SSM methods notes and writing

DECISION MADE: use fit\_smm(model=mp) approach --- this is what Gin did as well

36h gap, <25 locs is to short, 12h time step

* After making some test runs with different gap and short track values, 36h and <25 is the best (best convergence, least warnings)

TO DO:

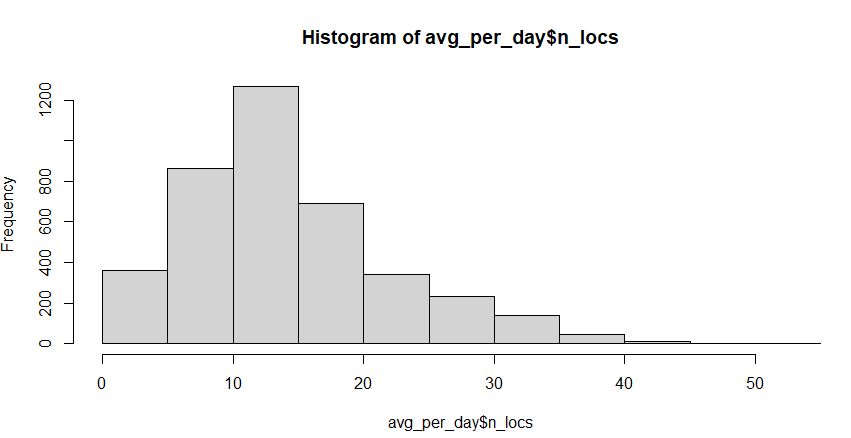
* Mia suggestion: calculate the average number of location fixes per day from data – that will help decide time step to use. USE SAME TIME STEP FOR EVERYTHING
  + Looking at NZ and OZ data, if get the avg locs/day for each whale it’s around 12-14, so use 12h step. Some tracks could work with shorter step, but others can’t and they need to all be the same (as what Mia said about inputting the data into GAM, if some tracks are 6hr then they have more influence on results)
* Test running eg all NZ data together, and running year cohorts
  + NO. Instead use fit\_smm(model=mp) (as per Gin’s advice), and according to the documentation this can only be run on one track segment at a time!

Number of locations per day

All NZ (2020, 2021, 2022) data combined – but some 2022 tags still going. Use file NZ\_SRW\_2020\_2021\_2022\_raw\_argos\_df\_20230706.rds

raw\_argos\_df %>% group\_by(Ptt,Date) %>% summarise(n\_locs = n())

hist(avg\_per\_day$n\_locs)



mean(avg\_per\_day$n\_locs)

14.62911

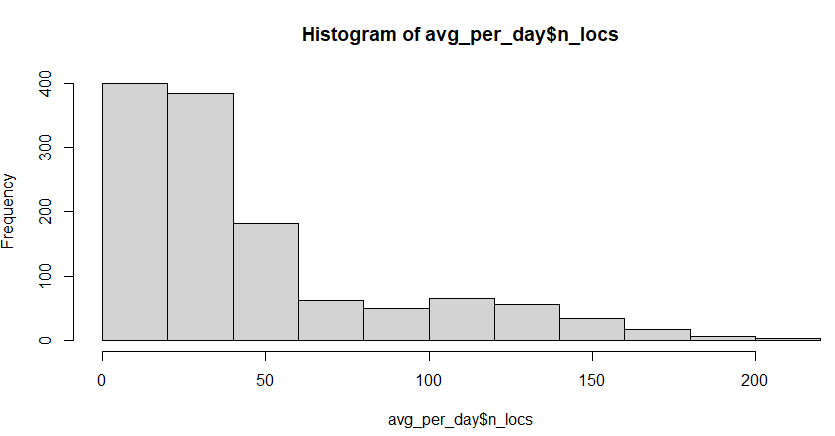
median(avg\_per\_day$n\_locs)

13

* So 12h step probably good

raw\_argos\_df %>% group\_by(cohort,Date) %>% summarise(n\_locs = n())

hist(avg\_per\_day$n\_locs)



mean(avg\_per\_day$n\_locs)

45.91753

median(avg\_per\_day$n\_locs)

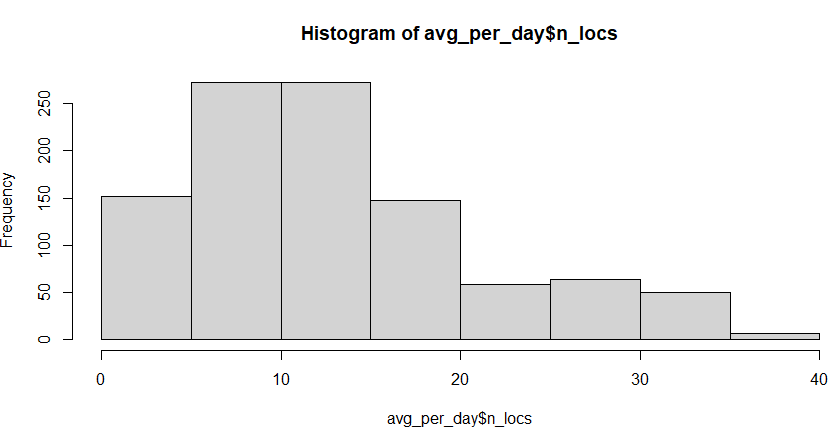
32

Australia data

Only 2022 data, some tags still going when file OZ\_SRW\_2022\_raw\_argos\_df\_20230713.rds created

avg\_per\_day <- raw\_OZ\_argos\_df %>% group\_by(Ptt,Date) %>% summarise(n\_locs = n())

hist(avg\_per\_day$n\_locs)



mean(avg\_per\_day$n\_locs)

13.4478

median(avg\_per\_day$n\_locs)

12

* So 12h probably good for everyone

Unsure if things work if all NZ data pooled together

Quite the difference if first run ssm then mp, v fit\_ssm(model=mp)

SOME TESTING DONE ON 20230904

NZ\_SRW\_2020\_2021\_2022\_raw\_argos\_df\_20230830.rds – as input

#speed filter threshold (vmax) of 5 ms−1

fit\_ssm\_12h\_model\_mp\_NZ\_all<- fit\_ssm(ssm\_df, vmax=5, model="mp", time.step=12, control = ssm\_control(verbose=0))

## actually this shouldn't work as model=mp should be only for running one track at a time?

#Gin says that it works

#if try to run on NZ 2020, 2021 and 2022 data, 36h gap, 20 locs is short, 12h ssm time step: has warning messages.

#try to identify which track causes the fail?

# View(fit\_ssm\_12h\_model\_mp\_NZ\_all)

#those that have converged == FALSE: 215262-1, 215262-14, 235399-4, 46635-1

# pdHess == FALSE: 215262-14, 46635-1 --- both are from 2021 cohort

fit\_ssm\_12h\_model\_mp\_NZ\_all\_p <- fit\_ssm\_12h\_model\_mp\_NZ\_all %>% grab(what="p")

# --> logit\_g.se == NA: 235399-4

## test mapping this particular out in QGIS

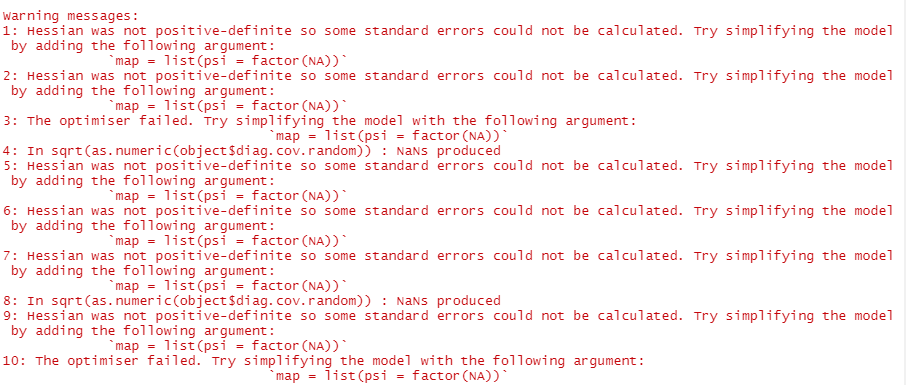
#write\_csv(fit\_ssm\_12h\_model\_mp\_NZ\_all\_p,here::here('SSM', 'data', 'ssm\_mpm\_together\_all\_NZ\_gap36h\_short20loc\_20230904.csv'))

fit\_ssm\_12h\_model\_mp\_NZ\_all\_p\_groupnormalised <- fit\_ssm\_12h\_model\_mp\_NZ\_all %>% grab(what="p",normalise = TRUE, group = TRUE)

#write\_csv(fit\_ssm\_12h\_model\_mp\_NZ\_all\_p\_groupnormalised,here::here('SSM', 'data', 'ssm\_mpm\_together\_all\_NZ\_gap36h\_short20loc\_GROUPNORMALISED\_20230904.csv'))

NO VISUAL DIFFERENCE ON QGIS MAP BETWEEN NORMALISED AND NON-NORMALISED FILES

FIT\_SSM(MODEL=MP) RAN “FINE” ON GROUPED DATA, ALTHOUGH DID HAVE SOME WARNING MESSAGES:



AND FEW CASES WHERE ‘CONVERGED’ OR ‘dpHess’ == FALSE

OR ONE CASE WITH g.se == NA

Visual check

PTT203572

If used the above ssm (12h, gap if >36h, 20loc is short)

Not normalised: Normalised:

A map of australia and the ocean

Description automatically generated A map of the ocean

Description automatically generated

If 50% g is cutoff for ARS

Not normalised: Normalised:

A map of the ocean

Description automatically generated A map of australia with a route

Description automatically generated

Pretty happy with that. Maybe could still adjust the cutoff of how many locs is too short of a track

Maybe could remove z locations first, and then do the gap stuff….

Maybe increase gap to 48h, and decrease short track to 15 locs??

20230906

Filter out lc==Z right at the start, before segmenting tracks.

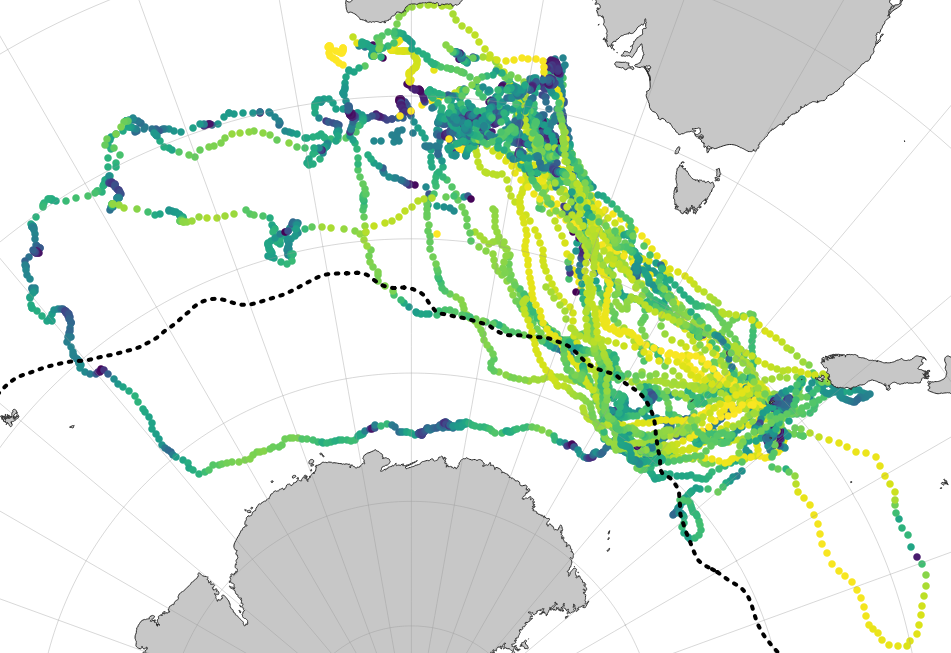
Keep data gap as 36h, and short tracks as 20 loc

Run SSM with 12h step

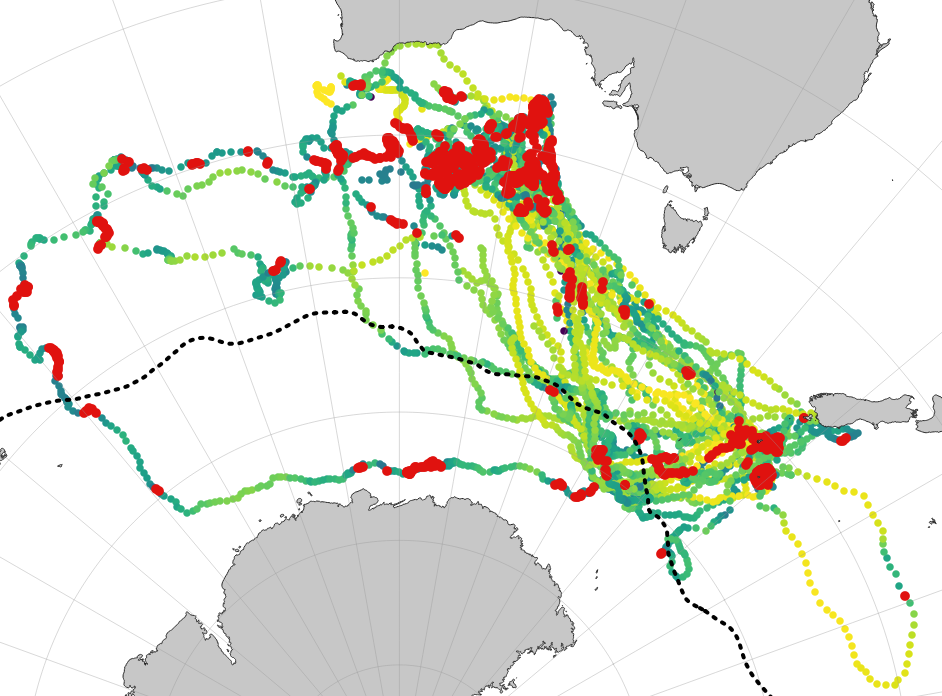
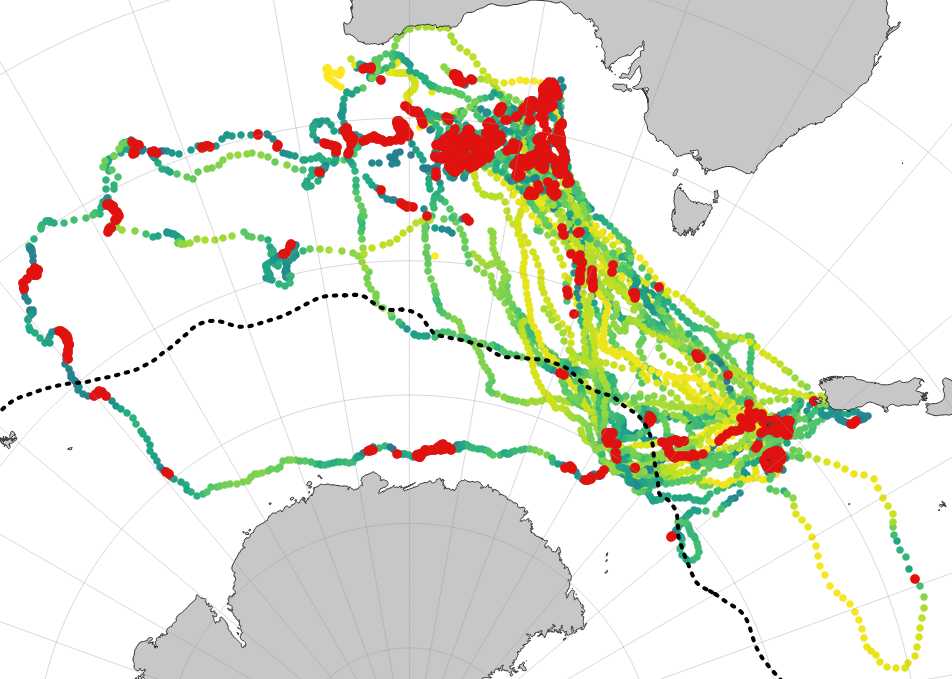
Compare to previous

* Doesn’t make a difference really. Couple individual point added ARS compared to previous version. But more or less the same

Lc==Z filtered later Lc==Z filtered earlier

 A map of the world

Description automatically generated

20230906

Good SSM (NZ TRACKS) if gap 36h, short track 25loc, 12h time step

**How much do things change if gap 48h?**

An extra case of converged=FALSE: 215262-12, otherwise same 🡪 so a tad worse

**How much do things change if gap 60?**

Some differences in converged=FALSE: 215262-1, 215262-3, 215262-10

No pdHess false though

But logit\_g\_se NA for 215262-3 🡪 so worse

--- so 36h seems pretty good

Current good SSM 36h gap, and short track <25 loc

**How much do things change if short is <20 loc?**

More issues. converged=FALSE: 215262-1, 215262-14, 235399-4, 46635-1

pdHess = false: 215262-14, 46635-1

logit\_g\_se NA 235399-4 🡪 so worse

**How much do things change if short is <15 loc?**

Even more issues 🡪 so worse

--- so <25 loc seems pretty good