Supporting information, supplementary analyses and R scripts

For the paper 'Radiocarbon dated trends and the prehistory of the Central Mediterranean'

Acknowledgements

Introduction

This document presents a workflow explaining how radiocarbon data for the Central Mediterranean was processed and used to generate the kernel density models discussed in the main paper. It also contains additional analyses and presentation of the results of our study. The source code to the custom functions called upon in this analysis is contained in the file `src.r`, which also contains scripts aggregated from the publications cited below, and further documentation.

This source code and the analysis that follows contains routines published previously by McLaughlin 2019 (`rowcal`, `MCmix`, `MCdensity`, `mixdensity` and `make_animated_map`), R versions of the python scripts published by Fernandez-Lopez de Pablo et al 2019 (`SPDboot`) and McLaughlin et al 2021 (`ggr`, `phasedensity`). Please cite to the appropriate paper if you plan to re-use the scripts in other publications.

For computational expediency, the analysis that follows uses a default 100 iterations of the Monte Carlo simulations. This can be increased, for example to 1000 iterations, although the results are very similar because convergence is typically achieved before 100 iterations. Interested readers can change the below scripts using the `N` parameter of the Monte Carlo simulations. See the file `src.r` for more details.

```
# load source code
# NB working directory must also contain the INTCAL calibration curves and source data
source('src.r')
```

Loading data

First we load the input data, i.e. the radiocarbon database held in csv format, and select samples that identified as passing the data screening. The third line excludes the small number of samples derived from marine reservoirs:

```
id_raw<-read.csv("/Supplementary_14C_data.csv")
id<-id_raw[id_raw$USE==1 & id_raw$MCODE!='M',1:18]
mt<-read.csv("Malta_dates.csv")
```

Calculating KDE models

The density estimates are computed using a Gaussian kernel and a bandwidth of 75 years. Other bandwidths can be chosen by changing the `bw` parameter in the following code, or this can be set to one of R's built-in
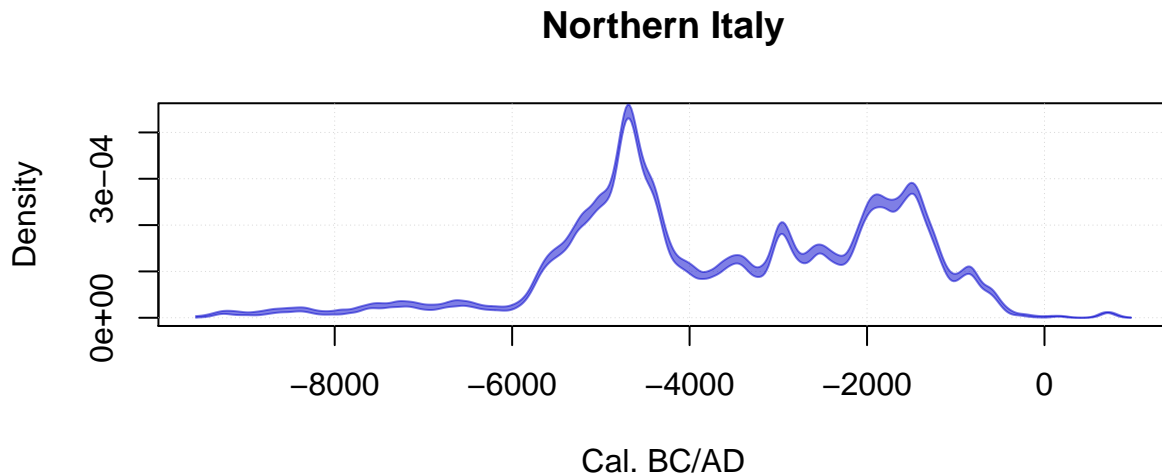
bandwidth selector algorithms, such as `"ucv"` for unbiased cross-validation or `"SJ"` for the method proposed by Sheather and Jones 1991. However, we have used trial-and-error to optimize the choice of bandwidth for this particular study, and found that 75 years achieves a good balance of fitting the data whilst maintaining interpretability. For certain research questions smaller bandwidths have proved useful and these are explored below.

The KDE models for the various regions are calculated using the following

```
NI<-MCdensity(id[id$Region=='Northern Italy',4:5],bw=75)
CI<-MCdensity(id[id$Region=='Central Italy',4:5],bw=75)
SI<-MCdensity(id[id$Region=='Southern Italy',4:5],bw=75)
SY<-MCdensity(id[id$Region=='Sicily',4:5],bw=75)
SD<-MCdensity(id[id$Region=='Sardinia',4:5],bw=75)
FR<-MCdensity(id[id$Region=='SE France',4:5],bw=75)
CO<-MCdensity(id[id$Region=='Corsica',4:5],bw=75)
MT<-MCdensity(mt[,2:3],bw=75)
```

Each model can be visualized by plotting it, for example:

```
plot(NI, main='Northern Italy')
```


To use only one date per site phase in the analysis, the function `phasedensity` is used. This follows Bevan and Crema's method for identifying unique site phases using hierarchical cluster analysis, and upon each Monte Carlo iteration of the KDE calculating process, one radiocarbon date is randomly picked from the database for each phase. This is intended to provide a simulation of even research pressure between sites.

```
NIp<-phasedensity(id[id$Region=='Northern Italy',c(15,4,5)],bw=75)
CIp<-phasedensity(id[id$Region=='Central Italy', c(15,4,5)],bw=75)
SIp<-phasedensity(id[id$Region=='Southern Italy',c(15,4,5)],bw=75)
SYp<-phasedensity(id[id$Region=='Sicily',        c(15,4,5)],bw=75)
SDp<-phasedensity(id[id$Region=='Sardinia',      c(15,4,5)],bw=75)
FRp<-phasedensity(id[id$Region=='SE France',     c(15,4,5)],bw=75)
COp<-phasedensity(id[id$Region=='Corsica',       c(15,4,5)],bw=75)
MTp<-phasedensity(mt[,c(4,2,3)],bw=75)
```

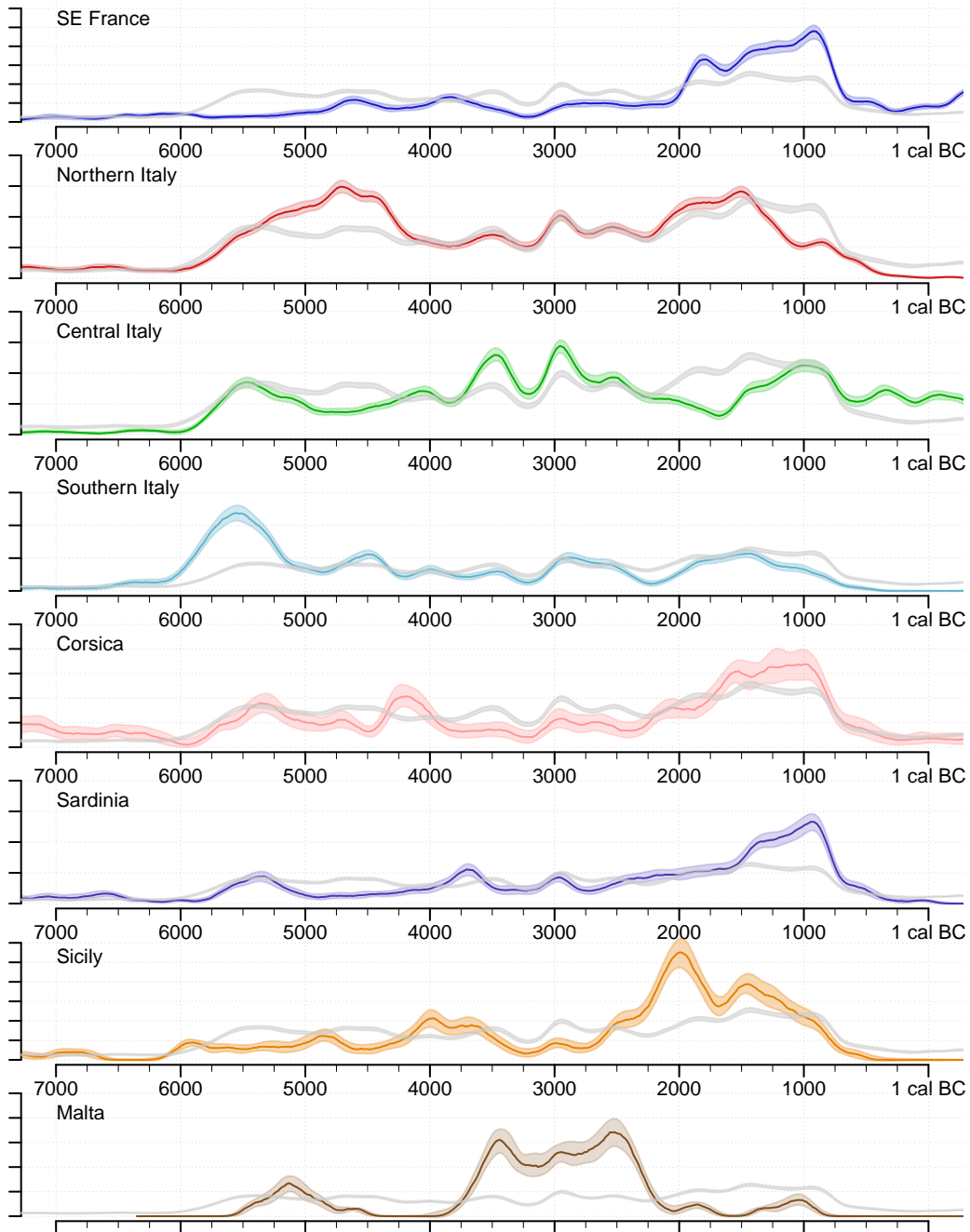To calculate an overall density model for the entire region:

```
all<-rbind(id[,c('Site','BP','SD')],mt[,c('Site','BP','SD')])
cmed<-phasedensity(all[all$SD>300,], bw=75)
```

This allows for Figure 4 in the main paper to be generated:

```
# Make a helper function to call all the custom plot parameters
PLT<-function(X,xlim=c(-7000,0), ylim, c1,c2,legend, at=seq(strt,0,1000),
              labels=c(seq(7000,1000,-1000),"1 cal BC"),lwd=1, ylabs=NA){
  plot(X,xlim=xlim,ylim=ylim,frame.plot=FALSE,xaxt='n',yaxt='n',xlab=NA,col=c1)
    lines(median(X),col=c2,lwd=lwd)
    legend("topleft",pch=NA,bty='n',legend=legend)
    strt<-round(xlim[1]/1000)*1000; stp<-round(xlim[2]/1000)*1000
    axis(1, at=at, labels=labels)
    rug(x=seq(xlim[1],xlim[2],250), side=1, ticksize=-0.05)
    axis(2, labels=ylabs)
    plot(cmed, add=T, col=rgb(0.8,0.8,0.8,0.5))
}


# Setup plot
par(mfrow=c(8,1))
par(mar=c(.5,2,.5,2))
par(mgp=c(3,.5,0))

# Plot density models
PLT(FRp,ylim=c(0,0.00065),c1=rgb(.1,.1,.8,.2),c2=rgb(.1,.1,.8,1),legend="SE France")
PLT(NIp,ylim=c(0,0.00040),c1=rgb(.8,.1,.1,.2),c2=rgb(.8,.1,.1,1),legend="Northern Italy")
PLT(CIp,ylim=c(0,0.00040),c1=rgb(.0,.7,.0,.2),c2=rgb(.0,.7,.0,1),legend="Central Italy")
PLT(SIp,ylim=c(0,0.00075),c1=rgb(.4,.7,.8,.3),c2=rgb(.4,.7,.8,1),legend="Southern Italy")
PLT(COp,ylim=c(0,0.00050),c1=rgb( 1,.6,.6,.3),c2=rgb( 1,.6,.6,1),legend="Corsica")
PLT(SDp,ylim=c(0,0.00080),c1=rgb(.3,.2,.7,.2),c2=rgb(.3,.2,.7,1),legend="Sardinia")
PLT(SYp,ylim=c(0,0.00063),c1=rgb(.9,.5,.0,.3),c2=rgb(.9,.5,.0,1),legend="Sicily")
PLT(MTp,ylim=c(0,0.00100),c1=rgb(.5,.3,.1,.2),c2=rgb(.5,.3,.1,1),legend="Malta")
```

To see whether the date thinning had any effect on the results, the models can be compared. This shows little significant difference between the resulting plots, the only case being Malta, where one site, the Xaghra hypogeum, causes a spuriously high peak due to the large number of available determinations (see Malone et al 2019). For visual simplicity, the following plot just considers the median of the density models:

```
# Set up plot window
par(mfrow=c(8,1))
par(mar=c(.5,2,.5,2))
par(mgp=c(3,.5,0))

# Helper function for this plot
```

```r
PLT2<-function(A, B, col=1, lty1=1, lty2=3, lwd=2, xlim=c(-7000,0), ylim=NULL, legend='',
               at=seq(strt,0,1000),labels=c(seq(7000,1000,-1000),"1 cal BC")) {
   plot(median(A), type='l',lwd=lwd, lty=lty1,xlim=xlim,ylim=ylim,frame.plot=FALSE,
               xaxt='n',yaxt='n',xlab=NA,col=col)
   lines(median(B),col=col,lwd=lwd,lty=lty2)
   legend("topleft",pch=NA,bty='n',legend=legend)
   strt<-round(xlim[1]/1000)*1000; stp<-round(xlim[2]/1000)*1000
   axis(1, at=at, labels=labels)
   rug(x=seq(xlim[1],xlim[2],250), side=1, ticksize=-0.05);axis(2, labels=NA)
}

# Plot the densitity models
PLT2(FR,FRp,ylim=c(0,0.00065),col=rgb(.1,.1,.8,1),legend="SE France")
PLT2(NI,NIp,ylim=c(0,0.00040),col=rgb(.8,.1,.1,1),legend="Northern Italy")
PLT2(CI,CIp,ylim=c(0,0.00040),col=rgb(.0,.7,.0,1),legend="Central Italy")
PLT2(SI,SIp,ylim=c(0,0.00075),col=rgb(.4,.7,.8,1),legend="Southern Italy")
PLT2(CO,COp,ylim=c(0,0.00050),col=rgb( 1,.6,.6,1),legend="Corsica")
PLT2(SD,SDp,ylim=c(0,0.00080),col=rgb(.3,.2,.7,1),legend="Sardinia")
PLT2(SY,SYp,ylim=c(0,0.00063),col=rgb(.9,.5,.0,1),legend="Sicily")
PLT2(MT,MTp,ylim=c(0,0.00100),col=rgb(.5,.3,.1,1),legend="Malta")
```
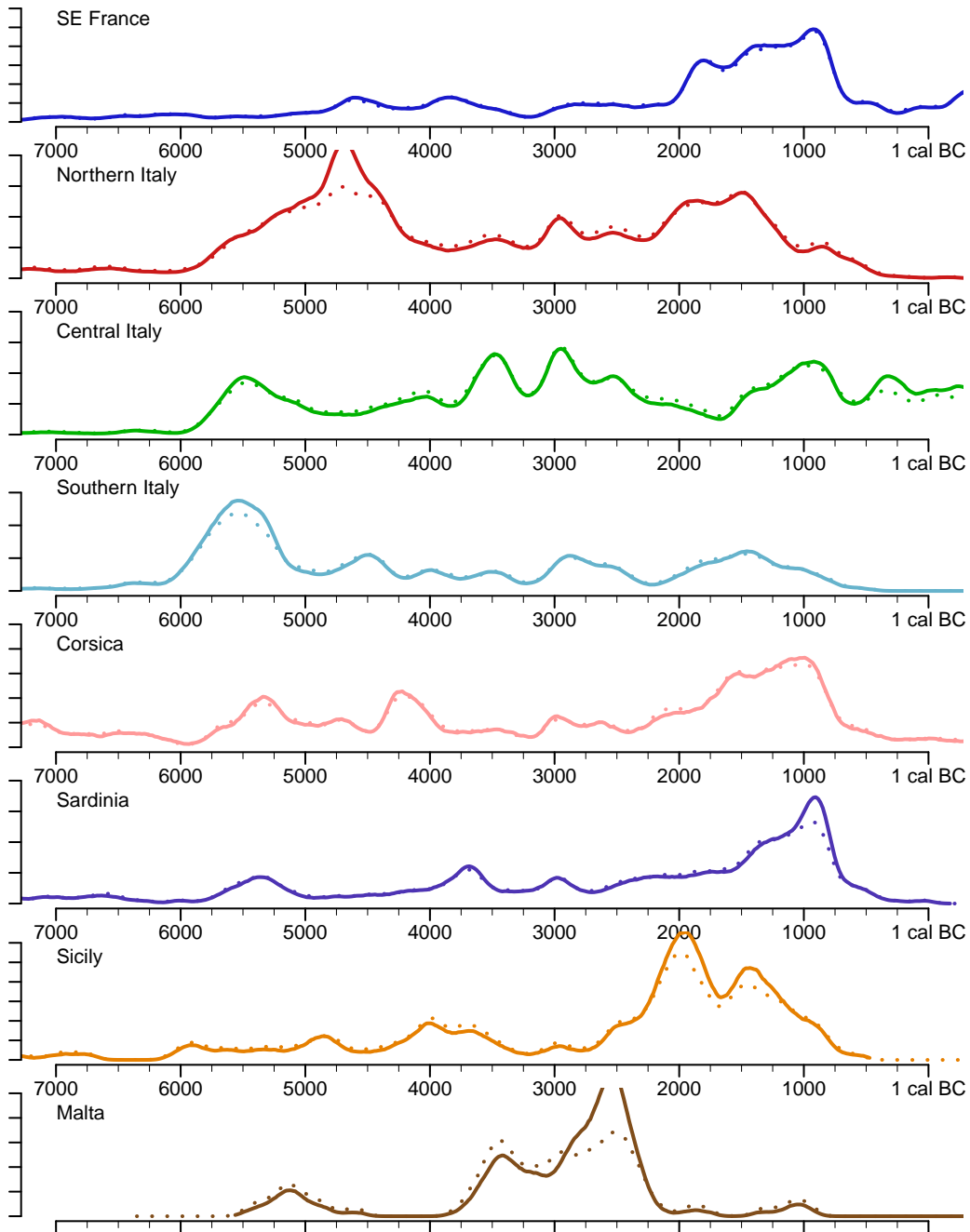
Examining the Mesolithic in detail

A detailed examination of trends in the Mesolithic period was not included in the main paper for reasons of space, but we can present it here using the following code to 'zoom' into the region of interest.
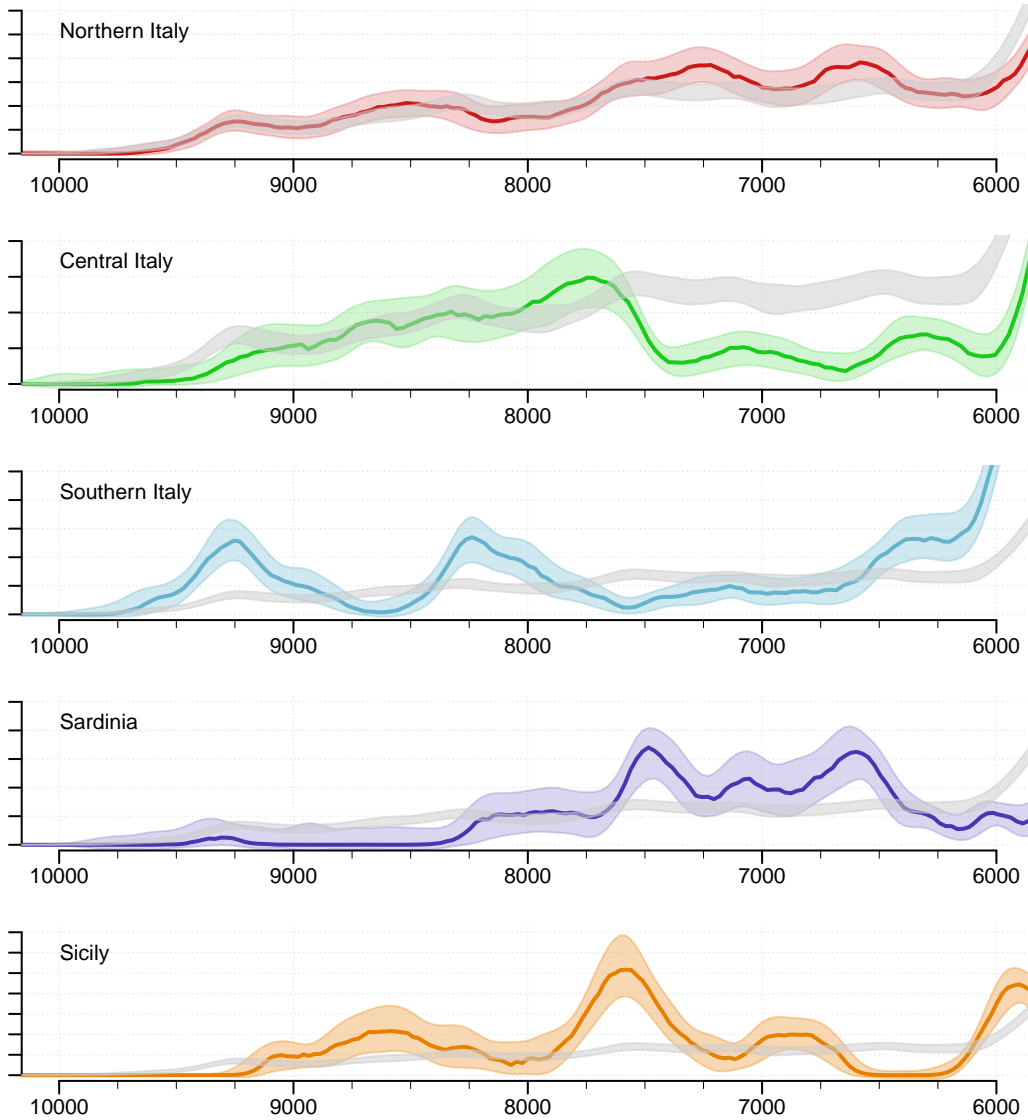
```
par(mfrow=c(5,1))
par(mar=c(2.5,2,.5,2))
par(mgp=c(3,.5,0))

ys<-seq(-10000,-6000,1000)
```

```
PLT(NIp,c1=rgb(.8,.1,.1,.2),c2=rgb(.8,.1,.1,1),xlim=c(-10000,-6000),ylim=c(0,0.00006),
    at=ys,labels=-ys,lwd=2,legend="Northern Italy")
PLT(CIp,c1=rgb(.1,.8,.1,.2),c2=rgb(.1,.8,.1,1),xlim=c(-10000,-6000),ylim=c(0,0.00004),
    at=ys,labels=-ys,lwd=2,legend="Central Italy")
PLT(SIp,c1=rgb(.4,.7,.8,.3),c2=rgb(.4,.7,.8,1),xlim=c(-10000,-6000),ylim=c(0,0.0001),
    at=ys,labels=-ys,lwd=2,legend="Southern Italy")
PLT(SDp,c1=rgb(.3,.2,.7,.2),c2=rgb(.3,.2,.7,1),xlim=c(-10000,-6000),ylim=c(0,0.0001),
    at=ys,labels=-ys,lwd=2,legend="Sardinia")
PLT(SYp,c1=rgb(.9,.5,.0,.3),c2=rgb(.9,.5,.0,1),xlim=c(-10000,-6000),ylim=c(0,0.00014),
    at=ys,labels=-ys,lwd=2,legend="Sicily")
```


The results suggest the Mesolithic period in Northern and Central Italy can be characterized by sustained, gradual growth, with fluctuating activity elsewhere. Although activity in Northern Italy also fluctuated somewhat, in general there was century-upon-century growth, although not anywhere near the growth observed during the Early Neolithic. In Central Italy there was a particularly strong monotonic growth trend until 7700 cal BC, at which point there was something of a population crash. A Late Mesolithic lull ensued in Central Italy until the introduction of agriculture. In Southern Italy, Sardinia and Sicily, fluctuating

activity over the millennia was the norm. An interesting peak occurred around 7600 cal BC in Sicily and a century later in Sardinia; whereas in Southern Italy the period is one of pronounced low activity. Certainly, it seems like the period 7700 to 7500 cal BC was an interesting and dynamic time across the region. The underlying explanation for this is illusive. There certainly is no simple environmental explanation; the most pronounced climate event in this period was a phase of cooling around 6200 cal BC; this seemingly had little effect but may have been a motivating factor behind the Neolithic colonization that began around this time from Southern Italy.
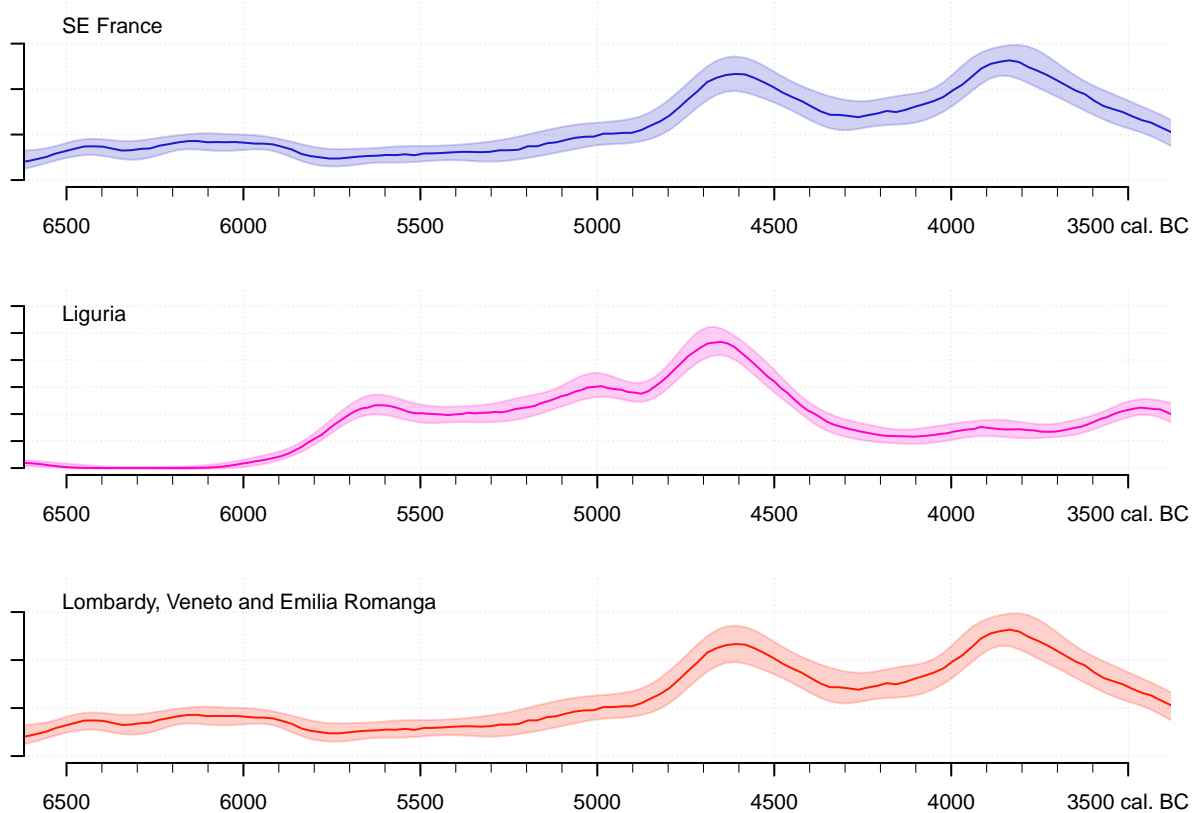
Sub-regional patterns

To compare sub-regions, for example in this case regions in the interior of the Alpine arc, we can sub-divide the dataset using the 'locality' column in the input data:

```r
lig<-phasedensity(id[id$Locality=='Liguria',c(15,4,5)], bw=75)
lomb_names<-c('Lombardia','Lombardy','Veneto','Emilia-Romagna','Emilia Romagna')
lomb<-phasedensity(id[id$Locality %in% lomb_names,c(15,4,5)], bw=75)


PLTr<-function(X,xlim=c(-6500,-3500), ylim=NULL, c1=rgb(.9,.9,.9,1),c2=rgb(0,0,0,1),legend,
            at=seq(-6500,-3500,500),labels=c(seq(6500,4000,-500),'3500 cal. BC'),lwd=1,ylabs=NA){
  plot(X,xlim=xlim,ylim=ylim,frame.plot=FALSE,xaxt='n',yaxt='n',xlab=NA,col=c1)
    lines(median(X),col=c2,lwd=lwd)
    legend("topleft",pch=NA,bty='n',legend=legend)
    axis(1, at=at, labels=labels)
    axis(2, labels=ylabs)
    rug(x=seq(xlim[1],xlim[2],100), side=1, ticksize=-0.05)
}

par(mfrow=c(3,1))
par(mar=c(3,3,1,1))
PLTr(FRp,ylim=c(0,0.00019),c1=rgb(.1,.1,.8,.2),c2=rgb(.1,.1,.8,1),legend="SE France")
PLTr(lig,ylim=c(0,0.00064),c1=rgb(1,.0,.8,.2),c2=rgb(1,.0,.8,1),legend="Liguria")
PLTr(FRp,ylim=c(0,0.00018),c1=rgb(1,.1,.0,.2),c2=rgb(1,.1,.0,1),
      legend="Lombardy, Veneto and Emilia Romanga")
```

Evaluating KDE predictive power using multiple permutations

It is possible to use cross validation methods (random sampling with replacement) to subset the data into random chunks. This is intended to express, via this process, an extra confidence envelope caused by the stochasticity of the data itself. The following code calculates density models with half of the input data, and repeats this process eight times, averaging the results. Ideally the process should be ran with many more iterations, but it is computationally expensive to do so.

```
NI2p<-phasedensity(id[id$Region=='Northern Italy',c(15,4,5)],
                   bw=75,perm_size=.5,perm_runs=8)
CI2p<-phasedensity(id[id$Region=='Central Italy', c(15,4,5)],
                   bw=75,perm_size=.5,perm_runs=8)
SI2p<-phasedensity(id[id$Region=='Southern Italy',c(15,4,5)],
                   bw=75,perm_size=.5,perm_runs=8)
SY2p<-phasedensity(id[id$Region=='Sicily',        c(15,4,5)],
                   bw=75,perm_size=.5,perm_runs=8)
SD2p<-phasedensity(id[id$Region=='Sardinia',      c(15,4,5)],
                   bw=75,perm_size=.5,perm_runs=8)
FR2p<-phasedensity(id[id$Region=='SE France',     c(15,4,5)],
                   bw=75,perm_size=.5,perm_runs=8)
CO2p<-phasedensity(id[id$Region=='Corsica',       c(15,4,5)],
                   bw=75,perm_size=.5,perm_runs=8)
MT2p<-phasedensity(mt[,c(4,2,3)],bw=75,perm_size=.5,perm_runs=8)
```

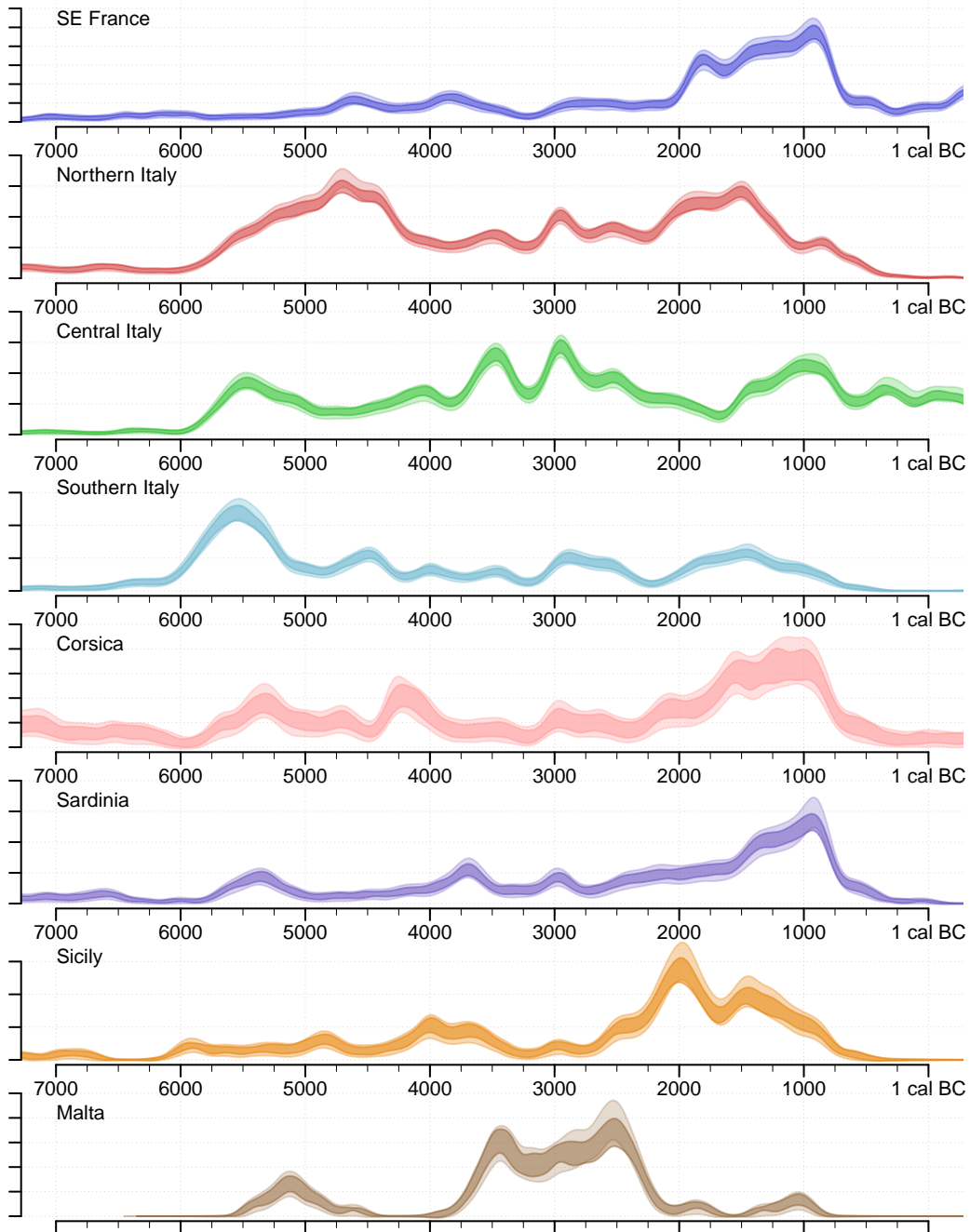The models may be compared to those generated using the full dataset using the following code:

```r
par(mfrow=c(8,1))
par(mar=c(.5,2,.5,2))
par(mgp=c(3,.5,0))

PLT3<-function(A,B,xlim=c(-7000,0), ylim, c1,c2,legend, at=seq(strt,0,1000),
               labels=c(seq(7000,1000,-1000),"1 cal BC"),lwd=1){
  plot(A,xlim=xlim,ylim=ylim,frame.plot=FALSE,xaxt='n',yaxt='n',xlab=NA,col=c1)
  plot(B,col=c2,add=TRUE)
   legend("topleft",pch=NA,bty='n',legend=legend)
   strt<-round(xlim[1]/1000)*1000; stp<-round(xlim[2]/1000)*1000
   axis(1, at=at, labels=labels)
   rug(x=seq(xlim[1],xlim[2],250), side=1, ticksize=-0.05);axis(2, labels=NA)
}

PLT3(FR2p,FRp,ylim=c(0,0.00065),c1=rgb(.1,.1,.8,.2),c2=rgb(.1,.1,.8,.4),
     legend="SE France")
PLT3(NI2p,NIp,ylim=c(0,0.00040),c1=rgb(.8,.1,.1,.2),c2=rgb(.8,.1,.1,.4),
     legend="Northern Italy")
PLT3(CI2p,CIp,ylim=c(0,0.00040),c1=rgb(.0,.7,.0,.2),c2=rgb(.0,.7,.0,.4),
     legend="Central Italy")
PLT3(SI2p,SIp,ylim=c(0,0.00075),c1=rgb(.4,.7,.8,.3),c2=rgb(.4,.7,.8,.5),
     legend="Southern Italy")
PLT3(CO2p,COp,ylim=c(0,0.00050),c1=rgb( 1,.6,.6,.3),c2=rgb( 1,.6,.6,.5),
     legend="Corsica")
PLT3(SD2p,SDp,ylim=c(0,0.00080),c1=rgb(.3,.2,.7,.2),c2=rgb(.3,.2,.7,.4),
     legend="Sardinia")
PLT3(SY2p,SYp,ylim=c(0,0.00075),c1=rgb(.9,.5,.0,.3),c2=rgb(.9,.5,.0,.5),
     legend="Sicily")
PLT3(MT2p,MTp,ylim=c(0,0.00100),c1=rgb(.5,.3,.1,.2),c2=rgb(.5,.3,.1,.4),
     legend="Malta")
```

The darker-colored part of the ribbons in the above plot represents the span of the confidence interval of the density model caused by measurement error and calibration; the paler area represents a combination of measurement error, calibration uncertainty and the sampling distribution (i.e. the error rate caused by the randomness in excavation and choice of samples for dating). There is however little significant extra uncertainty, although regions with smaller quantities of dates, such as Corsica and Malta, have larger confidence intervals and therefore the density models from those regions should treated with some extra caution. However, for the larger regions, this exercise demonstrates how well-powered the data are.
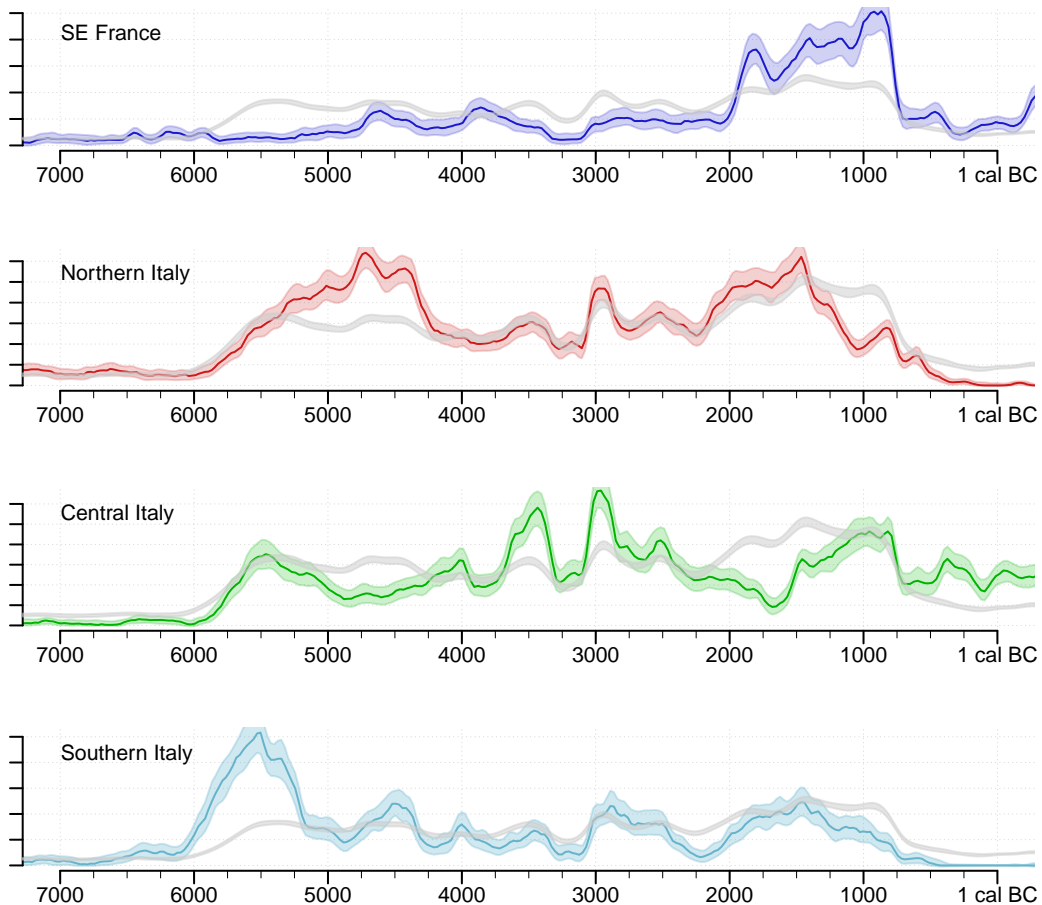
Examining short-term trends

We can calculate density models using a smaller, default bandwidth of 30 years (one generational interval) to examine short-term trends. Due to the limited power of data from the island regions, it is only worth doing this for the continental regions:

```
FR30<-phasedensity(id[id$Region=='SE France',c(15,4,5)])
NI30<-phasedensity(id[id$Region=='Northern Italy',c(15,4,5)])
CI30<-phasedensity(id[id$Region=='Central Italy', c(15,4,5)])
SI30<-phasedensity(id[id$Region=='Southern Italy',c(15,4,5)])
# Sadrinia not used in this plot, but in a plot further below
SD30<-phasedensity(id[id$Region=='Sardinia',c(15,4,5)])
```

The results can be plotted as per Figure 2:

```
par(mfrow=c(4,1))
par(mar=c(3.6,2,.2,2))
par(mgp=c(3,.45,0))
PLT(FR30, ylim=NULL, c1=rgb(.1,.1,.8,.2),c2=rgb(.1,.1,.8,1),legend="SE France")
PLT(NI30, ylim=NULL, c1=rgb(.8,.1,.1,.2),c2=rgb(.8,.1,.1,1),legend="Northern Italy")
PLT(CI30, ylim=NULL, c1=rgb(.0,.7,.0,.2),c2=rgb(.0,.7,.0,1),legend="Central Italy")
PLT(SI30, ylim=NULL, c1=rgb(.4,.7,.8,.3),c2=rgb(.4,.7,.8,1),legend="Southern Italy")
```

These models contain relatively more uncertainty, more noise and there are more difficult to 'read'. However, they nonetheless demonstrate the very rapid nature of the downturn around 800 BC in Southeast France and Central Italy and the stark contrast between Northern and Central Italy around 1800 BC.

Animated map

The animated map, also included with this paper as supplementary information, was plotted using the following code, which requires the R package `maptools` and a coastline map obtained from https://www.naturalearthdata.com:

```
require(maptools)
# Prepare the two datelists so that their column names are the same
mts<-mt[,c(1,2,3,9,8)]
ids<-id[,c(1,4:7)]
colnames(mts)<-colnames(ids)
# Load map (any coastline vector, the following is an exmaple):
earth<-readShapeSpatial("NaturalEarthCoastline/very_simplified.shp")
# Calibrate dates and store results in a large matrix
st<-time_matrix(na.omit(rbind(mts,ids)))

# The following parameters control the size and colour of the dots
make_animated_map(coast=earth,stmap=st,out='Med_animated_C14map.pdf',pt.cex=1,pt.
        col=rgb(1,0,0,0.4),threshold=0.002,width=6,height=6,xlim=c(3,19), ylim=c(37,46))
```

Comparing the Central Mediterranean to other regions

Other datasets can be downloaded using the `c14bazAAR` package Schmid et al 2018. In the following analysis we compare the Early Neolithic of various regions. For Britain and Ireland, the following code uses the file `archdates.csv` obtained from the repository available at https://discovery.ucl.ac.uk/id/eprint/10025178/. This should be downloaded and saved in the working directory. Similarly, it also uses data from the file C14Samples.txt from the reposatory of data from Greece by Katsianis et al https://doi.org/10.5522/04/12489137.v1, which have been prepared for this analysis by removing dates non terrestroal samples and other labeled as problematic in Katsianis's database. These data have been saved in a file 'Greece.csv' which readers should prepare in a similar way to above. Due to the variable provenance of some of these data, this analysis uses the whole available databases, without thinning to one date per phase. First, we download the data and calculate the density models:

```
bi<-read.csv('/Bevan_et_al_2017/dates/archdates.csv')
greece<-read.csv('Katsianis_et_al_2020/Greece.csv')
ire <-bi[bi$Region=='Ireland' & bi$CRA>1800 & bi$CRA<10000,c('SiteName','CRA','Error')]
brit<-bi[bi$Region!='Ireland' & bi$CRA>1800 & bi$CRA<10000,c('SiteName','CRA','Error')]
require(c14bazAAR)
radon<-get_radon()
radon<-as.data.frame(radon)

# Subset the RADON dataset
EAdr<-c("Croatia","Bosnia and Herzegovina","Serbia and Montenegro","Serbia","Kosovo",
        "Macedonia","former Yugoslavia","Albania")

adri<-radon[radon$country %in% EAdr,]
ch<-radon[which(radon$country=='Switzerland'),]
greece<-radon[radon$country=='Greece',]
```

```
iberia<-radon[radon$country %in% c('Spain','Portugal'),]
germany<-radon[radon$country=='Germany',]
scandan<-radon[radon$country %in% c('Denmark','Sweden'),]

# Calculate the density models (some data cleaning is required, omitting NAs etc)
chD<-MCdensity(na.omit(ch[ch$c14age<11000 & ch$c14std>2,4:5]),bw=75)
adriD<-MCdensity(adri[,4:5],bw=75)
greeceD<-MCdensity(na.omit(greece),bw=75)
iberiaD<-MCdensity(na.omit(iberia[iberia$c14std>1,4:5]),bw=75)
germanyD<-MCdensity(na.omit(germany[germany$c14std>1,4:5]),bw=75)
scandanD<-MCdensity(na.omit(scandan[scandan$c14std>1,4:5]),bw=75)
BrD<-MCdensity(brit[,2:3], bw=75)
IrD<-MCdensity(ire[,2:3], bw=75)
```
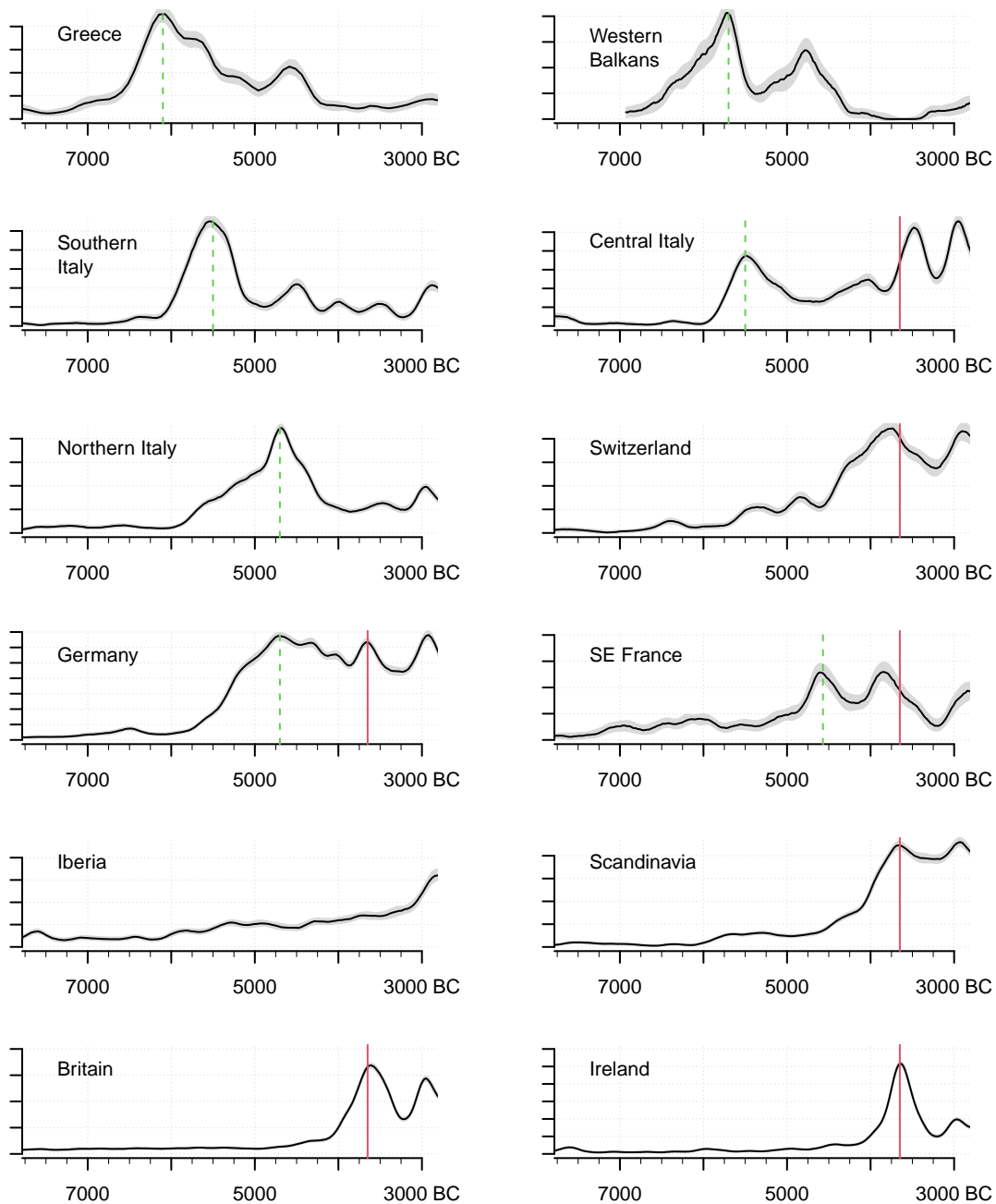
The following code reproduces the figure from the main paper:

```
PLT4<-function(X,xlim=c(-7600,-3000), ylim=NULL, c1=rgb(.85,.85,.85,1),c2=rgb(0,0,0,1),
               legend, at=seq(-8000,-3000,1000), labels=c(NA,7000,NA,5000,NA,'3000 BC'),
               lwd=1, ml=NA, l37=TRUE, ylabs=NA){
  plot(X,xlim=xlim,ylim=ylim,frame.plot=FALSE,xaxt='n',yaxt='n',xlab=NA,col=c1)
    lines(median(X),col=c2,lwd=lwd)
    legend("topleft",pch=NA,bty='n',legend=legend)
    strt<-round(xlim[1]/1000)*1000; stp<-round(xlim[2]/1000)*1000
    axis(1, at=at, labels=labels)
    rug(x=seq(-7750,-3250,250), side=1, ticksize=-0.05)
    if(l37) abline(v=-3650, col=2)
    abline(v=ml, col=3, lty=2)
    axis(2, labels=ylabs)
}


#pdf('Figure_5.pdf', width=6, height=6.5)
par(mfrow=c(6,2))
par(mar=c(2,3,2,2))
PLT4(greeceD, legend='Greece', ml=-6100, l37=FALSE)
PLT4(adriD,   legend='Western\nBalkans', ml=-5700, l37=FALSE)
PLT4(SI,      legend='Southern\nItaly',ml=-5500, l37=FALSE)
PLT4(CI,      legend='Central Italy',ml=-5500)
PLT4(NI,      legend='Northern Italy',ml=-4700, l37=FALSE)
PLT4(chD,     legend='Switzerland')
PLT4(germanyD,legend='Germany',ml=-4700)
PLT4(FR,      legend='SE France',ml=-4570,ylim=c(0,0.0002))
PLT4(iberiaD, legend='Iberia', l37=FALSE)
PLT4(scandanD,legend='Scandinavia')
PLT4(BrD,     legend='Britain',ylim=c(0,0.0002))
PLT4(IrD,     legend='Ireland',ylim=c(0,0.0003))
```

```
#dev.off()
```

Further data can be downloaded for the Bronze Age and processed in a similar way:

```
# Same for LBA, from RADON-B database

RB<-get_radonb(db_url="https://radon-b.ufg.uni-kiel.de/radondaily.txt")

ch<-as.matrix(RB[RB$country=='Switzerland',4:5])
ch<-ch[!is.na(ch[,1]) & !is.na(ch[,2]),]
```

```
ger<-as.data.frame(RB[RB$country=='Germany',4:5])
spain<-as.data.frame(RB[RB$country=='Spain',4:5])
gree<-as.data.frame(RB[RB$country=='Greece',4:5])
scandan2<-as.data.frame(RB[RB$country %in% c('Denmark','Sweden','Norway'),4:5])
adri2<-RB[RB$country %in% EAdr,4:5]

# Compute new density models with 30-year bandwidths
# (some data cleaning also performed here)
IrD30<-MCdensity(ire[,2:3])
BrD30<-MCdensity(brit[,2:3])

  chBA <-    MCdensity(ch)
adriBA <-    MCdensity(na.omit(adri2[adri2$c14std>1,]))
chBA <-      MCdensity(ch)
greeBA <-    MCdensity(na.omit(gree[gree$c14std>1,]))
gerBA <-     MCdensity(na.omit(ger[ger$c14std>1,]))
spainBA <-  MCdensity(na.omit(spain[spain$c14std>1 & spain$c14age<4000,]))
scandanBA <-MCdensity(na.omit(scandan2[scandan2$c14std>1,]))


#newversion

PLT5<-function(X,xlim=c(-1200,-400), ylim=NULL, c1=rgb(.9,.9,.9,1),c2=rgb(0,0,0,1),legend,
          at=seq(-1200,-400,200),labels=c(seq(1200,600,-200),'400 BC'),lwd=1,ylabs=NA){
  plot(X,xlim=xlim,ylim=ylim,frame.plot=FALSE,xaxt='n',yaxt='n',xlab=NA,col=c1)
   lines(median(X),col=c2,lwd=lwd)
   legend("topright",pch=NA,bty='n',legend=legend)
   strt<-round(xlim[1]/1000)*1000; stp<-round(xlim[2]/1000)*1000
   axis(1, at=at, labels=labels)
   rug(x=seq(xlim[1],xlim[2],50), side=1, ticksize=-0.05)
   abline(v=-800, col=2)
   axis(2, labels=ylabs)
}



#pdf('Figure_8.pdf', width=6, height=6.5)

par(mfrow=c(6,2))
par(mar=c(2,3,2,2))
PLT5(NI30, ylim=c(0,0.0002),   legend='Northern Italy')
PLT5(CO,   ylim=c(0,0.0004),   legend='Corsica')
PLT5(CI30, ylim=c(0,0.0003),   legend='Central Italy')
PLT5(SD30, ylim=c(0,0.001),    legend='Sardinia')
PLT5(SI30, ylim=c(0,0.0002),   legend='Southern Italy')
PLT5(chBA, ylim=c(0,0.00075),  legend='Switzerland')
PLT5(spainBA,ylim=c(0,.00075), legend='Iberia')
PLT5(FR30, ylim=c(0,0.0007),   legend='SE France')
PLT5(gerBA,ylim=c(0,0.0004),   legend='Germany')
PLT5(scandanBA,ylim=c(0,.0007),legend='Scandinavia')
PLT5(BrD30,ylim=c(0,0.0003),   legend='Britain   ')
PLT5(IrD30,ylim=c(0,0.0003),   legend='Ireland')
```
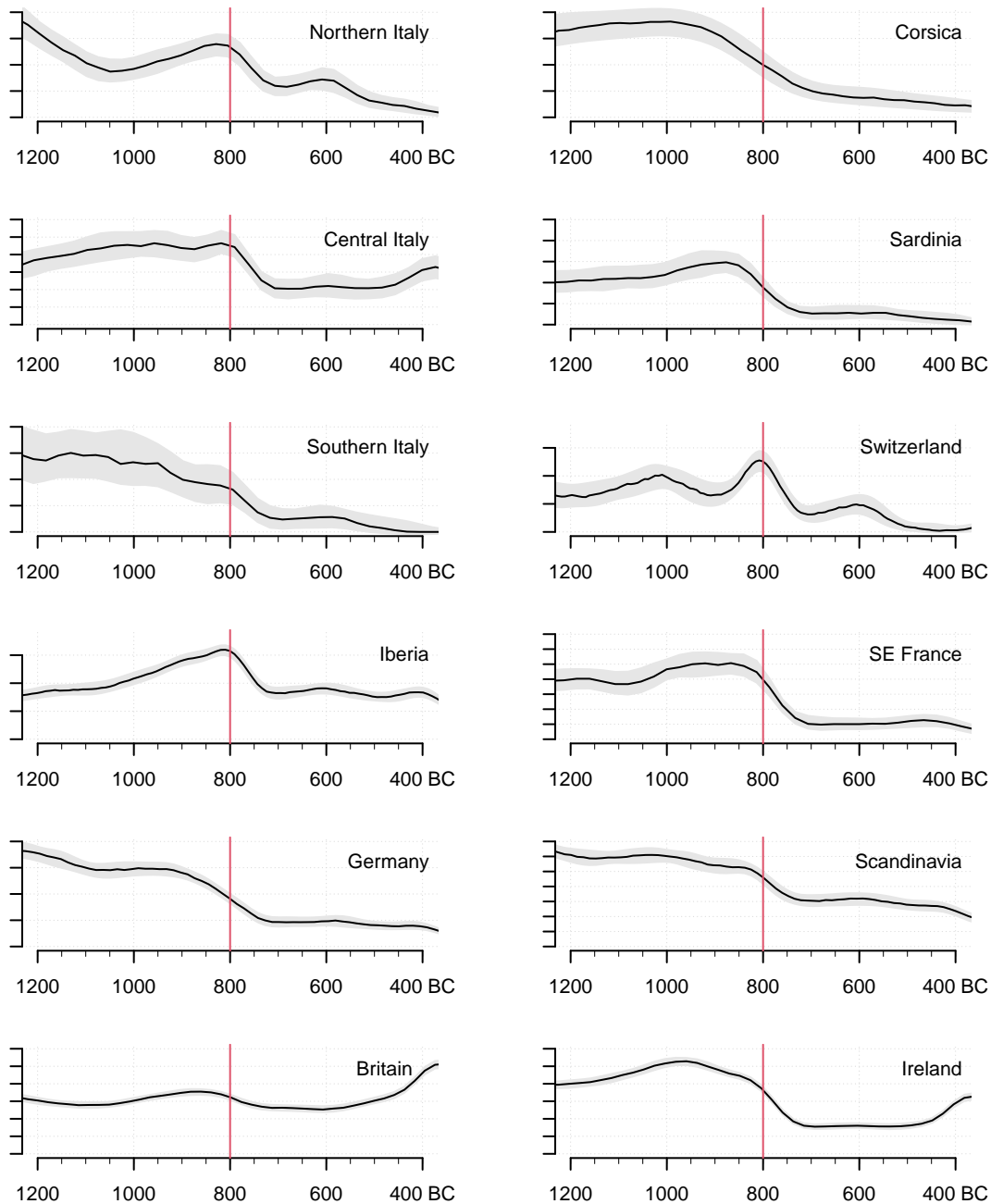
```
#dev.off()
```

The data from the study region can be compared to other regions, using a density model for the entire region.

```
# Combine tables for the central med
all<-rbind(id[,c('Site','BP','SD')],mt[,c('Site','BP','SD')])
# Obtain African data from INTERNET
MandT2014<-url('https://ars.els-cdn.com/content/image/1-s2.0-S0277379114002728-mmc6.csv')
af<-read.csv(MandT2014)[,c(1,5,6)]
```

```
af<-af[af[,2]<10000 & af[,2]>1800,]

# Combine RADON and RADON-B data for Iberia and do some data sle
iberia1<-radon[radon$country %in% c('Spain','Portugal'),]
iberia2<-as.data.frame(RB[RB$country%in% c('Spain','Portugal'),])
iball<-rbind(iberia1,iberia2)
idates<-iball[!(iball$material %in% c('shell','peat')),c('site','c14age','c14std')]
idates<-na.omit(idates[idates$c14age<10000 & idates$c14age>1800 & idates$c14std>0,])

# Calculate density models
africa<-phasedensity(af, bw=75)
IbD<-phasedensity(idates,bw=75)
```
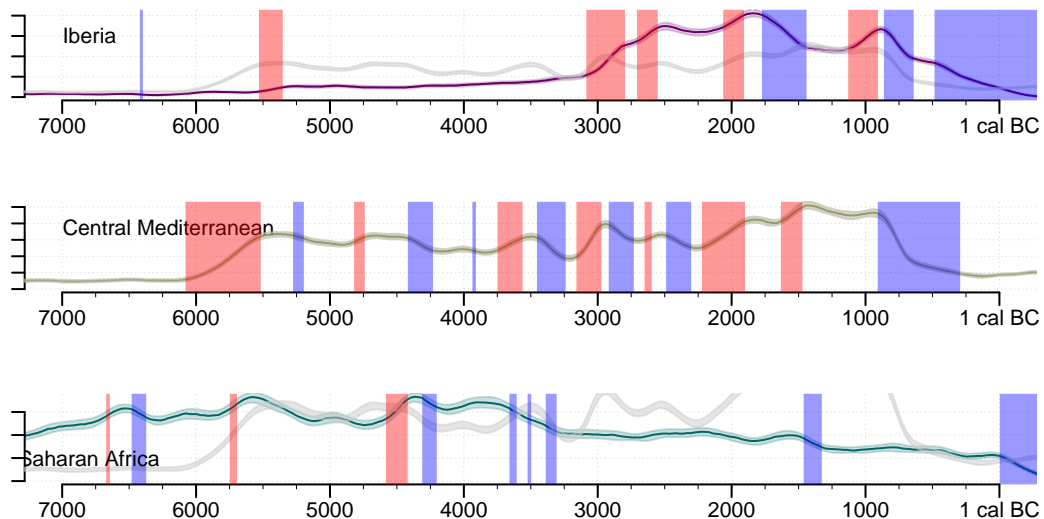
The following plot compares the two, highlighing periods where the growth rate was significantly high or low (see section on growth rate calculation below):

```
# Plot the results
par(mfrow=c(3,1))
par(mar=c(4,2,0,2))
par(mgp=c(3,.5,0))
PLT(IbD,  ylim=NULL,c1=rgb(.4,.0,.4,.2),c2=rgb(.4,.0,.4,1),legend="Iberia")
polygon(ggrsignif(ggr(IbD)))
PLT(cmed, ylim=NULL,c1=rgb(.4,.4,.0,.2),c2=rgb(.4,.4,.0,1),legend="Central Mediterranean")
polygon(ggrsignif(ggr(cmed)))
PLT(africa,ylim=NULL,c1=rgb(.0,.4,.4,.2),c2=rgb(.0,.4,.4,1),legend="")
polygon(ggrsignif(ggr(africa)))
text(-6800,0.00005,'Saharan Africa')
```


Models and simulations

Simulated radiocarbon datasets can be used to gain a perspective on how a model of past dynamics could be represented by density estimates of ensambles of radiocarbon dates. The process samples calendar dates,

weighted by a defined model of changing activity, 'back calibrates' the dates into simulated radiocarbon dates, and calculates the KDE for these.

First the models are defined:

```r
#set up date range
rng<- -1350:-250
#steady state
smod<-list(x=rng, y=rep(1,1101))
#drastic decline
dmod<-list(x=rng, y=c(rep(1,550),rep(0.1,551)))
#gradual (300 year) decline
gmod<-list(x=rng, y=c(rep(1,400),seq(1,0.1,-0.9/300),rep(0.1,400)))
```

Then the scrips simulate a set of back-calibrated radiocabron dates and compute KDE for four versions of these:

```r
drastic1<-MCsimulate(mod=dmod, bw=75)
drastic2<-MCsimulate(mod=dmod, bw=75)
drastic3<-MCsimulate(mod=dmod, bw=75)
drastic4<-MCsimulate(mod=dmod, bw=75)

steady1<-MCsimulate(smod, bw=75)
steady2<-MCsimulate(smod, bw=75)
steady3<-MCsimulate(smod, bw=75)
steady4<-MCsimulate(smod, bw=75)

gradual1<-MCsimulate(gmod, bw=75)
gradual2<-MCsimulate(gmod, bw=75)
gradual3<-MCsimulate(gmod, bw=75)
gradual4<-MCsimulate(gmod, bw=75)
```

The results are plotted side-by-side as follows:

```r
PLTS<-function(X,xlim=c(-1200,-500), c1=rgb(.9,.9,.9,1),c2=rgb(0,0,0,1),legend='',
           at=seq(-1200,-500,100),labels=seq(1200,500,-100),lwd=1, ylabs=NA, ...){
  plot(X,xlim=xlim,frame.plot=FALSE,xaxt='n',yaxt='n',xlab='Cal. BC',col=c1, ...)
   lines(median(X),col=c2,lwd=lwd)
   legend("topleft",pch=NA,bty='n',legend=legend)
   strt<-round(xlim[1]/1000)*1000; stp<-round(xlim[2]/1000)*1000
   axis(1, at=at, labels=labels)
   axis(2, labels=ylabs)
}


par(mfrow=c(3,5))

plot(dmod, frame=FALSE, axes=FALSE,ylab='', type='l', ylim=c(-0.2,1),xlim=c(-1200,-500),
           xlab='Cal. BC')
axis(1, at=seq(-1200,-500,100),  labels=seq(1200,500,-100))
title(main='Drastic decline\nmodel')
PLTS(drastic1); title(main='Drastic decline\nSimulation 1')
PLTS(drastic2); title(main='Drastic decline\nSimulation 2')
PLTS(drastic3); title(main='Drastic decline\nSimulation 3')
```

```
PLTS(drastic4); title(main='Drastic decline\nSimulation 4')

plot(gmod, frame=FALSE, axes=FALSE,ylab='', type='l', ylim=c(-0.2,1),xlim=c(-1200,-500),
        xlab='Cal. BC')
axis(1, at=seq(-1200,-500,100),  labels=seq(1200,500,-100))
title(main='Gradual decline\nmodel')

PLTS(gradual1); title(main='Gradual decline\nSimulation 1')
PLTS(gradual2); title(main='Gradual decline\nSimulation 2')
PLTS(gradual3); title(main='Gradual decline\nSimulation 3')
PLTS(gradual4); title(main='Gradual decline\nSimulation 4')

plot(smod, frame=FALSE, axes=FALSE,ylab='', type='l', ylim=c(0,1.2),xlim=c(-1200,-500),
        xlab='Cal. BC')
axis(1, at=seq(-1200,-500,100),  labels=seq(1200,500,-100))
title(main='Steady state\nmodel')

PLTS(steady1) ; title(main='Steady state\nSimulation 1')
PLTS(steady2) ; title(main='Steady state\nSimulation 2')
PLTS(steady3) ; title(main='Steady state\nSimulation 3')
PLTS(steady4) ; title(main='Steady state\nSimulation 4')
```
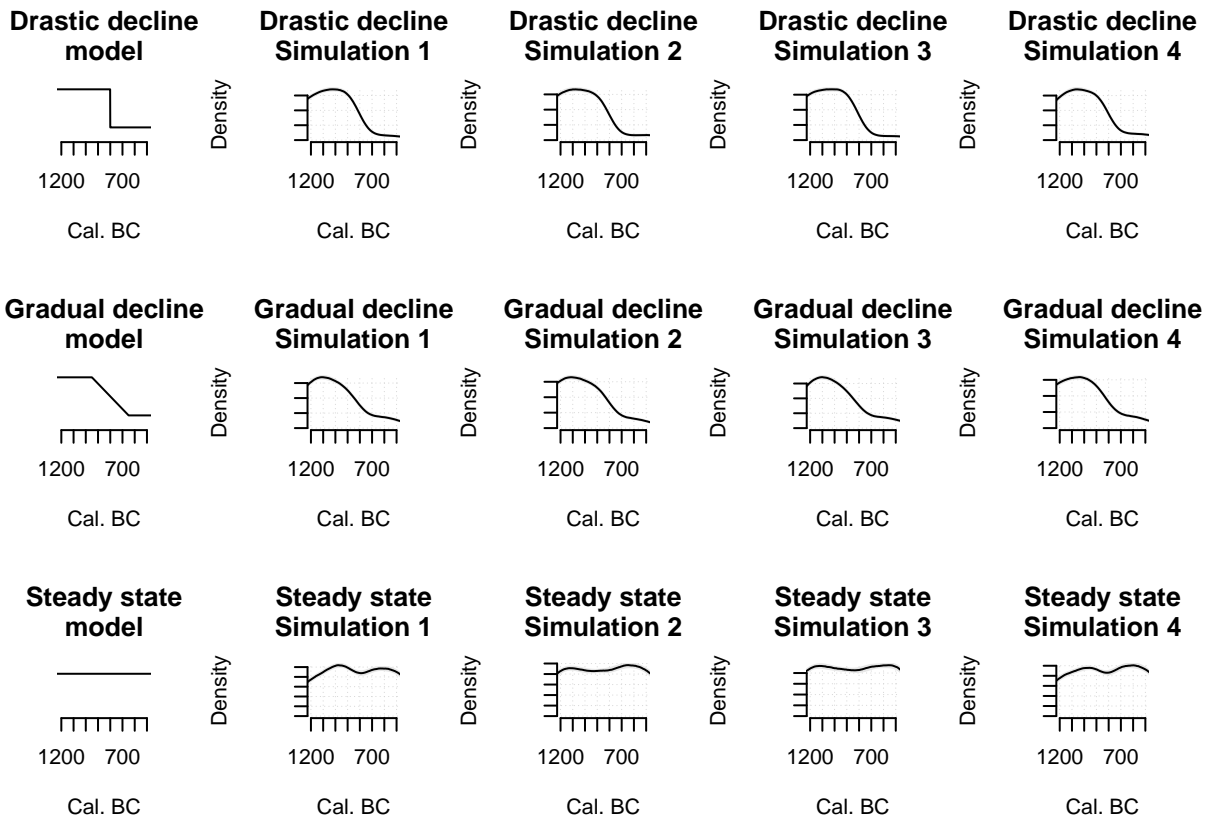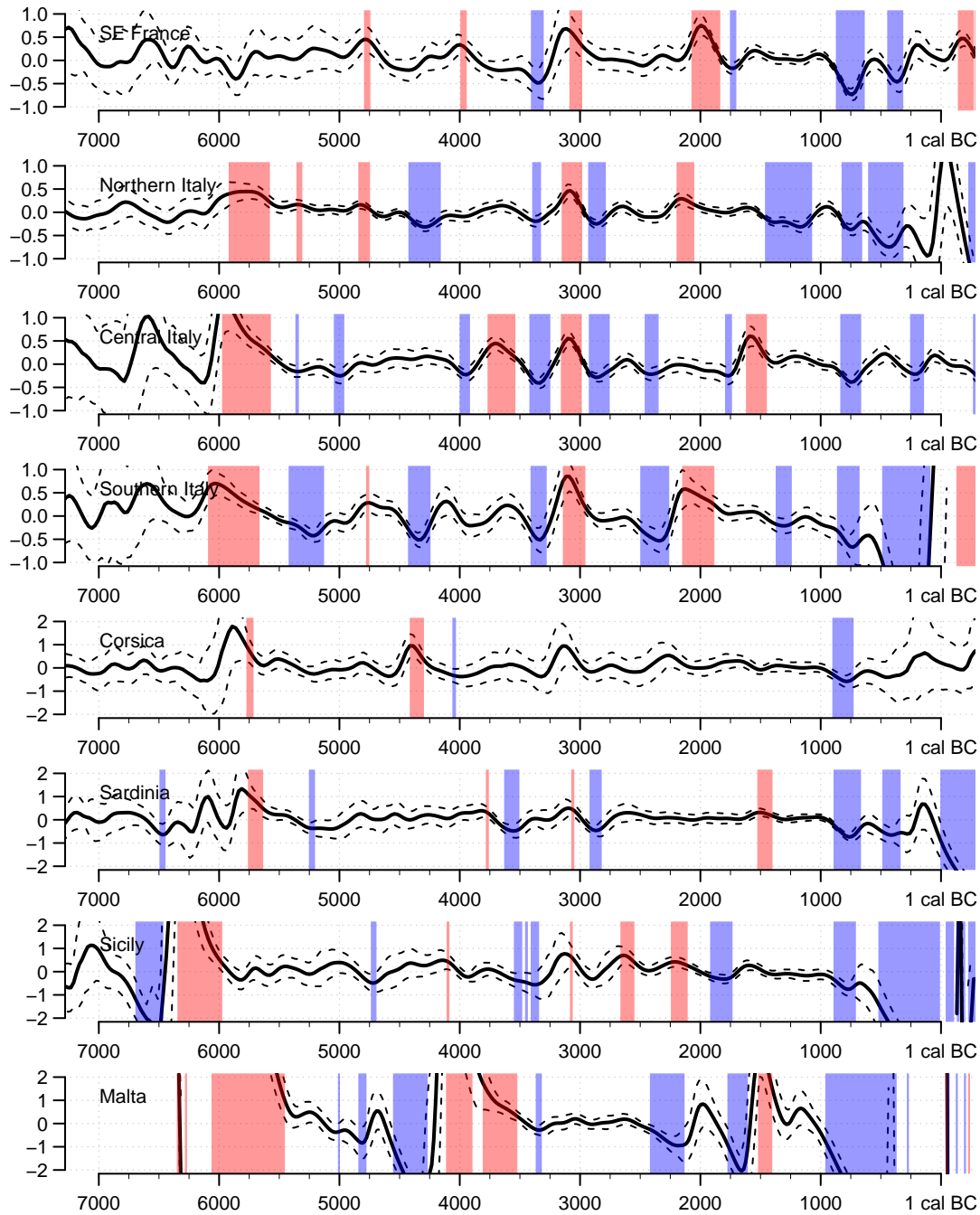
Growth rate calculations

Growth rates are obtained by differentiating the KDE models using the function `ggr`, with periods where growth was statistically significantly higher or lower than 0 highlighted using the function `ggrsignif`.

```r
# Helper function for plot
PLT6<-function(X,xlim=c(-7000,0), ylim=c(-2,2), col=1,legend, at=seq(strt,0,1000),
               labels=c(seq(7000,1000,-1000),"1 cal BC"),lwd=1) {
   plot(ggr(X),xlim=xlim,ylim=ylim,frame.plot=FALSE,xaxt='n',yaxt='n',xlab=NA,col=col,
        fill=NA)
   polygon(ggrsignif(ggr(X)))
   legend("topleft",pch=NA,bty='n',legend=legend)
   strt<-round(xlim[1]/1000)*1000; stp<-round(xlim[2]/1000)*1000
   axis(1, at=at, labels=labels)
   rug(x=seq(xlim[1],xlim[2],250), side=1, ticksize=-0.05)
   axis(2, las=1)
}
# Draw the figure
par(mfrow=c(8,1))
par(mar=c(2,3,.25,2))
par(mgp=c(3,.75,0))
 PLT6(FRp,legend="SE France", ylim=c(-1,1))
 PLT6(NIp,legend="Northern Italy", ylim=c(-1,1))
 PLT6(CIp,legend="Central Italy", ylim=c(-1,1))
 PLT6(SIp,legend="Southern Italy", ylim=c(-1,1))
 PLT6(COp,legend="Corsica")
 PLT6(SDp,legend="Sardinia")
 PLT6(SYp,legend="Sicily")
 PLT6(MTp,legend="Malta")
```

This plot illustrates the near-universal phase of growth at the start of the Neolithic, and the widespread decline / collapse in the early first millennium BC.

Investigating human burial

Dates in our database have been classified according to the kind of material that was submitted for radiocarbon measurement, so it is a simple matter to calculate density models for subsets of the data for each sample type. In this example we calculate density models for dates on human bone, and models for all the other samples in the database.

```
CIH<-MCdensity(id[which(id$Region=='Central Italy' & id$MCODE=='H'),4:5],bw=75)
CIO<-MCdensity(id[which(id$Region=='Central Italy' & id$MCODE!='H'),4:5],bw=75)

NIH<-MCdensity(id[which(id$Region=='Northern Italy' & id$MCODE=='H'),4:5],bw=75)
NIO<-MCdensity(id[which(id$Region=='Northern Italy' & id$MCODE!='H'),4:5],bw=75)

SIH<-MCdensity(id[which(id$Region=='Southern Italy' & id$MCODE=='H'),4:5],bw=75)
SIO<-MCdensity(id[which(id$Region=='Southern Italy' & id$MCODE!='H'),4:5],bw=75)
```

The resulting curves are independent from each other. The two sets of curves in each region can be compared:
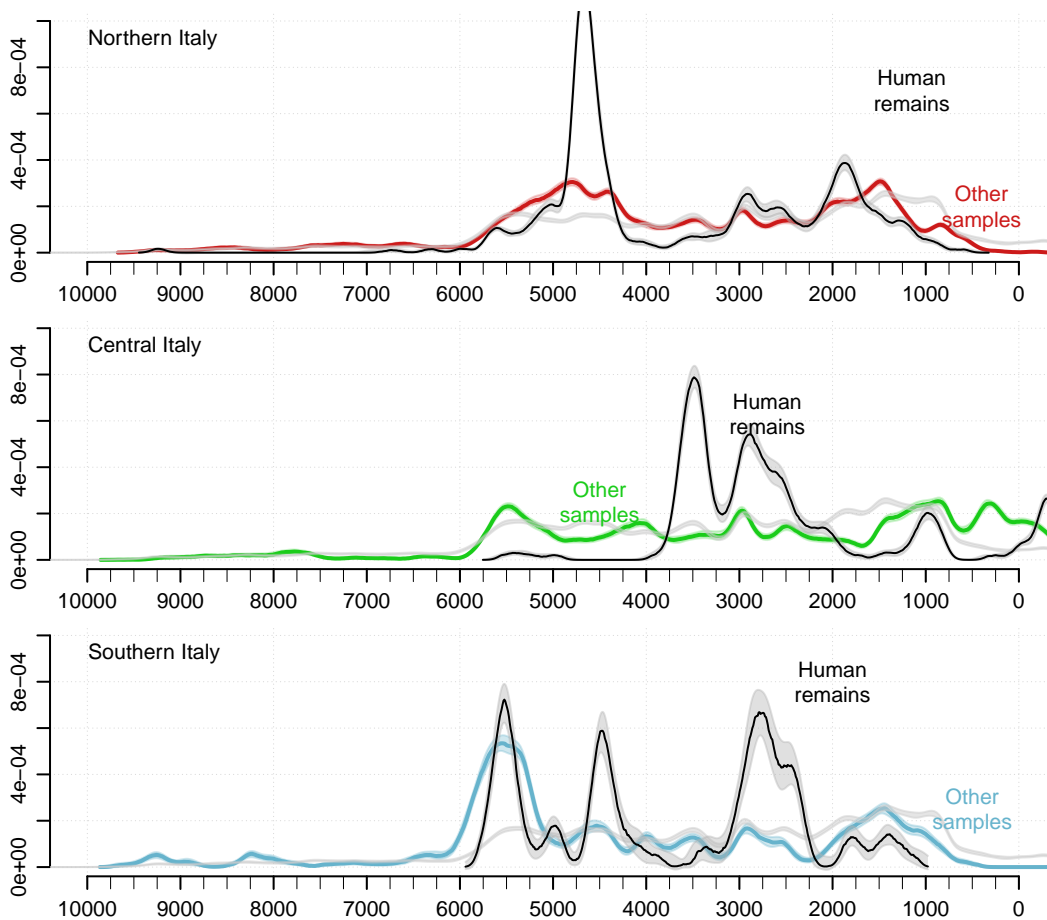
```
par(mfrow=c(3,1))
par(mar=c(2,3,.25,2))
par(mgp=c(3,.75,0))
ys<-seq(-10000,0,1000)
TX<-function(x,y,col) text(x,y,lab=c('Human\nremains','Other\nsamples'), col=c(1,col))

PLT(NIO,c1=rgb(.8,.1,.1,.2),c2=rgb(.8,.1,.1,1),xlim=c(-10000,0),ylim=c(0,0.001),
    ylabs=NULL,at=ys,labels=-ys,lwd=2,legend="Northern Italy")
plot(NIH, add=T, col=rgb(.7,.7,.7,.4)); lines(median(NIH))
TX(c(-1150,-400),c(0.0007,0.0002),rgb(.8,.1,.1,1))

PLT(CIO,c1=rgb(.1,.8,.1,.2),c2=rgb(.1,.8,.1,1),xlim=c(-10000,0),ylim=c(0,0.001),
    ylabs=NULL,at=ys,labels=-ys,lwd=2,legend="Central Italy")
plot(CIH, add=T, col=rgb(.7,.7,.7,.4)); lines(median(CIH))
TX(c(-2700,-4500),c(0.00063,0.00025),rgb(.1,.8,.1,1))

PLT(SIO,c1=rgb(.4,.7,.8,.3),c2=rgb(.4,.7,.8,1),xlim=c(-10000,0),ylim=c(0,0.001),
    ylabs=NULL,at=ys,labels=-ys,lwd=2,legend="Southern Italy")
plot(SIH, add=T, col=rgb(.7,.7,.7,.4)); lines(median(SIH))
TX(c(-2000,-500),c(0.0008,0.00025),rgb(.4,.7,.8,1))
```

Incoporating the Lucarini et al database

The Lucarini et al 2020 database of archaeological radiocarbon samples from Mediterranean Africa is availble from https://doi.org/10.5281/zenodo.3630619 which should be downloaded to the working directory.

```
Ldates<-read.csv('/Users/rowan/Dropbox/Global_C14/MedAfriCarbon/dateTable.csv',
                 stringsAsFactors = FALSE)
Lsites<-read.csv('/Users/rowan/Dropbox/Global_C14/MedAfriCarbon/siteTable.csv')
NAf<-merge(Ldates, Lsites)[,c(1,6,7,11,25,26)]
# Trim incomplete dates and dates outside study region (6th column is longitude)
NAf<-NAf[NAf$Error!='NULL' & NAf[,6]>8 & NAf[,6]<19,]
NAf<-NAf[NAf$Error!='NULL',]
NAf$Error<-as.integer(NAf$Error)

# Setup calibration curves, apply local delta R to marine curve
# NB The IntCal13 curves are referenced by name in the database but replaced here for
# analysis with IntCal20, labeled IntCal13 for convienence
IntCal13<-intcal
Marine13<-mixcurves(mix=1, delR = 60, error_delR = 70) # Actually dertived from Marine20

# Made density model
NAfr<-phasedensity(NAf[,1:4], bw=75)
```
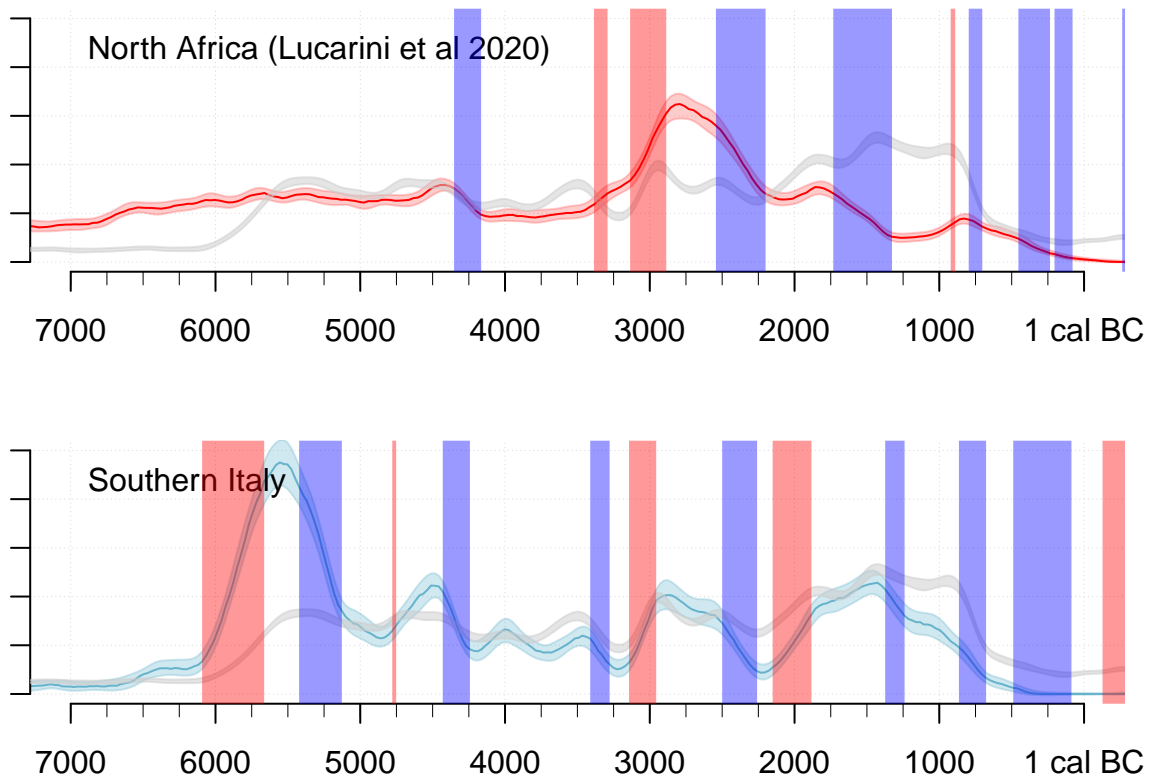
```
par(mfrow=c(2,1))
par(mar=c(3,3,1.4,1))
PLT(NAfr, c1=rgb(1,0,0,.2), c2=rgb(1,0,0,1), ylim=c(0,0.0005),
    legend='North Africa (Lucarini et al 2020)')
polygon(ggrsignif(ggr(NAfr)))
PLT(SIp,c1=rgb(.4,.7,.8,.3),c2=rgb(.4,.7,.8,1),legend="Southern Italy", ylim=c(0,0.0005))
polygon(ggrsignif(ggr(SIp)))
```
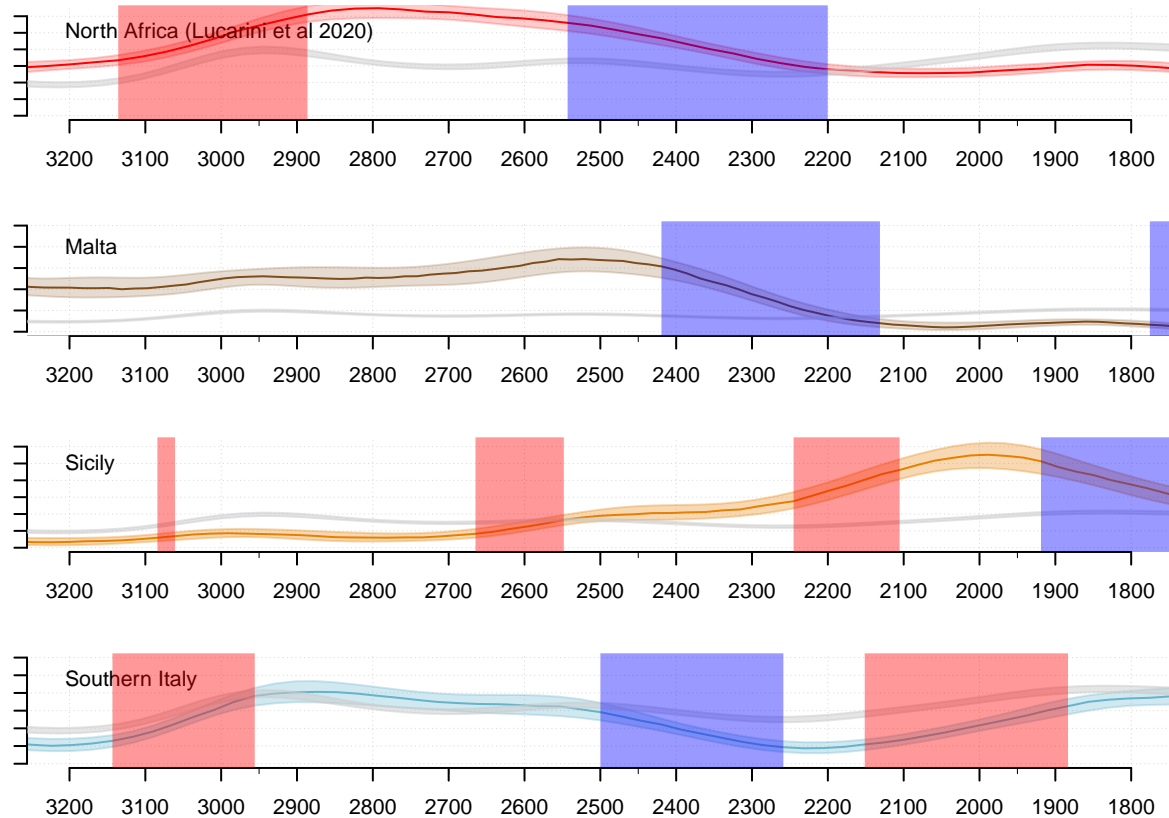

The results, shown above, indicate certain similarities, especially in the downturn that occured around 4250 cal BC (which was slight, but nonetheless significant) and the dynamics of the 3rd Millenium BC. To look at the 3rd Millenium in more detail, and to incoporate the islands between Southern Italy and North Africa:

```
par(mfrow=c(4,1))
par(mar=c(3,3,1,1))
ys<-seq(-3200,-1800,100)
PLT(NAfr, c1=rgb(1,0,0,.2), c2=rgb(1,0,0,1), ylim=NULL, xlim=c(-3200,-1800),
    at=ys,labels=-ys,legend='North Africa (Lucarini et al 2020)')
polygon(ggrsignif(ggr(NAfr)))
PLT(MTp,ylim=c(0,0.00100),c1=rgb(.5,.3,.1,.2),c2=rgb(.5,.3,.1,1),xlim=c(-3200,-1800),
    at=ys,labels=-ys,legend="Malta")
polygon(ggrsignif(ggr(MTp)))
PLT(SYp,ylim=c(0,0.00063),c1=rgb(.9,.5,.0,.3),c2=rgb(.9,.5,.0,1),xlim=c(-3200,-1800),
    at=ys,labels=-ys,legend="Sicily")
polygon(ggrsignif(ggr(SYp)))
PLT(SIp,c1=rgb(.4,.7,.8,.3),c2=rgb(.4,.7,.8,1), xlim=c(-3200,-1800),
    at=ys,labels=-ys,ylim=c(0,0.0003),legend="Southern Italy")
```

```
polygon(ggrsignif(ggr(SIp)))
```


This illustrates the different trajectory followed by activity on Sicily at the start of the Bronze Age, which was in a phase of rapid growth around the 2200 cal BC (i.e. during the 4.2kya event) whereas activity in the neighbouring African and European regions was in decline, including the collapse of the 'Temple Cultlure' on Malta.

Saving density models

The following code stores a copy of the density models to disc, thus expediting future analyses or visualization:

```
models<-c('NI','CI','SI','SY','SD','FR','CO','MT','NIp','CIp','SIp','SYp','SDp','FRp',
          'COp','MTp','NI2p','CI2p','SI2p','SY2p','SD2p','FR2p','CO2p','MT2p','FR30',
          'NI30','CI30','SI30','SD30','CIH','CIO','NIH','NIO','SIH','SIO','africa','cmed',
          'chD','adriD','greeceD','iberiaD','germanyD','scandanD','BrD','IrD','IrD30',
          'IbD','BrD30','chBA','adriBA','chBA','greeBA','gerBA','spainBA','scandanBA','NAfr',
          'drastic1','drastic2','drastic3','drastic4','steady1','steady2','steady3','steady4',
          'gradual1','gradual2','gradual3','gradual4', 'lig','lomb')
#heresOneImadeEarlier
save(list=models, file='Central_Med_14C_KDE_models_final_IntCal20.rData')
```

Notes

Compiled by T. Rowan McLaughlin, May 2021