

Etapa 2 Date satelitare privind temperatura de la nivelul suprafeței terestre

Alexandru Dumitrescu

Table of Contents

Introducere	2
Selectarea și preprocesarea datelor LST.....	2
Inventariere	2
Preprocesare.....	5
Extragerea variabilelor auxiliare	6
Modelul Numeric Altimetric.....	6
Corine Landcover.....	9
Skin temperature	10
Radiația solară globală	11
Completarea informației lipsă din datele LST	12
Date	12
Metode.....	13
Rezultate.....	15
Analiza de regresie	15
Validare modele.....	16
Îmbunătățirea rezoluției spațiale a datelor SEVIRI LST	19
Date	20
Metode.....	20
Rezultate.....	21
Concluzii.....	26
Bibliografie.....	27
Anexe	31
Anexa 1 Copernicus Toolbox script utilizat pentru extragerea datelor ERA5 Land (https://cds.climate.copernicus.eu)	31

Introducere

Regimul termic din imediata apropiere a suprafeței terestre poate fi analizat prin utilizarea informației privind temperatura suprafeței terestre active (land surface temperature LST), obținută prin prelucrarea imaginilor satelitare. LST este definită ca temperatura radiativă de la nivelul suprafeței terestre active, aflată pe direcția sistemului de senzori de teledetecție care pot fi instalați la bordul sateliților polari și geostaționari (<https://land.copernicus.eu/global/products/lst>).

Principalul avantaj al utilizării datelor LST este că furnizează informații, la o rezoluție spațio-temporală fină, în zonele unde nu sunt disponibile măsurători de suprafață (stații meteorologice). Calitatea datelor LST este însă condiționată în principal de prezența norilor sau a aerosolilor în atmosferă, care obstrucționează obținerea informației termice de la suprafața terestră de către instrumentele de teledetecție pasivă (din spectrul infraroșu termal).

Obiectivele acestei etape au fost inventarierea produselor LST disponibile pentru perioada 2009-2017 și testarea metodelor statistice de completare a datelor lipsă din produsele LST selectate, în vederea obținerii unei serii complete de date LST, la o rezoluție spațială de 1000m x 1000m, pentru teritoriul și imediata vecinătate a României. De asemenea, au fost derivați și predictorii potențial din Modelul Numeric Alitudinal (MNA) și din modul de utilizare a terenului (Corine Land Cover), utilizați ca variabile auxiliare de către modelele statistice de completare a datelor lipsă.

Selectarea și preprocesarea datelor LST

Au fost analizate produsele LST existente din punctul de vedere al disponibilității (lungime serie de timp), rezoluției spațiale și temporale. De asemenea, au fost testate metodele și instrumentele necesare obținerii și preprocesării produselor selectate pentru îndeplinirea obiectivelor din această etapă (descărcare, conversie fișiere, reproiectarea datelor etc.).

Inventariere

Există o multitudine de metode utilizate în extragerea informației LST din spectrul infraroșu termal, din imaginile furnizate de către sateliții geostaționari și polari, pentru regiunile care nu sunt afectate de nebulozitate. Informații LST se pot obține și pentru zonele acoperite cu nori, la o rezoluție spațială redusă, prin utilizare instrumentelor de teledetecție pasivă bazate pe detectarea microundelor. Un sumar al principalelor produse LST existente este prezentat în tabelul 1.

Caracteristicile produselor LST sunt determinate de tipul sateliților și senzorilor utilizați în scanarea suprafeței terestre. Astfel, imaginile furnizate de satelitul geostaționar MSG sunt disponibile la fiecare 15 minute, în timp ce sateliții cu orbită polară scanează de 2 ori pe zi aceeași zonă. Sateliții care oferă imagini de înaltă rezoluție (Landsat, Terra/ASTER), dar cu o scenă satelitară mai mică (footprint), obțin informații pentru aceeași regiune la un interval de 16 zile. Imaginile din spectrul infraroșu termal oferă informații LST la o rezoluție spațială fină doar pentru regiunile care nu sunt acoperite cu nori, față de informațiile din domeniul microundelor (ex. AMSR2), care pot fi utilizate pentru estimarea LST și în regiunile cu nebulozitate atmosferică, dar la o rezoluție spațială de 15 km.

Tabel 1: Exemple de produse LST analizate.

Satelit	Senzor	Rezoluție spațială	Rezoluție temporală	Perioada	Sursa
MSG	Seviri	3.1 km	15 min	1983- prezent	https://cmsaf.eu , https://landsaf.ipma.pt/en/
Terra/Aqua	Modis	1km	12 ore	2000- prezent	https://modis.gsfc.nasa.gov/data/data_prod/mod11.php
Terra	ASTER	100 m	16 day	2000- prezent	https://lpdaac.usgs.gov/products/ast_08v003/
METOP	AVHRR	1.1km	12 ore	1978- prezent	https://landsaf.ipma.pt/en/
Suomi-NPP	VIIRS	750m	12 ore	2012- prezent	https://data.nodc.noaa.gov/cgi-bin/iso?id=gov.noaa.ncdc:C01412
Landsat 7	ETM+	60m	16 zile	2000- prezent	https://www.usgs.gov/land-resources/nli/landsat/landsat-collection-1
Landsat 8	TIR	100m	16 zile	2013- prezent	https://www.usgs.gov/land-resources/nli/landsat/landsat-collection-1
GCOM-W	AMSR2	15km	12 ore	2012- prezent	https://suzaku.eorc.jaxa.jp/GCOM_W/research/resdist.html
FY-3A	VIRR	1km	12 ore	2009- prezent	https://satellite.nsmc.org.cn

Pentru o mai bună înțelegere a caracteristicilor produselor LST existente, au fost reprezentate grafic valorile de temperatură pentru teritoriul României, extrase din șase arhive, pentru diferite intervale de timp dintr-o zi cu nebulozitate redusă (figura 1). Senzorul care extrage informații din domeniul microundelor (AMSR2) oferă date complete pentru întreaga suprafață scanată, indiferent de gradul de acoperire al cerului cu nori. Se poate observa că această modalitate de extragere a datelor LST întâmpină dificultăți în estimarea temperaturilor din zona suprafețelor acvatice mari (Delta Dunării, lacurile din vecinătatea Bucureștiului). De asemenea, pentru scanarea completă a suprafeței terestre la latitudini medii este nevoie de un interval mai mare de 24 de ore (nu sunt disponibile date pentru o suprafață importantă din zona de interes). Senzorul de la bordul satelitelui geostaționar MSG

(Seviri) oferă informații LST la o rezoluție mai redusă, iar cea mai detaliată imagine privind distribuția spațială LST este oferită de senzorii existenți pe sateliții cu orbită polară (VIIRS, AVHRR).

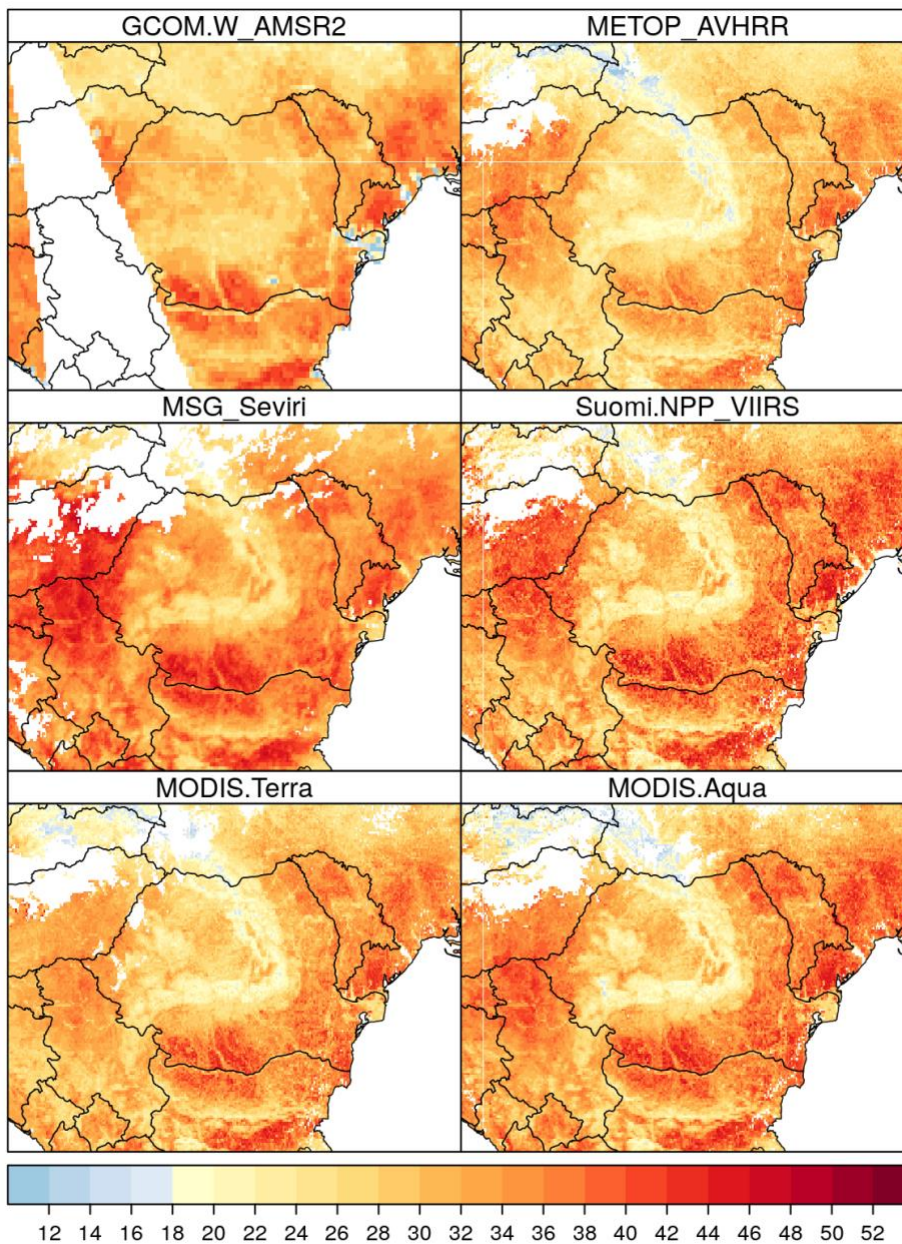


Figura 1: Temperatura (°C) de la nivelul suprafeței terestre active pentru teritoriul României, obținută prin procesarea a patru produse LST (ziua 2017-08-25).

Având în vedere caracteristicile produselor analizate, pentru îndeplinirea obiectivelor din această etapă s-au utilizat produse furnizate de senzorul satelitului geostaționar MSG, ca informații de înaltă rezoluție temporală (15 min). Datele LST extrase

din informațiile de la sateliții polari Terra/Aqua au fost utilizate pentru îmbunătățirea rezoluției spațiale (1km) a datelor orare Seviri LST.

Preprocesare

Datele LST MSG/Seviri au fost obținute din arhiva [https://landsaf.ipma.pt/en/seçiunea Get Data](https://landsaf.ipma.pt/en/seçiunea%20Get%20Data). Informația LST este furnizată în format binar HDF5, pentru întreg discul Meteosat (Full Disk), având extinderea -70° latitudine sudică, 70° latitudine nordică, -70° longitudine vestică, 70° longitudine estică. Datele SEVIRI sunt disponibile într-un sistem de proiecție cartografică geostaționară normalizată numit "Geostationary Satellite View". Distanța dintre satelit și centrul Pământului este de 35785,831 km. În acest tip de proiecție, pământul este reprezentat ca un elipsoid perfect, cu o rază de ecuator de 6378,169 km și o rază polară de 6356,583984 km (Combal și Noel 2009).

Pentru transformarea datelor LST Seviri, s-au utilizat funcțiile disponibile în aplicația **GDAL** *gdal_translate* și *gdalwarp*. **GDAL** (GDAL – Geospatial Data Abstraction Library) este o librerie opensource pentru manipularea datelor geospațiale atât în format raster, cât și în vector (GDAL/OGR contributors 2019). Aceasta este compusă dintr-o suită de funcții apelabile din linie de comandă. Funcția *gdal_translate* a fost utilizată pentru conversia obiectelor raster HDF5 în formatul GeoTIFF și declararea în obiect a proiecției, iar *gdalwarp* pentru transformarea datelor raster LST din proiecția geostaționară (Geostationary Satellite) în proiecție geografică (WGS 84):

```
gdal_translate -a_srs '+proj=geos +h=35785831 +a=6378169 +b=6356583.8 +no_defs'
-a_ullr -5570248.832537 5570248.832537 5570248.832537 -5570248.832537 HDF5:HDF5_LSASA
F_MSG_LST_MSG-Disk_201703010000.h5://LST temp.tif
gdalwarp -t_srs EPSG:4326 -wo SOURCE_EXTRA=100 temp.tif LSASAF_MSG_LST_MSG-Disk_20170
3010000.h5.tif
```

Datele LST MODIS Aqua/Terra au fost descărcate utilizându-se funcții disponibile în biblioteca suplimentară **R MODIS** (Mattiuzzi și Detsch 2019). Pachetul oferă posibilitatea de a obține date din arhivele globale online DAAC (<https://lpdaac.usgs.gov/>), LAADS (<https://ladsweb.modaps.eosdis.nasa.gov/>) și NSIDC (<https://nsidc.org/>), precum și capacități de procesare (ex: conversia fișierelor, mozaicarea, subsetarea și filtrarea seriei de timp). Utilizându-se funcția bibliotecii MODIS *getHdf()*, au fost extrase datele LST disponibile în versiunea 06 a produselor MOD11A1/MYD11A1, care conțin date LST zilnice, la o rezoluție spațială de 1 km. Datele sunt furnizate într-un sistem de proiecție Sinusoidală, în subseturi de 1200 km × 1200 km. Subseturile produselor MODIS corespunzătoare teritoriului României au următorii indecși în gridul sinusoidal: *h20v04* și *h19v04*.

Deoarece formatul fișierelor Modis HDF-EOS este recunoscut de mediul **R**, iar prelucrările statistice s-au bazat pe funcțiile implementate în acest mediu de programare, transformarea datelor din proiecție Sinusoidală în WGS84 s-a realizat utilizându-se bibliotecile **rgdal** și **raster** (Bivand et al. 2019; Hijmans 2019):

```
# lista cu fisierele *.hdf descarcate
files <- list.files(pattern = ".hdf$", full.names=T)
```

```

# citirea datelor MODiS in R
ls1 <- rgdal::readGDAL(paste('HDF4_EOS:EOS_GRID:\'', files[1], '\':MODIS_Grid_Daily_1km
_LST:LST_Day_1km', sep = ""))
ls2 <- rgdal::readGDAL(paste('HDF4_EOS:EOS_GRID:\'', files[2], '\':MODIS_Grid_Daily_1km
_LST:LST_Day_1km', sep = ""))
# mozaicarea si decuparea datelor MODIS
rm <- raster::merge(raster(ls1), raster(ls2))
rm <- raster::crop(rm, extent(1400000, 2500000, 4600000, 5500000))
# transforma din grade Kelvin in grade Celsius
rm <- rm - 273.15
# proiecteaza din proiectie Sinusoidală in WGS84
srm <- raster::projectRaster(rm, crs = "+init=epsg:4326", method = "ngb")

```

Extragerea variabilelor auxiliare

Distribuția spațială a temperaturii aerului este în cea mai mare parte determinată de relief, principalul factor determinant al particularităților climatului local și regional. Dintre caracteristicile morfologice și morfografice ale reliefului, altitudinea are rolul cel mai important în condiționarea spațială a temperaturii aerului. Caracteristicile suprafeței subiacente active (modul de acoperire a terenului) constituie, de asemenea, un important factor de condiționare a regimului termic.

Temperatura suprafeței subiacente, extrasă din datele de reanaliză (Skin temperature - Skt), reprezintă o altă variabilă explicativă utilizată în metodele estimarea spațio-temporală a datelor LST.

Modelul Numeric Altimetric

Deoarece metodele de interpolare multi-variate permit utilizarea a uneia sau a mai multor variabile auxiliare, s-a decis extragerea informației privind caracteristicile altimetrice și derivarea altor predictorii potențial din Modelul Numeric Altimetric (MNA). Reprezentarea digitală a reliefului (MNA) reprezintă punctul de plecare în calcularea elementelor morfometrice în mediul SIG₁. MNA-urile sunt de obicei întocmite prin digitizarea curbilor de nivel de pe hărțile topografice sau prin utilizarea unui sistem radar.

În această lucrare s-a utilizat modelul altimetric SRTM30 (Shuttle Radar Topography Mission)², obținut în cadrul proiectului coordonat de NASA (National Aeronautics și Space Administration)³ și NGA (National Geospatial-Intelligence Agency)⁴, care a urmărit realizarea unui MNA la nivel global cu o rezoluție spațială de 30 secunde de arc. Întrucât datele SRTM30 prezintă anumite erori mai ales în zonele foarte abrupte, s-a utilizat

¹ <http://earth.unibuc.ro/download/modele-digitale-altimetrice-si-geostatistica>

² <http://www2.jpl.nasa.gov/srtm/>

³ <http://www.nasa.gov/>

⁴ <https://www.nga.mil/About/History/NGAinHistory/Pages/SRTM.aspx>

versiunea 4.1 a datelor arhivate de consorțiul CGIAR-CSI. Acestea prezintă avantajul că au fost supuse procedurii de umplere a arealelor cu date lipsă, acolo unde radarul nu a extras informația privind altimetria de la nivelul suprafeței terestre (Reuter et al. 2007).

Descărcarea și procesarea datelor SRTM s-a realizat direct din mediul R. Pentru descărcare s-a aplicat funcția *getData()* din biblioteca **R raster**:

```
library(raster)
r1 <- getData('SRTM', lon=26, lat=45, path = "../grids/lst")
r2 <- getData('SRTM', lon=23, lat=45, path = "../grids/lst")
r3 <- getData('SRTM', lon=26, lat=47, path = "../grids/lst")
r4 <- getData('SRTM', lon=23, lat=47, path = "../grids/lst")
```

Datele SRTM sunt disponibile în scene cu o extindere spațială de 1×1 grade lonlat, teritoriului României corespunzând scenele cu indecșii: 41_03, 41_04, 42_03 și 42_04. Scenele au fost transformate într-un singur fișier cu ajutorul funcției *raster::mosaic()*.

```
par(mar = c(0.1, 0.1, 0.1, 0.4))
mna <- mosaic(r1, r2, r3, r4, fun = mean)
mna.crop <- crop(mna, extent(18,32,42, 49.5))
plot(mna.crop, col = (terrain.colors(15)))
```

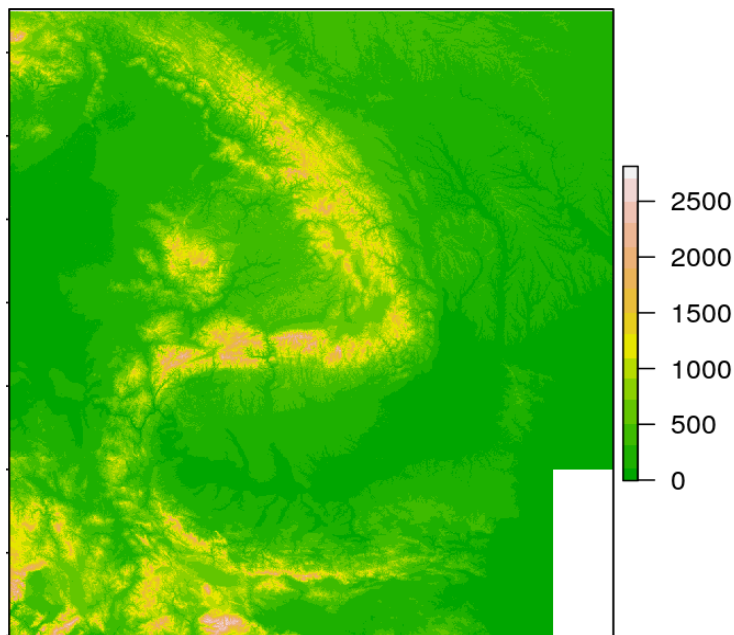


Figura 2: Modelul numeric altitudinal SRTM30 (CGIAR-CSI 4.1).

⁵ <https://cgiarcsi.community/data/srtm-90m-digital-elevation-database-v4-1/>

Utilizând MNA, au fost derivați alți patru predictorii pentru a se analiza influența distribuției spațiale a temperaturii aerului în funcție de extinderea în latitudine a teritoriului analizat, influența suprafețelor acvatice mari aflate în imediata vecinătate (Marea Neagră) și delimitarea suprafețelor cu înclinare mare sau redusă a versanților.

Câmpurile de longitudine (lon) și de latitudine (lat) au fost obținute prin utilizarea bibliotecii **sp** a limbajului R, prin extragerea coordonatelor fiecărui pixel care compune MNA-ul (Pebesma și Bivand 2005). Rasterele cu distanța față de Marea Neagră (`adri_w_dist`) au rezultat în urma utilizării funcției **`rasster::gridDistance()`**. Altitudinea medie și minimă pe o rază de 20 km (`focal_mean` și `focal_min`) au fost obținute cu ajutorul funcției **`rasster::focal()`**.

```
dem.a <- raster::aggregate(mna.crop, fact=12, fun=mean)

library(rnaturalearth)
borders <- ne_countries(scale = 10, type = "countries", continent = "Europe")
borders <- crop(borders, extent(18,32,42, 49.5))
dem.b <- rasterize(borders, dem.a, mask=TRUE, progress = 'text')

xyz <- rasterToPoints(dem.b)
lat <- rasterFromXYZ(cbind(xyz[,1:2], xyz[,2]))
focal.mean <- focal(dem.a, w = matrix(1, nrow = 5, ncol = 5), fun = mean)
focal.min <- focal(dem.a, w = matrix(1, nrow = 5, ncol = 5), fun = min)

library(igraph)
sea_dist <- gridDistance(dem.b, origin=NA, progress = 'text')

par(mfrow = c(2,2), mar = c(1, 2, 3, 2), oma =c(0,0,0,1))
plot(lat, col=heat.colors(21), main = "latitude")
plot(borders, add = T)
plot(focal.mean,col = terrain.colors(15), main = "focal_mean")
plot(borders, add = T)
plot(focal.min, col = terrain.colors(15), main = "focal_min")
plot(borders, add = T)
plot(sea_dist/1000,col=heat.colors(21), main = "sea_dist")
plot(borders, add = T)
```

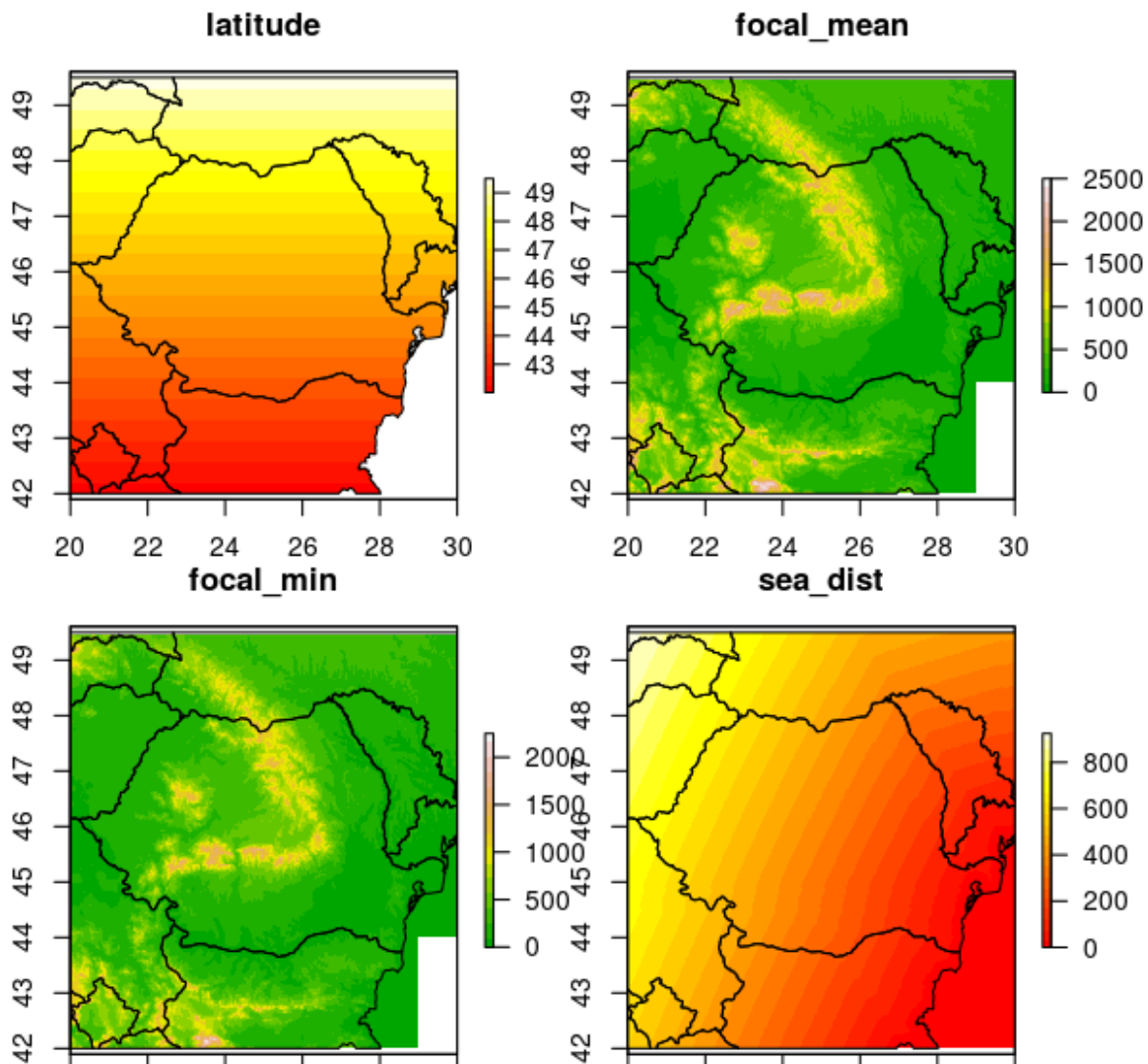


Figura 3: Predictorii derivați din MNA – rezoluție spațială 1000 m (altitudine, altitudinea medie și minimă pe o rază de 20 km, distanța față de Marea Neagră) .

Corine Landcover

Fișierul raster cu numărul pixelilor încadrați ca suprafață urbană pe o rază de 3 km a fost utilizat pentru analiza influenței orașelor asupra distribuției spațiale a temperaturii. Această informație a fost extrasă din setul de date european de referință pentru modul de acoperire al terenului (Land Cover CORINE CLC2012) (Büttner 2014). Fișierul a fost obținut cu ajutorul funcțiilor **rasterize()** și **focal()** din biblioteca **R raster**

```
library(rgdal, quietly = T)
urb <- readOGR("../shp","urban_areas", verbose = F)
urb.geo <- spTransform(urb, CRS("+init=epsg:4326"))
urb.r <- rasterize(urb.geo,dem.b,mask=T)
```

```

urb.r[urb.r > 0]<-1
urb.r[is.na(urb.r)]<-0
fun <- function(x) { return(length(x[x>0])) }
rf <- focal(urb.r, w = matrix(1, nrow = 3, ncol = 3), fun=fun, padValue=T)
plot(rf, col = rev(heat.colors(10)), main = "urban areas")
plot(borders, add = T)

```

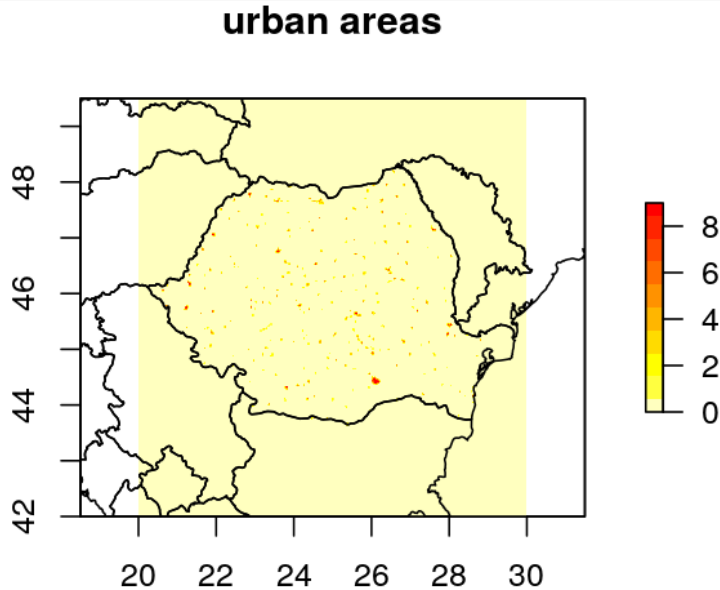


Figura 4: Numărul pixelilor încadrați ca suprafață urbană pe o rază de 3 km – rezoluție spațială 1000 m.

Skin temperature

Temperatura suprafeței subiacente poate fi obținută și în urma experimentelor numerice rulate în modele climatice de generație nouă. Astfel, pornindu-se de la starea inițială a atmosferei, se estimează parametri meteorologici pe baza unui set complex de ecuații matematice, obținându-se datele de reanaliză.

ERA5-Land este un set de date de reanaliză, care conține estimări ale parametrilor meteorologici la un interval de timp orar, pe parcursul a câteva decenii (2001 - prezent), la o rezoluție spațială de $0.1^\circ \times 0.1^\circ$. ERA5-Land este ultima generație de date de reanaliză, realizate de Centrul European de Prognoză Meteorologică pe Termen Mediu (European Centre for Medium-Range Weather Forecasts - ECMWF), care conține o serie de îmbunătățiri pentru utilizarea în modelarea proceselor mediului înconjurător (Vitart et al. 2019).

Datele ERA5-Land sunt puse la dispoziția utilizatorilor prin intermediul serviciului Copernicus Climate Change Service (C3S), serviciul european care furnizează informații pentru a crește baza de cunoștințe în sprijinul politicilor de adaptare și atenuare a efectelor schimbările climatice (<https://climate.copernicus.eu/>). Datele pot fi procesate și descărcate direct pe infrastructura serviciului, prin utilizarea funcțiilor disponibile în secțiunea Toolbox

a site-ului (https://cds.climate.copernicus.eu/toolbox-editor/9924/get_era5_land). Datele de reanaliză pot fi descărcate atât în formatul original GRIB2, cât și în formatul binar NetCDF.

Temperatura orară a suprafeței subiacente (Skin temperature) din arealul corespunzător teritoriului României, a fost extrasă pentru perioada de interes utilizându-se un script rulat direct pe infrastructura Copernicus (Anexa 1). Datele au fost salvate în format NetCDF și utilizate ca variabile explicative în procedurile de completare cu date în arealele în care imaginile satelitare nu furnizează informații LST (figura 5).

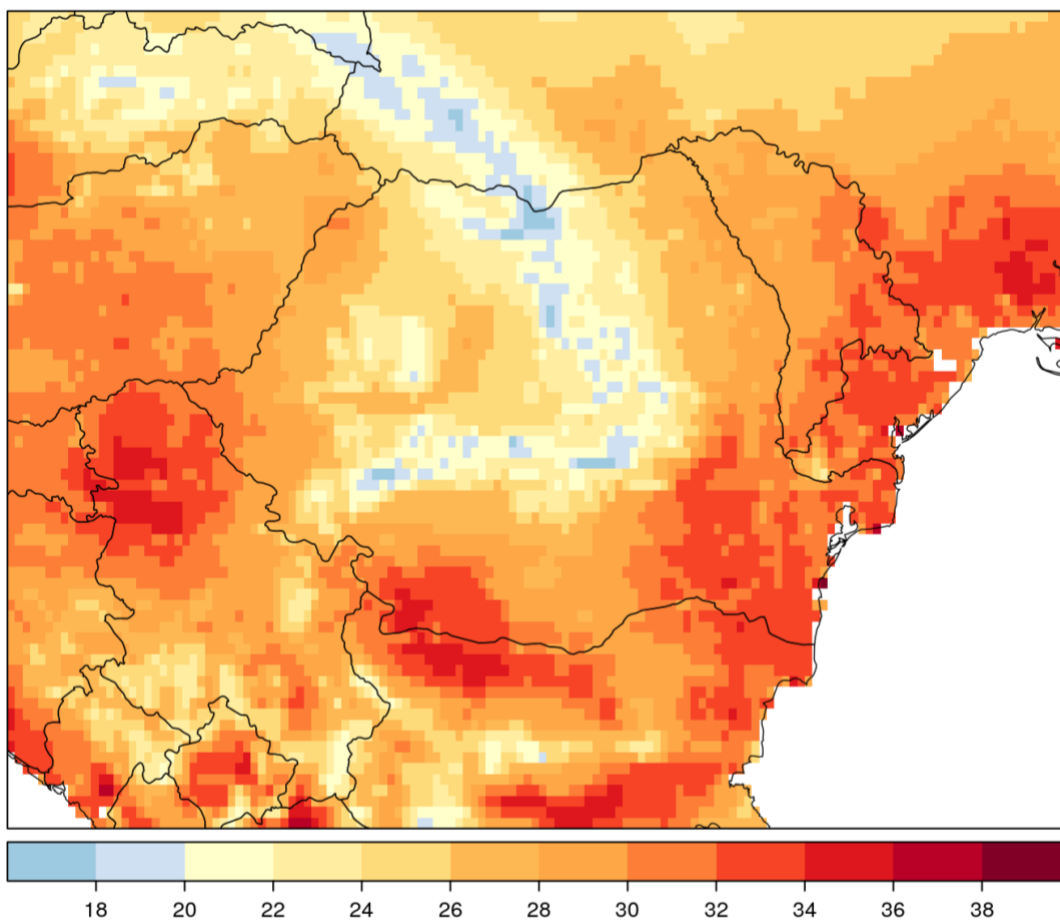


Figura 5: Temperatura (°C) de la nivelul suprafeței terestre active pentru teritoriul României, obținută prin procesarea datelor de reanaliză ERA5 Land (ziua 2017-08-25, ora 09 UTC).

Radiația solară globală

Radiația solară globală estimată din prelucrarea imaginilor satelitare MSG, a constituit o altă variabilă utilizată în estimarea spațio-temporală a datelor LST. În acest studiu s-au utilizat datele obținute din arhiva CM SAF (The Satellite Application Facility on Climate Monitoring - www.cmsaf.eu), mai exact estimările realizate cu modelul *Heliosat*, versiunea 2.1 (Pfeifroth et al. 2017).

Organizația CM SAF are ca scop întocmirea unei arhive cu produse satelitare, obținută prin prelucrarea informațiilor furnizate de senzorii instalați la bordul sateliților geostaționari (SEVIRI) și polari (AVHRR, METOP). Produsele sunt grupate în două categorii: *operational products* - seturi de date obținute în timp real, care pot fi folosite în monitorizarea climatică, și *climate data sets* - seturi de date climatice care sunt statistic omogene și cu un șir de observații suficient de lung pentru analize privind variabilitatea și schimbarea climei (Schulz et al. 2009).

Completarea informației lipsă din datele LST

Disponibilitatea datelor LST este strict condiționată de prezența norilor, umbrelor și a altor condiții atmosferice, care împiedică observarea directă a suprafeței terestre de către senzorii instalați la bordul sateliților. Pentru îmbunătățirea acoperirii spațiale cu date LST, cea mai aplicată metodă este agregarea temporală a datelor, la o rezoluție de până la 8 zile (Li et al. 2018). Aceste produse nu oferă informații suficient de detaliate pentru analiza impactului variațiilor bruște de temperatură asupra activităților socio-economice sau a mediului înconjurător. Prin urmare, au fost dezvoltate tehnici de completare a informației lipsă din produsele LST zilnice și sub-zilnice.

Utilizarea informației primită de la diferiți senzori, care scanează aceeași suprafață la diferite intervale de timp, reprezintă o metodă aplicată frecvent în completarea datelor lipsă (Coops et al. 2007; Duan et al. 2017). Modele de regresie multiplă (MLR), bazate pe variabile auxiliare obținute din prelucrarea MNA sau a altor produse rezultate din prelucrarea imaginilor satelitare (ex.: NDVI), constituie o altă categorie de metode concepute pentru estimarea informației lipsă din datele LST (Metz et al. 2017). Tehnici de interpolare spațială, temporală sau spațio-temporală, univariate și multivariate, sunt aplicate frecvent pentru obținerea informației lipsă din imaginile satelitare (Kilibarda et al. 2014; Neteler 2010; Weiss et al. 2014).

Având în vedere volumul mare de date utilizat, dar și necesitatea utilizării variabilelor auxiliare în procedura de estimare spațio-temporală, în acest studiu au fost testate două categorii de metode statistice utilizate cu succes în estimarea spațială a datelor de temperatura aerului obținute de la stațiile meteorologice (Aalto et al. 2013; Ruiz-Álvarez et al. 2019; Szymanowski et al. 2013): regresia multiplă liniară, prin metoda celor mai mici pătrate (MLR), și modelul de regresie aditiv generalizat (GAM).

Date

Pentru testarea metodelor de estimare a valorilor lipsă, s-a utilizat datele orare LST Seviri disponibile în intervalul 2016 - 2017 (17544 de imagini). Întrucât scopul acestui studiu este alegerea metodei optime de estimare spațială, validarea metodelor s-a realizat prin utilizarea unui set de date independent, extras spațial în mod aleatoriu (300 de puncte pentru fiecare pas de timp), care nu au fost utilizate în calibrarea modelelor.

Pentru a se asigura calitatea și omogenitatea datelor LST și a se elimina eventuale erori de prelevare și procesare, a fost aplicat un filtru de 1-99%: din fiecare imagine, au fost excluse din analiză valorile extreme, respectiv cele mai mici de percentila 1% și mai mari de percentila 99% a setului de valori (Hyndman și Fan 1996). De asemenea, au fost eliminați pixelii cu date incerte, conform informațiilor auxiliare care însoțesc produsul de bază (banda Q_FLAGS din fișierul original NetCDF)(Land SAF Project Team 2018).

Variabile auxiliare derivate din MNA, radiația solară CMSAF și temperatura suprafeței subiacente ERA5 Land au fost utilizate în testarea diferitelor calibrări ale modelor de regresie.

Metode

Regresia liniară aproximează setul de date printr-o dependență liniară, care minimizează suma pătratelor dintre dreapta de aproximare a norului de puncte și punctele date (Thom 1966). Regresia multiplă evidențiază relația dintre o variabilă dependentă (explicată) și o mulțime de variabile independente (explicative):

$$Z(s) = \beta_0 + \beta_1\chi_1(s) + \beta_2\chi_2(s) + \dots + \beta_p\chi_p(s) + \delta$$

unde: $Z(s)$ reprezintă valoarea estimată a variabilei dependente în punctul s , $\chi_1(s)$, $\chi_2(s)$, ..., $\chi_p(s)$ variabilele explicative, $\beta_0, \beta_1, \beta_2 \dots \beta_p$ sunt coeficienții de regresie parțială, iar δ eroarea standard de estimare a variabile dependente. Prin procedura automată pas cu pas (stepwise regression), utilizându-se criteriul Akaike, se poate alege cea mai bună combinație a regresiei, incluzând numai variabilele auxiliare care sunt statistic semnificative pentru modelul respectiv (Venables și Ripley 2013). Astfel, la un pas ulterior al regresiei prospective se permite eliminarea unei variabile. O variabilă eliminată din model devine candidată pentru includerea în model, iar o variabilă inclusă în model devine candidată la excludere. După ce se testează toate combinațiile posibile ale modelului, se aleg acele variabile care explica cel mai bine din punct de vedere statistic fenomenul analizat.

Modelele aditive generalizate (GAM), dezvoltate de Hastie și Tibshirani (1990), nu presupun existența unui tip special de relații între variabilele analizate (liniare, cuadratiche, logaritmice), acestea fiind determinate prin analiza caracteristicilor statistice ale seturilor de date analizate. GAM implică o sumă de funcții posibil neliniare (din categoria spline), una pentru fiecare predictor (Wood și Augustin 2002). Derivat din regresia liniară, modelul GAM poate fi scris:

$$Z(s) = \beta_0 + s_1(\chi_1) + s_2(\chi_2) + \dots + s_p(\chi_p) + \delta$$

unde fiecărei variabile χ_p i se aplică o funcție non-parametrică s_p . Calibrarea generală a modelului se realizează printr-un proces iterativ de penalizare a complexității modelului, în care un rol important îl au reziduurile parțiale, folosind criteriul erorii de predicție prin validarea încrucișată. Procesul de selecție al proprietăților funcțiilor non-parametrice

(gradul de netezire) se finalizează în momentul în care modelul nu mai poate fi îmbunătățit (Wood 2006).

Ambele metode de regresie prezintă avantajul că pot fi folosite la calibrarea unor modele utilizându-se subseturi de date, în funcție de caracteristicile variabilei analizate. Ținând cont de variabilitatea spațio-temporală a temperaturii aerului specifică regiunii geografice analizate, modelele de regresie au fost calibrate utilizându-se subseturi de date orare, pentru fiecare lună și an analizat.

Cuantificarea diferențelor dintre estimări (Ps_i) și datele măsurate (Os_i) s-a realizat prin utilizarea indicatorilor de măsurare a erorilor (Fox 1981):

- eroarea medie (ME) reprezintă mediile diferențelor dintre valorile estimate și cele măsurate:

$$ME = \frac{1}{N} \sum_{i=1}^N (Ps_i - Os_i)$$

- eroarea medie absolută (MAE) reprezintă mediile diferențelor absolute dintre valorile estimate și cele măsurate:

$$MAE = \frac{1}{N} \sum_{i=1}^N |(Ps_i - Os_i)|$$

- eroarea pătratică medie (RMSE) poate fi sensibilă la prezența erorilor mari, procesul de ridicare la pătrat atribuind ponderi disproporționate reziduurilor:

$$RMSE = \left(\frac{1}{N} \sum_{i=1}^N (Ps_i - Os_i)^2 \right)^{\frac{1}{2}}$$

- coeficientul de corelație Pearson (CORR)

$$r_{x,y} = \frac{\sum_{i=1}^N (x_i - \bar{x})(y_i - \bar{y})}{s_x \cdot s_y}$$

unde N este mărimea eșantionului format din cele două variabile, x_i = măsurile individuale ale variabilei x , y_i = măsurile individuale ale variabilei y , \bar{x} = media aritmetică a variabilei x , \bar{y} = media aritmetică a variabilei y , s_x = deviația standard pentru valorile x și s_y = deviația standard pentru valorile y . Valoarea coeficientului de corelație Pearson poate varia între +1 și -1. Valoare apropiată de +1 indică o corelație pozitivă puternică, o valoare apropiată de -1 ne indică o corelație negativă puternică, iar o valoare apropiată de zero ne indică faptul că între cele două variabile nu există nici o corelație.

Rezultate

Analiza de regresie

Interdependența dintre variabila analizată și temperatura suprafeței subiacente din modelul ERA5 Land este demonstrată în figura 6, unde se poate observa o intensitate puternică a relației statistice liniare între cele două variabile. Exceptând sezonul rece, coeficientul de corelație dintre cele două variabile este mai mare de 0.9.

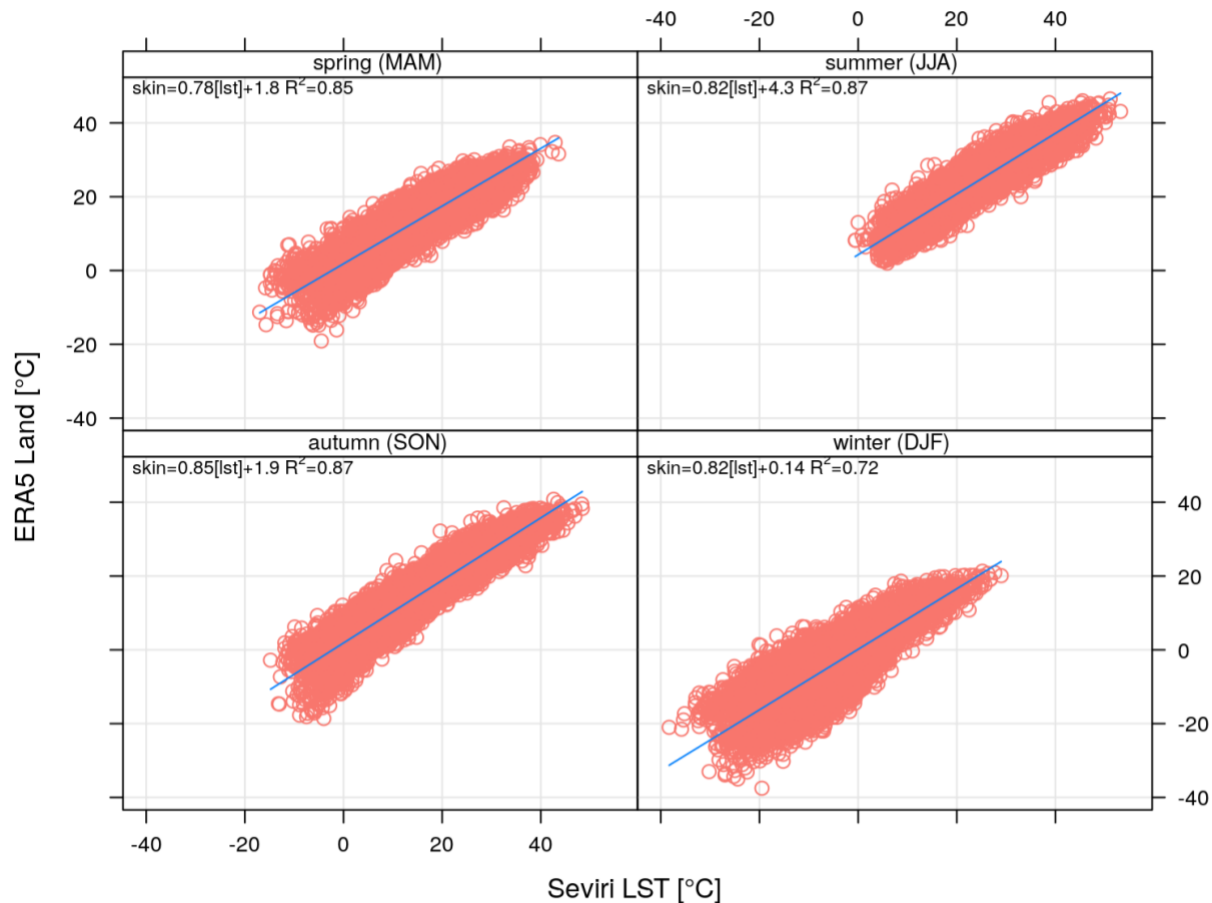


Figura 6: Temperatura ($^{\circ}C$) suprafeței active pe teritoriul României, obținută prin procesarea datelor de reanaliză ERA5 Land versus Seviri LST (anul 2017).

Similaritatea dintre variabila analizată și ceilalți predictori (radiația solară globală și altitudine) a fost analizată prin calcularea coeficienților de corelație Pearson (tabelul 5). Având în vedere volumul mare de date, pentru această analiză s-au selectat datele din anul 2017, iar pentru fiecare pas de timp (oră) au fost selectate în mod aleator valorile corespunzătoare a 10 pixeli.

Tabel 2: Coeficienți de corelație Pearson calculați între datele Seviri LST și variabilele explicative.

Variabilă	Iarna	Primăvara	Vara	Toamna
Altitudine	-0.1400900	-0.2341505	-0.3195518	-0.2001000
Radiație solară	0.5984169	0.8197970	0.8170176	0.7364474
Temperatura ERA5	0.8594143	0.9256755	0.9251502	0.9361422

Cu toate că au fost identificate corelații semnificative dintre LST și alte variabile auxiliare (extinderea în latitudine, suprafețele urbane), pentru simplificarea modelului și evitarea efectului de coliniaritate (proprietățile acestora fiind înglobate în variabilele dinamice temperatura suprafeței active și radiația solară) s-a decis eliminarea lor din analiză. Prin urmare, modelele de regresie au fost calibrate utilizându-se ca variabile auxiliare date (1) altitudinea, (2) radiația solară CMSAF și (3) temperatura suprafeței subiacente ERA5 Land.

Validarea modelelor

Pentru alegerea metodei optime de completare a datelor lipsă, s-a ales ca strategie calibrarea modelelor de regresie utilizându-se subseturi de date lunare pentru fiecare an luat în calcul (2016 – 2017, 24 de luni). Astfel, au fost calibrate trei modele de regresie: două modele (GAM și MLR) având ca predictor temperatura ERA5, radiație solară și altitudine. Deoarece s-a constatat că similaritatea dintre datele analizate și variabila auxiliară principală utilizată (temperatura ERA5) este mai puțin evidentă în pixelii din apropierea datelor lipsă LST (figura 7), s-a decis și verificarea unui model de regresie GAM care, alături de predictorii menționați mai sus, să țină cont de variabila distanța față de datele lipsă LST (GAM_dist).

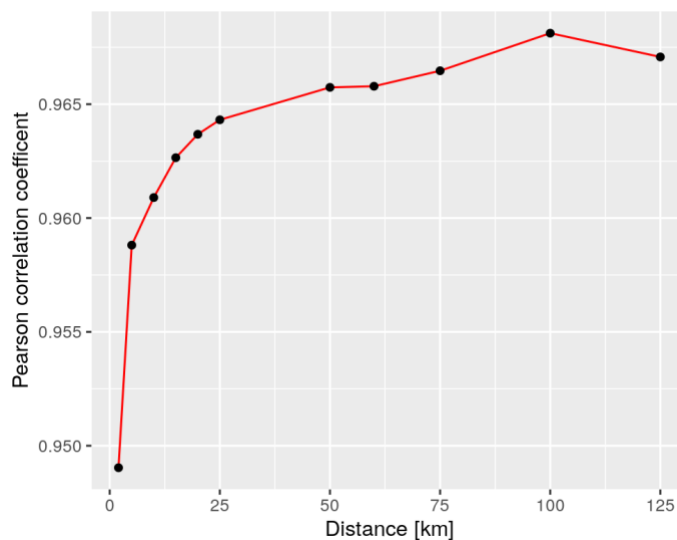


Figura 7: Coeficientul de corelație Pearson (Seviri LST și temperatura suprafeței subiacente ERA5) calculat pe praguri de distanțe față de zonele cu date lipsă LST (anul 2017).

Verificarea estimărilor s-a realizat utilizându-se setul de date independent, obținut conform descrierii din secțiunea [Date](#). Într-o primă etapă, s-au calculat indicatorii sezonieri de măsurare a erorilor MAE, RMSE și CORR (tabelul 3). Erorile de estimare ale modelului GAM_dist sunt în toate cazurile mai mici decât modelele GAM și MLR, fie că ne referim la MAE sau RMSE, aceste rezultate fiind confirmate și prin calculul coeficienților de corelație (mai mari în toate cazurile decât celelalte două modele de regresie). Analizând performanțele celorlalte două modele de regresie, s-a constatat că estimările modelului GAM sunt superioare modelului MLR, în toate cazurile analizate.

Tabel 3: Valorile medii sezoniere ale erorilor medii absolute (MAE), erorilor pătratice medii (RMSE) și ale coeficienților de corelație Pearson (CORR) (2016 - 2017).

Model	Sezon	MAE	RMSE	CORR
MLR	DJF	2.795	3.745	0.888
GAM	DJF	2.583	3.511	0.902
GAM_dist	DJF	2.493	3.351	0.911
MLR	MAM	2.173	2.900	0.950
GAM	MAM	1.979	2.696	0.957
GAM_dist	MAM	1.926	2.603	0.960
MLR	JJA	1.934	2.491	0.953
GAM	JJA	1.811	2.351	0.958
GAM_dist	JJA	1.796	2.328	0.959
MLR	SON	2.117	2.799	0.958
GAM	SON	1.900	2.564	0.965
GAM_dist	SON	1.848	2.462	0.968

Performanțele celor trei modele au fost analizate și cu graficele de corelație de tip puncte (scatterplot) (figura 8). Rezultatele foarte bune obținute de cele trei modele sunt confirmate de gruparea punctelor în jurul liniei de regresie, fără poziționarea vreunui punct în afara norului de puncte. Diferențierea s-a realizat prin calcularea pantei regresiei liniare dintre datele măsurate și cele estimate, cele mai mari valori rezultând din analiza estimărilor GAM_dist (în cazul de față cu cât valorile pantelor sunt mai mari, cu atât similaritatea datelor analizate este mai mare).

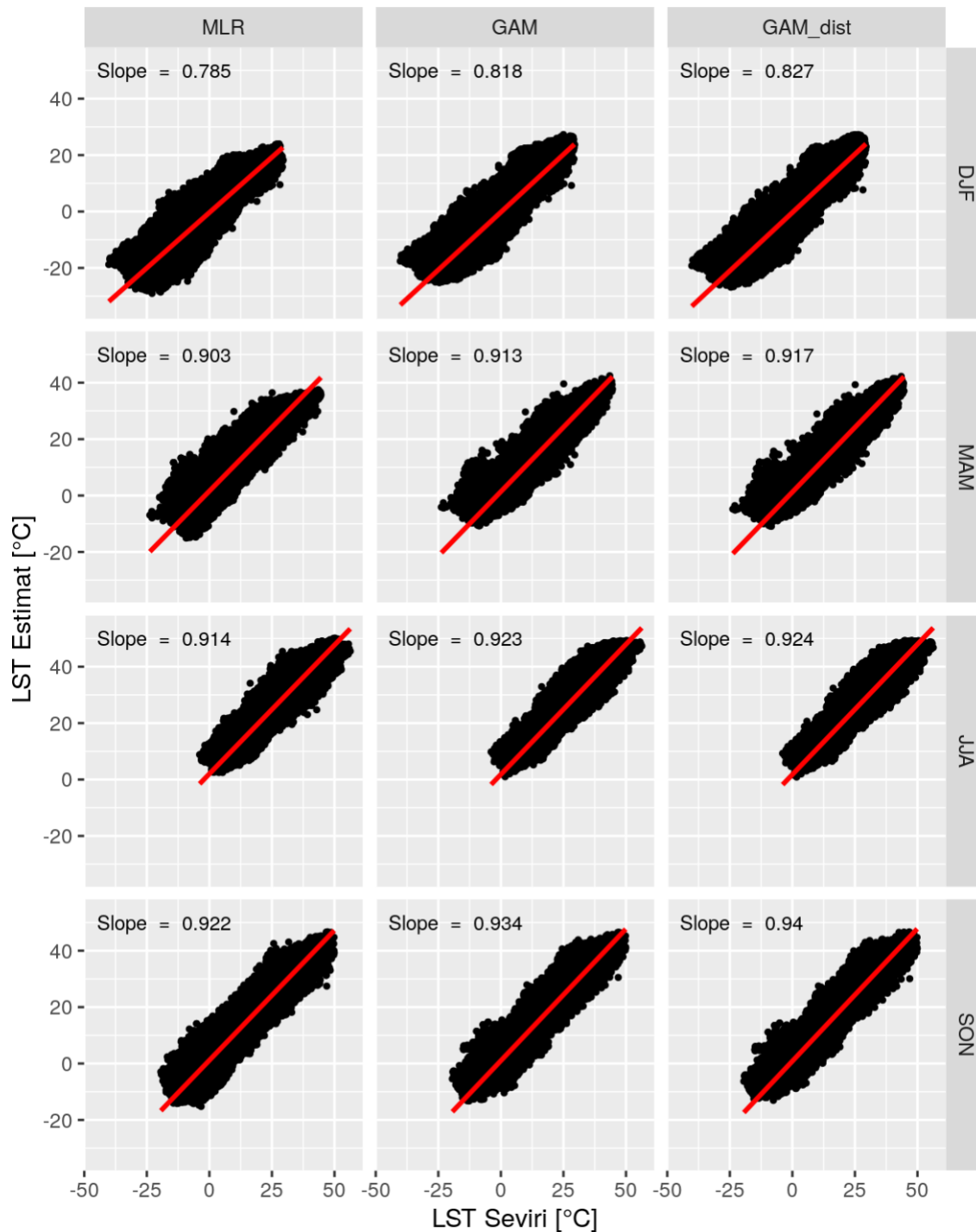


Figura 8: Seviri LST vs. LST estimată utilizând trei modele de regresie (2016 - 2017).

Având în vedere că după toate criteriile utilizate în analiza rezultatelor celor trei modele s-a constatat că modelul GAM_dist furnizează estimările cele mai apropiate de datele măsurate, s-a luat decizia ca această metodă să fie aplicată în completarea datelor lipsă Seviri LST. De asemenea, rezultatele foarte bune obținute de toate modele analizate sugerează

faptul că predictorii luați în calcul satisfac cerințele necesare obținerii unui set complet de date LST.

Pentru obținerea setului de date complet de date Seviri LST, s-au calibrat 96 de modele GAM_dist pentru fiecare an și lună din perioada analizată (2009 - 2017). Un exemplu de estimare a datelor lipsă Seviri LST este prezentat în figura 9, în care se observă că estimările LST sunt din punct de vedere calitativ acceptabile, indiferent de regiunea geografică pentru care s-a realizat completarea. Datele au fost arhivate în fișiere multidimensionale binare NetCDF zilnice, care conține valorile LST corespunzătoare pentru 24 de ore.

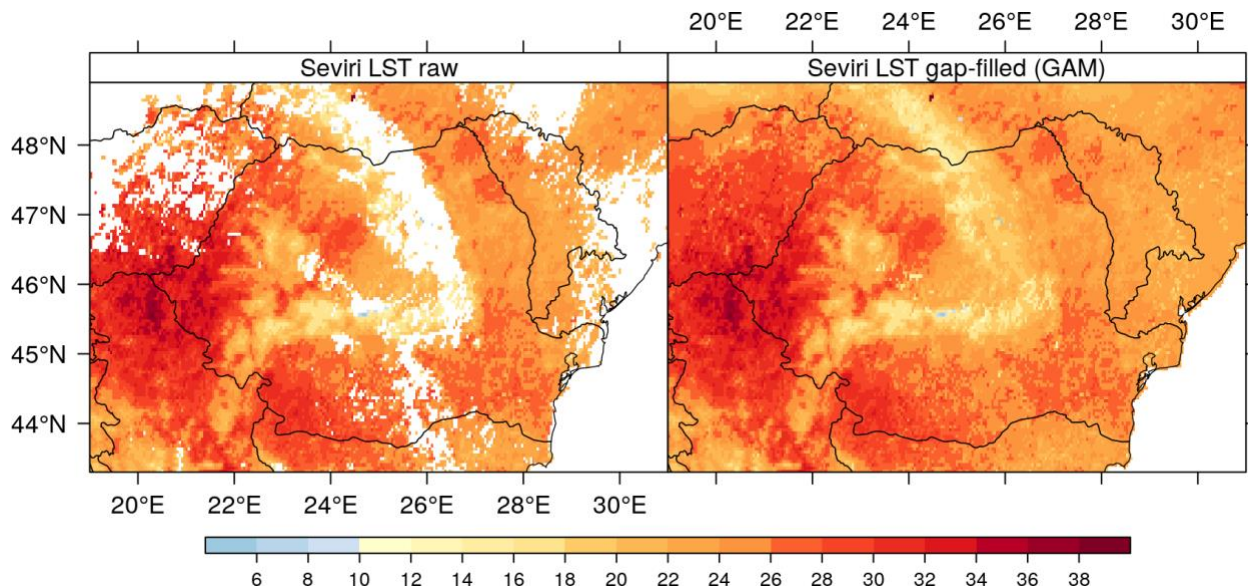


Figura 9: Completarea datelor Seviri LST (°C) din 4 iulie 2017, ora 14:00 UTC: imaginea originală (stânga) și datele completate (dreapta)

Îmbunătățirea rezoluției spațiale a datelor SEVIRI LST

Deoarece datele Seviri LST sunt la rezoluție spațială care nu permite studiul variației spațio-temporale a temperaturii la scară mare (ex.: variabilitatea diurnă a insulei de căldură urbană), s-au analizat câteva metode utilizate în mod frecvent pentru îmbunătățirea rezoluției spațiale datelor satelitare (downscaling spatial).

Metodele de regresie bazate pe variabile explicative care au o rezoluție spațială mai detaliată față de datele analizate sunt utilizate frecvent pentru îmbunătățirea rezoluției spațiale a produselor rezultate din prelucrarea imaginilor satelitare (Keramitsoglou et al. 2013). În cazul datelor LST, s-au dezvoltat modele de regresie calibrate cu ajutorul predictorilor derivați din MNA, din datele privind modul de utilizare a terenurilor, de emisivitate și prin utilizarea altor produse LST disponibile la o rezoluție spațială detaliată comparativ cu datele analizate (Amazirh et al. 2019; Bechtel et al. 2012; Sismanidis et al. 2016; Wu et al. 2019). În mai multe studii recente, au fost corectate erorile de estimare ale

metodelor de regresie prin utilizarea metodelor de interpolare spațială (Bartkowiak et al. 2019; Kwak et al. 2018; Pereira et al. 2018). Park et al. (2019) realizează o recenzie a metodelor aplicate în downscalingul spațial al datelor satelitare LST, în care sunt prezentate și unele aspecte practice. Teleconecțiile ortogonale empirice (EOT) au fost utilizate cu succes în creșterea gradului de detaliere spațială a datelor satelitare (Detsch et al. 2016a; Detsch et al. 2016b). EOT a fost introdus ca o alternativă la funcțiile empirice ortogonale (EOF), având avantajul că metoda este relativ simplu de implementat, iar rezultatele sale sunt ușor de interpretat (Appelhans et al. 2015). Această metodă presupune identificarea unui număr de componente din domeniul predictorului, care explică într-o anumită măsură variabilitatea datelor dependente. Domeniul predictorului este redus la un punct de grilă, care explică cel mai bine varianța datelor analizate. Procedura este repetată prin identificarea următoarelor puncte importante de grilă, până când se obține un număr predefinit de puncte din domeniul analizat.

Considerându-se specificul datelor analizate (date orare în puncte de grilă), în lucrarea de față s-a utilizat o metodă hibridă, care presupune parcurgerea mai multor etape: (1) realizarea climatologiei lunare a datelor Seviri LST la o rezoluție spațială de 1000m × 1000m, (2) aplicarea metodei anomaliilor pentru obținerea setului final de date orare LST, utilizându-se ca referință climatologia lunară. Metoda de interpolare spațială bazată pe climatologie este de obicei aplicată în situațiile când se procesează un volum mare de date la scară spațio-temporală fină (Dumitrescu et al. 2016; Dumitrescu et al. 2017; Haylock et al. 2008; Krähenmann et al. 2011).

Date

Pentru calibrarea modelelor de regresie, s-au utilizat ca variabile explicative datele MODIS LST și variabile derivate din MNA și Corine Landcover (vezi secțiunea [Extragerea variabilelor auxiliare](#)). Datele MODIS LST zilnice MOD11A1 și MYD11A1 au fost mediate lunar, utilizându-se toate disponibile între anii 2016 - 2017. S-au descărcat datele MODIS LST obținute cu versiunea 6.3.0 a codului de generare a produselor (LST Product Generation Executive code), care au deviația standard a erorilor de estimare de 0.5 K (Wan 2014).

Metode

În prima etapă a metodei de downscaling (climatologia lunară) s-a utilizat regresia multiplă liniară, utilizându-se ca date de intrare valorile medii lunare Seviri LST și variabilele explicative reprezentate de mediile lunare MODIS LST (Modis), altitudine (Alt), altitudinea medie pe o rază de 20 km, câmpul de latitudine (Lat) și numărul pixelilor încadrați ca suprafață urbană pe o rază de 3 km (Urb).

Pentru a se elimina posibilul efect de coliniaritate (variabile independente corelate între ele), predictorii au fost supuși procesului de filtrare prin analiza componentelor principale. Filtrarea predictorilor prin analiza componentelor principale (PCA) se realizează prin transformarea variabilelor inițiale într-un set nou de variabile, necorelate între ele și de dimensiune mai mică. Noul set de date astfel obținut conține cea mai mare parte a variabilității setului de date original.

PN-III-P1-1.1-PD-2016-1579 Interpolarea spațio-temporală la scară fină a temperaturii aerului 20 pentru teritoriul României

Calibrarea modelelor de regresie s-a realizat prin utilizarea variabilelor explicative, agregate la rezoluția datelor Seviri LST (aprox. 4.5 km), estimările obținându-se la rezoluția nativă a predictorilor (1000m × 1000m). Erorile de estimare ale modelelor de regresie au fost spațializate printr-o metodă simplă de interpolare (biliniară), specifică tipurilor de date gridate (Press et al. 1992). O schemă simplificată a metodei de interpolare este prezentată în figura 10. Pentru a se ține cont de ciclul diurn al temperaturii și având în vedere că datele Modis LST sunt disponibile atât ziua, cât și noaptea, pentru fiecare perioadă din zi și lună, au fost calibrate două modele de regresie (pentru zi și noapte). Unghiul de elevație solară, calculat pentru fiecare oră și zi, a fost utilizat pentru selectarea imaginilor orare Seviri LST care au fost mediate lunar pentru zi sau noapte.

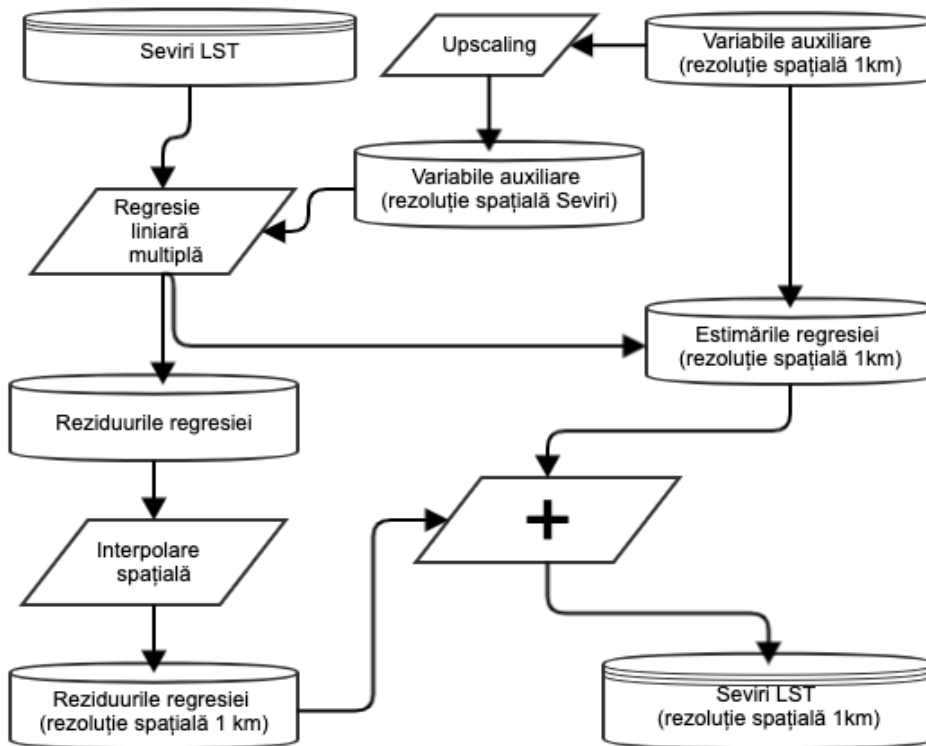


Figura 10: Schemă simplificată a metodei utilizate în îmbunătățirea spațială a datelor medii lunare Seviri LST.

Rezultate

Într-o primă etapă, pentru fiecare lună au fost identificate relațiile statistice dintre Seviri LST și variabilele auxiliare (predictorii PCA). Prin aplicarea metodei stepwise regression de tip retrograd, au fost selectați, pentru fiecare caz în parte (lună), predictorii semnificativi din punct de vedere statistic. După cum se poate observa în tabelul 4, prin utilizarea predictorilor selectați în lunile de vară putem explica și peste 70% din variabilitatea spațială a datelor Seviri LST. Există și unele situații cu corelații mai puțin intense dintre predictorii selectați și variabila analizată (lunile septembrie și octombrie

2017), posibil generate de gradul mai mare de acoperire cu nori care condiționează calitatea datelor LST analizate. De asemenea, în majoritatea cazurilor au fost selectați automat în modelele de regresie toți predictorii, fapt ce sugerează că au fost incluși predictorii optimi în această analiză.

Tabel 4: Predictorii selectați utilizând stepwise regression și coeficienții de determinare R^2 (2016 - 2017).

An și lună	R^2 zi	R^2 noapte	Predictori zi	Predictori noapte
201601	0.584	0.507	PC1 + PC2 + PC3 + PC4 + PC5	PC1 + PC2 + PC4 + PC5
201602	0.609	0.494	PC1 + PC2 + PC3 + PC4 + PC5	PC1 + PC2 + PC3 + PC4 + PC5
201603	0.541	0.732	PC1 + PC2 + PC3 + PC4 + PC5	PC1 + PC2 + PC4 + PC5
201604	0.592	0.680	PC1 + PC2 + PC3 + PC4 + PC5	PC1 + PC2 + PC3 + PC4 + PC5
201605	0.663	0.666	PC1 + PC2 + PC3 + PC4 + PC5	PC1 + PC2 + PC3 + PC4 + PC5
201606	0.736	0.773	PC1 + PC2 + PC3 + PC5	PC1 + PC2 + PC3 + PC4 + PC5
201607	0.874	0.749	PC1 + PC2 + PC3 + PC4 + PC5	PC1 + PC2 + PC3 + PC4 + PC5
201608	0.807	0.763	PC1 + PC2 + PC3 + PC4 + PC5	PC1 + PC2 + PC3 + PC4 + PC5
201609	0.858	0.702	PC1 + PC2 + PC3 + PC4 + PC5	PC1 + PC2 + PC3 + PC4 + PC5
201610	0.793	0.523	PC1 + PC2 + PC3 + PC4 + PC5	PC1 + PC2 + PC3 + PC4 + PC5
201611	0.627	0.419	PC1 + PC2 + PC3 + PC4 + PC5	PC1 + PC2 + PC3 + PC4 + PC5
201612	0.576	0.534	PC1 + PC2 + PC3 + PC4 + PC5	PC1 + PC2 + PC3 + PC4 + PC5
201701	0.554	0.369	PC1 + PC2 + PC3 + PC4 + PC5	PC1 + PC2 + PC3 + PC4 + PC5
201702	0.530	0.487	PC1 + PC2 + PC3 + PC4 + PC5	PC1 + PC2 + PC3 + PC4 + PC5
201703	0.298	0.432	PC1 + PC2 + PC3 + PC4 + PC5	PC1 + PC2 + PC3 + PC4 + PC5
201704	0.516	0.415	PC1 + PC2 + PC3 + PC4 + PC5	PC1 + PC3 + PC4 + PC5
201705	0.567	0.578	PC1 + PC2 + PC3 + PC4 + PC5	PC1 + PC2 + PC3 + PC4
201706	0.773	0.776	PC1 + PC2 + PC4 + PC5	PC1 + PC2 + PC3 + PC4 + PC5
201707	0.645	0.626	PC1 + PC2 + PC3 + PC4 + PC5	PC1 + PC2 + PC3 + PC4 + PC5
201708	0.699	0.589	PC1 + PC2 + PC3 + PC4 + PC5	PC1 + PC2 + PC3 + PC4 + PC5
201709	0.399	0.316	PC1 + PC2 + PC3 + PC4 + PC5	PC1 + PC2 + PC4 + PC5
201710	0.359	0.521	PC1 + PC2 + PC3 + PC4 + PC5	PC1 + PC2 + PC3 + PC5
201711	0.724	0.445	PC1 + PC2 + PC3 + PC4 + PC5	PC1 + PC2 + PC3 + PC4 + PC5
201712	0.559	0.512	PC1 + PC2 + PC3 + PC4 + PC5	PC1 + PC2 + PC3 + PC4 + PC5

Pentru exemplificare, în figura 11 au fost reprezentate datele medii lunare Seviri LST noaptea. Se observă foarte bine zonele cu inversiune termică, care se produc frecvent în lunile de iarnă (ianuarie și februarie) în depresiunile intracarpatică din centrul țării. De asemenea, este vizibilă și insula de căldură urbană din Municipiul București, care este foarte bine conturată în lunile de vară.

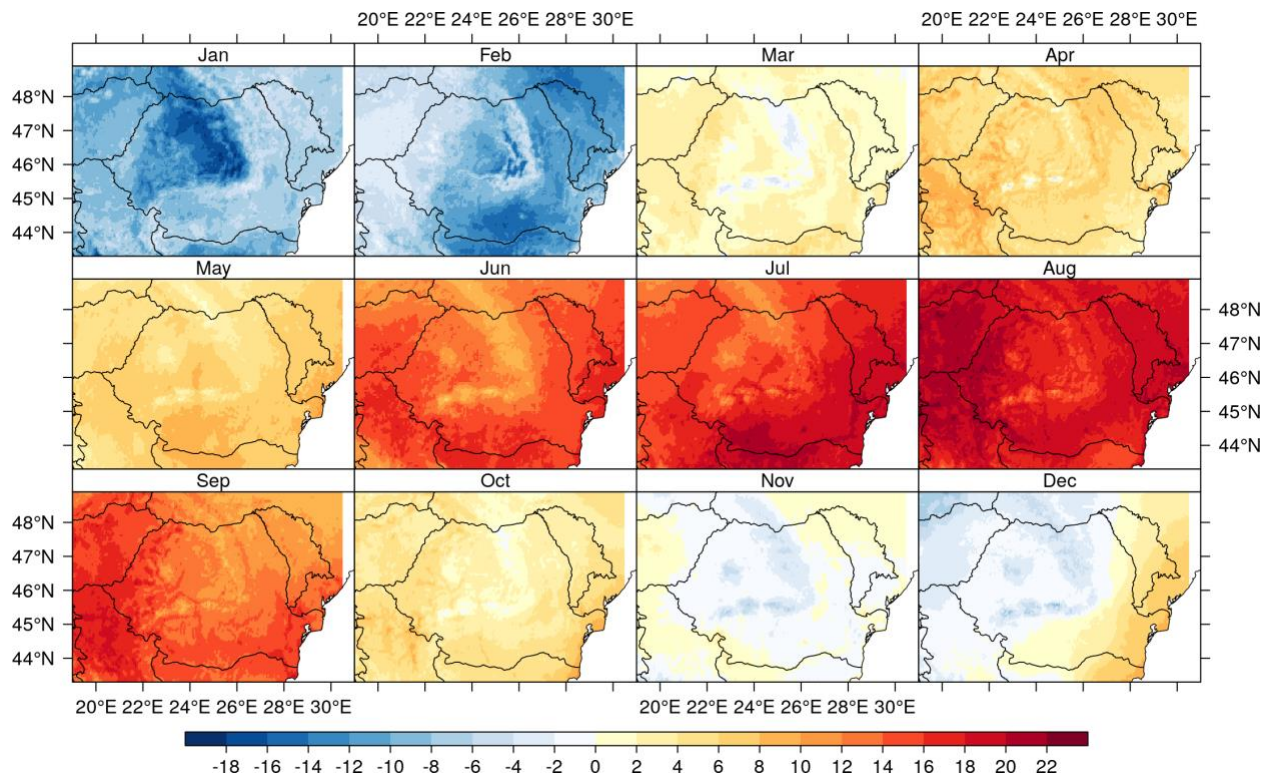


Figura 11: Medii lunare Seviri LST (°C), 2017 noaptea (1000m x 1000m rezoluție spațială).

Datele finale Seviri LST au rezultat prin îmbinarea anomaliilor spațializate la o rezoluție de 1000m x 1000m cu climatologia Seviri LST obținută în etapa precedentă. Anomaliile au fost calculate luându-se în considerare climatologia lunară atât pentru perioada din zi, cât și din noapte (în funcție de ora la care datele LST au fost prelevate), și interpolate spațial cu metoda biliniară. Diferențele au fost calculate utilizându-se datele la rezoluția nativă a datelor Seviri LST (aprox. 4.5 km). La fel ca în etapa precedentă, datele orare au fost clasificate ca făcând parte din zi și noapte prin calcularea unghiului de elevație solară. Astfel au fost obținute pentru perioada de interes (2009 - 2017), 78888 de hărți orare. Datele au fost arhivate în fișiere zilnice multidimensionale NetCDF (3287 de fișiere).

Ciclul diurn al temperaturii suprafeței subiacente din 24 iulie 2016, obținută din datele Seviri LST, având 1000m x 1000m rezoluție spațială, este reprezentat în figura 12. Rezultatele bune ale metodelor de completare a datelor lipsă și de downscaling spațial sunt confirmate de consistența datelor, mai exact de omogenitatea evoluției spațio-temporală a parametrului analizat pe parcursul unei zile. Performanțele metodologiilor aplicate în această etapă sunt demonstrate și prin analiza comparativă a datelor inițiale (figura 13) cu cele finale (figura 12). Se constată că indiferent de gradul de disponibilitate al datelor inițiale, prin utilizarea modelelor de regresie și a predictorilor adecvați, se poate obține un set omogen și complet de date satelitare LST.

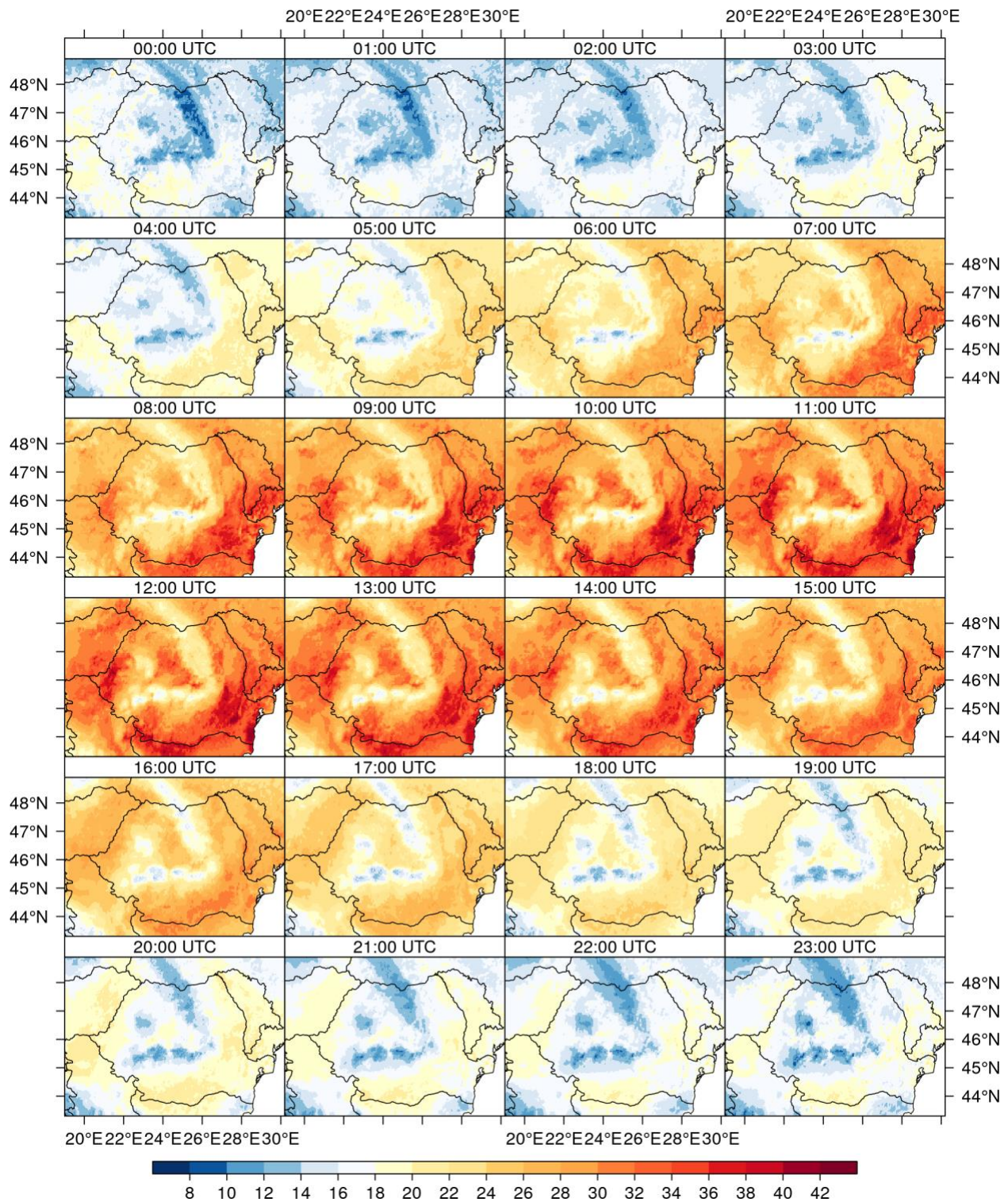


Figura 12: Date orare Seviri LST ($^{\circ}\text{C}$), 25 iulie 2016 (1000m x 1000m rezoluție spațială).

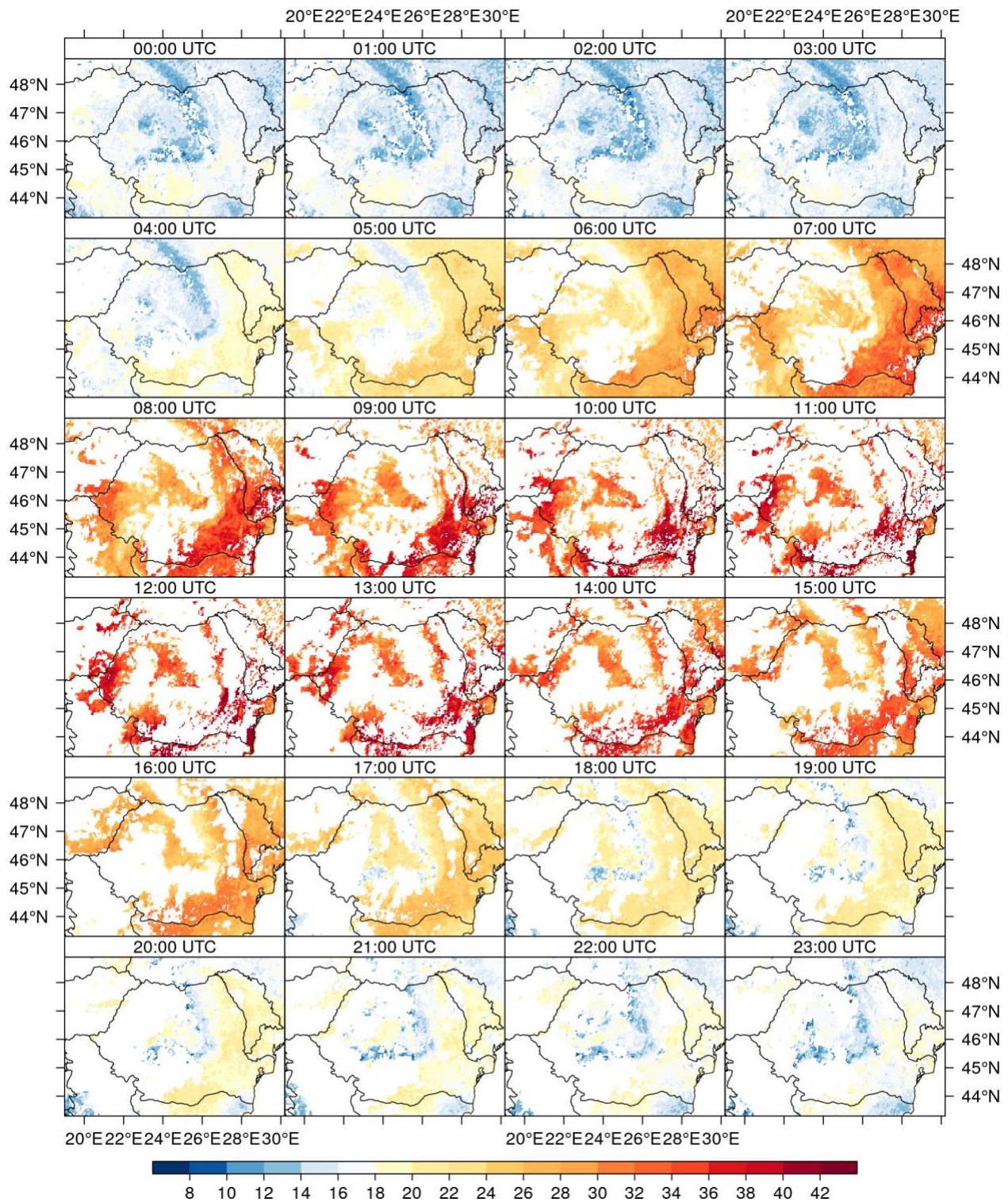


Figura 13: Date orare brute Seviri LST (°C), 25 iulie 2016 (rezoluție spațială aprox.4.5 km).

Concluzii

În această etapă au fost analizate produsele LST existente din punctul de vedere al disponibilității (lungimea seriei de timp), rezoluției spațiale și temporale. De asemenea, au fost testate metodele și instrumentele necesare obținerii și preprocesării produselor selectate. S-a decis utilizarea produselor furnizate de senzorul satelitelui geostaționar MSG (Seviri LST), ca informații de înaltă rezoluție temporală (15 min). Datele LST extrase din informațiile de la sateliții polari Terra/Aqua (produsele zilnice MOD11A1 și MYD11A1) au fost utilizate pentru îmbunătățirea rezoluției spațiale (1km) a datelor Seviri LST orare.

Având în vedere caracteristicile datelor selectate, dar și necesitatea utilizării variabilelor auxiliare în procedura de estimare spațio-temporală, au fost analizate rezultatele a două categorii de metode statistice utilizate cu succes în estimarea spațială a parametrilor meteorologici: regresia multiplă liniară, prin metoda celor mai mici pătrate (MLR), și modelul de regresie aditiv generalizat (GAM). S-au utilizat datele orare LST Seviri disponibile în intervalul 2016 - 2017 (17544 de imagini). Validarea metodelor s-a realizat prin utilizarea unui set de date independent, extras spațial în mod aleatoriu (300 de puncte pentru fiecare pas de timp), care nu au fost utilizate în calibrarea modelelor de regresie.

Modelul GAM, care ține cont și de variabila distanța față de datele lipsă LST (GAM_dist), furnizează estimările cele mai apropiate de datele măsurate, această metodă fiind aplicată în completarea datelor orare lipsă Seviri LST. Rezultatele foarte bune obținute de toate modele analizate sugerează faptul că predictorii selectați satisfac cerințele necesare obținerii unui set complet de date LST.

Îmbunătățirea rezoluției spațiale a datelor Seviri LST s-a realizat printr-o procedură care presupune parcurgerea mai multor etape: realizarea climatologiei lunare a datelor Seviri LST la o rezoluție spațială de 1000m × 1000m și aplicarea metodei anomalilor pentru obținerea setului final de date orare LST, utilizându-se ca referință sintezele LST lunare. Estimările metodei utilizate sunt satisfăcătoare și în situațiile extreme (volum redus de date inițiale Seviri LST), rezultând un set omogen de LST, cu rezoluție temporară orară și cu acoperire spațială completă pentru perioada de interes (2009 - 2017).

Bibliografie

Aalto J, Pirinen P, Heikkinen J, Venäläinen A. Spatial interpolation of monthly climate data for Finland: Comparing the performance of kriging and generalized additive models. *Theoretical and Applied Climatology*. Springer; 2013;112(1-2):99–111.

Amazirh A, Merlin O, Er-Raki S. Including Sentinel-1 radar data to improve the disaggregation of MODIS land surface temperature data. *ISPRS journal of photogrammetry and remote sensing*. Elsevier; 2019;150:11–26.

Appelhans T, Detsch F, Nauss T. Remote: Empirical orthogonal teleconnections in R. *Journal of Statistical Software*. Foundation for Open Access Statistics; 2015;65(i10).

Bartkowiak P, Castelli M, Notarnicola C. Downscaling land surface temperature from Modis dataset with random forest approach over alpine vegetated areas. *Remote Sensing*. Multidisciplinary Digital Publishing Institute; 2019;11(11):1319.

Bechtel B, Zakšek K, Hoshyaripour G. Downscaling land surface temperature in an urban area: A case study for Hamburg, Germany. *Remote Sensing*. Molecular Diversity Preservation International; 2012;4(10):3184–200.

Bivand R, Keitt T, Rowlingson B. Rgdal: Bindings for the 'geospatial' data abstraction library. R package version 1.4-4. 2019;

Büttner G. CORINE land cover and land cover change products. *Land use and land cover mapping in Europe*. Springer; 2014. pp. 55–74.

Combal B, Noel J. Projection of Meteosat Images Into World Geodetic System WGS-84 Matching Spot-vegetation Grid. EUR-OP; 2009.

Coops NC, Duro DC, Wulder MA, Han T. Estimating afternoon Modis land surface temperatures (LST) based on morning Modis overpass, location and elevation information. *International Journal of Remote Sensing*. Taylor & Francis; 2007;28(10):2391–6.

Detsch F, Otte I, Appelhans T, Hemp A, Nauss T. Seasonal and long-term vegetation dynamics from 1-km gims-based NDVI time series at mt. Kilimanjaro, Tanzania. *Remote Sensing of Environment*. Elsevier; 2016a;178:70–83.

Detsch F, Otte I, Appelhans T, Nauss T. A comparative study of cross-product NDVI dynamics in the Kilimanjaro region—a matter of sensor, degradation calibration, and significance. *Remote Sensing*. Multidisciplinary Digital Publishing Institute; 2016b;8(2):159.

Duan S-B, Li Z-L, Leng P. A framework for the retrieval of all-weather land surface temperature at a high spatial resolution from polar-orbiting thermal infrared and passive microwave data. *Remote Sensing of Environment*. Elsevier; 2017;195:107–17.

Dumitrescu A, Birsan M-V, Manea A. Spatio-temporal interpolation of sub-daily (6 h) precipitation over Romania for the period 1975–2010. *International Journal of Climatology*. Wiley Online Library; 2016;36(3):1331–43.

PN-III-P1-1.1-PD-2016-1579 Interpolarea spațio-temporală la scară fină a temperaturii aerului 27 pentru teritoriul României

Dumitrescu A, Birsan M-V, Nita I-A. A Romanian daily high-resolution gridded dataset of snow depth (2005-2015). *Geofizika*. Andrija Mohorovicic Geophysical Institute; 2017;34(2):275-95.

Fox D. Judging Air Quality Model Performance. *Bulletin of the American Meteorological Society*. 1981;62:599-609.

GDAL/OGR contributors. GDAL/OGR geospatial data abstraction software library [Internet]. Open Source Geospatial Foundation; 2019. Available from: <https://gdal.org>

Hastie T, Tibshirani R. *Generalized additive models*. Chapman & Hall/CRC; 1990.

Haylock M, Hofstra N, Klein Tank A, Klok E, Jones P, New M. A european daily high-resolution gridded data set of surface temperature and precipitation for 1950-2006. *Journal of Geophysical Research: Atmospheres*. Wiley Online Library; 2008;113(D20).

Hijmans RJ. *Raster: Geographic data analysis and modeling*. 2019; Available from: <https://CRAN.R-project.org/package=raster>

Hyndman RJ, Fan Y. Sample quantiles in statistical packages. *The American Statistician*. Taylor & Francis; 1996;50(4):361-5.

Keramitsoglou I, Kiranoudis CT, Weng Q. Downscaling geostationary land surface temperature imagery for urban analysis. *IEEE Geoscience and Remote Sensing Letters*. IEEE; 2013;10(5):1253-7.

Kilibarda M, Hengl T, Heuvelink GB, Gräler B, Pebesma E, Perčec Tadić M, et al. Spatio-temporal interpolation of daily temperatures for global land areas at 1 km resolution. *Journal of Geophysical Research: Atmospheres*. Wiley Online Library; 2014;119(5):2294-313.

Krähenmann S, Bissolli P, Rapp J, Ahrens B. Spatial gridding of daily maximum and minimum temperatures in Europe. *Meteorology and Atmospheric Physics*. Springer; 2011;114(3-4):151.

Kwak G-H, Park N-W, Kyriakidis P. Development of an R-based spatial downscaling tool to predict fine scale information from coarse scale satellite products. *The Korean Society of Remote Sensing*; 2018;

Land SAF Project Team (IPMA K. Product user manual land surface temperature (lst), products: LSA-001 (mlst), lsa-050 (mlst-r), lsa-003 (dlst), lsa-002 (elst). version 3.0. EUMETSAT LSA SAF (Satellite Application Facility on Land Surface Analysis); 2018.

Li X, Zhou Y, Asrar GR, Zhu Z. Creating a seamless 1 km resolution daily land surface temperature dataset for urban and surrounding areas in the conterminous united states. *Remote sensing of environment*. Elsevier; 2018;206:84-97.

Mattiuzzi M, Detsch F. MODIS: Acquisition and processing of Modis products. R package version 1.1.5. 2019;

Metz M, Andreo V, Neteler M. A new fully gap-free time series of land surface temperature from MODIS LST data. Remote sensing. Multidisciplinary Digital Publishing Institute; 2017;9(12):1333.

Neteler M. Estimating daily land surface temperatures in mountainous environments by reconstructed Modis LST data. Remote sensing. Molecular Diversity Preservation International; 2010;2(1):333–51.

Park N-W, Kim Y, Kwak G-H. An overview of theoretical and practical issues in spatial downscaling of coarse resolution satellite-derived products. Korean Journal of Remote Sensing. The Korean Society of Remote Sensing; 2019;35(4):589–607.

Pebesma E, Bivand RS. S classes and methods for spatial data: The sp package. R news. 2005;5(2):9–13.

Pereira O, Melfi A, Montes C, Lucas Y. Downscaling of ASTER thermal images based on geographically weighted regression kriging. Remote Sensing. Multidisciplinary Digital Publishing Institute; 2018;10(4):633.

Pfeifroth U, Kothe S, Müller R, Trentmann J, Hollmann R, Fuchs P, et al. Surface radiation data set–Heliosat (Sarah)–edition 2, satellite application facility on climate monitoring. 2017.

Press WH, Teukolsky SA, Vetterling WT, Flannery BP. Numerical recipes in C++. The art of scientific computing. 1992;2:1002.

Reuter HI, Nelson A, Jarvis A. An evaluation of void-filling interpolation methods for SRTM data. International Journal of Geographical Information Science. Taylor & Francis; 2007;21(9):983–1008.

Ruiz-Álvarez M, Alonso-Sarri'a F, Gomariz-Castillo F. Interpolation of instantaneous air temperature using geographical and Modis derived variables with machine learning techniques. Preprints. org; 2019;

Schulz J, Albert P, Behr H-D, Caprion D, Deneke H, Dewitte S, et al. Operational climate monitoring from space: the EUMETSAT Satellite Application Facility on Climate Monitoring (CM-SAF). Atmospheric Chemistry and Physics. 2009;9(5):1687–709.

Sismanidis P, Keramitsoglou I, Kiranoudis C, Bechtel B. Assessing the capability of a downscaled urban land surface temperature time series to reproduce the spatiotemporal features of the original data. Remote Sensing. Multidisciplinary Digital Publishing Institute; 2016;8(4):274.

Szymanowski M, Kryza M, Spallek W. Regression-based air temperature spatial prediction models: An example from Poland. Meteorologische Zeitschrift. E. Schweizerbart'sche Verlagsbuchhandlung; 2013;22(5):577–85.

Thom HCS. Some methods of climatological analysis. World Meteorological Organisation; 1966 p. 53. Report No.: 81.

Venables WN, Ripley BD. Modern applied statistics with S-PLUS. Springer Science & Business Media; 2013.

Vitart F, Balsamo G, Bidlot J, Lang S, Tsonevsky I, Richardson D, et al. Use of ERA5 to Initialize Ensemble Re-forecasts. European Centre for Medium Range Weather Forecasts; 2019.

Wan Z. New refinements and validation of the collection-6 Modis land-surface temperature/emissivity product. Remote sensing of Environment. Elsevier; 2014;140:36–45.

Weiss DJ, Atkinson PM, Bhatt S, Mappin B, Hay SI, Gething PW. An effective approach for gap-filling continental scale remotely sensed time-series. ISPRS Journal of Photogrammetry and Remote Sensing. Elsevier; 2014;98:106–18.

Wood S. Generalized additive models: An introduction with r. Chapman & Hall/CRC; 2006.

Wood SN, Augustin NH. GAMs with integrated model selection using penalized regression splines and applications to environmental modelling. Ecological Modelling. 2002;157(2-3):157–77.

Wu J, Zhong B, Tian S, Yang A, Wu J. Downscaling of urban land surface temperature based on multi-factor geographically weighted regression. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing. IEEE; 2019;12(8):2897–911.

Anexe

Anexa 1 Copernicus Toolbox script utilizat pentru extragerea datelor ERA5 Land (<https://cds.climate.copernicus.eu>)

```
import cdstoolbox as ct
variables = {
    'Skin Temperature': 'skin_temperature',
    #'Eastward Near-Surface Wind': '10m_u_component_of_wind',
}

#https://gist.github.com/matteodefelice/c324f9c72005db8bbb867da658ec6066
@ct.input.dropdown('variable', label='Variable', values=variables.keys())
@ct.input.dropdown('show_year', label = 'Year', default = '2016', values=range(2009,
2018))

@ct.application(title='Retrieve Data Era Land Date')
@ct.output.download()
#def application(variable, show_year):
def download(variable, show_year):

    data = ct.catalogue.retrieve(
        'reanalysis-era5-land',
        {
            'format': 'grib',
            'variable': variables[variable],
            'year': show_year,
            'month': [
                '01', '02', '03',
                '04', '05', '06',
                '07', '08', '09',
                '10', '11', '12'
            ],
            'day': [
                '01', '02', '03',
                '04', '05', '06',
                '07', '08', '09',
                '10', '11', '12',
                '13', '14', '15',
                '16', '17', '18',
                '19', '20', '21',
                '22', '23', '24',
                '25', '26', '27',
                '28', '29', '30',
                '31'
            ],
            'time': [
                '00:00', '01:00', '02:00',
                '03:00', '04:00', '05:00',
                '06:00', '07:00', '08:00',
```

```

        '09:00', '10:00', '11:00',
        '12:00', '13:00', '14:00',
        '15:00', '16:00', '17:00',
        '18:00', '19:00', '20:00',
        '21:00', '22:00', '23:00'
    ],
    'area': [
        '55./15./40/35'
    ]
    #'area'          : [60, -10, 50, 2], # North, West, South, East.
})
return data

```