.enne),puuliik==puuliik.enne,aasta-aasta.enne==
5)|>
  mutate(V_muut=V-V.enne)|>
  select(-puuliik.enne,-plgrp.enne,-em_id.enne,-onsaar.enne,-aasta.enne,-vanus.enne,-valepu
u.enne,-koord_e.enne,-koord_n.enne,-maakategooria.enne)
```

3.5.2. Imputeerimine kluppimise andmetel

Viime imputeerimise läbi ka kluppimise andmestikul. Selleks kõigepealt tekitame andmestiku, kus on olemas vajalikud andmed, sh prognoositud kõrgused.

```
klupp2=klupp|>filter(arv_grp=='ES',aasta>2004,!is.na(puuliik),!is.na(d_13uus))|>
  select(klupp_id,atrakt_id,d_13uus,aasta,prto_aid,puuliik,em_id)|>
  inner_join(prtos|>filter(maakategooria %in% c("M","MM"))|>select(prto_aid,peapuuliik,van
us,inv_diameeter,inv_bon,inv_g1,inv_g2,inv_korgus,onsaar,prto_aid,maakategooria,majandkatego
oria,arenguklass),by="prto_aid")|>
  mutate(plgrp=puugrupp(puuliik),
         valepuu=1*(puuliik!=peapuuliik),
         hprognoos=prognoos_h(k6rgusmudel, d_13uus, puuliik, peapuuliik, inv_diameeter, inv
_korgus, inv_g1, inv_g2),
         relaskoop=pmax((coalesce(inv_g1,inv_g2,5)+coalesce(inv_g2,inv_g1,5))/2,5)
  )|>
  select(-inv_g1,-inv_g2)|>
  inner_join(prt|>select(prt_aid,koord_n,koord_e))

## Joining with `by = join_by(prt_aid)`

save(klupp2,file="klupp_ES_korgusega.RData")
```

Fikseerime ajaperioodi, mille jaoks imputeerida ja tekitame vastava andmestiku

```
aasta_alg=2020
aasta_lopp=2024
veerud=c('puu_id','h_puu','d_13','puuliik','aasta','peapuuliik','plgrp','inv_diameeter','re
laskoop','inv_korgus','onsaar','inv_bon','hprognoos','valepuu','vanus','koord_n','koord_e',
'maakategooria','majandkategooria','h_puu.enne','d_13.enne','inv_diameeter.enne','relaskoop
.enne','inv_korgus.enne','inv_bon.enne','hprognoos.enne')
klupp_imputeerimiseks=rbind(
  mudelpuud_teada|>mutate(puu_id=paste0("m",amudelpuud_id))|> select(all_of(veerud)),
  klupp2|>filter(aasta>=aasta_alg,aasta<=aasta_lopp)|>
  left_join(klupp2,by=c("em_id","klupp_id"),suffix=c("", ".enne"))|>
  mutate(puu_id=paste0("k",klupp_id),h_puu=NA,h_puu.enne=NA,d_13=d_13uus,d_13.enne=d_13uu
s.enne)|>
  select(all_of(veerud))
)
```

Imputeerime väärtused:

```
m=20 #mitu komplekti andmetest tekitada
seed=20251031 #juhuslikkuse fikseerimine korratavuse huvides

klupp_imp <- mice::mice(klupp_imputeerimiseks,m=m,method="rf",seed=seed,maxit=10)

klupp_imp_long=complete(klupp_imp,action="long")
save(klupp_imp_long,file="klupp_imp_long_2020_2024.RData")
```

Leiame, kui täpselt klupitud puu keskmise juurdekasvu võiksime saada.

```
load("klupp_imp_long_2020_2024.RData")
m=max(klupp_imp_long$.imp)
klupp_imp_long_mahtudega=klupp_imp_long|>filter(grepl("^k",puu_id))|>select(puu_id,h_puu,h_puu.enne,d_13,d_13.enne,puuliik,plgrp,.imp,.id,koord_n,koord_e)|>
  mutate(V=V_maht_koik(puuliik,d_13,h_puu),V.enne=V_maht_koik(puuliik,d_13.enne,h_puu.enne),
  V_muut=V-V.enne)
```

Leiame hinnangud ja veahinnangud vähemalt 8cm diameetriga (viimasel mõõtmisel) puude keskmise juurdekasvu

```
hinnang_imp_klupp=klupp_imp_long_mahtudega|>
  filter(d_13>=80)|>
  group_by(.imp,plgrp)|>summarize(kesk=mean(V_muut),halve=var(V_muut)/n(),mitu=n())|>arrange(p,plgrp)|>
  ungroup()|>
  group_by(plgrp)|>
  summarize(hinnang=mean(kesk),sd=sqrt(mean(halve)+(1+1/m)*var(kesk)),mitu=mean(mitu))

## `summarise()` has grouped output by `.imp`. You can override using the
## `.groups` argument.

hinnang_imp_klupp
```

Agregeerime juurdekasvud proovitükiosadele ja normeerime ümber hektari ja aasta kohta.

```
prtosa_imp=klupp_imp_long_mahtudega|>
  filter(d_13 >= 80)|>
  mutate(klupp_id=as.numeric(substring(puu_id,2))|>
  inner_join(klupp|>select(klupp_id,prto_aid))|>
  group_by(prto_aid,.imp)|>
  summarize(muudu_maht=sum(V_muut))|>
  ungroup()|>
  inner_join(prtosa|>select(prto_aid,pindosak,pindala))|>
  mutate(muudu_maht_ha=muudu_maht/pindala*10000/5)

## Joining with `by = join_by(klupp_id)`
## `summarise()` has grouped output by 'prto_aid'. You can override using the
## `.groups` argument.
## Joining with `by = join_by(prto_aid)`
```

3.5.3. Hinnangu arvutamine kogu Eesti metsamaa kohta

See on varasema projekti kood, kohandatud juurdekasvudele.

Defineerime abifunktsioonid

```
arvuta_sd3=function(andmestik,kv){
  kohandatud=andmestik%>%mutate(hinnang=tinglik_kv-pind_sees*kv)%>%dplyr::select(i,j,hinnang)
  kohandatud2=kohandatud%>%mutate(i=i+1,h2=hinnang)%>%dplyr::select(i,j,h2)
  kohandatud3=kohandatud%>%mutate(i=i+1,j=j+1,h3=hinnang)%>%dplyr::select(i,j,h3)
  kohandatud4=kohandatud%>%mutate(j=j+1,h4=hinnang)%>%dplyr::select(i,j,h4)
  vead=kohandatud%>%
    inner_join(kohandatud2,by=c("i","j"))%>%
    inner_join(kohandatud3,by=c("i","j"))%>%
    inner_join(kohandatud4,by=c("i","j"))%>%
    mutate(vead=(hinnang-h2+h3-h4)/2)%>%pull(vead)
  return(sqrt(sum(vead^2))/sum(andmestik$pind_sees))
}
arvuta_tingimustel=function(cond_A,cond_y,y_string,pind_A,max_per_trakt=16){
```

```
#cond_A on tmp_prtosa jaoks arvesse tulevate proovitükkide tingimus
#cond_y on tingimus, mis määrab, kas summeeritava suuruse nullitakse või mitte
#y_string annab summeeritava väärtuse veerunime stringina
#samm on selleks, et saaks hõrendatud võrgul korduvaid arvutusi teha
#pind_A on tingimusele A vastava Eesti osa teadaolev pindala
#max_per_trakt annab, mitu proovitükki trakti kohta maksimaalselt arvesse läheb
taiendatud_trakt=tmp_prtosa%>%mutate(mm_osak=if_else(cond_y,get({y_string})),0),pind_o
sak=if_else(cond_A,pindosa,0)%>%
group_by(atrakt_id)%>%summarize(mm_proc=sum(mm_osak)/max_per_trakt,pind_sees=sum(pind_osa
k)/max_per_trakt)
piirkond=taiendatud_trakt%>%inner_join(trakt_piirkond,by="atrakt_id")%>%group_by(i,j)%>%s
ummarize_at(vars(mm_proc,pind_sees),.funs=~mean(.x))%>%ungroup()
#Laiendame nullidega
piirkond_laiendatud=koik_indeksid%>%left_join(piirkond,by=c("i","j"))%>%mutate(tinglik_kv
=coalesce(mm_proc,0),pind_sees=coalesce(pind_sees,0))%>%select(i,j,tinglik_kv,pind_sees)
tulem=tibble(mu=rep(NA,1),sd=rep(NA,1))#,sd2=rep(NA,mitu))
osakaal=sum(piirkond_laiendatud$tinglik_kv)/sum(piirkond_laiendatud$pind_sees)
tulem$mu[1]=osakaal*pind_A
sd3_tul=arvuta_sd3(piirkond_laiendatud,osakaal)
#tulem$sd[rida]=sd3_tul$sd*pind_tykk*samm^2
tulem$sd[1]=sd3_tul*pind_A
return(tulem)
}

tmp_prt=prt%>%dplyr::filter(aasta>=aasta_alg,aasta<=aasta_lopp,psees=="J")
tmp_prtosa_orig=tmp_prt%>%dplyr::select(aproovitykk_id)%>%inner_join(prtosa,by="aproovitykk
_id")%>%filter(on_sees=="J")
#tmp_klupp=tmp_prt%>%dplyr::select(aproovitykk_id)%>%inner_join(klupp)%>%filter(kaugus<=100
,arv_d13>=8)
tmp_trakt=trakt%>%filter(aasta>=aasta_alg,aasta<=aasta_lopp)
```

Kasutame varem pakutud meetodit vaatluste teisendamiseks regulaarsele võrgule

```
theta=0 #Kui palju on regulaarne võrk pööratud horisontaaltelje suhtes
kyljepikkus=5000 #meetrites, regulaarse võrgustiku ühe ruudu küljepikkus
vec=c(cos(theta),sin(theta))
vec2=c(-sin(theta),cos(theta))
M=matrix(c(vec,vec2),byrow=TRUE,nrow=2)
abi_nihe=tmp_prt%>%pivot_wider(names_from = c(pkylg,pkaugus),values_from=c(koord_e,koord_n)
,id_cols =atrakt_id)%>%
mutate(v1x=(koord_e_N_08-koord_e_W_08)/2,v1y=(koord_n_N_08-koord_n_W_08)/2,v2x=(koord_e_N
_08-koord_e_E_08)/2,v2y=(koord_n_N_08-koord_n_E_08)/2)
vecs=abi_nihe %>%summarize(across(v1x:v2y,~mean(.x,na.rm=TRUE)))
v1=c(vecs$v1x[1],vecs$v1y[1])
v2=c(vecs$v2x[1],vecs$v2y[1])
teisendus=tibble(pkylg=rep(c("W","N","E","S"),each=4),pkaugus=rep(c("02","04","06","08"),
times=4),
muut_x=rep(c(v1[1],-v2[1],-v1[1],v2[1]),each=4)+rep(c(0.5,0,-0.5,-1),tim
es=4)*rep(c(v2[1],v1[1],-v2[1],-v1[1]),each=4),
muut_y=rep(c(v1[2],-v2[2],-v1[2],v2[2]),each=4)+rep(c(0.5,0,-0.5,-1),tim
es=4)*rep(c(v2[2],v1[2],-v2[2],-v1[2]),each=4))
trakt_keskpunktiga=tmp_prt%>%inner_join(teisendus,by=c("pkylg","pkaugus"))%>%group_by(atr
akt_id)%>%filter(row_number()==1)%>%
ungroup()%>%mutate(koord_e=koord_e+muut_x,koord_n=koord_n+muut_y)%>%mutate(id=row_number(
))%>%dplyr::select(atrakt_id,koord_e,koord_n,id)
teisendatud=as.matrix(trakt_keskpunktiga%>%dplyr::select(koord_e,koord_n))%*%t(M)
colnames(teisendatud)=c("x","y")
trakt_teisendatud=trakt_keskpunktiga%>%dplyr::bind_cols(as_tibble(teisendatud))

xlimits=range(trakt_teisendatud$x)
ylimits=range(trakt_teisendatud$y)
```

```
xp=seq(xlimits[1]-2.5*kyljepikkus,xlimits[2]+2.5*kyljepikkus,by=kyljepikkus)
yp=seq(ylimits[1]-2.5*kyljepikkus,ylimits[2]+2.5*kyljepikkus,by=kyljepikkus)
i_x=as.numeric(cut(trakt_teisendatud$x,breaks=xp))
i_y=as.numeric(cut(trakt_teisendatud$y,breaks=yp))
trakt_piirkond=mutate(trakt_teisendatud,i=i_x,j=i_y)%>%dplyr::select(atrakt_id,i,j)
koik_indeksid=tibble(i=rep(1:max(i_x),each=max(i_y)),j=rep(1:max(i_y),times=max(i_x)))
```

3.5.4. Keskmise juurdekasvu hindamine kogu metsamaal

Arvutame hinnangu koos veahinnanguga imputeeritud andmete baasil, kasutades mitmese imputeerimise metoodikat

```
kv_imp=rep(NA,m)
sd_imp=rep(NA,m)
for(i in 1:m){
  #tekitame i-nda andmestiku, kus kasvukoha proovitükkide mahud on imputeeritud
  tmp_prtosa=tmp_prtosa_orig|>filter(prtliik==1)|>left_join(prtosa_imp|>dplyr::filter(.imp=
=i)|>dplyr::select(-.imp,-pindosak),by = "prto_aid")|>mutate(maht=coalesce(muudu_maht_ha*pi
ndosak,0))
  cond_A=tmp_prtosa$maakategooria %in% c("M","MM")
  cond_y=cond_A
  tmp=arvuta_tingimustel(cond_A,cond_y,"maht",1,8)

  kv_imp[i]=tmp$mu[1]
  #arvutame standardhälbe hinnangu
  sd_imp[i]=tmp$sd[1]
}
viga_imp=sqrt(mean(sd_imp^2)+(m+1)/m*var(kv_imp))
hinnang_imp=mean(kv_imp)
paste0("Imputeerimise abil leitud hinnang: ",hinnang_imp,"", hinnangu standardhälve: ",viga_
imp)
```

Märkus: Üaltoodud mitmese imputeerimise kood annab üksnes kogujuurdekasvu selle komponendi, mis vastab elusate üle 8 cm klupitud puude juurdekasvule - puudub peenpuude ja väljalangenud (surnud ja raiutud) puude juurdekasvu arvestus.

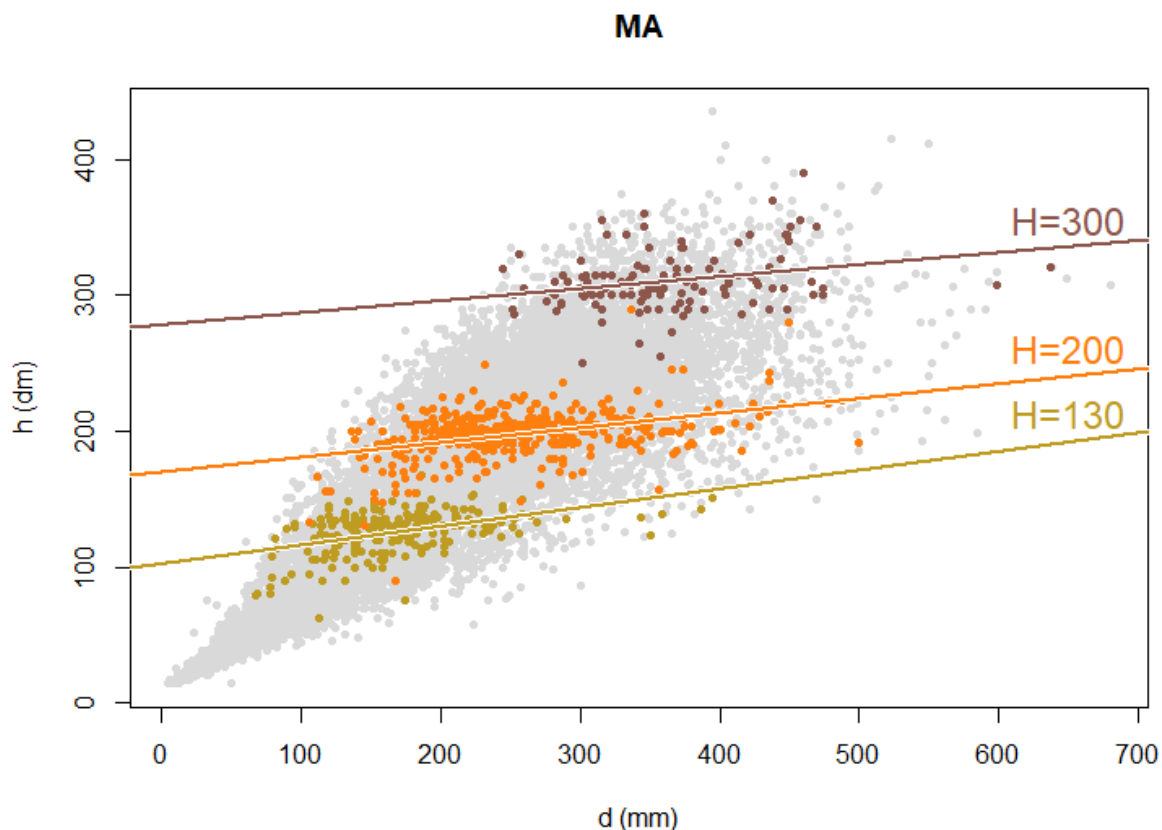
4. Kirjanduse loetelu

1. Statistiline metsainventeerimine. Arvutusmetoodiline juhend. 2018.
2. Keskkonnaministri 16.jaan. 2009 määruse nr. 2 “Metsa korraldamise juhend“ lisa 11 (Tagavara ja täiuse arvutamiseks kasutatavad matemaatilised mudelid)
3. Statistilise metsainventuuri (SMI) traktivõrgu analüüs ja arvutusmeetoditega seotud uuring. Lepingu aruanne, TÜ MSI (2023).
4. Statistilise metsainventuuri (SMI) arendamine. Lepingu aruanne, TÜ MSI (2024).
5. KHG inventuuri LULUCF sektori heite arvutamiseks vajalike metsakasvumudelite väljatöötamine. Lepingu nr 2-1/27/2021 aruanne (2023). Keskkonnaagentuur.
6. Zhang, P. (2003), Multiple Imputation: Theory and Method. *International Statistical Review / Revue Internationale de Statistique*, **71** (3), 581-592.
7. Breiman, L. (2001). Random Forests. *Machine Learning*. **45** (1), 5–32.
8. Cochran W. G. (1977) Sampling techniques, 3rd edition. Wiley, New York. 428 p.
9. Fridman J., Holm S., Nilsson M., Nilsson P., Ringvall A.H., Ståhl G. (2014). Adapting National Forest Inventories to changing requirements – the case of the Swedish National Forest Inventory at the turn of the 20th century. *Silva Fennica*, vol. **48** no.3 article id 1095. <https://doi.org/10.14214/sf.1095>
10. Ranney B., Cruse T., Hägglund B., Jonasson H., Swärd J. (1987). Designing a new national forest survey for Sweden. *Studia Forestalia Suecica* 177. 29 p.
11. Kangas A. & Maltamo M. (Eds.) (2006). *Forest Inventory. Methodology and Applications* (2006). Springer.
12. Tomppo E., Heikkinen J., Henttonen H.M., Ihalainen A., Katila M., Mäkelä H., Tuomainen T., Vainikainen N. (2011). Designing and Conducting a Forest Inventory – case: 9th National Forest Inventory of Finland. *Managing Forest Ecosystems* 22. Springer, 305 p. DOI: 10.1007/978-94-007-1652-0.
13. Korhonen K. T., Ahola A., Heikkinen J., Henttonen H. M., Hotanen J.-P., Ihalainen A., Melin M., Pitkänen J., Rättyä M., Sirviö M., Strandström M. (2021). Forests of Finland 2014–2018 and their development 1921–2018. *Silva Fennica* vol. 55 no. 5 article id 10662. <https://doi.org/10.14214/sf.10662>
14. Korhonen K.T., Rättyä M., Haakana H., Heikkinen J., Hotanen J.-P., Kuronen M., Pitkänen J. (2024). Forests of Finland 2019–2023 and their development 1921–2023. *Silva Fennica* vol. 58 no. 5 article id 24045. 47 p. <https://doi.org/10.14214/sf.24045>
15. Tomppo E., Gschwantner Th., Lawrence M., McRoberts R.E. (eds.). (2010). In: *National forest inventories: pathways for common reporting*. Springer. 612 p. <http://dx.doi.org/10.1007/978-90-481-3233-1>.
16. Sims A. (2022). *Principles of National Forest Inventory Methods. Theory, Practice, and Examples from Estonia*. Springer.
17. Traat I., Inno J. (1997). *Tõenäosuslik valikuuring*. TÜ Kirjastus, 211 lk.
18. Kaev T. (2024). Puu kõrguskõvera mudelite analüüs Eesti metsa kasvukäigu proovitükkide andmetel. Magistritöö. Eesti Maaülikool.

Lisa 1. Uue kõrgusmudeli detailne põhjendus

Siinkohal antakse punktis 2.2 esitatud uue kõrguse mudeli (5) – (8) detailsem põhjendus.

Meid huvitab seos puu diameetri d ja puu kõrguse h vahel. Osutub, et olulist rolli selle seose juures mängib proovitüki inventeeritud kõrgus H . Olukorda peegeldab hästi järgmine joonis, mis kujutab mändide rinnasdiameetri ja kõrguse vahelist seost sõltuvalt proovitüki keskmisest kõrgusest 2. Vaatame näiteks mände, mis kasvavad proovitükkidel inventeeritud kõrgusega $H=200$ (oranžid punktid). Selliste puude diameetrid võivad muutuda üsna suures vahemikus (100dm...500dm), kuid sellele vaatamata on peaaegu kõigi nende puude kõrgus ligilähedasel 200dm. Sõltuvus puu diameetrist on võrdlemisi väike ja lihtsalt modelleeritav sirge abil.



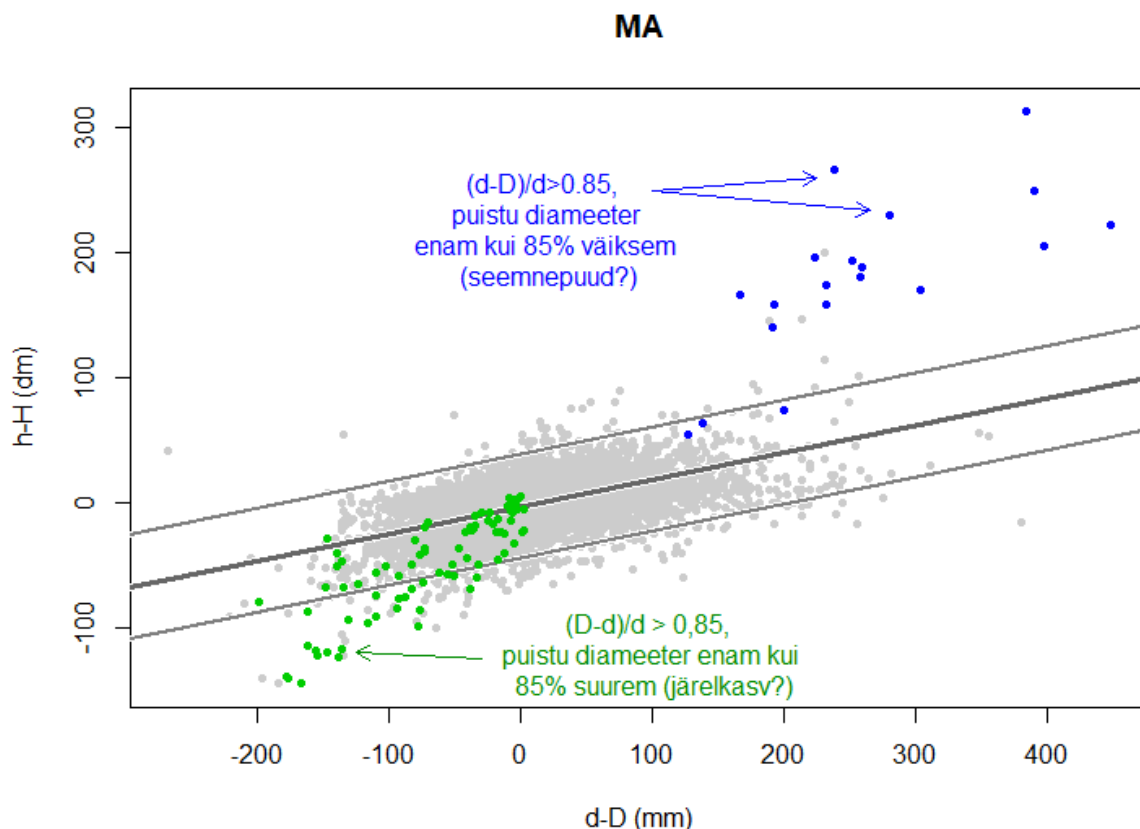
Joonis 1. Mäni diameetri d ja kõrguse h vaheline seos inventeeritud kõrguse H fikseeritud tasemetel korral.

Kui sama metsa kõrguse (inventeeritud diameetri H) korral on seos üksikpuu enda diameetri ja kõrguse vahel lineaarne, siis

² Lisas 2 on toodud nende kolme tunnuse - diameetri, kõrguse ja inventeeritud kõrguse - vahelise seose graafikud üksikute puuliikide (MA, KU, KS) kaupa.

$h = c(H) + c \cdot d$. Kui sama mudel kirjeldab ka puistu keskmist puud, siis peaks ka kehtima $H = c(H) + c \cdot D$ ja järelikult $c(H) = H - c \cdot D$ ning saame, kõrguse prognoosimiseks lihtsa mudeli: $h = H + c \cdot (d - D)$. Saadud mudel sisuliselt väidab, et kõrvuti kasvavatel puudel toimub kõrguste konvergens – üksikpuu peab pingutama, et mitte naaberpuude varju jääda, kuid naabritest märkimisväärselt kõrgemaks kasvamine olulist eelist ei anna.

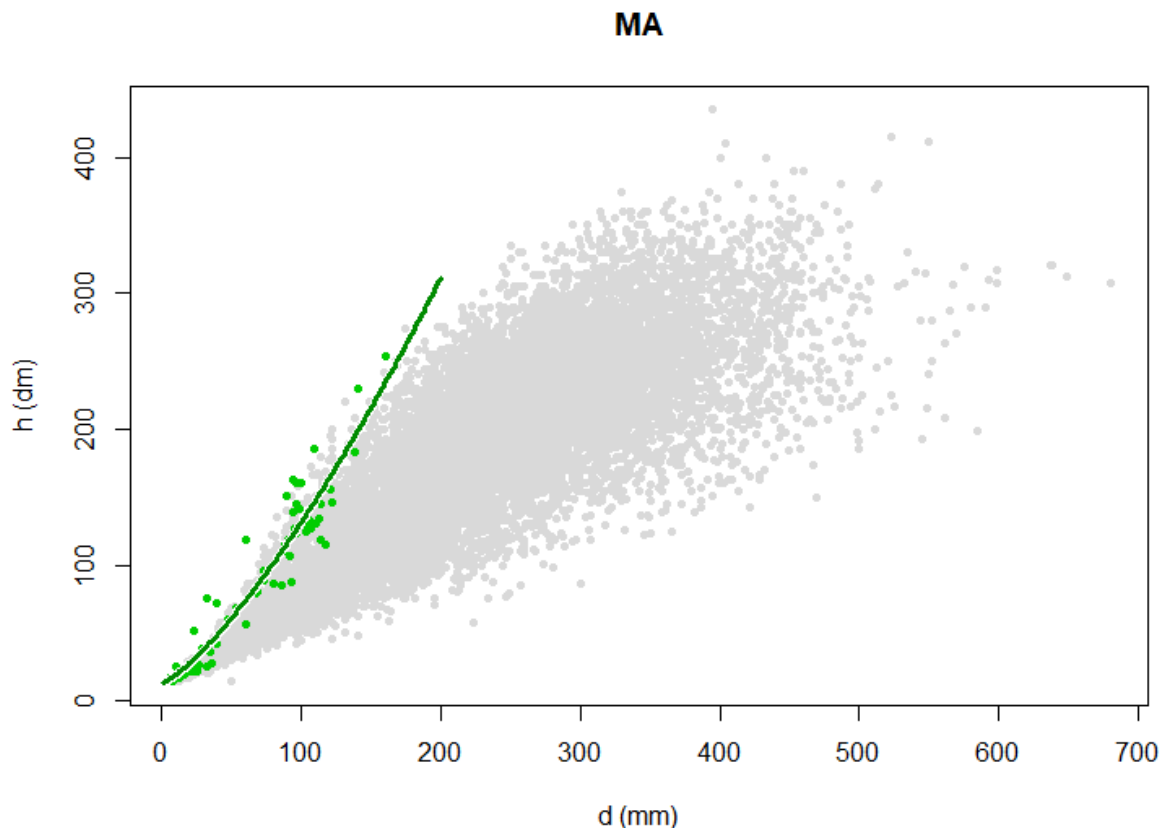
Täiesti ilmselt ei suuda taoline ülilihtne mudel kirjeldada näiteks seemnepuude kõrgust või ka teise rinde/järelkasvu käitumist. Seetõttu uurime järgmisena, kuidas oleks võimalik tuvastada seemnepuid või järelkasvu hulka kuuluvaid puid (noorpuid). Selliste puude tuvastamisel võiksime nende kõrgust prognoosida teistsuguse, puistu hetkekõrgusest vähem sõltuva mudeli abil. Selleks vaatame lihtsa mudeli prognoosivigu, $h-H$:



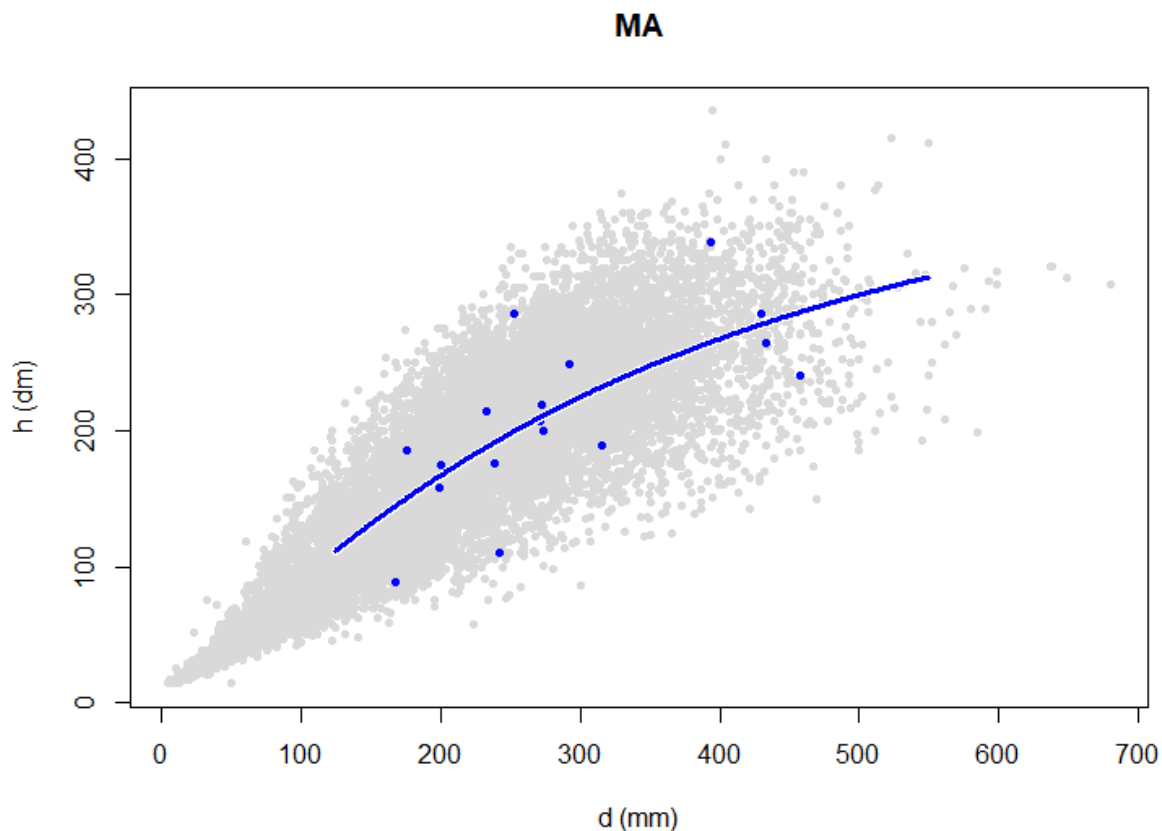
Kui puistu inventeeritud diameeter D on märkimisväärselt väiksem puu diameetrist (15% puu diameetrist või vähem), siis puistu kõrgus enam ei prognoosi puu kõrgust (arvatavasti on tegemist seisma jäetud seemnepuuga, sinised punktid ülaltoodud joonisel). Ka siis, kui puistu inventeeritud kõrgus on märkimisväärselt suurem puu diameetrist (85% suurem või enam – rohelised punktid ülaltoodud joonisel), ei ole puistu kõrgus enam adekvaatne mõõdik, mille abil puu kõrgust prognoosida (arvatavasti on tegemist järelkasvuga või teise rinde puudega).

Juhul, kui kahtlustame, et puu mille kõrgust prognoosime ei pruugi olla tüüpiline puu (on kas teise rinde puu või seemnepuu), siis prognoosime tema kõrgust lihtsa, paari hinnatavat parameetrit sisaldava kõrguskõvera mudeli abil, näiteks Hossfeld'i tüüp IV mudeli abil ($c=1,3$). Kasutatav mudel peab olema üsna lihtne ja sisaldama maksimaalselt vaid paari hinnatavat parameetrit, sest mingit liiki seemnepuid on mõõdetud üsna vähe ja peab olema võimalik ka nende väheste seemnepuude andmeid kasutades mudeli parameetreid hinnata.

Järgneval joonisel on kujutatud järelkasvuks klassifitseeritud männid (inventeeritud diameeter 85% või enam suurem kui üksikpuu diameeter) ning nende jaoks sobitatud Hossfeld'i kõrguskõver. Väärrib märkimist, et vanemas metsas kasvavad pisemad puud panustavad kõrguse kasvu rohkem kui diameetri kasvu, st sageli on tegemist peenikeste kuid oma diameetri kohta üllatavalt kõrgete puudega.



Sama loogika abil võime prognoosida ka seemnepuuks klassifitseeritud puude kõrguseid. Seemnepuude kõrgused on aga diameetri järgi märksa halvemini prognoositavad - sama diameetriga võime kohata nii kõrgeid kui ka madalaid puid. Nähtavasti mängib siin rolli ka aeg, kuna ümbritsev mets ära raiuti – peale ümbritseva metsa raiumist võis ka seemnepuude edasine kasvukiirus aeglustuda. Kuna aga raie aja määramisega esineb raskuseid, siis ilmselt peame nende puude puhul leppima ebatäpsemalt prognoositud kõrgustega :



Seemnepuu kõrguste prognoosi täpsust on teoreetiliselt võimalik tõsta, kasutades näiteks kasvukoha boniteedi, informatsiooni raie toimumise aja ja antud alal enne raiet kasvanud metsa kohta. Samas on praktikas selle lisainformatsiooni kasutamine problemaatiline – hiljuti raiutud metsade boniteedi/H100 määrangud ei ole alati kõige usaldusväärsemad; varasem informatsioon sama ala kohta puudub ajutistel proovitükkidel; osade liikide puhul esineb seemnepuid mudelpuude seas sedavõrd harva, et isegi ühe täiendava mudeli parameetri hindamine andmete põhjal on problemaatiline.

$$P_s = \frac{1 - \exp\left(-10\left(\frac{d-D}{d} \cdot I(d > D)\right)\right)}{1 + \exp\left(-10\left(\frac{d-D}{d} \cdot I(d > D) - c_s\right)\right)} \quad (6)$$

Mitme mudeli koos kasutamisel kerkib oluline küsimus: kuhu kohta täpselt tõmmata piirid? Millal kasutada puu kõrguse määramiseks ümbritseva metsa kõrgust (inventeeritud kõrgust) ja millal seemnepuude või järelkasvu puude jaoks hinnatud kõrguskõverat? Selle üle otsustamiseks saab kasutada andmete põhjal hinnatud parameetreid c_s ja c_{NP} – need määravad, milliste inventeeritud diameetri ja puu enda diameetri erinevuste korral muutub

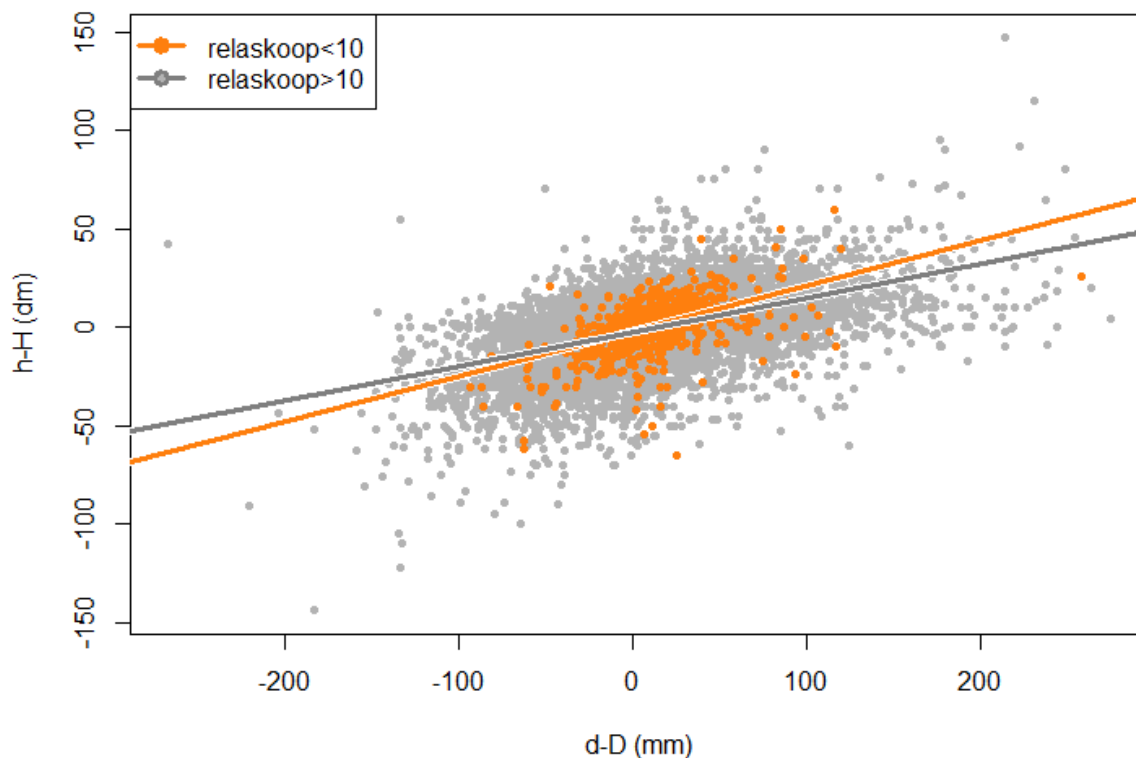
seemnepuude/järelkasvu kõrguskõvera abil saadud prognoos täpsemaks võrreldes ümbritseva metsa (inventeeritud kõrguse) abil saadud prognoosist. Seemnepuuks olemise tõenäosust P_s hindame logistilise kõvera abil, $f(D/d) = 1/(1 + \exp(-10 \cdot (1 - D/d - c_s)))$, kus suhe D/d iseloomustab kui palju väiksem on inventeeritud diameeter uuritava puu diameetrist (seemnepuude korral $\frac{D}{d} < 1$). Andmete põhjal hinnatav kordaja c_s iseloomustab D/d suhet, mille korral seemnepuuks olemise tõenäosus on 50%. Toodud logistiline kõver aga ei taga, et suhte $\frac{D}{d} = 1$ korral seemnepuuks olemise tõenäosus oleks 0. Sestap teeme täiendava teisenduse, peale mida seemnepuuks olemise tõenäosus hinnatakse matemaatiliselt täpselt 0-ks siis, kui puu diameeter on võrdne või väiksem kui inventeeritud diameeter (sellisel juhul tahame puu kõrguse prognoosimiseks kasutada inventeeritud kõrgust – või vajadusel järelkasvu/noorpuu mudelit):

$$P_s\left(\frac{D}{d}\right) = \frac{f\left(\frac{D}{d}\right) - f(0)}{1 - f(0)} \cdot I(D < d).$$

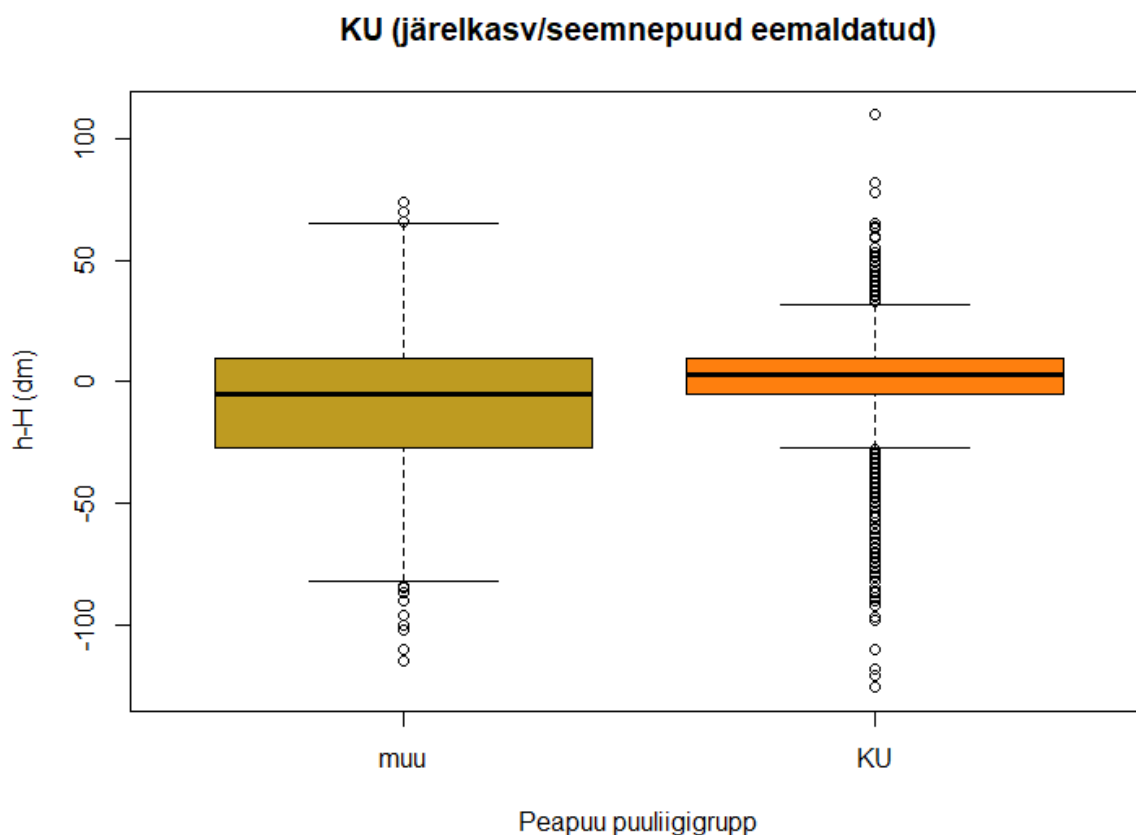
Mudeli peenhäälestus

Kuigi ümbritseva puistu kõrgus määrab suuresti ka üksikpuu kõrguse, esineb siiski väike seos puu enda diameetri ja puu kõrguse vahel – sama kõrgusega puistus on jämedamad puud ikkagi keskmiselt veidi kõrgemad. Milline on aga puu enda diameetri mõju, see muutub mõnevõrra – tihedas metsas on kõigi puude kõrgused sarnasemad, nii jämedad kui ka peenikesed puud kipuvad olema üsnagi sama kõrgusega. Hõredamas metsas aga on puu enda diameetril ka mõju, jämedamad puud on ikkagi veidi kõrgemad võrreldes peenikeste puudega. Sestap kasutame puu kõrguse modelleerimisel $(d-D)$ kordajat, mis sõltub metsa tihedusest (relaskoobi mõõtmistulemusest).

MA, seemnepuud ja järelkasv eemaldatud



Kuused, mis kasvavad männikutes või kaasikutes, on keskmiselt veidi väiksemad kui metsa inventeeritud kõrguse järgi võiks arvata. Seos jääb paika isegi siis, kui puude diameetrite põhjal selgelt järelkasvu kuuluvad kuused (mida esineb ka palju) eelnevalt andmestikust eemaldada. Osalt võib põhjuseks olla ka see, et teistest puudest ehk aasta või paar hiljem kasvama hakanud kuuskede puhul ei pruugi olla lihtne otsustada, kas neid peaks prognoosima järelkasvu mudeli järgi või peaks nende kõrguse prognoosimiseks siiski kasutama puistu kõrgust. Vaata ka alljärgnevat graafikut:



Lahenduseks on mudelisse sisse kirjutatud parandusliige c_2 , mis võimaldab korrigeerida puude ennustatavat kõrgust siis, kui enamuspuuliik on teistsugune (st. lubab näiteks kaasikutes kasvavatel kuuskedel olla keskmiselt veidi neid ümbritsevatest kaskedest madalamad). Mõne puuliigigrupi korral hinnati vastav parandusliige ühest suuremaks – st teist liiki puude seas kasvades olid antud liiki puud pigem ümbritsevast metsast veidi suuremad.

Metsa kõrgust hinnates (inv_korgus) võivad suuremad ja kõrgemad puud omada hindaja silmis suuremat kaalu. Üksikult mõõdetud puude kõrguste keskmine (iseги peale järelkasvu ja II rinde puude eemaldamist) on veidikene väiksem inventeeritud kõrguste keskmisest (kuigi erinevus on väga väike). Siiski on võimalik prognooside täpsust veidi suurendada korrutades inventeeritud kõrgused läbi andmete põhjal hinnatud kordajaga c_3 , mis enamike puuliikide puhul hinnatakse pisut väiksemaks kui 1 (kasel 0,99; männil 0,995 jne)

Mudeli puudused

Antud mudel kasutab üksikpuu kõrguse prognoosimisel ka proovitüki osa üldandmeid (inventeeritud kõrgus, inventeeritud diameeter, peapuuliik). Võrreldes ühe ja sama puu kõrguse muutust erinevatel ajahetkedel, võivad muutused taustaandmetes tuua endaga kaasa ootamatuid



muutuseid prognoosides. Näiteks kui puu enda diameeter kasvab, aga ka ümbritsevate puude diameetrid kasvavad, siis võib puu kõrguse prognoos kahaneda. Sest viis aastat tagasi määras antud puu kõrguse suuresti seemnepuude kõrguskõver, aga tänu ümbritseva metsa kiirele kasvule prognoositakse nüüd puu kõrgust juba metsa inventeeritud kõrguse kaudu – mis siiski võib olla madalam, kui puu kõrgus viis aastat tagasi. Seetõttu kõrguse (ja mahu) muutuste prognoosimiseks/jälgimiseks soovitame spetsiaalselt selleks mõeldud mudelit (9).

Probleemid, mis tekivad siis, kui uus kõrguse mudel on kalibreeritud liiga pika ajavahemiku (nt 20a) pealt, leiavad käsitlemist Lisas 3. Samuti on seal välja selgitatud optimaalne ajaaken mudeli hindamiseks.

Lisa 2. Diameetri, kõrguse ja inventeeritud kõrguse seoste graafikud

Järgnevaid graafikuid võib vaadelda kui lähteinfot uue kõrguse mudeli väljatöötamiseks.

Mõnevõrra lihtsustatult väljendades näitab alljärgnev analüüs seda, et kui proovitüki keskmine kõrgus (inv_korgus) on fikseeritud, siis puude (eriti kehtib see mändide kohta) kõrgus diameetrist väga palju ei sõltu - kõik puud ühel ja samal proovitükil kipuvad olema laias laastus ühekõrgused. Täpsemalt, see käib keskmiste ja suuremate diameetrite kohta - väikeste diameetrite korral on olukord teistsugune.

Uurime lähemalt, kuidas sõltub mudelpuu kõrgus diameetrist, kui vastava proovitüki inventeeritud kõrgus on fikseeritud (täpsemalt, see asub mingis kitsas vahemikus).

Lisame mudelpuudele proovitüki andmetest inventeeritud kõrguse ja peapuuliigi

```
andmed_min=mudelpuud|>select(amudelpuud_id,prto_aid,h_puu,puuliik,d_13)|>inner_join(prtosa|>select(prto_aid,peapuuliik,inv_korgus),by="prto_aid")|>mutate(inv_vahemik=cut(inv_korgus,breaks=20))
```

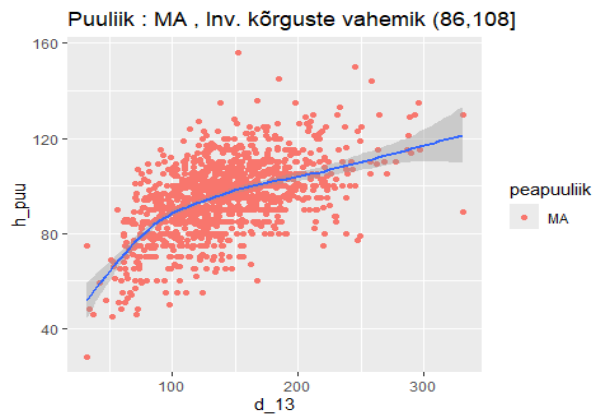
Defineerime abifunktsiooni joonise tegemiseks

```
joonista=function(pliik,vahemiku_nr,on_peapuu=TRUE){  
  vahemiku_nimi=levels(andmed_min$inv_vahemik)[vahemiku_nr]  
  andmed_pliik=andmed_min|>filter(puuliik==pliik)  
  if(on_peapuu){  
    andmed_pliik=filter(andmed_pliik,peapuuliik==pliik)  
  } else {  
    andmed_pliik=filter(andmed_pliik,peapuuliik!=pliik)  
  }  
  ggplot(andmed_pliik|>filter(as.numeric(inv_vahemik)==vahemiku_nr),  
    aes(x=d_13,y=h_puu))+geom_point(aes(color=peapuuliik))+geom_smooth()+ggtitle(paste('Puuliik:',pliik,', Inv. kõrguste vahemik',vahemiku_nimi))  
}
```

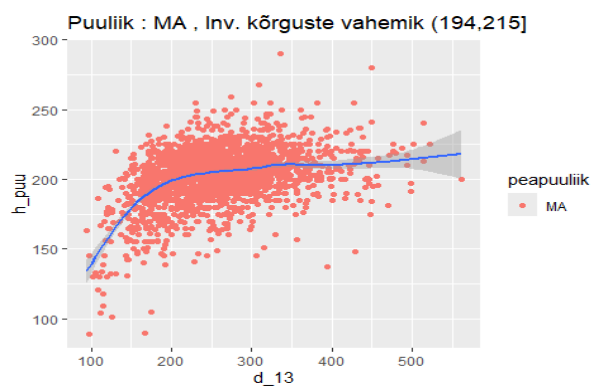
Mänd

Juht, kus peapuuliik on ka mänd

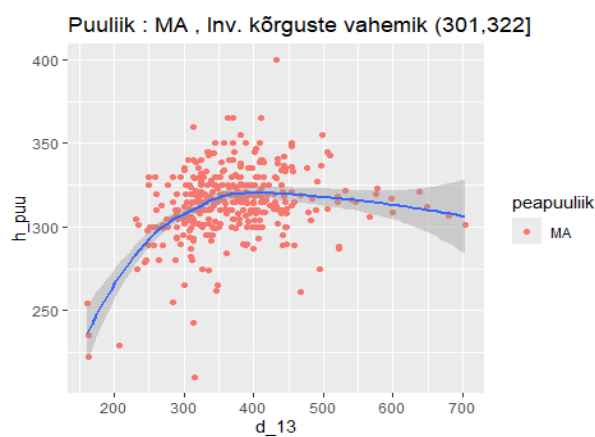
```
joonista('MA',5)
```



```
joonista('MA',10)
```

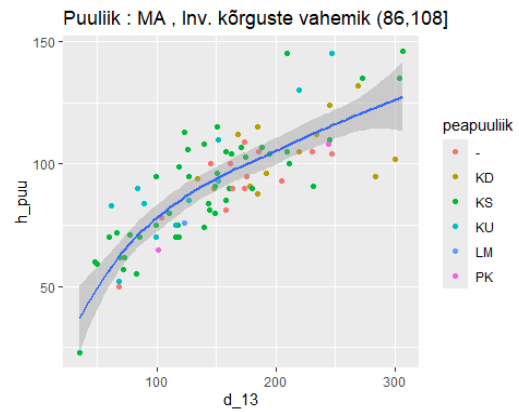


```
joonista('MA',15)
```

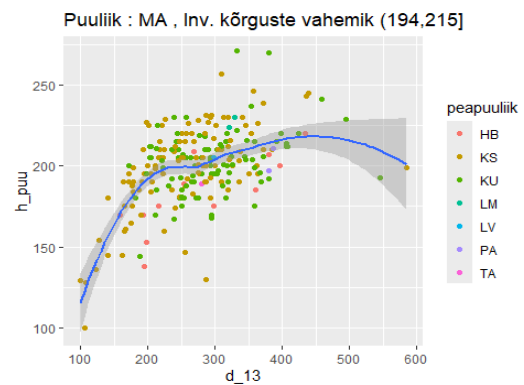


Peapuuliik ei ole mänd

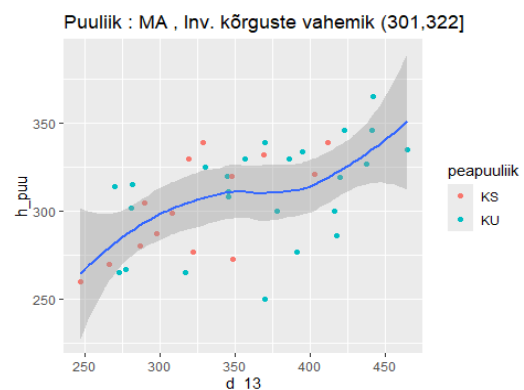
```
joonista('MA',5,FALSE)
```



```
joonista('MA',10,FALSE)
```



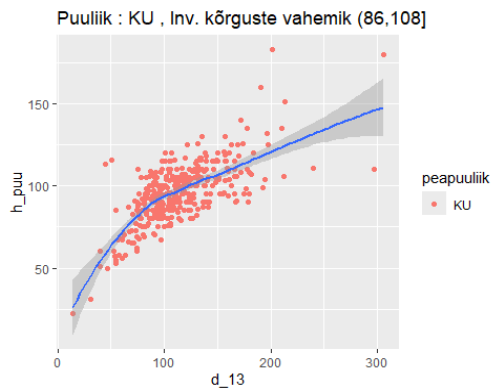
```
joonista('MA',15,FALSE)
```



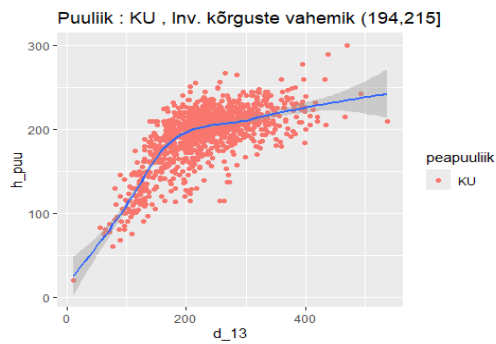
Kuusk

Peapuuliik on kuusk

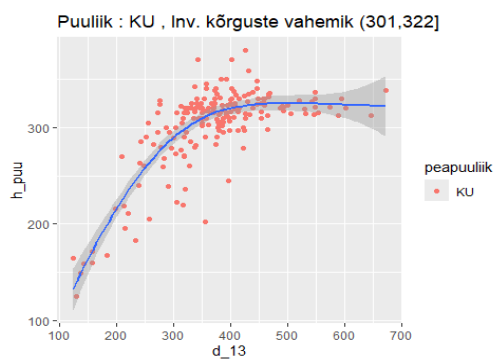
```
joonista('KU',5)
```



```
joonista('KU',10)
```

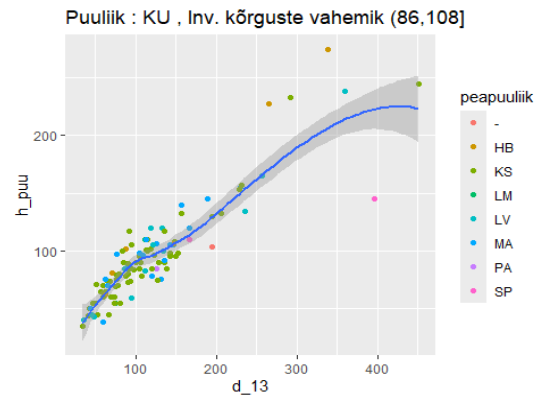


```
joonista('KU',15)
```

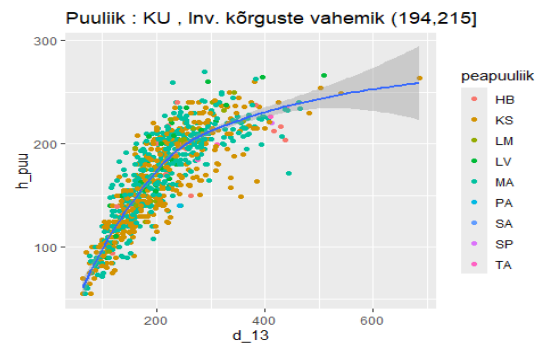


Peapuuliik ei ole kuusk

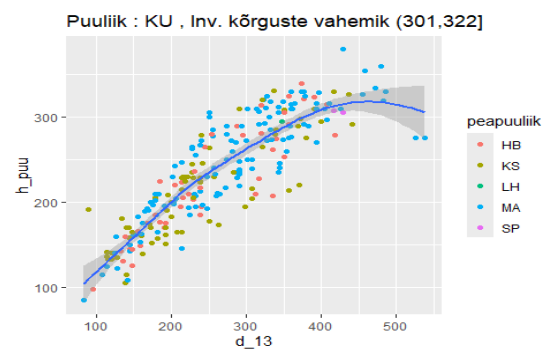
```
joonista('KU',5,FALSE)
```



```
joonista('KU',10, FALSE)
```



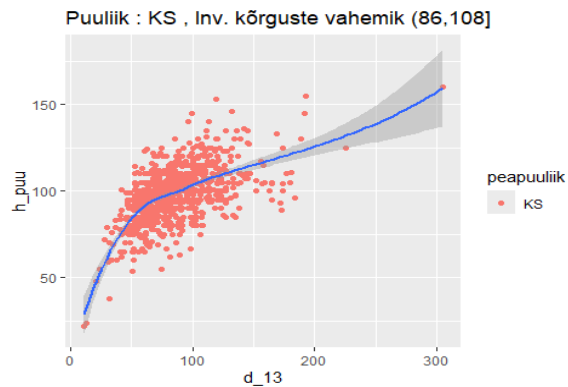
```
joonista('KU',15, FALSE)
```



Kask

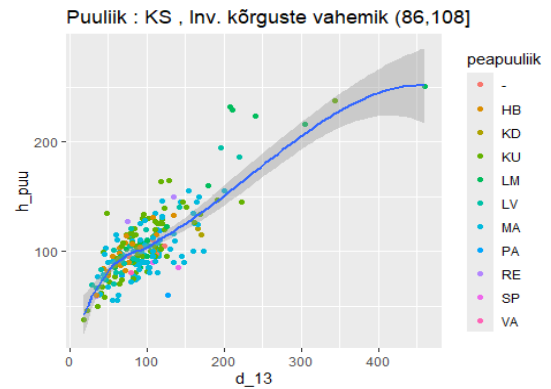
Peapuuliik on kask

```
joonista('KS',5)
```

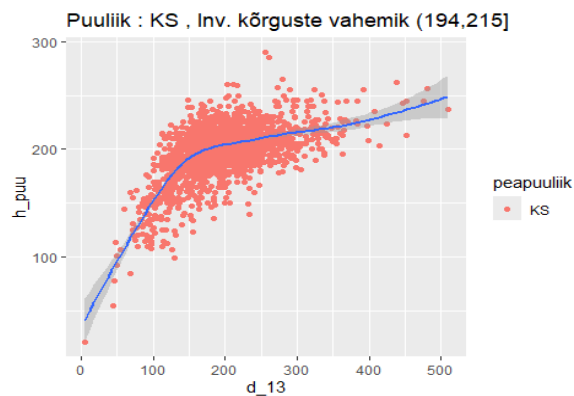


Peapuuliik ei ole kask

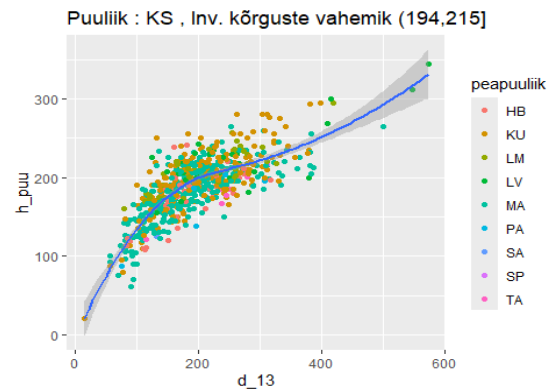
```
joonista('KS',5,FALSE)
```



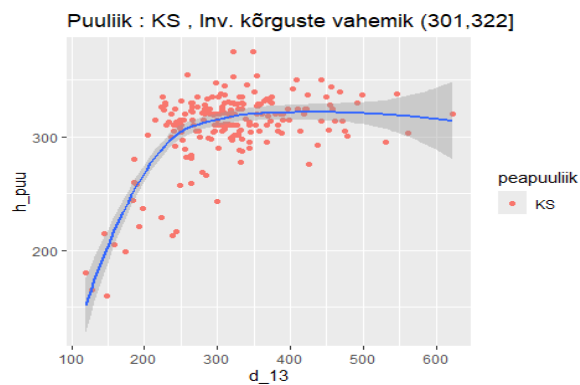
```
joonista('KS',10)
```



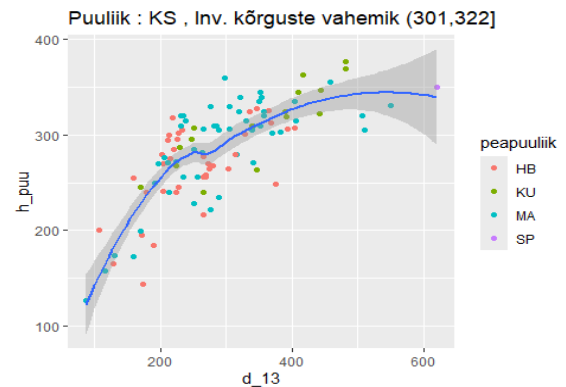
```
joonista('KS',10,FALSE)
```



```
joonista('KS',15)
```



```
joonista('KS',15, FALSE)
```





Kokkuvõte: Joonistelt on näha, et seos puu rinnasdiameetri ja kõrguse vahel sõltub suuresti proovitüki inventeeritud kõrgusest. Kui $inv_kõrgus$ on suur, siis seos diameetri ja kõrguse vahel on nõrk (kõrgus sõltu diameetrist vähe); kui $inv_kõrgus$ on väike, siis diameetri ja kõrguse vahel on märkimisväärne seos. See on põhjuseks, miks uus kõrguse mudel koosneb mitmest erinevast komponendist.

Lisa 3. Optimaalne ajaaken uue kõrgusmudeli hindamiseks

Nagu punktides 2.4 ja 2.5 läbi viidud analüüs näitas, on uus kõrgusmudel (5) suhteliselt hea täpsusega (rmse mõttes). Mudeli töökindluse täiendavaks uurimiseks vaatame, kas prognoosivead sõltuvad mõnest olulisest muutujast, näiteks puuliik ja aasta. Alljärgnevast tabelist nähtub, et kõrgusmudel (5), mis on kalibreeritud võimalikult pika andmestiku (2005-2024) pealt, töötab teatud ajaperioodidel märgatava süstemaatilise veaga. Tabeli viimases veerus on toodud t-statistiku väärtused (=kahe eelneva veeru suhe), kusjuures roheline värv vastab olukorrale, kus prognoos ületab tunduvalt tegelikku kõrgust.

Tabel 1. Mudeliga (5) leitud kõrgusprognooside vead mudelpuudel, mudel hinnatud andmetel 2005-2024.

Rohelisega märgitud read vastavad üleprognoosile, punased alaprognoosile.

plgrp	aasta	kesk_h	kesk_hprognoos_viga	sd_kesk_hprognoos_viga	t statistic
HB	2010	254,68	3,69	4,33	0,852
HB	2011	252,14	-2,90	3,10	-0,936
HB	2012	270,87	7,35	3,35	2,196
HB	2013	249,14	4,71	3,06	1,538
HB	2014	277,79	4,90	3,09	1,588
HB	2015	265,44	3,43	5,62	0,610
HB	2016	254,81	2,45	2,62	0,934
HB	2017	264,64	0,46	3,80	0,120
HB	2018	247,43	1,62	3,24	0,499
HB	2019	250,27	-1,33	2,43	-0,548
HB	2020	262,67	-0,32	4,07	-0,078
HB	2021	236,37	-4,15	2,40	-1,732
HB	2022	253,85	1,81	2,28	0,791
HB	2023	228,78	-2,28	2,17	-1,049
KS	2010	201,49	2,77	0,89	3,119
KS	2011	196,69	1,31	1,06	1,245
KS	2012	197,08	0,73	1,15	0,637
KS	2013	200,03	2,48	1,38	1,801
KS	2014	209,62	2,14	1,28	1,681
KS	2015	209,94	2,56	1,08	2,367
KS	2016	205,43	2,13	1,21	1,770
KS	2017	187,35	-4,73	1,17	-4,055
KS	2018	189,43	-4,17	1,09	-3,821
KS	2019	198,05	-3,45	0,90	-3,842
KS	2020	196,70	-3,22	0,92	-3,501
KS	2021	191,89	-4,73	0,94	-5,046
KS	2022	188,11	-3,82	0,87	-4,402
KS	2023	187,83	-5,83	0,92	-6,358

KU	2010	195,83	-0,49	1,44	-0,339
KU	2011	199,03	3,72	1,41	2,641
KU	2012	196,16	3,30	1,46	2,252
KU	2013	205,54	3,76	1,47	2,559
KU	2014	198,25	4,98	1,40	3,558
KU	2015	204,58	-0,09	1,88	-0,050
KU	2016	206,24	5,59	1,43	3,901
KU	2017	190,95	-4,08	1,47	-2,783
KU	2018	195,31	-3,58	1,35	-2,649
KU	2019	193,61	-3,99	1,06	-3,767
KU	2020	193,39	-2,73	1,35	-2,025
KU	2021	197,99	1,02	1,26	0,808
KU	2022	189,75	-1,55	1,20	-1,294
KU	2023	190,58	-1,26	1,22	-1,029
LM	2010	211,19	6,18	2,51	2,457
LM	2011	195,05	0,85	2,19	0,387
LM	2012	204,00	-5,73	3,28	-1,744
LM	2013	211,78	4,37	1,89	2,311
LM	2014	203,71	2,31	2,52	0,915
LM	2015	228,30	6,44	2,49	2,590
LM	2016	199,67	2,03	2,77	0,732
LM	2017	204,38	-9,29	3,54	-2,625
LM	2018	204,38	-2,95	1,58	-1,860
LM	2019	201,82	-0,99	2,35	-0,420
LM	2020	211,55	1,81	1,88	0,961
LM	2021	199,62	-1,99	2,04	-0,977
LM	2022	193,41	-7,33	2,42	-3,031
LM	2023	193,68	-2,76	1,73	-1,595
LV	2010	169,65	1,33	2,33	0,570
LV	2011	182,10	1,28	1,25	1,027
LV	2012	181,16	0,07	2,35	0,031
LV	2013	181,31	0,56	1,76	0,319
LV	2014	181,74	4,69	2,46	1,907
LV	2015	166,81	-1,61	3,02	-0,535
LV	2016	185,53	2,66	1,39	1,914
LV	2017	166,65	-6,24	2,49	-2,512
LV	2018	167,15	-6,84	1,84	-3,720
LV	2019	167,96	-3,26	1,73	-1,888
LV	2020	161,05	-2,69	1,87	-1,436
LV	2021	171,20	0,17	1,34	0,128
LV	2022	150,15	-3,15	1,40	-2,244
LV	2023	163,13	-3,36	1,52	-2,209
MA	2010	198,37	4,58	0,87	5,276
MA	2011	203,46	1,76	0,99	1,783
MA	2012	207,86	-0,29	0,79	-0,363

MA	2013	183,44	1,62	0,80	2,020
MA	2014	204,18	1,79	0,84	2,139
MA	2015	202,30	2,53	1,05	2,404
MA	2016	198,56	3,31	0,94	3,533
MA	2017	200,73	-5,59	0,83	-6,731
MA	2018	177,75	-3,41	0,80	-4,281
MA	2019	201,55	-3,65	0,74	-4,944
MA	2020	192,98	-5,69	0,83	-6,888
MA	2021	190,54	-4,54	0,89	-5,073
MA	2022	196,66	-5,49	0,67	-8,151
MA	2023	177,37	-3,35	0,69	-4,835
RE	2010	157,57	4,52	7,04	0,643
RE	2011	168,00	-19,94	31,57	-0,631
RE	2012	170,33	4,04	11,40	0,354
RE	2013	147,00	4,45	3,88	1,149
RE	2014	169,71	11,80	12,15	0,971
RE	2015	163,50	-16,09	12,84	-1,252
RE	2016	154,20	-1,96	20,84	-0,094
RE	2017	169,33	-14,17	17,65	-0,803
RE	2018	160,33	2,15	5,73	0,376
RE	2019	146,64	2,12	6,42	0,330
RE	2020	162,08	7,19	13,65	0,527
RE	2021	172,17	-7,23	14,03	-0,516
RE	2022	159,67	0,60	6,11	0,098
RE	2023	152,50	-6,02	3,78	-1,593
XX	2010	203,50	9,67	5,45	1,774
XX	2011	200,65	10,68	3,35	3,183
XX	2012	189,75	1,46	7,33	0,199
XX	2013	172,71	0,78	4,56	0,171
XX	2014	204,83	19,38	7,10	2,730
XX	2015	205,67	9,66	7,37	1,311
XX	2016	204,24	12,83	4,26	3,012
XX	2017	184,35	-1,40	7,25	-0,194
XX	2018	175,14	-6,09	4,94	-1,234
XX	2019	183,20	7,51	4,41	1,703
XX	2020	177,14	-4,21	4,54	-0,925
XX	2021	183,00	-0,93	4,48	-0,208
XX	2022	172,18	-11,20	5,80	-1,932
XX	2023	176,66	-7,78	3,77	-2,067
keskmise					
nihe			-0,18		

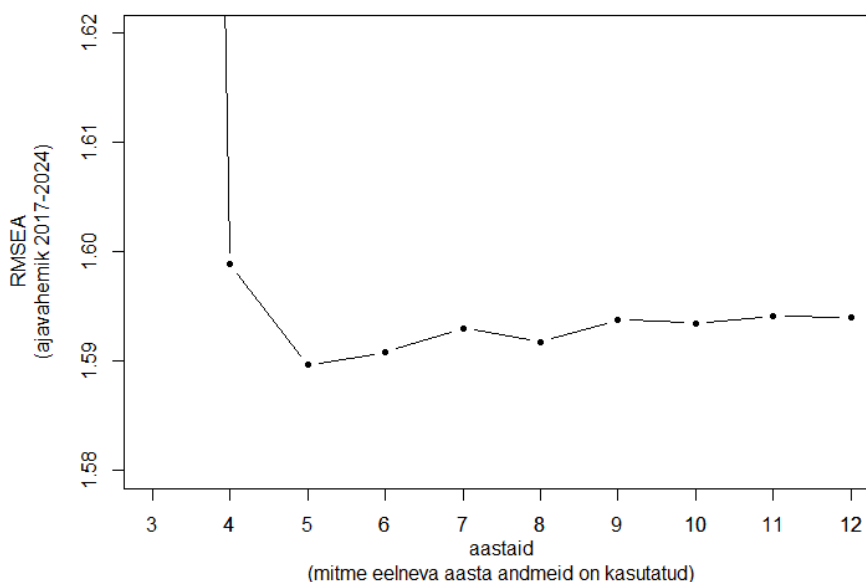
Vaadeldes ülaltoodud tabelit (eriti selle värvitud tulpa) on näha, et kõikide suuremate puuliikide (KU, KS, MA) korral ilmneb kõrguse märgatav ülehindamine aastatel 2017-2022. Ilmselt on

ülehindamise põhjuseks nimetatud ajaperioodil puude kõrguse juurdekasvu järsk langus. Et kõrguse mudel seda langust arvestaks, tuleks mudel (5) kalibreerida lühema ajaperioodi pealt, kuid sel juhul peab arvestama sellega, et väiksema arvukusega puuliikide jaoks mudeli parameetrid muutuvad ebatäpsemaks. Järgnevas püüame välja selgitada, kui mitme eelneva aasta andmeid tuleks kasutada, et nende pealt kõrgusmudeli (5) parameetreid hinnates mudeli täpsus oleks suurim.

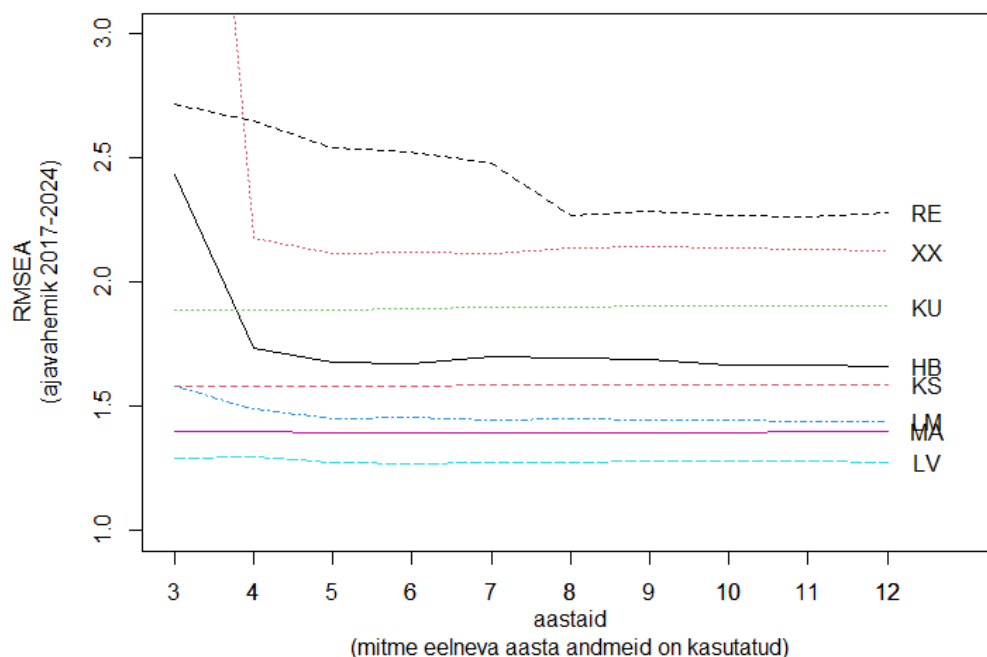
Optimaalse andmeakna pikkuse leidmine uue kõrgusmudeli jaoks

Uurime, kui täpselt suudame puu kõrgust prognoosida siis, kui kasutame 3, 4, 5, ... eelneva aasta andmete pealt hinnatud kõrguskõverat. Prognoositäpsuse hindamiseks kasutame aastaid 2017...2024. Näiteks, kui soovime näha, millise täpsusega suudame puu kõrgust prognoosida, kasutades näiteks 10 eelneva aasta andmeid, siis prognoosime 2017.a mõõdetud puude kõrguseid kasutades aastatel 2006..2016 mõõdetud puude pealt hinnatud mudelit; 2018.a mõõdetud puid prognoosime 2007...2017 andmete põhjal hinnatud mudelit kasutades jne. Lõpuks leiame keskmise prognoosivea (RMSEA) aastatel 2017..2024 mõõdetud puude kõrguse prognoosimisel.

Esmane vaade näib soovivat **viimase 5 aasta** mõõtmiste kasutamist mudeli hindamisel:



Kui aga vaatame kõrguse prognoosi täpsust puuliikide kaupa, näeme mõnevõrra keerukamat pilti.



Näeme, et harvemini mudelpuude sekka sattuvate puude korral, näiteks remmelgate korral, ei kogune 5 aasta jooksul piisavalt mudelpuid, et nende põhjal usaldusväärselt kõrguskõvera mudelit hinnata. Tasuks vaadata ka alljärgnevat tabelit ja pöörata tähelepanu mudelpuude seas vähem esindatud puuliigigruppidel (HB, RE, XX):

Tabel 2. Prognoosivead (RMSEA) puuliigigrupi kaupa. Kõrgust prognoosiva mudeli hindamiseks on kasutatud kas 3, 5 või 10 eelneva aasta andmeid.

aastaid	HB	KS	KU	LM	LV	MA	RE	XX
3	2.435	1.582	1.886	1.580	1.292	1.401	2.719	5.846
5	1.678	1.580	1.891	1.454	1.277	1.395	2.541	2.116
10	1.668	1.590	1.905	1.445	1.278	1.397	2.270	2.137

Kui juhuse tahtel mõne puuliigigrupi esindajate arv peaks mudelpuude seas veelgi langema, siis võib kõrguskõverate prognoositäpsus nende puuliigigruppide jaoks muutuda ebarahuldavaks. Vaatame näiteks 3 aasta andmete pealt hinnatud mudelite prognoositäpsuseid. See rida võiks kirjeldada seda, mis juhtub prognoositäpsusega siis, kui mudelpuude seas antud puuliigi osatähtsus veidi kahaneb. Näeme, et niigi harvem esinevate puuliigigruppide puhul oleks mudelpuude väikegi kadu väga suure mõjuga mudeli prognoositäpsusele (vt näiteks, mis juhtub haabade või remmelgate prognoositäpsusega). Seetõttu võiks pakkuda välja kaks erinevat strateegiat.



Strateegia 1: kasutada enamesinevate puuliikide korral (nagu MA, KS, KU) kõrguskõvera hindamiseks viimase 5 aasta andmeid ja harvemini esinevate puuliikide korral 10 aasta andmeid.

Strateegia 2: kasutada kõigi puuliikide korral viimase 10 aasta andmeid kõrguskõvera hindamiseks. Sagedasemate puuliikide korral kaotame selle valiku korral õige pisut prognoositäpsuses (kuna mudel kohaneb liiga aeglaselt looduses aset leidvate muutustega), aga väldime tekkida võivaid probleeme harvemini vaatluse alla sattuvate puuliikide puhul, mis võivad liiga väheste mudelpuude vaatluste korral viia dramaatiliste vigadeni nende puuliigigruppide korra.

Lisa 4. Uue kõrgusmodeli täiendusvõimaluse uurimine

Siinkohal uurime üht hüpoteetilist võimalust uue kõrgusmodeli täpsuse edasiseks suurendamiseks, mis seisneb prognoosivigade vaheliste korrelatsioonide ärakasutamises.

Loomulik küsimus on: kas vaadeldava mudelpuuga samal proovitükil mõõdetud teiste mudelpuude kõrgused annavad täiendavat infot, mida proovitüki üldandmete põhjal ei saa või on keeruline kirjeldada? Kui see oleks nii, siis peaksid samal proovitükil olevate mudelpuude kõrguste prognoosivead olema omavahel korreleeritud. Seetõttu uurimegi kõigepealt, kui tugevad korrelatsioonid prognoosivigadel on. Siin on oluline märkida, et testandmestikul mõõdetud mudelpuid me kõrguse prognoosimudeli väljatöötamisel ei kasutanud. Sõltuvuse hindamiseks kasutame esmalt ICC (intra class correlation) mõõdikut.

```
model_vead=lme4::lmer(viga~(1|prto_aid),data=progn_test|>#inner_join(head_test_andmed|>select(amudelpuud_id))|>
mutate(viga=(prognoos_uus -h_puu))|>
inner_join(mudelpuud|>select(amudelpuud_id,prto_aid,aasta)|>filter(
aasta>2004))

)

## Joining with `by = join_by(amudelpuud_id)`
## Joining with `by = join_by(amudelpuud_id)`
## Joining with `by = join_by(amudelpuud_id)`
## Joining with `by = join_by(amudelpuud_id)`

var_components <- as.data.frame(lme4::VarCorr(model_vead))
icc <- var_components$vcov[1] / sum(var_components$vcov)
icc

## [1] 0.08937364
```

Alternatiivne võimalus korrelatsioonide hindamiseks on järgmine: valime igast proovitükiosast, kus on vähemalt kaks mudelpuud, kaks juhuslikku ja arvutame nende vahelise korrelatsiooni.

```
#Valime kaks vaatlust igast proovitükiosast
katse=progn_test|>
mutate(viga=prognoos_uus-h_puu)|>
select(amudelpuud_id,viga)|>
inner_join(mudelpuud|>select(amudelpuud_id,aasta,puuliik,prto_aid))|>
filter(aasta>2004)|>
group_by(prto_aid)|>
mutate(mitu=n())|>
filter(mitu>1)|>
slice_sample(n=2)|>
mutate(rea_nr=row_number())|>ungroup()

## Joining with `by = join_by(amudelpuud_id)`

#paneme valitud kaks vaatlust erinevatesse veergudesse, et oleks parem korrelatsioone leida
katse_lai=katse|>filter(rea_nr==1)|>inner_join(katse|>filter(rea_nr==2),by="prto_aid")
cor(katse_lai$viga.x,katse_lai$viga.y)

## [1] 0.08431133
```

Kommentaari: tulemus on väga lähedane eelnevale hinnangule 0.089.

```
cor.test(katse_lai$viga.x, katse_lai$viga.y)

##
## Pearson's product-moment correlation
##
## data: katse_lai$viga.x and katse_lai$viga.y
## t = 6.7308, df = 6328, p-value = 1.835e-11
## alternative hypothesis: true correlation is not equal to 0
## 95 percent confidence interval:
## 0.05980004 0.10872101
## sample estimates:
## cor
## 0.08431133
```

Siit järeldub, et sama proovitüki mudelpuude kõrgusproгноoside vead on nõrgalt korreleeritud. Nende korrelatsioonide arvestamine mudelis tähendaks mudeli täiendavat keerukust, kuid olulist võitu arvatavasti ei saa – näiteks lihtviisiline prognoosivea parandamine ühe teadaoleva mudelpuu prognoosivea abil peaks teoreetiliselt andma RMSE-s võitu sama palju kui kordajaga $\sqrt{1 - \rho^2}$ korrutamine (praegusel juhul on see umbes 0.995).

Eelneva arutelu illustreerimiseks uurime numbriliselt, kui palju saab parandada RMSE väärtust, kui proovitükil on ühe puu kõrguse prognoosiviga teada ja seda kasutatakse teiste puude prognooside korrigeerimisel. Efekt saab mõnevõrra ülehinnatud, kuna kasutame parandusmudeli sobitamist samadel andmetel, kus headust mõõdame (testandmestikul).

Määrame igal proovitükil, kus on mitu mudelpuud, esimese mudelpuu kõrguse teadaolevaks ja uurime teiste puude kõrgusproгноoside parandamist seda teadaolevat mõõtmist kasutades.

```
teada=testandmed|>inner_join(mudelpuud|>select(amudelpuud_id,prto_aid,aasta))|>
  filter(aasta>2004)|>
  group_by(prto_aid)|>
  mutate(mitu=n(),nr=row_number())|>
  filter(mitu>1,nr==1)|>
  select(prto_aid,amudelpuud_id,mitu)

## Joining with `by = join_by(amudelpuud_id, aasta)`

teada=teada|>inner_join(progn_test|>mutate(viga_teada=prognoos_uus-h_puu)|>select(amudelpuud_id,viga_teada),by="amudelpuud_id")
proovimiseks=progn_test|>
  select(amudelpuud_id,h_puu,prognoos_uus)|>
  mutate(viga=prognoos_uus-h_puu)|>
  inner_join(mudelpuud|>select(amudelpuud_id,prto_aid))|>
  inner_join(teada|>select(prto_aid,viga_teada,amudelpuud_id),by="prto_aid")|>
  filter(amudelpuud_id.x!=amudelpuud_id.y) #eemaldame need mudelpuud, mille viga kasutame p
arandamiseks

## Joining with `by = join_by(amudelpuud_id)`
```

```
head(proovimiseks)
```

```
##      amudelpuud_id.x h_puu prognoos_uus      viga      prto_aid viga_teda
## 1          34419    213    181.20754 -31.792464 2005_1130E082    6.895664
## 2          12571    147    156.88316   9.883159 2005_1134N040   -1.751595
## 3          12569    246    299.68084  53.680845 2005_1134E040   -8.541162
## 4          12568    310    313.39418   3.394180 2005_1134E040   -8.541162
## 5          34434     73     86.95951  13.959513 2005_1134E080   12.531647
## 6          34444    125    112.54932 -12.450684 2005_1136N080    4.851522
##      amudelpuud_id.y
## 1          34420
## 2          12572
## 3          12567
## 4          12567
## 5          34433
## 6          34443
```

Sobitame mudeli parandatud prognoosi jaoks.

```
m_parandus=lm(viga~viga_teda,data=proovimiseks)
summary(m_parandus)

##
## Call:
## lm(formula = viga ~ viga_teda, data = proovimiseks)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -105.220   -9.179   -0.812    8.443   116.724
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)  0.369991   0.173768   2.129   0.0333 *
## viga_teda    0.092944   0.009826   9.459   <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 18.06 on 10799 degrees of freedom
## Multiple R-squared:  0.008217, Adjusted R-squared:  0.008125
## F-statistic: 89.47 on 1 and 10799 DF, p-value: < 2.2e-16

proovimiseks|>mutate(parandatud=prognoos_uus-predict(m_parandus))|>summarize(rmse_enne=rmse_
_vec(truth=h_puu,estimate=prognoos_uus)/10,rmse_parast=rmse_vec(truth=h_puu,estimate=parand
atud)/10)

##      rmse_enne      rmse_parast
## 1    1.813459      1.805565
```

Sama analüüs puuliigiti:

```
proovimiseks|>mutate(parandatud=prognoos_uus-predict(m_parandus))|>inner_join(mudelpuud|>se
lect(amudelpuud_id,puuliik),by=c("amudelpuud_id.x"="amudelpuud_id"))|>
  group_by(puuliik)|>summarize(rmse_enne=rmse_vec(truth=h_puu,estimate=prognoos_uus)/10,rms
e_parast=rmse_vec(truth=h_puu,estimate=parandatud)/10)

## # A tibble: 14 × 3
##   puuliik rmse_enne rmse_parast
```

##	<chr>	<dbl>	<dbl>
##	1 HB	1.92	1.92
##	2 JA	4.79	4.77
##	3 KP	4.22	3.87
##	4 KS	1.71	1.70
##	5 KU	2.16	2.14
##	6 LH	1.43	1.46
##	7 LM	1.78	1.78
##	8 LV	1.67	1.66
##	9 MA	1.57	1.56
##	10 PN	4.54	4.52
##	11 PP	0.995	0.812
##	12 SA	2.03	2.06
##	13 TA	2.53	2.51
##	14 VA	2.39	2.39

Kokkuvõte: Võrreldes rmse väärtusi enne ja pärast nii kõigi puude kui ka üksikute puuliikide lõikes võib tõdeda, et rmse kahanemine kõrgusproгноoside korrigeerimise tõttu on vaevumärgatav (kui seda üldse on). Seega prognoositava puuga samal proovitükil asuvate teiste mudelpuude kõrguste kasutamine täiendava infona praktiliselt ei suurenda kõrgusproгноosi täpsust.

Lisa 5. Uue kõrgusmodeli hindamise ja mahu veaarvutuse R-kood

Alljärgnevas on toodud 2 koodi:

- 1) uue kõrgusmodeli (5) arvutamise ja kasutamise R-kood
(faili nimi **proгноosifunktsioon_h_ver3.R**)
- 2) veaarvutuse kood (faili nimi **abifunktsioonid_maht_viga.R**)

proгноosifunktsioon_h_ver3.R

```
library("RSQLite")
library(gslnlsl)
library(dplyr)
library(data.table)

# -----
# Funktsioonide kirjeldused - algus
# -----

# Abifunktsioonid

# -----
# puugrupp - teisendab puuliigi puuliigi grupiks (eraldi mudel hinnatakse vaid igale puuliigi grupile)
puugrupp=function(puuliik){
  return( as.character( factor(puuliik,
    levels=
      c("MA","LH","SD", "KU","NU","TO","TS", "KS","PN","TL","EL", "HB","PP", "LM", "LV",
        "TA","SA","VA","JA","KP", "RE","PI","TM","SP","OP","TY"),
    labels=
      c("MA","MA","MA", "KU","KU","KU","KU", "KS","KS","KS","KS", "HB","HB", "LM", "LV",
        "XX","XX","XX","XX","XX", "RE","RE","RE","RE","RE","RE"))))
}

# -----
# Puude kõrguste mudel, mudeli kirjeldus:

# NB! mudel hindab 6 esimese parameetri väärtuseid logaritmilisel skaalal

mudel_h_kirjeldus <-function(seemnepuu_piiir_ln, jarelkasvu_piiir_ln, a_NP_ln, a_S_ln, b_NP_l
n, b_S_ln, c0, c1, c2, c3,
  d_13, inv_korgus, valepuu, inv_diameeter, relaskoop){

  # Lihtne mudel diameetri põhjal kõrguse prognoosimiseks, Hossfeldt c=1.3
  f=function(a,b,d_13){13+d_13^1.3/(a+b*d_13^1.3)}

  koef=10
  # Logistiline kõver
  P_seemnepuu_0 = 1/(1+exp(- ((d_13>inv_diameeter)*(d_13-inv_diameeter)/d_13-exp(seemnepu
u_piiir_ln))*koef ))
  P_jarelkasv_0 = 1/(1+exp(- ((d_13<=inv_diameeter)*(inv_diameeter-d_13)/inv_diameeter-ex
p(jarelkasvu_piiir_ln))*koef ))
```

```
# Normeerime nii, et d_13->inv_diameeter korral mõlema tõenäosuse piirväärtused oleksid 0
-d ja et maksimaalne
# võimalik tõenäosus jääks üheks.
P_seemnepuu = (P_seemnepuu_0-1/(1+exp(exp(seemnepuu_piir_ln)*koef))) / (1-1/(1+exp(exp(seemnepuu_piir_ln)*koef)))
P_jarelkasv = (P_jarelkasv_0-1/(1+exp(exp(jarelkasvu_piir_ln)*koef))) / (1-1/(1+exp(exp(jarelkasvu_piir_ln)*koef)))

h= ( ( c2^valepuu*f(exp(a_NP_ln),exp(b_NP_ln),d_13)*P_jarelkasv+ f(exp(a_S_ln),exp(b_S_ln),d_13)*P_seemnepuu) +
      c2^valepuu*(1-P_seemnepuu-P_jarelkasv)*(c3*inv_korgus + (c0+c1*(1/relaskoop))*(d_13-inv_diameeter)))
# Puuduvad väärtused (peale aastat 2004) peaaegu alati metsata metsamaal - seal inv_korgu
s peaaegu alati ebausaldusväärne
h[is.na(h)]=f(exp(a_S_ln),exp(b_S_ln),d_13[is.na(h)])
# Teoreetiliselt võimalikud prognoositud kõrgused, mis väiksemad kui 1.3m.
# Kui selliseid prognoose peaks tegelikult tekkima, siis asendame need väärtusega 1.3m:
h[h<13]=13
return(h)
}

# -----
# Funktsioon mis hindab kõrguskõvera mudeli

hindamudel_h = function(h, d, puuliik, peapuu, inv_diameeter, inv_korgus, rel1, rel2){
# h - mudelpuu mõõdetud kõrgus, dm
# d - mudelpuu rinnasdiameeter, mm
# puuliik - puu liik (MA/KS/..) või puuliigigrupp (teisendatakse nagunii puuliigigrupiks)
# peapuu - peapuuliik (MA/KS/...)
# inv_diameeter - metsatüki inventeeritud (nn keskmine) diameeter
# inv_korgus - metsatüki inventeeritud kõrgus
# rel1, rel2 - relaskoobi mõõtmistulemused inv_g1 ja inv_g2

stabel= table(puugrupp(puuliik))
loetelu=names(stabel[stabel>0])
gruppe=length(loetelu)

# Algandmete töötlus:

# Looime töödeldud/kombineeritud relaskoobimõõtmise ...
relaskoop = (rel1+rel2)/2
ind=is.na(relaskoop)
relaskoop[ind] = rel1[ind]
ind=is.na(relaskoop)
relaskoop[ind] = rel2[ind]
# Puudu enamasti hõredatel aladel. Samuti asendame kõige hõredamatel aladel väärtuse nn miinimumväärtusega.
ind=is.na(relaskoop) | relaskoop<5
relaskoop[ind] = 5

# Kas antud puu kasvab teistsorti metsas?
valepuu = 1*(puugrupp(puuliik)!=puugrupp(peapuu) )
# Tegeleme puudega, mil peapuu puuliik pole märgitud (puudu)
# Loogika: kus peapuid pole märgitud, seal arvatavasti tegemist üksikute puudega ja puu ei kasva teistest erinevate puude seas
valepuu[is.na(valepuu)] = 0
```

```

andmed_mudelisse=data.frame(h, d, plgrp=puugrupp(puuliik), valepuu, inv_diameeter, inv_korgus, relaskoop)

# Hindame mudelid iga puuliigigrupi jaoks:
mudelid=list(gruppe)

for (i in 1:gruppe){
  jooksevgrupp=loetelu[i]
  print(paste("Hindan mudelit: ", jooksevgrupp))
  print(paste("n=", sum(andmed_mudelisse$plgrp==jooksevgrupp, na.rm=T), sep=""))

  mudel_h_kirjeldus <-function(seemnepuu_piir_ln, jarelkasvu_piir_ln, a_NP_ln, a_S_ln, b_NP_ln, b_S_ln, c0, c1, c2, c3,
    d_13, inv_korgus, valepuu, inv_diameeter, relaskoop){

    hinnatud_mudel = gsl_nls(fn= h~mudel_h_kirjeldus(seemnepuu_piir_ln,jarelkasvu_piir_ln, a_NP_ln, a_S_ln, b_NP_ln, b_S_ln, c0, c1, c2, c3,
      d, inv_korgus,valepuu, inv_diameeter, relaskoop),
      start = list(seemnepuu_piir_ln=-0.3,jarelkasvu_piir_ln=-1.12, a_NP_ln=0.95, a_S_ln=-2.32, b_NP_ln=-6,
        b_S_ln=-7.5, c0=0.99, c1=0.97, c2=0.11, c3=0.95),
      data= andmed_mudelisse[andmed_mudelisse$plgrp==jooksevgrupp & !is.na(andmed_mudelisse$plgrp), ],
      control=gsl_nls_control(maxiter = 500))

    mudelid[[i]]= hinnatud_mudel

    # Vahetulemuste väljutamine
    print( paste("Jääkide standardhälve:", round(summary(hinnatud_mudel)$sigma,2) ) )
  }

  return(list(loetelu=loetelu, mudelid=mudelid))
}
# -----
# Funktsioon puude kõrguste prognoosimiseks. Kasutab juba hinnatud mudelit ja andmeid uute puude kohta
prognoos_h=function(mudel, d, puuliik, peapuu, inv_diameeter, inv_korgus, rel1, rel2){

  stabel= table(puugrupp(puuliik))
  loetelu=names(stabel[stabel>0])
  gruppe=length(loetelu)

  # Algandmete töötlus:

  # Looma töödeldud/kombineeritud relaskoobimõõtmise .....
  relaskoop = (rel1+rel2)/2
  ind=is.na(relaskoop)
  relaskoop[ind] = rel1[ind]
  ind=is.na(relaskoop)
  relaskoop[ind] = rel2[ind]
  # Puudu enamasti hõredatel aladel. Samuti asendada kõige hõredamatel aladel väärtuse nn m
  iinimumväärtusega.
  ind=is.na(relaskoop) | relaskoop<5
  relaskoop[ind] = 5

  # Kas antud puu kasvab teistsorti metsas?
  valepuu = 1*(puugrupp(puuliik)!=puugrupp(peapuu) )

```

```

# Tegeleme puudega, mil peapuu puuliik pole märgitud (puudu)
# Loogika: kus peapuud pole märgitud, seal arvatavasti tegemist üksikute puudega ja puu ei kasva teistest erinevate puude seas
valepuu[is.na(valepuu)] = 0

prognoositavad = data.frame(d, plgrp=puugrupp(puuliik), valepuu, inv_diameeter, inv_korgus, relaskoop)

abi=sum(is.na(puuliik)|is.na(d))
if (abi>0) print(paste("Hoiatus! Puudele, mille diameeter või puuliik pole teada kõrgust ei prognoosita!\n Selliseid puid:", abi))

prognoositud_h=rep(NA, length(d))
for (i in 1:grupe){
  jooksevgrupp=loetelu[i]
  print(paste("Prognoosime puuliiki:", jooksevgrupp))
  valimudel=(jooksevgrupp==mudel$loetelu)
  if (sum(valimudel)<1) {print (paste("Hoiatus!!! Puuliigi:", jooksevgrupp, " jaoks pole mudelit!!!"))} else {
    ind=(!is.na(prognoositavad$plgrp) & prognoositavad$plgrp==jooksevgrupp )
    prognoositud_h[ind]= predict(mudel$mudelid[[ 1:length(mudel$loetelu)][valimudel] ], newdata=prognoositavad[ind, ])
  }
}

return(prognoositud_h)
}

# -----
# Funktsioonide kirjeldused - lõpp
# -----

```

Fail **abifunktsioonid_maht_viga.R**

```
olemas=c('MA','LH','SD','KU','NU','TS','TO','KS','PN','HB','LM','LV',
         'PP','RE','TL','OP','TA','VA','SA','TM','JA','KP','SP','PI','EL','TY')
kordajad_maht=data.frame(puuliigid=c('MA','LH','SD','KU','NU','TS','TO','KS','PN','HB','LM','LV',
                                     'PP','RE','TL','OP','TA','VA','SA','TM','JA','KP','SP','PI','EL','TY'),a=c(0.3571,0.4216,0.4080,0.4723,0.4033),b=c(0.
660,0.181,0.757,-0.608,0),c=c(2.156,1.190,0.801,0,1.586),d=c(-8.312,-1.309,-10.707,12.724,1.440))
kordajad_maht_maatriks=matrix(NA,nrow=(length(olemas)),ncol=4)
row.names(kordajad_maht_maatriks)=olemas
colnames(kordajad_maht_maatriks)=c('a','b','c','d')
for(i in 1:length(olemas)){
  kordajad_maht_maatriks[i,]=kordajad_maht|>filter(grepl(olemas[i], puuliigid))|>select(a:d)|>as.matrix()
}
V_puu=function(d13,h,a=0.3571,b=0.66,c=2.156,d=-8.312){
  #eeldame, et d on millimeetrites ja h on detsimeetrites
  #puu mahu valem metsakorralduse eeskirja kohaselt, vaikeparameetrid
  #männi, lehise ja seedermänni jaoks
  D=d13/10
  H=h/10
  return(0.0000785*D^2*H*(a+b/D+c/H+d/(D*H)))
}

V_maht_koik=function(puuliigid,d13,h,kordajadM=kordajad_maht_maatriks){
  #eeldus on, et puuduvaid andmeid pole
  D=d13/10
  H=h/10
  return(0.0000785*D*H*(rowSums(kordajadM[puuliigid,]*cbind(D,1,D/H,1/H))))
}
puugrupp=function(puuliik){
  return( as.character( factor(puuliik,
                              levels=
                                c("MA","LH","SD", "KU","NU","TO","TS", "KS","PN","TL","EL", "HB","PP", "LM", "LV",
                                "TA","SA","VA","JA","KP", "RE","PI","TM","SP","OP","TY"),
                              labels=
                                c("MA","MA","MA", "KU","KU","KU","KU", "KS","KS","KS","KS", "HB","HB", "LM", "LV",
                                "XX","XX","XX","XX","XX", "RE","RE","RE","RE","RE","RE"))))
}

arvuta_sd3=function(andmestik,kv){
  kohandatud=andmestik%>%mutate(hinnang=tinglik_kv-pind_sees*kv)%>%dplyr::select(i,j,hinnang)
  kohandatud2=kohandatud%>%mutate(i=i+1,h2=hinnang)%>%dplyr::select(i,j,h2)
  kohandatud3=kohandatud%>%mutate(i=i+1,j=j+1,h3=hinnang)%>%dplyr::select(i,j,h3)
  kohandatud4=kohandatud%>%mutate(j=j+1,h4=hinnang)%>%dplyr::select(i,j,h4)
  vead=kohandatud%>%
    inner_join(kohandatud2,by=c("i","j"))%>%
    inner_join(kohandatud3,by=c("i","j"))%>%
    inner_join(kohandatud4,by=c("i","j"))%>%
    mutate(vead=(hinnang-h2+h3-h4)/2)%>%pull(vead)
  return(sqrt(sum(vead^2))/sum(andmestik$pind_sees))
}

arvuta_tingimustel=function(cond_A,cond_y,y_string,pind_A,max_per_trakt=16){
  #Eeldab, et tmp_prtosa, trakt_piirkond, koin_indeksid on defineeritud globaalselt
  #cond_A on tmp_prtosa jaoks arvesse tulevate proovitükkide tingimus
```



#cond_y on tingimus, mis määrab, kas summeeritav suurus nullitakse või mitte
 #y_string annab summeeritava väärtuse veerunime stringina
 #pind_A on tingimusele A vastava Eesti osa teadaolev pindala
 #max_per_trakt annab, mitu proovitükki trakti kohta maksimaalselt arvesse läheb

```
taiendatud_trakt=tmp_prtosa%>%mutate(mm_osak=if_else(cond_y,get({y_string}),0),pind_osak=if_else(cond
_A,pindosak,0))%>%
```

```
group_by(atrakt_id)%>%summarize(mm_proc=sum(mm_osak)/max_per_trakt,pind_sees=sum(pind_osak)/ma
x_per_trakt)
```

```
piirkond=taiendatud_trakt%>%inner_join(trakt_piirkond,by="atrakt_id")%>%group_by(i,j)%>%summarize_at(v
ars(mm_proc,pind_sees),.funs=~mean(.x))%>%ungroup()
#laiendame nullidega
```

```
piirkond_laiendatud=koik_indeksid%>%left_join(piirkond,by=c("i","j"))%>%mutate(tinglik_kv=coalesce(mm_pr
oc,0),pind_sees=coalesce(pind_sees,0))%>%select(i,j,tinglik_kv,pind_sees)
tulem=tibble(mu=rep(NA,1),sd=rep(NA,1))#,sd2=rep(NA,mitu))
osakaal=sum(piirkond_laiendatud$tinglik_kv)/sum(piirkond_laiendatud$pind_sees)
tulem$mu[1]=osakaal*pind_A
sd3_tul=arvuta_sd3(piirkond_laiendatud,osakaal)
#tulem$sd[rida]=sd3_tul$sd*pind_tykk*samm^2
tulem$sd[1]=sd3_tul*pind_A
return(tulem)
```



Aruande koostaja: **Kalev Pärna**

Aruande valmimisse panustasid:

Raul Kangro, Märt Möls (analüüsid ja arvutuseksperimendid)

19.12.2025

Kalev Pärna

TÜ MSI emeriitprofessor

Projekti vastutav täitja