Using R for analysing spatio-temporal datasets: a satellite-based precipitation case study

Session "R's deliberate role in Earth sciences"

EGU2017-18343, Wien, Austria

Mauricio Zambrano-Bigiarini[1,2]

[1]Universidad de La Frontera, Temuco, Chile

[2]Center for Climate and Resilience Research, Santiago, Chile

mauricio.zambrano @ ufrontera.cl

April 25th, 2017

Outline

Problem description

- Precipitation is a **key driver** of the water and energy cycles.

Problem description

- Precipitation is a **key driver** of the water and energy cycles.
- Traditional (i.e., station-based) representation of the **spatio-temporal** variability of precipitation has several limitations.

Problem description

- Precipitation is a **key driver** of the water and energy cycles.
- Traditional (i.e., station-based) representation of the **spatio-temporal** variability of precipitation has several limitations.
- In the last decades, several satellite-based rainfall estimates (SREs) have provided an **unprecedented opportunity** for improving the spatio-temporal representation of precipitation.

Problem description

- Precipitation is a **key driver** of the water and energy cycles.
- Traditional (i.e., station-based) representation of the **spatio-temporal** variability of precipitation has several limitations.
- In the last decades, several satellite-based rainfall estimates (SREs) have provided an **unprecedented opportunity** for improving the spatio-temporal representation of precipitation.
- State-of-the-art SREs are provided in different **file formats** (e.g., .bin, .nc, .tiff), with different **spatial extents** and different **temporal frequencies** (e.g., half-hourly, 3-hours, daily, monthly).

Problem description

- Precipitation is a **key driver** of the water and energy cycles.
- Traditional (i.e., station-based) representation of the **spatio-temporal** variability of precipitation has several limitations.
- In the last decades, several satellite-based rainfall estimates (SREs) have provided an **unprecedented opportunity** for improving the spatio-temporal representation of precipitation.
- State-of-the-art SREs are provided in different **file formats** (e.g., .bin, .nc, .tiff), with different **spatial extents** and different **temporal frequencies** (e.g., half-hourly, 3-hours, daily, monthly).
- Hydrological models usually require **long time series** (e.g., 30 years) of precipitation to run and explore climate impacts on streamflows.

Problem description

- Precipitation is a **key driver** of the water and energy cycles.
- Traditional (i.e., station-based) representation of the **spatio-temporal** variability of precipitation has several limitations.
- In the last decades, several satellite-based rainfall estimates (SREs) have provided an **unprecedented opportunity** for improving the spatio-temporal representation of precipitation.
- State-of-the-art SREs are provided in different **file formats** (e.g., .bin, .nc, .tiff), with different **spatial extents** and different **temporal frequencies** (e.g., half-hourly, 3-hours, daily, monthly).
- Hydrological models usually require **long time series** (e.g., 30 years) of precipitation to run and explore climate impacts on streamflows.

$\therefore$ it is **computationally challenging** to read and analyse **hundreds/thousands** of station-based time series and SRE files.

My R solution:

1. To develop R scripts to **automatically download** daily SRE files for a **user-defined time period** and **clip** them to the desired **spatial extent** (if necessary).

My R solution:

1. To develop R scripts to **automatically download** daily SRE files for a **user-defined time period** and **clip** them to the desired **spatial extent** (if necessary).
2. To use the `raster` package to **read**, **plot**, and carry out an **EDA**, in order to **detect unexpected problems** (e.g., rotated spatial domains, wrong order of variables in NetCDF files, missing `NA` flags).

My R solution:

1. To develop R scripts to **automatically download** daily SRE files for a **user-defined time period** and **clip** them to the desired **spatial extent** (if necessary).
2. To use the `raster` package to **read**, **plot**, and carry out an **EDA**, in order to **detect unexpected problems** (e.g., rotated spatial domains, wrong order of variables in NetCDF files, missing `NA` flags).
3. To use `raster` along with the `hydroTSM` package to **aggregate** SRE files and rain gauge time series into different **temporal scales** (daily, monthly, seasonal, annual).

My R solution:

1. To develop R scripts to **automatically download** daily SRE files for a **user-defined time period** and **clip** them to the desired **spatial extent** (if necessary).
2. To use the `raster` package to **read**, **plot**, and carry out an **EDA**, in order to **detect unexpected problems** (e.g., rotated spatial domains, wrong order of variables in NetCDF files, missing `NA` flags).
3. To use `raster` along with the `hydroTSM` package to **aggregate** SRE files and rain gauge time series into different **temporal scales** (daily, monthly, seasonal, annual).
4. To use `hydroTSM` along with the `hydroGOF` package to carry out a **point-to-pixel** comparison between ts observed at 366 stations and the corresponding grid cell of each SRE.

My R solution:

1. To develop R scripts to **automatically download** daily SRE files for a **user-defined time period** and **clip** them to the desired **spatial extent** (if necessary).
2. To use the `raster` package to **read**, **plot**, and carry out an **EDA**, in order to **detect unexpected problems** (e.g., rotated spatial domains, wrong order of variables in NetCDF files, missing `NA` flags).
3. To use `raster` along with the `hydroTSM` package to **aggregate** SRE files and rain gauge time series into different **temporal scales** (daily, monthly, seasonal, annual).
4. To use `hydroTSM` along with the `hydroGOF` package to carry out a **point-to-pixel** comparison between ts observed at 366 stations and the corresponding grid cell of each SRE.

Are you curious about specific functions and results? → go to **Spot A.4**.

Outline

Precipitation (rainfall, snow, hail, ...)

- It is a key component of the **water** and **energy** cycles, that contributes to moderate the climate.

Precipitation (rainfall, snow, hail, ...)

- It is a key component of the **water** and **energy** cycles, that contributes to moderate the climate.
- Several **ecosystems** and **economic activities** depend on it, in particular silviculture and agriculture.

Precipitation (rainfall, snow, hail, ...)

- It is a key component of the **water** and **energy** cycles, that contributes to moderate the climate.
- Several **ecosystems** and **economic activities** depend on it, in particular silviculture and agriculture.
- In contrast to other meteorological variables (e.g., Temp), precipitation presents a **low correlation in time and space**. In particular, its distribution might be **fractal** in space and **discontinuous** in time.

Precipitation (rainfall, snow, hail, ...)

- It is a key component of the **water** and **energy** cycles, that contributes to moderate the climate.
- Several **ecosystems** and **economic activities** depend on it, in particular silviculture and agriculture.
- In contrast to other meteorological variables (e.g., Temp), precipitation presents a **low correlation in time and space**. In particular, its distribution might be **fractal** in space and **discontinuous** in time.
- Moreover, local variations of **topography** might have an important effect on the total amount of an event.

Precipitation (rainfall, snow, hail, ...)

- It is a key component of the **water** and **energy** cycles, that contributes to moderate the climate.
- Several **ecosystems** and **economic activities** depend on it, in particular silviculture and agriculture.
- In contrast to other meteorological variables (e.g., Temp), precipitation presents a **low correlation in time and space**. In particular, its distribution might be **fractal** in space and **discontinuous** in time.
- Moreover, local variations of **topography** might have an important effect on the total amount of an event.

∴ The correct assessment of its **amount**, **distribution** and **intensity** it is of utmost importance for the **integrated water resources management** of a basin.

Type of precipitation data

In general, available data on precipitation can be classified in:

1. **Station-based** (**In situ**) only: e.g., raingauges, CRU TS, GPCC, APHRODITE, PREC/L.

Type of precipitation data

In general, available data on precipitation can be classified in:

1. **Station-based (In situ)** only: e.g., raingauges, CRU TS, GPCC, APHRODITE, PREC/L.
2. **Satellite-based** only: e.g., PERSIANN, CMORPH, CHOMPS, etc.

Type of precipitation data

In general, available data on precipitation can be classified in:

1. **Station-based** (**In situ**) only: e.g., raingauges, CRU TS, GPCC, APHRODITE, PREC/L.
2. **Satellite-based** only: e.g., PERSIANN, CMORPH, CHOMPS, etc.
3. **Combination** of *in situ* and satellite: e.g., GPCP, CMAP, TRMM 3B42, etc.

Some limitations of station-based precipitation:

- **Incomplete** time series $\rightarrow$ gap filling (from other incomplete time series).

Some limitations of station-based precipitation:

- **Incomplete** time series $\rightarrow$ gap filling (from other incomplete time series).
- **Low spatial density** of stations in high-elevation areas (installation and maintenance costs), where usually most of the precipitation happens.

Some limitations of station-based precipitation:

- **Incomplete** time series → gap filling (from other incomplete time series).
- **Low spatial density** of stations in high-elevation areas (installation and maintenance costs), where usually most of the precipitation happens.
- **Underestimation** of the precipitation amount in high-elevation areas → high uncertainties in hydrological modelling applications (as input data)

Some limitations of station-based precipitation:

- **Incomplete** time series → gap filling (from other incomplete time series).
- **Low spatial density** of stations in high-elevation areas (installation and maintenance costs), where usually most of the precipitation happens.
- **Underestimation** of the precipitation amount in high-elevation areas → high uncertainties in hydrological modelling applications (as input data)
- Moreover, *in situ* measurements of precipitation are affected by wind, installation errors, and other **systematic and random errors**.

Why using satellite-based rainfall estimates (SREs)?

- SREs were developed to **overcome** many of the **limitations** of *in situ* measurements

Why using satellite-based rainfall estimates (SREs)?

- SREs were developed to **overcome** many of the **limitations** of *in situ* measurements
- Several SREs have become **operational** in last decades, with **quasi-global** spatial coverage and relatively **high temporal and spatial** resolution.

Why using satellite-based rainfall estimates (SREs)?

- SREs were developed to **overcome** many of the **limitations** of *in situ* measurements
- Several SREs have become **operational** in last decades, with **quasi-global** spatial coverage and relatively **high temporal and spatial** resolution.
- SREs have opened **unprecedent opportunities** for hydrological applications in areas with scarce or inexistent data.

Why using satellite-based rainfall estimates (SREs)?

- SREs were developed to **overcome** many of the **limitations** of *in situ* measurements
- Several SREs have become **operational** in last decades, with **quasi-global** spatial coverage and relatively **high temporal and spatial** resolution.
- SREs have opened **unprecedent opportunities** for hydrological applications in areas with scarce or inexistent data.
- Many satellite-based precipitation products **combine** information coming from different satellites (i.e., **multi-satellite**).

Outline

Selected satellite-based rainfall estimates (SREs)

SRE	Full name (with hyperlink)	Latitudinal Coverage	Spatial Resol.	Temporal Coverage	Temporal Resol.	References
CMORPH	NOAA **C**limate Prediction Center (CPC) **MORPH**ing technique	60°N-60°S	0.07°, 0.25°	Dec-2002 - present	half-hourly, 3-hourly, daily	Joyce et al. 2004; CPC-NCEP-NWS-NOAA-USDC 2011
PERSIANN-CDR	PERSIANN **C**limate **D**ata **R**ecord, Version 1 Revision 1	60°N-60°S	0.25°	Jan-1983 - present	daily	Sorooshian et al. 2014; Ashouri et al. 2015
PERSIANN-CCS-adj	**P**recipitation **E**stimation from **R**emotely **S**ensed **I**nformation using **A**rtificial **N**eural **N**etworks	50°N-50°S	0.04°	Jan-2003 - present	daily	Yang et al. 2016; Hong et al. 2004
3B42v7	**T**RMM **M**ulti-satellite **P**recipitation **A**nalysis research product 3B42 Version 7	50°N-50°S	0.25°	Jan-1998 - present	3-hourly, daily	Huffman et al. 2007, 2010
CHIRPSv2	**C**limate **H**azards group **I**nfrared **P**recipitation with **S**tations Version 2.0	50°N-50°S	0.05°	Jan-1981 - present	daily, pentadal, monthly	Funk et al. 2015
MSWEPv1.1	**M**ulti-**S**ource **W**eighted-**E**nsemble **P**recipitation Version 1.1	90°N-90°S	0.25°	Jan-1979 Dec-2014	3-hourly, daily	Beck et al. 2016
PGFv3	**P**rinceton University **G**lobal Meteorological **F**orcing Version 3	17°S-57°S	0.25°	Jan-1979 Dec-2010	daily	Peng et al. 2016; Sheffield et al. 2006

Outline

Comparison SREs vs rain gauges

Procedure to compare SRE against rain gauge data:

Comparison SREs vs rain gauges

Procedure to compare SRE against rain gauge data:

1. **Download** satellite images for each selected SRE.

Comparison SREs vs rain gauges

Procedure to compare SRE against rain gauge data:

1. **Download** satellite images for each selected SRE.
2. **Re-project** and apply a zonal mask.

Comparison SREs vs rain gauges

Procedure to compare SRE against rain gauge data:

1. **Download** satellite images for each selected SRE.
2. **Re-project** and apply a zonal mask.
3. To **aggregate** raster files into different temporal resolutions (daily $\rightarrow$ monthly $\rightarrow \rightarrow$ annual).

Comparison SREs vs rain gauges

Procedure to compare SRE against rain gauge data:

1. **Download** satellite images for each selected SRE.
2. **Re-project** and apply a zonal mask.
3. To **aggregate** raster files into different temporal resolutions (daily $\rightarrow$ monthly $\rightarrow \rightarrow$ annual).
4. Point-to-pixel comparison: SRE vs raingauge (Thiemig et al., 2012), using continuous and categorical **performance indices**.

All the previous steps were carried out with **R** (R Core Team, 2016), **"the"** open source software for statistic computations and graphics.

Outline

1) Automatic downloading of SRE files

```r
library(hydroTSM) # dip

PERSIANN_CCS_adj.drty.out <- "mypath"

# ftp folder location on the CHIRPS server.  0.05° spatial resolution
ftp.drty <-  "http://www.climatedatalibrary.cl/SOURCES/.UCIrvine/.CHRS/.PERSIANN-CCS-adj/.Daily/.precip/T"

# Example of the name of a single file
fname.example <- "http://www.climatedatalibrary.cl/SOURCES/.UCIrvine/.CHRS/.PERSIANN-CCS-adj/.Daily/.precip/T/20673.5/VALUE/Y/-56.98/-16.02/RANGEEDGES/%5

CDL.StartDate <- "1960-01-01" # Initial date for the Climate Data Library
Date.Ini <- "2003-01-01"
Date.Fin <- "2015-12-31"
days <- dip(Date.Ini, Date.Fin)

for (day in days) {
  # For the Data Climate Library, 2003-Jan-01 is the day 15706.5 since 1960-Jan-01
  ndays  <- dip(CDL.StartDate, day,,"nmbr")  - 0.5
  fname <- paste0(ftp.drty, "/", ndays, "/VALUE/Y/-56.98/-16.02/RANGEEDGES/%5BX/Y/%5D/data.tiff?filename=data", format(as.Date(day), "%Y%m%d"), ".tiff")

  message("getting file '", basename(fname), "'")
  fname.out <- paste0(PERSIANN_CCS_adj.drty.out, "/", basename(fname))
  download.file(url=fname, destfile=fname.out, method="auto")

  # renaming the file
  fname.out.new <- paste0(PERSIANN_CCS_adj.drty.out, "/PERSIANN_CCS_adj_", format(as.Date(day), "%Y_%m_%d"), ".tif" )
  file.copy(from=fname.out, to=fname.out.new, overwrite=TRUE)

  # delete the original .gz file (~ 3 or 4 Mb)
  unlink(fname.out)
} # FOR 'day' end
```

2) Main `raster` functions used in the analysis - I

- **raster**: it reads **any single raster file** supported by GDAL (and the `ncdf4` pkg) into a `RasterLayer` object.
  $x <- raster("path\_to\_my\_file")$

2) Main raster functions used in the analysis - I

- **raster**: it reads **any single raster file** supported by GDAL (and the ncdf4 pkg) into a RasterLayer object.
  $x <- raster("path\_to\_my\_file")$
- **stack**: it reads **all the file(s) stored in a directory** into a (multi-band) RasterStack object.
  $s <- stack("path\_to\_my\_directory")$

2) Main `raster` functions used in the analysis - I

- **raster**: it reads **any single raster file** supported by GDAL (and the `ncdf4` pkg) into a `RasterLayer` object.
  $x < -raster(''path\_to\_my\_file'')$

- **stack**: it reads **all the file(s) stored in a directory** into a (multi-band) `RasterStack` object.
  $s < -stack(''path\_to\_my\_directory'')$

- **brick**: it reads a **single (multi-band) file** into a (multi-layer) `RasterBrick` object. **Processing time should be shorter** when using a `RasterBrick` object.
  $b < -brick(''path\_to\_my\_multiband\_file'')$

2) Main `raster` functions used in the analysis - I

- **raster**: it reads **any single raster file** supported by GDAL (and the `ncdf4` pkg) into a `RasterLayer` object.
  $x < -raster("path\_to\_my\_file")$
- **stack**: it reads **all the file(s) stored in a directory** into a (multi-band) `RasterStack` object.
  $s < -stack("path\_to\_my\_directory")$
- **brick**: it reads a **single (multi-band) file** into a (multi-layer) `RasterBrick` object. **Processing time should be shorter** when using a `RasterBrick` object.
  $b < -brick("path\_to\_my\_multiband\_file")$
- **plot**: it plots **any raster object** already read with `raster/stack/brick`.
  $plot(x)$; $plot(s)$; $plot(b)$

2) Main `raster` functions used in the analysis - II

- **crop**: it returns a geographic subset of a `Raster*` object as specified by an `Extent` object.
  $e < -extent(-160, 10, 30, 60)$
  $rc < -crop(x, e)$

2) Main `raster` functions used in the analysis - II

- **crop**: it returns a geographic subset of a `Raster*` object as specified by an `Extent` object.
  $e < -extent(-160, 10, 30, 60)$
  $rc < -crop(x, e)$

- **extract**: it extracts values from a `Raster*` object at the locations of other spatial data.
  $stations < -readOGR(".", "raingauges")$
  $rp < -extract(x, stations)$

2) Main `raster` functions used in the analysis - II

- **crop**: it returns a geographic subset of a `Raster*` object as specified by an `Extent` object.
  $e < -extent(-160, 10, 30, 60)$
  $rc < -crop(x, e)$

- **extract**: it extracts values from a `Raster*` object at the locations of other spatial data.
  $stations < -readOGR(".", "raingauges")$
  $rp < -extract(x, stations)$

- **writeRaster**: it writes a `Raster*` object into any format supported by GDAL (and the `ncdf4` pkg).
  $x < -writeRaster(x, filename = "my\_file.tif", format = "GTiff")$

3) Main hydroTSM functions used in the analysis

- **daily2monthly**: it transforms a daily (sub-daily or weekly) regular time series into a monthly one.
  $data(SanMartinoPPts)$
  $d < -SanMartinoPPts$
  $m < -daily2monthly(d, FUN = sum, na.rm = TRUE)$

3) Main `hydroTSM` functions used in the analysis

- **daily2monthly**: it transforms a daily (sub-daily or weekly) regular time series into a monthly one.
  $data(SanMartinoPPts)$
  $d < -SanMartinoPPts$
  $m < -daily2monthly(d, FUN = sum, na.rm = TRUE)$

- **daily2annual**: it transforms a (sub)daily/monthly (weekly and quarterly) regular time series into an annual one.
  $a < -daily2annual(d, FUN = sum, na.rm = TRUE)$

3) Main `hydroTSM` functions used in the analysis

- **daily2monthly**: it transforms a daily (sub-daily or weekly) regular time series into a monthly one.
  $data(SanMartinoPPts)$
  $d < -SanMartinoPPts$
  $m < -daily2monthly(d, FUN = sum, na.rm = TRUE)$

- **daily2annual**: it transforms a (sub)daily/monthly (weekly and quarterly) regular time series into an annual one.
  $a < -daily2annual(d, FUN = sum, na.rm = TRUE)$

- **dm2seasonal**: it computes a seasonal value for every year of a sub-daily/daily/weekly/monthly time series.
  $dm2seasonal(d, FUN = sum, season = "DJF")$

4) Continuous performance indices (hydroGOF)

Modified Kling-Gupta efficiency (KGE')

It was used along with its three individual components; linear correlation (r), bias (β) and variability (γ); to identify possible sources of **systematic errors** in each SRE.

4) Continuous performance indices (hydroGOF)

Modified Kling-Gupta efficiency (KGE')

It was used along with its three individual components; linear correlation (r), bias (β) and variability (γ); to identify possible sources of **systematic errors** in each SRE.

1. $KGE' = 1 - \sqrt{(r-1)^2 + (\beta-1)^2 + (\gamma-1)^2}$

4) Continuous performance indices (hydroGOF)

Modified Kling-Gupta efficiency (KGE')

It was used along with its three individual components; linear correlation (r), bias (β) and variability (γ); to identify possible sources of **systematic errors** in each SRE.

1. $KGE' = 1 - \sqrt{(r-1)^2 + (\beta-1)^2 + (\gamma-1)^2}$
2. $r = \frac{Cov(S,O)}{\sigma_S \cdot \sigma_O}$

4) Continuous performance indices (hydroGOF)

Modified Kling-Gupta efficiency (KGE')

It was used along with its three individual components; linear correlation (r), bias (β) and variability (γ); to identify possible sources of **systematic errors** in each SRE.

1. $KGE' = 1 - \sqrt{(r-1)^2 + (\beta-1)^2 + (\gamma-1)^2}$
2. $r = \frac{Cov(S,O)}{\sigma_S \cdot \sigma_O}$
3. $\beta = \frac{\mu_s}{\mu_o}$

4) Continuous performance indices (hydroGOF)

Modified Kling-Gupta efficiency (KGE')

It was used along with its three individual components; linear correlation (r), bias (β) and variability (γ); to identify possible sources of **systematic errors** in each SRE.

1. $KGE' = 1 - \sqrt{(r-1)^2 + (\beta - 1)^2 + (\gamma - 1)^2}$
2. $r = \frac{Cov(S,O)}{\sigma_S \cdot \sigma_O}$
3. $\beta = \frac{\mu_s}{\mu_o}$
4. $\gamma = \frac{CV_s}{CV_o} = \frac{\sigma_s / \mu_s}{\sigma_o / \mu_o}$

4) Continuous performance indices (hydroGOF)

Modified Kling-Gupta efficiency (KGE')

It was used along with its three individual components; linear correlation (r), bias (β) and variability (γ); to identify possible sources of **systematic errors** in each SRE.

1. $KGE' = 1 - \sqrt{(r-1)^2 + (\beta-1)^2 + (\gamma-1)^2}$
2. $r = \frac{Cov(S,O)}{\sigma_S \cdot \sigma_O}$
3. $\beta = \frac{\mu_s}{\mu_o}$
4. $\gamma = \frac{CV_s}{CV_o} = \frac{\sigma_s/\mu_s}{\sigma_o/\mu_o}$

where:

- S: Satellite-based precipitation values, [mm].
- O: Precipitation values observed at the raingauge, [mm].

5) Categorical performance indices (hydroGOF)

Rainfall event	Intensity (i), [mm d^{-1}]
No rain	[0 , 1)
Light rain	[1 , 5)
Moderate rain	[5 , 20)
Heavy rain	[20 , 40)
Violent rain	≥ 40

Satellite-product	Observed rainfall		
	Yes	No	**Total**
Yes	Hit (H)	False Alarm (FA)	$H + FA$
No	Miss(M)	Correct Negative (CN)	$M + CN$
Total	$H + M$	$FA + CN$	Ne

5) Categorical performance indices (hydroGOF)

Rainfall event	**Intensity (i)**, [mm d^{-1}]
No rain	[0 , 1)
Light rain	[1 , 5)
Moderate rain	[5 , 20)
Heavy rain	[20 , 40)
Violent rain	≥ 40

Satellite-product	**Observed rainfall**		
	Yes	No	**Total**
Yes	Hit (H)	False Alarm (FA)	$H + FA$
No	Miss(M)	Correct Negative (CN)	$M + CN$
Total	$H + M$	$FA + CN$	Ne

1. **Percent correct**: $PC = \frac{H+CN}{Ne}$

5) Categorical performance indices (hydroGOF)

Rainfall event	Intensity (i), [mm d^{-1}]
No rain	[0 , 1)
Light rain	[1 , 5)
Moderate rain	[5 , 20)
Heavy rain	[20 , 40)
Violent rain	$\geq$ 40

Satellite-product	Observed rainfall		
	Yes	No	**Total**
Yes	Hit (H)	False Alarm (FA)	$H + FA$
No	Miss(M)	Correct Negative (CN)	$M + CN$
Total	$H + M$	$FA + CN$	Ne

1. **Percent correct**: $PC = \frac{H+CN}{Ne}$
2. **Probability of detection**: $POD = \frac{H}{H+M}$

5) Categorical performance indices (hydroGOF)

Rainfall event	Intensity (i), [mm d^{-1}]
No rain	[0 , 1)
Light rain	[1 , 5)
Moderate rain	[5 , 20)
Heavy rain	[20 , 40)
Violent rain	$\geq$ 40

Satellite-product	Observed rainfall		
	Yes	No	**Total**
Yes	Hit (H)	False Alarm (FA)	$H + FA$
No	Miss(M)	Correct Negative (CN)	$M + CN$
Total	$H + M$	$FA + CN$	Ne

1. **Percent correct**: $PC = \frac{H+CN}{Ne}$
2. **Probability of detection**: $POD = \frac{H}{H+M}$
3. **False alarm ratio**: $FAR = \frac{FA}{H+FA}$

5) Categorical performance indices (hydroGOF)

Rainfall event	Intensity (i), [mm d^{-1}]
No rain	[0 , 1)
Light rain	[1 , 5)
Moderate rain	[5 , 20)
Heavy rain	[20 , 40)
Violent rain	$\geq$ 40

Satellite-product	Observed rainfall		
	Yes	No	**Total**
Yes	Hit (H)	False Alarm (FA)	$H + FA$
No	Miss(M)	Correct Negative (CN)	$M + CN$
Total	$H + M$	$FA + CN$	Ne

1. **Percent correct**: $PC = \frac{H+CN}{Ne}$
2. **Probability of detection**: $POD = \frac{H}{H+M}$
3. **False alarm ratio**: $FAR = \frac{FA}{H+FA}$
4. **Equitable threat score**: $ETS = \frac{H-H_e}{(H+F+M)-H_e}$

5) Categorical performance indices (hydroGOF)

Rainfall event	Intensity (i), [mm d^{-1}]
No rain	[0 , 1)
Light rain	[1 , 5)
Moderate rain	[5 , 20)
Heavy rain	[20 , 40)
Violent rain	$\geq$ 40

Satellite-product	**Observed rainfall**		
	Yes	No	**Total**
Yes	Hit (H)	False Alarm (FA)	$H + FA$
No	Miss(M)	Correct Negative (CN)	$M + CN$
Total	$H + M$	$FA + CN$	Ne

1. **Percent correct**: $PC = \frac{H+CN}{Ne}$
2. **Probability of detection**: $POD = \frac{H}{H+M}$
3. **False alarm ratio**: $FAR = \frac{FA}{H+FA}$
4. **Equitable threat score**: $ETS = \frac{H-H_e}{(H+F+M)-H_e}$
5. **Frequency bias**: $fBias = \frac{H+F}{H+M}$

Outline

Results

You can see the results of this work on **Thursday 27th, 16:00 hrs**:

- **Session**: *HS7.1/AS1.11/NH1.15/NP10.11 - Precipitation: from measurement to modelling and application in catchment hydrology* (co-organized), **room B**.
- **EGU2017-10425**: *Assessing the temporal and spatial performance of satellite-based rainfall estimates across the complex topographical and climatic gradients of Chile*.

Outline

Ongoing work

- To upload to CRAN the new stable version of `hydroTSM` package, with many new features and source code on Github.
- To upload to CRAN the new stable version of `hydroGOF` package, with many new functions and source code on Github.
- To release the **first beta** of a new package (under-development) for automatic processing of different SRE files.

References I

Ashouri, H., Hsu, K.L., Sorooshian, S., Braithwaite, D.K., Knapp, K.R., Cecil, L.D., Nelson, B.R., Prat, O.P., 2015. PERSIANN-CDR: Daily precipitation climate data record from multisatellite observations for hydrological and climate studies. Bulletin of the American Meteorological Society 96, 69–83. doi:10.1175/BAMS-D-13-00068.1.

Beck, H.E., van Dijk, A.I.J.M., Levizzani, V., Schellekens, J., Miralles, D.G., Martens, B., de Roo, A., 2016. MSWEP: 3-hourly 0.25° global gridded precipitation (1979-2015) by merging gauge, satellite, and reanalysis data. Hydrology and Earth System Sciences Discussions , 1doi:10.5194/hess-2016-236.

CPC-NCEP-NWS-NOAA-USDC, 2011. NOAA CPC Morphing Technique (CMORPH) Global Precipitation Analyses. Technical Report. Boulder CO. doi:10.5065/D6CZ356W. [Last Accessed: 25.Jan.2016].

Funk, C., Peterson, P., Landsfeld, M., Pedreros, D., Verdin, J., Shukla, S., Husak, G., Rowland, J., Harrison, L., Hoell, A., Michaelsen, J., 2015. The climate hazards infrared precipitation with stations-a new environmental record for monitoring extremes. Sci Data 2, 150066. doi:10.1038/sdata.2015.66.

Hijmans, R.J., 2016. raster: Geographic Data Analysis and Modeling. URL: https://CRAN.R-project.org/package=raster. r package version 2.5-8.

Hong, Y., Hsu, K.L., Sorooshian, S., Gao, X., 2004. Precipitation estimation from remotely sensed imagery using an artificial neural network cloud classification system. Journal of Applied Meteorology 43, 1834–1853.

References II

Huffman, G.J., Adler, R.F., Bolvin, D.T., Gu, G., Nelkin, E.J., Bowman, K.P., Hong, Y., Stocker, E.F., Wolff, D.B., 2007. The TRMM multisatellite precipitation analysis (TMPA): Quasi-global, multiyear, combined-sensor precipitation estimates at fine scales. Journal of Hydrometeorology 8, 38. doi:10.1175/JHM560.1.

Huffman, G.J., Adler, R.F., Bolvin, D.T., Nelkin, E.J., 2010. The TRMM multi-satellite precipitation analysis (TMPA), in: Gebremichael, M., Hossain, F. (Eds.), Satellite Rainfall Applications for Surface Hydrology. Springer Dordrecht Heidelberg, London New York, pp. 3–22. doi:10.1007/978-90-481-2915-7_1.

Joyce, R.J., Janowiak, J.E., Arkin, P.A., Xie, P., 2004. CMORPH: A method that produces global precipitation estimates from passive microwave and infrared data at high spatial and temporal resolution. Journal of Hydrometeorology 5, 487–503. doi:10.1175/1525-7541(2004)005<0487:CAMTPG>2.0.CO;2.

Peng, L., Sheffield, J., Verbist, K.M.J., 2016. Merging station observations with large-scale gridded data to improve hydrological predictions over Chile, in: 2016 AGU Fall Meeting Abstract, 12-16 December 2016, San Francisco, CA, USA.

R Core Team, 2016. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing. Vienna, Austria. URL: https://www.R-project.org/.

Sheffield, J., Goteti, G., Wood, E.F., 2006. Development of a 50-year high-resolution global dataset of meteorological forcings for land surface modeling. Journal of Climate 19, 3088. doi:10.1175/JCLI3790.1.

References III

Sorooshian, S., Hsu, K., Braithwaite, D., Ashouri, H., NOAA CDR Program , 2014. NOAA Climate Data Record (CDR) of Precipitation Estimation from Remotely Sensed Information using Artificial Neural Networks (PERSIANN-CDR), Version 1 Revision 1. [2003-2014]. Technical Report. NOAA National Centers for Environmental Information. doi:10.7289/V51V5BWQ. [access date: 30-Jan-2016].

Thiemig, V., Rojas, R., Zambrano-Bigiarini, M., Levizzani, V., De Roo, A., 2012. Validation of satellite-based precipitation products over sparsely gauged African river basins. Journal of Hydrometeorology 13, 1760–1783. doi:10.1175/JHM-D-12-032.1.

Yang, Z., Hsu, K., Sorooshian, S., Xu, X., Braithwaite, D., Verbist, K.M.J., 2016. Bias adjustment of satellite-based precipitation estimation using gauge observations-a case study in Chile. Journal of Geophysical Research: Atmospheres doi:10.1002/2015JD024540.

Zambrano-Bigiarini, M., 2016a. hydroGOF: Goodness-of-fit functions for comparison of simulated and observed hydrological time series. URL: http://CRAN.R-project.org/package=hydroGOF. R package version 0.4-0 [under-development].

Zambrano-Bigiarini, M., 2016b. hydroTSM: Time Series Management, Analysis and Interpolation for Hydrological Modelling. URL: http://CRAN.R-project.org/package=hydroTSM. r package version 0.5-0 [under-development].