

SoilMercury data example

ruoyong

2023-04-22

```
library('terra')

## terra 1.7.65
load("hgmTerra.RData")
hgm = unwrap(hgmWrap)
myBgMap = mapmisc::openmap(hgm, zoom=2)

load("covListWrap.RData")
covList = lapply(covListWrap, unwrap)
names(covList) = gsub("Wrap", "", names(covList))

world1 = crop(vect(rnaturalearth::ne_countries(scale='medium', returnclass='sf')),
              ext(-20, 100, 0, 85))
worldMap = project(world1, mapmisc::crsLL)

[hg] [elevation] [night] [evi] [land]
```

Figure 1: data plot

Get the MLEs on CPU

```
hgm$land = extract(covList$land, project(hgm, crs(covList$land)))[,2]
hgm$land = releval(hgm$land, 'Rainfed croplands')
library('geostatsp')
hgRes = lgm(HG ~ elevation + land + night + evi, data = hgm,
            grid = 20, covariates = covList, fixBoxcox=FALSE,
            fixShape=FALSE, fixNugget = FALSE,
            reml=FALSE, aniso=TRUE)

## Warning in krigelgm(formula = formula, data = data, grid = grid, covariates =
## covariates, : covariates and grid aren't compatible

dataFromLgmWrap = wrap(hgRes$data)

hgRes2 = lgm(HG ~ elevation + land + night + evi, data = hgm,
            grid = 20, covariates = covList, shape=1,
            fixBoxcox=FALSE, fixShape=TRUE, fixNugget = FALSE,
            reml=FALSE, aniso=TRUE)

## Warning in krigelgm(formula = formula, data = data, grid = grid, covariates =
## covariates, : covariates and grid aren't compatible
```

```

hgRes3 = lgm(HG ~ elevation + land + night + evi, data = hgm,
             grid = 20, covariates = covList, shape=0.8,
             fixBoxcox=FALSE, fixShape=TRUE, fixNugget = FALSE,
             reml=FALSE, aniso=TRUE)

## Warning in kriggLgm(formula = formula, data = data, grid = grid, covariates =
## covariates, : covariates and grid aren't compatible

hgRes4 = lgm(HG ~ elevation + land + night + evi, data = hgm,
             grid = 20, covariates = covList, shape=0.5,
             fixBoxcox=FALSE, fixShape=TRUE, fixNugget = FALSE,
             reml=FALSE, aniso=TRUE)

## Warning in kriggLgm(formula = formula, data = data, grid = grid, covariates =
## covariates, : covariates and grid aren't compatible

```

geostatsp's estimates

```

hgRes$summary[,c('estimate', 'ci0.025', 'ci0.975')]

##
## (Intercept)
## elevation
## landbroadleaved deciduous forest
## landRainfed croplands
## landMosaic cropland / vegetation
## landOpen needleleaved deciduous or evergreen forest
## landmixed broadleaved and needleleaved forest
## landMosaic grassland / forest or shrubland
## landMosaic vegetation / cropland
## landSparse vegetation
## landherbaceous vegetation
## landMosaic forest or shrubland / grassland
## landneedleleaved evergreen forest
## landshrubland
## landgrassland or woody vegetation on regularly flooded or waterlogged soil
## landArtificial surfaces and associated areas
## landWater bodies
## night
## evi
## sdNugget
## sdSpatial
## range/1000
## shape
## anisoRatio
## anisoAngleRadians
## anisoAngleDegrees
## boxcox
##
## ci0.025
## (Intercept)
## elevation
## landbroadleaved deciduous forest
## landRainfed croplands
## landMosaic cropland / vegetation

```

estimate	-7.182301e+00
	3.397573e-04
	0.000000e+00
	3.147825e-01
	2.202481e-01
	-3.700127e-01
	-1.956695e-01
	1.654901e-01
	-9.354922e-02
	-7.019042e-01
	2.796480e-01
	4.725326e-02
	2.669799e-01
	4.647602e-01
	-5.117302e-01
	-1.118267e-01
	-8.485128e-01
	1.770369e-03
	2.556537e+00
	8.178370e-01
	1.489135e+00
	1.374379e+03
	1.764062e-01
	1.181663e+00
	-9.056046e-01
	-5.188732e+01
	-2.305009e-01
ci0.025	
(Intercept)	-8.419758e+00
elevation	-3.698501e-05
landbroadleaved deciduous forest	NA
landRainfed croplands	-4.093047e-02
landMosaic cropland / vegetation	-1.178287e-01

```

## landOpen needleleaved deciduous or evergreen forest -8.598177e-01
## landmixed broadleaved and needleleaved forest -6.073267e-01
## landMosaic grassland / forest or shrubland -2.246200e-01
## landMosaic vegetation / cropland -4.872423e-01
## landSparse vegetation -1.180212e+00
## landherbaceous vegetation -2.753262e-01
## landMosaic forest or shrubland / grassland -5.093636e-01
## landneedleleaved evergreen forest -2.695206e-01
## landshrubland -2.112822e-01
## landgrassland or woody vegetation on regularly flooded or waterlogged soil -1.327508e+00
## landArtificial surfaces and associated areas -1.973410e+00
## landWater bodies -3.331836e+00
## night -1.485800e-04
## evi 1.427065e+00
## sdNugget NA
## sdSpatial NA
## range/1000 NA
## shape NA
## anisoRatio NA
## anisoAngleRadians NA
## anisoAngleDegrees NA
## boxcox NA
## ci0.975
## (Intercept) -5.9448432136
## elevation 0.0007164996
## landbroadleaved deciduous forest NA
## landRainfed croplands 0.6704955499
## landMosaic cropland / vegetation 0.5583249062
## landOpen needleleaved deciduous or evergreen forest 0.1197923205
## landmixed broadleaved and needleleaved forest 0.2159877222
## landMosaic grassland / forest or shrubland 0.5556001174
## landMosaic vegetation / cropland 0.3001438847
## landSparse vegetation -0.2235959111
## landherbaceous vegetation 0.8346222640
## landMosaic forest or shrubland / grassland 0.6038701034
## landneedleleaved evergreen forest 0.8034803390
## landshrubland 1.1408026123
## landgrassland or woody vegetation on regularly flooded or waterlogged soil 0.3040473941
## landArtificial surfaces and associated areas 1.7497567284
## landWater bodies 1.6348101131
## night 0.0036893174
## evi 3.6860089165
## sdNugget NA
## sdSpatial NA
## range/1000 NA
## shape NA
## anisoRatio NA
## anisoAngleRadians NA
## anisoAngleDegrees NA
## boxcox NA
hgRes$optim$logL

## m2logL.ml logL.ml
## -3793.079 1896.539

```

Configure the params

```
library(gpuLik)
set.seed(66)
alpha1=c(0.00001, 0.01, 0.1, 0.2, 0.25, 0.3, 0.5, 0.8, 0.9, 0.95, 0.99, 0.999)
alpha2=c(0.00001, 0.01, 0.1, 0.2, 0.5, 0.8, 0.9, 0.95, 0.99, 0.999)

model_list <- list(hgRes, hgRes2, hgRes3, hgRes4)
A <- gpuLik::configParams(model_list,
                           alpha=alpha1,
                           alphasecond = alpha2,
                           data = terra::unwrap(dataFromLgmWrap))

paramsUse <- A$representativeParamaters[,1:5]
b <- A$boxcox
```

total number of correlation parameter sets

```
nrow(paramsUse)
```

```
## [1] 12316
```

```
#1
intercept<- sort(c(hgRes$summary['(Intercept)', 'estimate'], seq(-10, -4, len=199)))
#2
elevation <- sort(c(hgRes$summary['elevation', 'estimate'], seq(-3, 8, len=199))*1e-04)
#3
landMosaic_cropland <- sort(c(hgRes$summary['landMosaic cropland / vegetation', 'estimate'], seq(-0.5, 0
#4
landMosaic_vegetation <- sort(c(hgRes$summary['landMosaic vegetation / cropland', 'estimate'], seq(-1, 0
#5
landbroadleaved_deciduous_forest <- sort(c(hgRes$summary['landbroadleaved deciduous forest', 'estimate']
#6
landneedleleaved_evergreen_forest <- sort(c(hgRes$summary['landneedleleaved evergreen forest', 'estimate
#7
landOpen_needleleaved_deciduous_or_evergreen_forest <- sort(c(hgRes$summary['landOpen needleleaved deci
#8
landmixed_broadleaved_and_needleleaved_forest <- sort(c(hgRes$summary['landmixed broadleaved and needle
#9
landMosaic_forest_or_shrubland <- sort(c(hgRes$summary['landMosaic forest or shrubland / grassland', 'es
#10
landMosaic_grassland <- sort(c(hgRes$summary['landMosaic grassland / forest or shrubland', 'estimate'], s
#11
landshrubland <- sort(c(hgRes$summary['landshrubland', 'estimate'], seq(-0.6, 0.95, len=199)))
#12
landherbaceous_vegetation <- sort(c(hgRes$summary['landherbaceous vegetation', 'estimate'], seq(-0.8, 0.
#13
landSparse_vegetation <- sort(c(hgRes$summary['landSparse vegetation', 'estimate'], seq(-1.8, 0.1, len=19
#14
landgrassland_or_woody_vegetation <- sort(c(hgRes$summary['landgrassland or woody vegetation on regular
#15
```

```

landArtificial_surfaces_and_associated_areas <- sort(c(hgRes$summary['landArtificial surfaces and asso
#16
landWater_bodies <- sort(c(hgRes$summary['landWater bodies','estimate'], seq(-5, 3, len=199)))
#17
night <- sort(c(hgRes$summary['night','estimate'],seq(-0.5, 5, len=199))*1e-03)
#18
evi <- sort(c(hgRes$summary['evi','estimate'],seq(0, 5, len=199)))

Betas <- cbind(intercept, elevation, landbroadleaved_deciduous_forest,
               landMosaic_cropland,
               landOpen_needleleaved_deciduous_or_evergreen_forest,
               landMosaic_grassland, landMosaic_vegetation,
               landmixed_broadleaved_and_needleleaved_forest, landSparse_vegetation,
               landherbaceous_vegetation, landMosaic_forest_or_shrubland,
               landneedleleaved_evergreen_forest, landshrubland,
               landgrassland_or_woody_vegetation,
               landArtificial_surfaces_and_associated_areas,
               landWater_bodies, night, evi)
sdSpatial <- sort(c(seq(0.7, 2.8, len=59), hgRes$summary['sdSpatial','estimate']))

```

Get all estimates in a short-cut way (may take 20-25 minutes)

```

result3<-gpuLik::likfitLgmGpu(model=hgRes,
                             params=paramsUse,
                             data = terra::unwrap(dataFromLgmWrap),
                             paramToEstimate=c('range','combinedRange',
                                                  'sdNugget',
                                                  'shape','nugget', 'aniso1',
                                                  'aniso2','boxcox'),
                             boxcox = seq(b[1],b[9],len=31),
                             Betas = Betas,
                             sdSpatial = sdSpatial,
                             cilevel=0.95, # decimal
                             type = "double",
                             NparamPerIter=256,#400,
                             Ngloba=c(64,64),
                             Nlocal=c(16, 8),#c(16,16),
                             NlocalCache=2000,#2800,
                             verbose=c(1,0))

## workgroupSize
## 64 64
## localSize
## 16 8
##
## Nparams 12316 NparamsPerIter 256 Niter 49 Ncovariates 18 Ndatasets 32 NlocalCached 936
## did not find lower ci for sdNugget
## did not find lower ci for nugget

rownames(result3$summary) <- substr(rownames(result3$summary), 1, 25)

```

```
result3$summary
```

##	estimate	lower95ci	upper95ci
## (Intercept)	-7.212121e+00	-9.856346e+00	-4.949217e+00
## elevation	3.388889e-04	-4.700000e-05	7.450000e-04
## landRainfed croplands	3.146465e-01	-4.737119e-02	4.000000e-01
## landMosaic cropland / veg	2.202481e-01	-1.214862e-01	5.000000e-01
## landOpen needleleaved dec	-3.636364e-01	-8.937740e-01	0.000000e+00
## landmixed broadleaved and	-1.909091e-01	-6.269624e-01	2.198982e-01
## landMosaic grassland / fo	1.000000e-01	-2.356722e-01	1.000000e-01
## landMosaic vegetation / c	-9.404040e-02	-4.988000e-01	-1.000000e-02
## landSparse vegetation	-7.019042e-01	-1.242521e+00	-2.123728e-01
## landherbaceous vegetation	2.796480e-01	-2.821208e-01	7.000000e-01
## landMosaic forest or shru	4.725326e-02	-5.197152e-01	4.000000e-01
## landneedleleaved evergree	2.621212e-01	-2.792106e-01	8.000000e-01
## landshrubland	4.646465e-01	-2.178218e-01	9.500000e-01
## landgrassland or woody ve	-5.060606e-01	-1.376544e+00	3.108508e-01
## landArtificial surfaces a	-1.118267e-01	-2.019217e+00	1.600000e+00
## landWater bodies	-8.485128e-01	-3.433398e+00	1.657199e+00
## night	1.777778e-03	-1.700000e-04	3.845000e-03
## evi	2.575758e+00	1.373725e+00	3.916356e+00
## sdSpatial	1.489135e+00	1.006572e+00	2.800000e+00
## range	1.374379e+06	3.931983e+05	6.901446e+07
## combinedRange	1.264327e+06	3.446417e+05	6.437687e+07
## sdNugget	8.311206e-01	0.000000e+00	1.784101e+00
## shape	1.764062e-01	7.388702e-02	1.140939e+00
## nugget	3.016235e-01	0.000000e+00	1.402165e+00
## aniso1	-1.014845e-01	-7.541779e-01	6.683817e-01
## aniso2	-4.139615e-01	-1.024663e+00	5.915883e-01
## boxcox	-2.305009e-01	-2.883087e-01	-1.740460e-01

```
result3$reml
```

```
## [1] FALSE
```

Or, you can do it step by step, in the following way, showing only the codes here:

Get the profile log likelihoods only

```
result1 <- gpuLik::getProfLogL(data=unwrap(dataFromLgmWrap),
  formula=hgRes$model$formula,
  coordinates=dataFromLgmWrap@coordinates,
  params=paramsUse,
  boxcox = seq(b[1],b[9],len=31),
  type = "double",
  NparamPerIter=256,
  gpuElementsOnly = FALSE,
  reml=FALSE,
  Nglob=c(256,256),
  Nlocal=c(16,16),
  NlocalCache=1600,
  verbose=c(1,0))
```

```

result1$Infixindex

result1$predictors

# gpuLik and geostatsp have same optimal value
max(result1$LogLik)

hgRes$optim$logL

```

Estimates and plots for correlation params

```

result2<-gpuLik::prof1dCov(LogLik = result1$LogLik, # cpu matrix
  XXYXVX = result1$XXYXVX, # cpu matrix
  ssqResidual = result1$ssqResidual, # cpu matrix
  paramToEstimate = c('range','combinedRange','shape', 'sdNugget',
    'nugget','aniso1', 'aniso2','boxcox'),
  cilevel=0.95, # decimal
  params = result1$paramsRenew, # cpu matrix,
  boxcox = result1$boxcox,
  Ndata = result1$Ndata,
  Nobs = result1$Nobs,
  Ncov = result1$Ncov,
  reml = FALSE,
  predictors = result1$predictors, # character string
  verbose=FALSE)

# table of estimates
result2$summary
# index for the MLE parameter set
result2$mleIndex

```

Geostatsp's estimates

```

rownames(hgRes$summary) <- substr(rownames(hgRes$summary), 1, 25)
hgRes$summary[,c('estimate','ci0.025', 'ci0.975')]

```

##	estimate	ci0.025	ci0.975
## (Intercept)	-7.182301e+00	-8.419758e+00	-5.9448432136
## elevation	3.397573e-04	-3.698501e-05	0.0007164996
## landbroadleaved deciduous	0.000000e+00	NA	NA
## landRainfed croplands	3.147825e-01	-4.093047e-02	0.6704955499
## landMosaic cropland / veg	2.202481e-01	-1.178287e-01	0.5583249062
## landOpen needleleaved dec	-3.700127e-01	-8.598177e-01	0.1197923205
## landmixed broadleaved and	-1.956695e-01	-6.073267e-01	0.2159877222
## landMosaic grassland / fo	1.654901e-01	-2.246200e-01	0.5556001174
## landMosaic vegetation / c	-9.354922e-02	-4.872423e-01	0.3001438847
## landSparse vegetation	-7.019042e-01	-1.180212e+00	-0.2235959111
## landherbaceous vegetation	2.796480e-01	-2.753262e-01	0.8346222640
## landMosaic forest or shru	4.725326e-02	-5.093636e-01	0.6038701034
## landneedleleaved evergree	2.669799e-01	-2.695206e-01	0.8034803390
## landshrubland	4.647602e-01	-2.112822e-01	1.1408026123
## landgrassland or woody ve	-5.117302e-01	-1.327508e+00	0.3040473941

```
## landArtificial surfaces a -1.118267e-01 -1.973410e+00 1.7497567284
## landWater bodies -8.485128e-01 -3.331836e+00 1.6348101131
## night 1.770369e-03 -1.485800e-04 0.0036893174
## evi 2.556537e+00 1.427065e+00 3.6860089165
## sdNugget 8.178370e-01 NA NA
## sdSpatial 1.489135e+00 NA NA
## range/1000 1.374379e+03 NA NA
## shape 1.764062e-01 NA NA
## anisoRatio 1.181663e+00 NA NA
## anisoAngleRadians -9.056046e-01 NA NA
## anisoAngleDegrees -5.188732e+01 NA NA
## boxcox -2.305009e-01 NA NA
```

Or, get estimates for betas first, manually set the configuration

```
#1
intercept<- sort(c(-6.626632,seq(-10, -4, len=199)))
#2
elevation <- sort(c(2.7,seq(-3, 8, len=199))*1e-04)
#3
landbroadleaved_deciduous_forest <- sort(c(-2.764583e-01,seq(-0.9, 0.4, len=199)))
#4
landMosaic_cropland <- sort(c(-8.473104e-02,seq(-0.5, 0.5, len=199)))
#5
landOpen_needleleaved_deciduous_or_evergreen_forest <- sort(c(-6.289483e-01,seq(-1.5, -0, len=199)))
#6
landMosaic_grassland <- sort(c(-8.093025e-02,seq(-0.7, 0.5, len=199)))
#7
landMosaic_vegetation <- sort(c(-3.748334e-01,seq(-1, 0.1, len=199)))
#8
landmixed_broadleaved_and_needleleaved_forest <- sort(c(-5.063632e-01,seq(-10.5e-01, -0.1e-01, len=199)))
#9
landSparse_vegetation <- sort(c(-1.019370e+00,seq(-1.8, 0.1, len=199)))
#10
landherbaceous_vegetation <- sort(c(4.877336e-03,seq(-0.8, 0.7, len=199)))
#11
landMosaic_forest_or_shrubland <- sort(c(-2.842735e-01, seq(-0.98, 0.4, len=199)))
#12
landneedleleaved_evergreen_forest <- sort(c(-8.919963e-03, seq(-0.7, 0.8, len=199)))
#13
landshrubland <- sort(c(2.023893e-01, seq(-0.6, 0.95, len=199)))
#14
landgrassland_or_woody_vegetation <- sort(c(-6.918658e-01, seq(-1.9, 0.4, len=199)))
#15
landArtificial_surfaces_and_associated_areas <- sort(c(-4.358506e-01, seq(-2.6, 1.6, len=199)))
#16
landWater_bodies <- sort(c(-1.165791, seq(-5, 3, len=199)))
#17
night <- sort(c(1.560504,seq(-0.5, 5, len=199))*1e-03)
#18
evi <- sort(c(2.066869,seq(0, 5, len=199)))

Betas <- cbind(intercept, elevation, landbroadleaved_deciduous_forest,
```



```

landMosaic_cropland,
landOpen_needleleaved_deciduous_or_evergreen_forest,
landMosaic_grassland, landMosaic_vegetation,
landmixed_broadleaved_and_needleleaved_forest, landSparse_vegetation,
landherbaceous_vegetation, landMosaic_forest_or_shrubland,
landneedleleaved_evergreen_forest, landshrubland,
landgrassland_or_woody_vegetation,
landArtificial_surfaces_and_associated_areas,
landWater_bodies, night, evi)

```

Or, automatically set the configuration

```

ConfigBetas<-gpuLik::ParamsFromLgm(hgRes$summary,      #lgm model
    covariates=result1$predictors)

Betas <- ConfigBetas$Betas

```

Get estimates and profile plots for betas'

```

output<-gpuLik::Prof1dBetas(Betas=Betas,
    cilevel=0.95,
    Nobs = result1$Nobs,
    Ndata = result1$Ndata,
    Nparam = result1$Nparam,
    Ncov = result1$Ncov,
    detVar = result1$detVar,
    detReml = result1$detReml,
    ssqY = result1$ssqY,
    XXYX = result1$XXYX,
    jacobian = result1$jacobian,
    reml=FALSE,
    convexHull = FALSE)

output$estimates
output$breaks

```

Set sigma manually

```

#sdSpatial <- sort(c(1.478384, seq(0.7, 2.8, len=59)))
sdSpatial <- sort(seq(0.7, 2.8, len=59) hgRes$summary['sdSpatial','estimate']))

```

Or, configure it automatically

```

sdSpatial <- ConfigBetas$sdSpatial

```

Estimates and profile plots for sdSpatial

```

sdoutput <- gpuLik::profVariance(sdSpatial,
    cilevel=0.95,
    Nobs = result1$Nobs,

```

```

Ndata = result1$Ndata,
Nparam = result1$Nparam,
Ncov = result1$Ncov,
detVar = result1$detVar,
detReml = result1$detReml,
ssqResidual = result1$ssqResidual,
jacobian = result1$jacobian)

sdoutput$estimates
breaks <- sdoutput$breaks

temp <- qchisq(0.95, df = 1)/2
LogLik <- sdoutput$LogLik
breaks <- sdoutput$breaks

plot(sdSpatial,LogLik-breaks-temp, cex=0.6, ylab='profileLogL')
lines(sdSpatial,LogLik-breaks-temp, col='blue')
abline(h=-temp, lty = 2)
abline(v=sdoutput$estimates[c(1,2,3)], lty = 2)
text(sdoutput$estimates[1], -9, round(sdoutput$estimates[1],digits = 3))
text(sdoutput$estimates[2], -9, round(sdoutput$estimates[2],digits = 3))
text(sdoutput$estimates[3], -9, round(sdoutput$estimates[3],digits = 3))

```