HELLO, WELCOME

About Me:

Nasrin Fathollahzadeh Attar

PhD of Water Resources Engineering, Interested in Stochastic and Time-series Models, Numerical Modelling, Hydroinformatics, Data-driven Techniques

@attar.nasrin91

@NasrinAttar

Nasrin Fathollahzadeh Attar

Nasrin Fathollahzadeh Attar

FOLLOWING

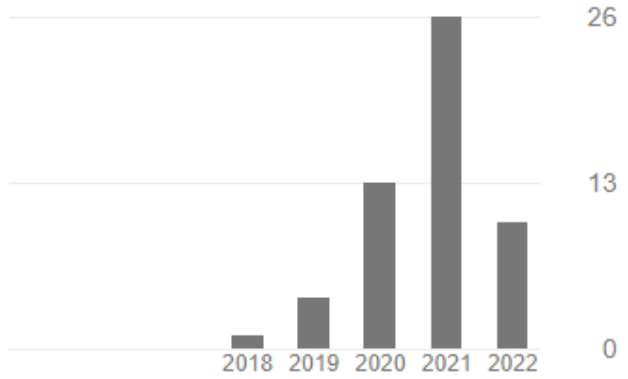
Ph.D. of Water Resources Engineering, Faculty of Agriculture, Urmia University, Urmia, IRAN
Verified email at urmia.ac.ir - Homepage

Stochastic Models Hydroinformatics Time series Modelling R Statistical Hydrology

TITLE	CITED BY	YEAR
On the reliability of soft computing methods in the estimation of dew point temperature: The case of arid regions of Iran NF Attar, K Khalili, J Behmanesh, N Khanmohammadi Computers and electronics in agriculture 153, 334-346	18	2018
Physicochemical parameters data assimilation for efficient improvement of water quality index prediction: Comparative assessment of a noise suppression hybridization approach M Rezaie-Balf, NF Attar, A Mohammadzadeh, MA Murti, AN Ahmed, ... Journal of Cleaner Production 271, 122576	15	2020

Cited by

	All	Since 2017
Citations	54	54
h-index	4	4
i10-index	3	3

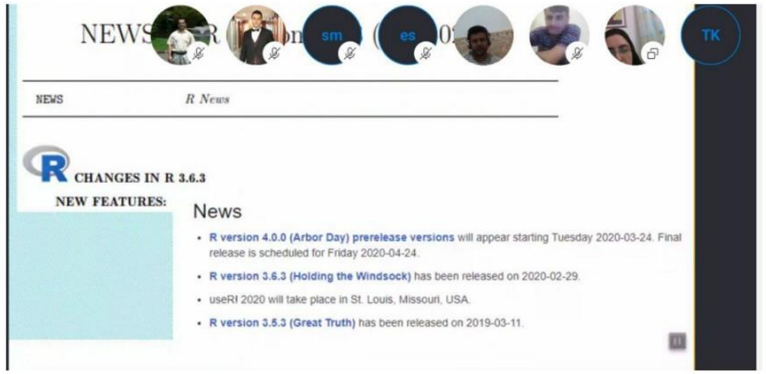


n.fatolahzadeh@urmia.ac.ir


THE UNIVERSITY OF ADELAIDE


THE UNIVERSITY
OF ADELAIDE
adelaide.edu.au


حضور میهمانان محترم را در پژوهشکده
مطالعات دریاچه ارومیه گرامی می داریم.


NEWS
R News
CHANGES IN R 3.6.3
NEW FEATURES:
News
R version 4.0.0 (Arbor Day) prerelease versions will appear starting Tuesday 2020-03-24. Final release is scheduled for Friday 2020-04-24.
R version 3.6.3 (Holding the Windsock) has been released on 2020-02-29.
useR! 2020 will take place in St. Louis, Missouri, USA.
R version 3.5.3 (Great Truth) has been released on 2019-03-11.

Rladies Urmia:

- Email address: urmia@rladies.org
- LinkedIn: R-Ladies Urmia
- Twitter: @RladiesUrmia
- Meetup: https://www.meetup.com/rladies-Urmia/
- YouTube: @RladiesUrmia

Our Team

https://user2021.r-project.org/about/global-team/

Meet Our Team

Fortune Teller

VAHID NAGHSHIN, PH. D

Senior Data Scientist

NASRIN ATTAR, PH. D

Environmental Postdoctoral Research Fellow
R-Ladies Urmia Founder

MARYAM ALIZADEH

Data Analyst
R-Ladies Urmia Co-Founder

HOMAYOON KHADIVI

AI/ML Cloud Technologist
Mentor at DeepLearning.ai

Alibaba Cloud

Global AI Innovation Challenge Series 2021

Intelligent Weather Forecast for Better Life

Innovation Awards:

Siddharth_Deshpande

Clouders

Fortune Teller

Beau&#39

G Force

Tech Phantoms

TeamAS

Sam&Ben

THE APPLICATION OF R IN HYDROLOGY

NASRIN FATHOLLAHZADEH ATTAR

Outline:

- ❑ Environmental Data science?
- ❑ Computational Hydrology?
- ❑ Hydroinformatics?
- ❑ What is R?
- ❑ Why R?
- ❑ Important Communities?
- ❑ R Pubs?
- ❑ JOSS?
- ❑ Hydrological R packages?
- ❑ Workflow of R in Hydrology
- ❑ Data Retrieval packages in R
- ❑ Data tidying packages in R
- ❑ Hydrologic Time Series Analysis tools in R
- ❑ Machine Learning tools in R
- ❑ Spatial Analysis Projects and packages (GIS in R)

❑Environmental Data science?

- **Earth and environmental science, derived data technologies provide comprehensive information about the earth system. Research on environmental science has a long tradition because of continuously changing phenomena.**

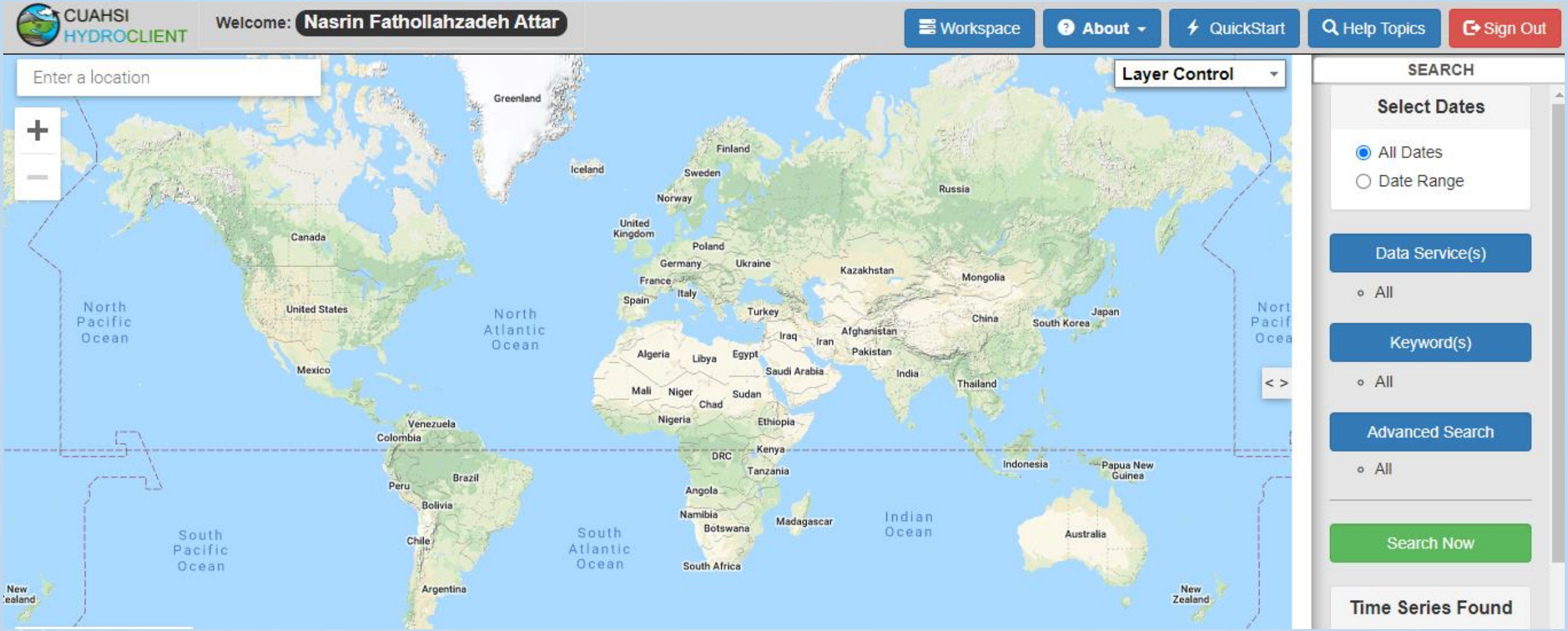
- **Recent theoretical developments have revealed that to deal with extensive environmental observations, data-driven methods such as machine learning and deep learning methods can be applied.**

❑ Computational Hydrology?

- **The Computational Hydrology group develops tools to simulate and investigate the planetary hydrological cycle, applied to many hydrologic research projects.**

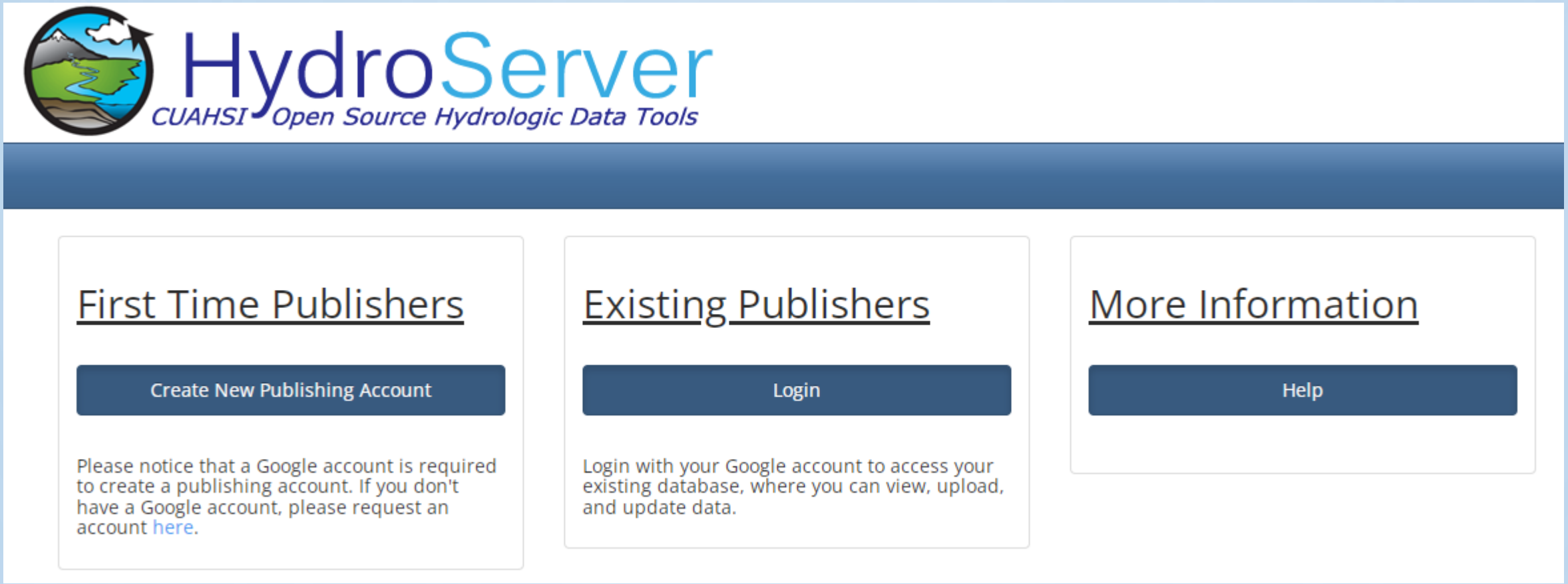
- **The computational hydrology group develops simulation and prediction tools.**

- **Computational hydrology is a discipline that helps researchers to get data, do some preprocessing and data cleaning, analyses big data, model them in specific workflows, document them, visualize and share them in specific web-based clouds such as ….**

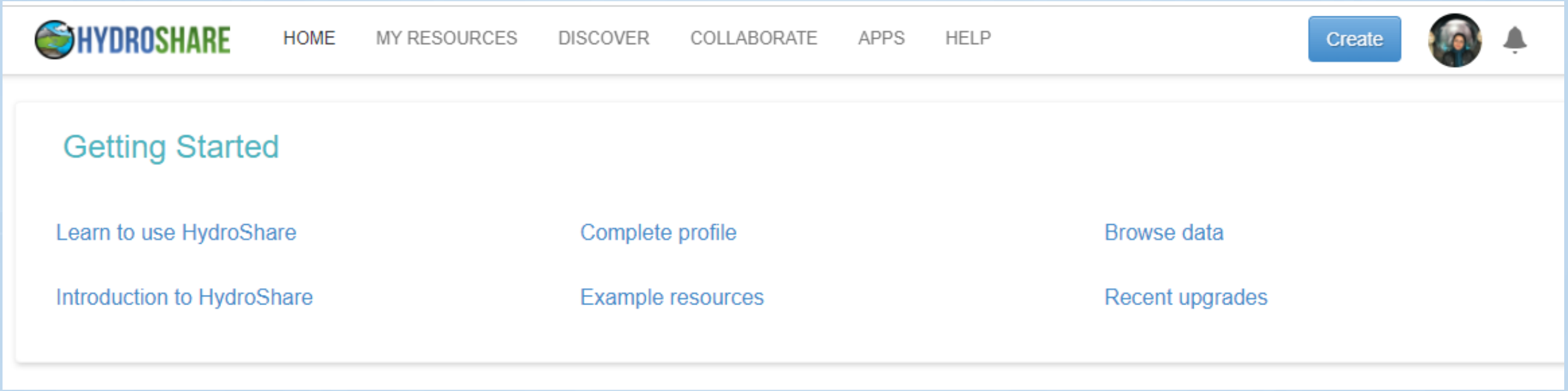


Hydroclient

https://data.cuahsi.org/Home/Index



Hydro Server https://hydroserver.cuahsi.org/



Hydro Share https://www.hydroshare.org/home/

❑Hydroinformatics?

- Hydroinformatics primarily focuses on black-box methods to address water-related problems using famous data-driven techniques. Hydro informatics consists of two words of "Hydro" and "information," which carries us into the water world and gives us the knowledge and information using some tools and technologies

- From this point of view, the new term of "Data science" and its frameworks comes to mind.

- In hydrology, data science plays a primary role, especially in subfields such as flood, precipitation, groundwater, water quality modeling, and prediction.

- In addition, to apply different data science algorithms, programming languages are essential for computational hydrology and Hydroinformatics.

Free, open-source software (FOSS)

Different programming languages are used in hydrology, such as Fortran, C, C++, Java Matlab, Python, R, Julia.

https://github.com/raoulcollenteur/Python-Hydrology-Tools

We choose R

R is free & open source software (FOSS)

❑What is R?

- **R was created by Ross Ihaka and Robert Gentleman at the University of Auckland, New Zealand**
- **R first appeared in 1993**
- **Currently developed by the R Development Core Team**

Ross Ihaka
(New Zealand Statistician)

Robert Gentleman
(Canadian Statistician)

R: A Language for Data Analysis and Graphics

Ross IHAKA and Robert GENTLEMAN

In this article we discuss our experience designing and implementing a statistical computing language. In developing this new language, we sought to combine what we felt were useful features from two existing computer languages. We feel that the new language provides advantages in the areas of portability, computational efficiency, memory management, and scoping.

Key Words: *Computer language; Statistical computing.*

1. INTRODUCTION

This article discusses some issues involved in the design and implementation of a computer language for statistical data analysis. Our experience with these issues occurred while developing such a language. The work has been heavily influenced by two existing languages—Becker, Chambers, and Wilks' S (1985) and Steel and Sussman's Scheme (1975). We felt that there were strong points in each of these languages and that it would be interesting to see if the strengths could be combined. The resulting language is very similar in appearance to S, but the underlying implementation and semantics are derived from Scheme. In fact, we implemented the language by first writing an interpreter for a Scheme subset and then progressively mutating it to resemble S.

Ross Ihaka and Robert Gentleman

❑Why R?

Open Source

Advanced Statistical language

Outstanding Graphs

Flexible ‘n’ Fun

Extremely Comprehensive

Support Extensions

Vast Community

Relates to other languages

THINGS TO KNOW

You can’t “learn R”

You learn the basics

And learn how to learn more

❑Important Communities?

Communities

https://www.r-bloggers.com/
https://stackoverflow.com/questions/tagged/r
https://community.rstudio.com/
https://stats.stackexchange.com/questions/tagged/r
https://stat.ethz.ch/mailman/listinfo/r-help
https://r-forge.r-project.org/forum/forum.php?forum_id=78&group_id=34
https://rviews.rstudio.com/
https://r.789695.n4.nabble.com/
https://rladies.org/
#rstats

❑R Pubs?

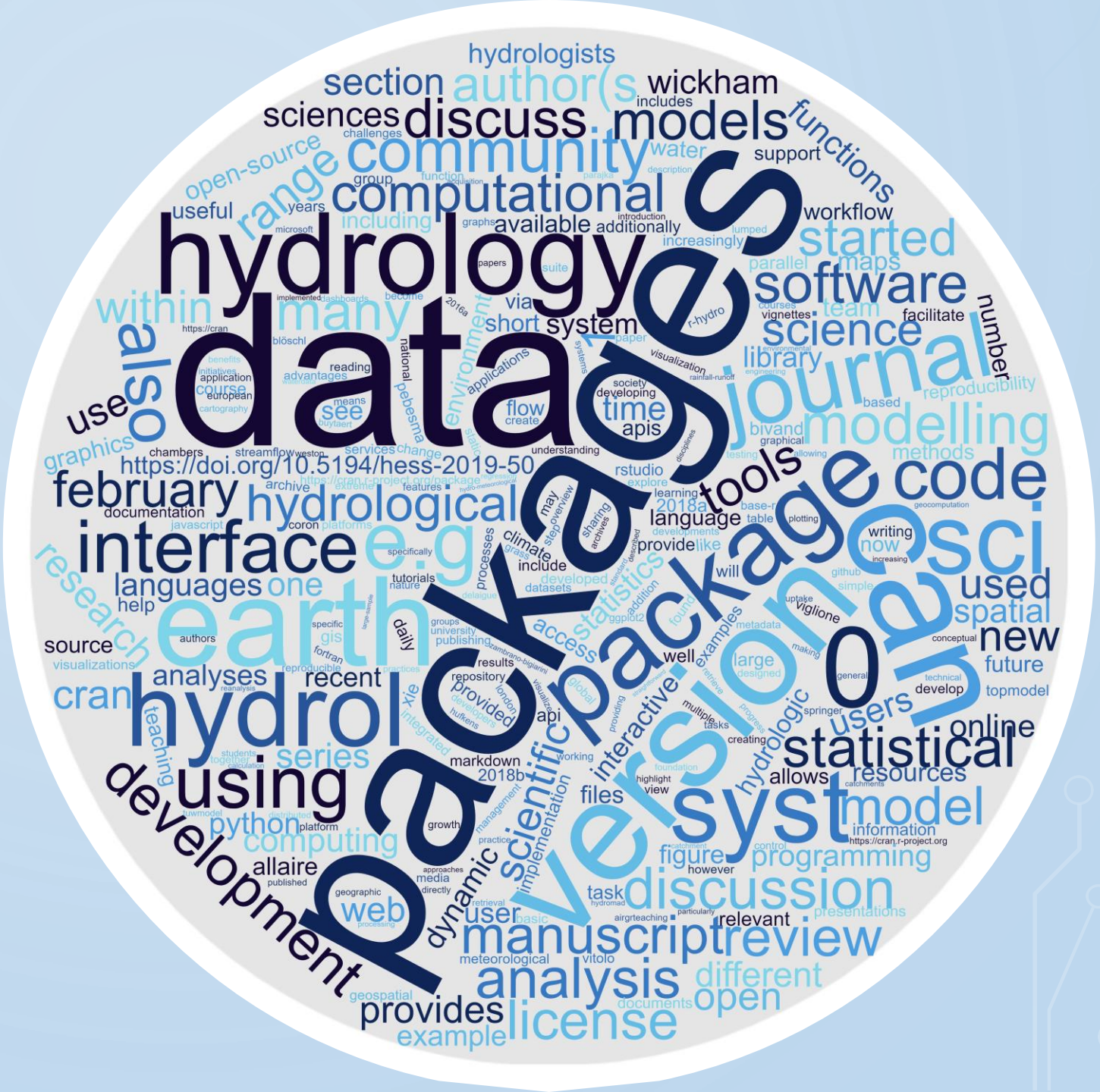
RPubs (https://rpubs.com/) is an open-source place for publishing presentations and scientific resources created in R Markdown. R Markdown documents or presentations simply created within RStudio, and by clicking the "Publish button,"; files can be published for the readers and their feedback.

❑ JOSS?

The Joss stands for Journal of Open Source Software, a peer-reviewed, developer-friendly, open-access scientific journal covering open-source software from any research field in any programming language. The journal uses GitHub as a publishing platform (https://github.com/openjournals/joss). The papers can be searched by the title, tag, author, or language in the search button.

Famous Packages

https://www.r-pkg.org/



❑Hydrological R packages?

- R packages are also growingly used in hydrological studies such as surface water, groundwater quantity and quality, meteorological data, data tiding packages (gap-filling, data cleaning, missing data, organization), hydrograph analysis (function to work with flow data, flow trends, flow statistics and indices), spatial data and GIS application packages, statistical modeling and other packages. The whole list can be obtained from here (https://cran.r-project.org/web/views/Hydrology.html). All packages which are utilized in R are citable. The code below shows how to refer to a package in a publication by giving the title, author's name, journal name, volume, number, doi, and URL of the package, which can be added in reference manager software such as Mendeley or endnote.
- Code: writeLines(toBibtex(citation("package name")))

❑ Workflow of R in Hydrology

- To have a reproducible computational hydrological model in R; these stages should be presented in every model.
- 1) Retrieving data sets
- 2) Preprocessing scripts
- 3) Defining input parameters
- 4) Splitting data sets into calibration and validation stages
- 5) Model building
- 6) Reporting the results in more innovative ways.

❑ Workflow of R in Hydrology

1-Retrieving data sets

2-Pre-processing data

3-Defining input parameters

4-Splitting data sets in calibration and validation stages

5- Model building

6-Reporting the results

❑Data Retrieval packages in R

Package name	Details
AWAPer	Catchment Area Climate Data Anywhere in Australia
dataRetrieval	water quality and hydrology data from EPA and USGS
echor	download discharge records
FedData	Downloading Geospatial Data Available from Several Federated Data Sources
metScanR	metadata from over 157,000 environmental monitoring stations among 219 countries
nhdR	functions for downloading, and networking United States Geological Survey
rnrfa	functions to retrieve data from the UK National River Flow Archive
tidyhydat	Provides functions to access historical and real-time national 'hydrometric' data from Water Survey of Canada data sources

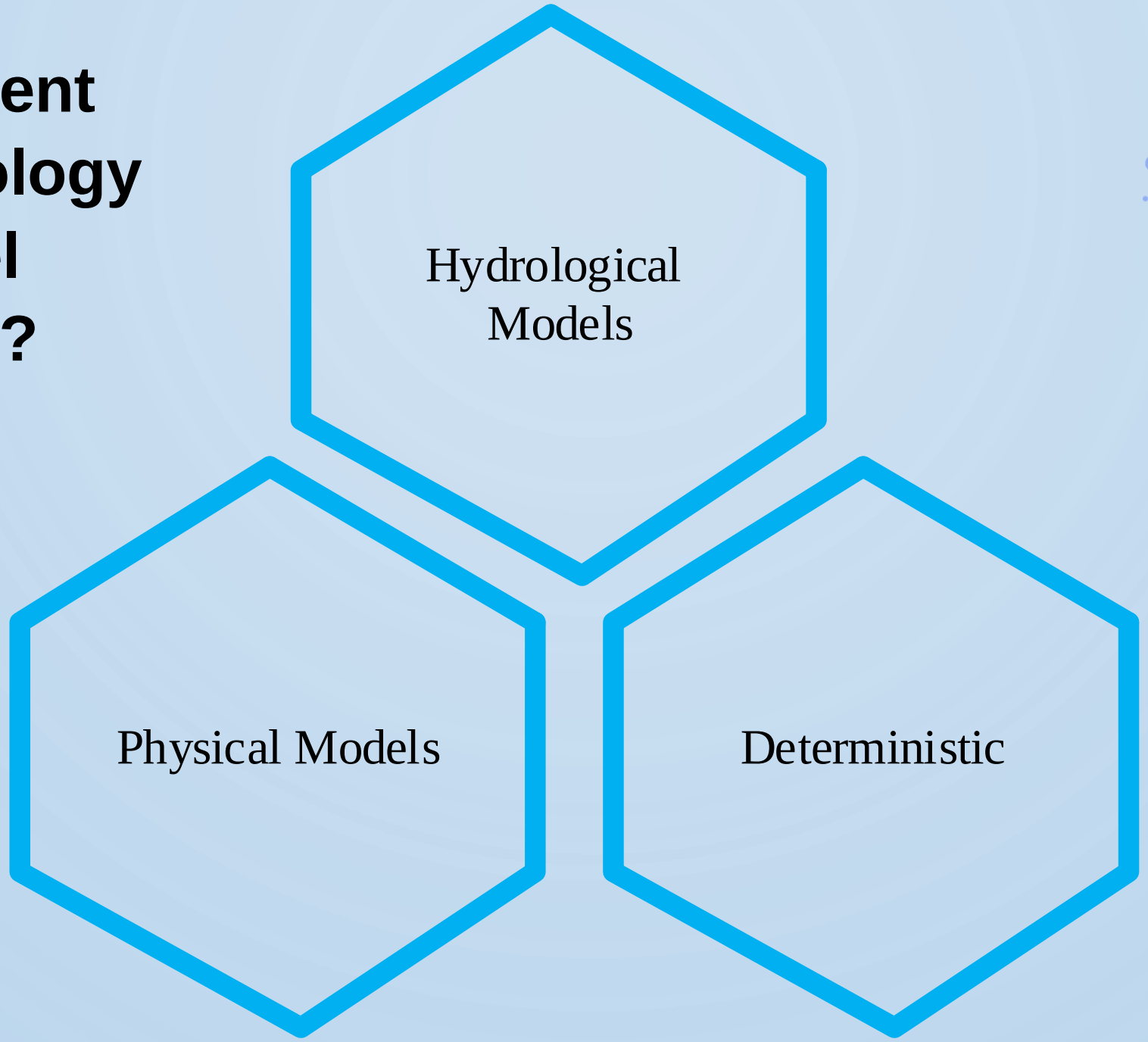
Package name	Details
waterData	plots the data, addresses some common data problems, and calculates and plots anomalies of (USGS) daily hydrologic data
climate	downloading of meteorological and hydrological data from publicly available repositories
clifro	Website of New Zealand National Climate Database of around 6,500 climate stations
getMet	The ability to source, format, and edit meteorological data for hydrologic models
GSODR	automated downloading, parsing, cleaning, unit conversion and formatting from USA
MODISTools	Allows for easy downloads of 'MODIS' time series
nasapower	daily meteorology, interannual and 30-year climatology
metR	commonly used analysis methods in the atmospheric sciences
prism	Data are presented as gridded rasters at four different temporal scales: daily, monthly, annual, and 30 years normal

Package name	Details
rdwd	Download observational time series from meteorological stations from German
RNCEP	functions to retrieve, organize, and visualize weather data
rnoaa	Data, data sets, types, locations, and stations from NOAA data sources in API
rwunderground	getting historical weather information and forecasts
smapr	acquire, and extract NASA Soil Moisture
stationaRy	A global network of weather stations provides hourly weather data.
worldmet	Download over 30,000 surface meteorological sites around the world
FAOSTAT	(Food and Agricultural Organization of the United Nations) database

❑ Preprocessing retrieved hydrological data (Data tidying) in R

Package name	Details
driftR	cleaning and correcting of water quality data
climdex.pcic	computation of extreme climate indices
climatol	Functions to quality control, homogenization, and missing data infilling of climatological series
plyr	visualization of data
tidyr	Cleaning your data by identifying the variables in your dataset and using the tools provided
janitor	Cleaning dirty data able to find duplicates
hyfo	data processing and visualization in hydrology Data of precipitation extracted, data downscaled and resampled

❑ Different hydrology model types?



Physically-based models	
Package name	Details
airGR	conceptual rainfall-runoff models and snow accumulation and melt model
bigleaf	Calculation of physical and physiological ecosystem properties
boussinesq	Boussinesq Equation (ground-water modelling)
Ecohydmod	soil water balance simulation
EcoHydRology	SWAT calibration functions
geotopbricks	Hydrological Distributed Model GEOtop
hydroPSO	Particle Swarm Optimisation (PSO) algorithm for the calibration of environmental models
HBV.IANIGLA	the HBV hydrological model

Physically-based models	
Package name	Details
kwb.hantush	Calculation groundwater mounding beneath an infiltration basin
RavenR	Raven Hydrological Modelling Framework
reservoir	Analysis, Design, and Operation of Water Supply Storages
RHMS	construction, simulation, visualization, and calibration of hydrologic systems
RSAlgaeR	Empirical Remote Sensing Models of Water Quality Variables
streamDepletr	calculate the impacts of groundwater pumping
swmmr	Storm Water Management Model (SWMM)
telemac	modeling of free surface flow
topmodel	hydrological functions TOPMODEL
TUWmodel	Lumped Hydrological Model for Education Purposes
WRSS	Water resources system simulator

Deterministic (statistical) models	
Package name	Details
CoSMoS	generates univariate/multivariate non-Gaussian time series
hydroApps	regional analysis of hydrological applications
hydroGOF	goodness-of-fit measures between observed and simulated values
HydroMe	parameters in infiltration and water retention models
LPM	Long Memory Models to hydrological data sets
nsRFA	Regional Frequency Analysis methods in hydrology
RMAWGEN	stochastic generation of daily time series of temperature and precipitation
SCI	Functions for generating Standardized Climate Indices
soilwater	soil water retention or conductivity curve
synthesis	Generate synthetic time series

Deterministic (statistical) models	
Package name	Details
SPEI	Standardized Precipitation-Evapotranspiration Index (SPEI)
WASP	A wavelet-based variance transformation method
Evapotranspiration	Functions to calculate potential evapotranspiration (PET) and actual evapotranspiration (AET)
MBC	Multivariate Bias Correction of Climate Model Output
meteoland	Functions to estimate weather variables
musica	Multiscale Climate Model Assessment
openair	Tools to analyze, interpret and understand air pollution data
qmap	climate model simulations using quantile mapping
MODIStsp	MODIS satellite data can be downloaded along with preprocessing Land Products Data

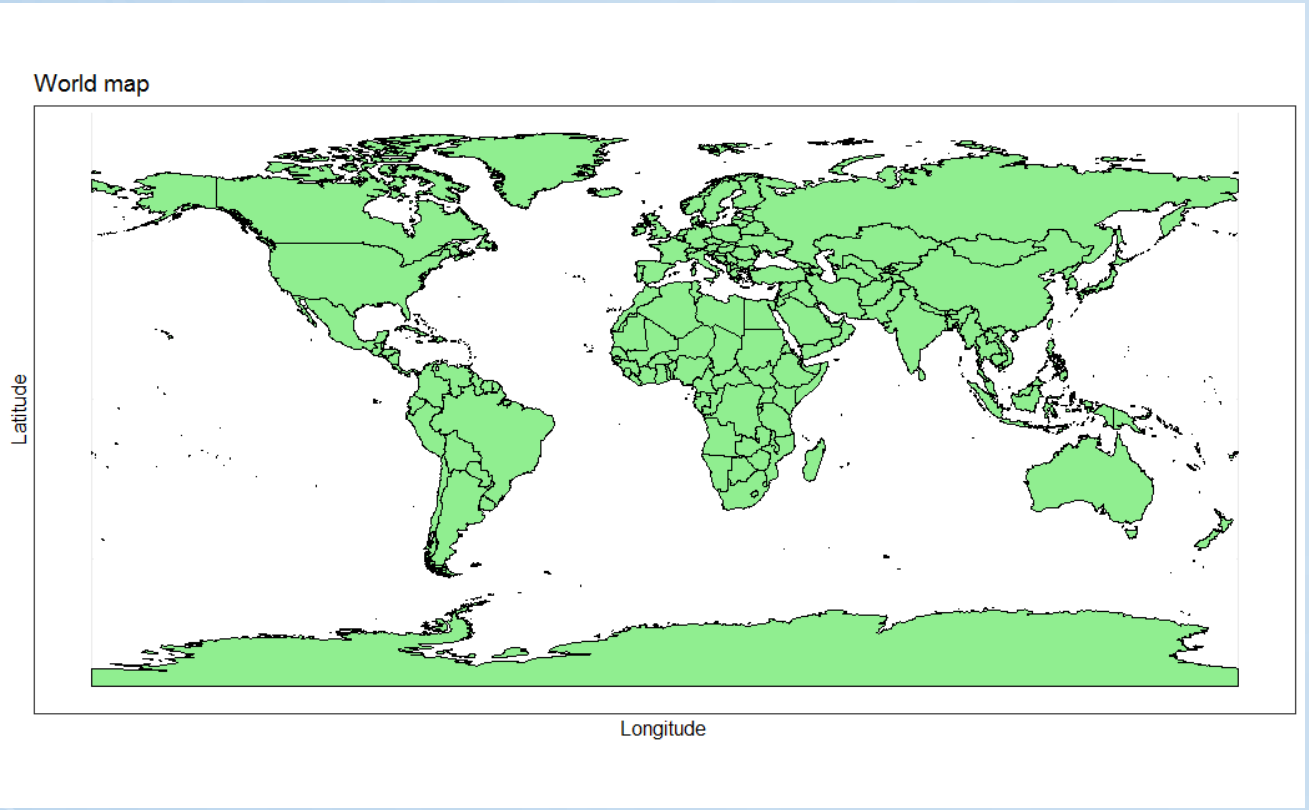
❑ Hydrologic Time Series Analysis tools in R

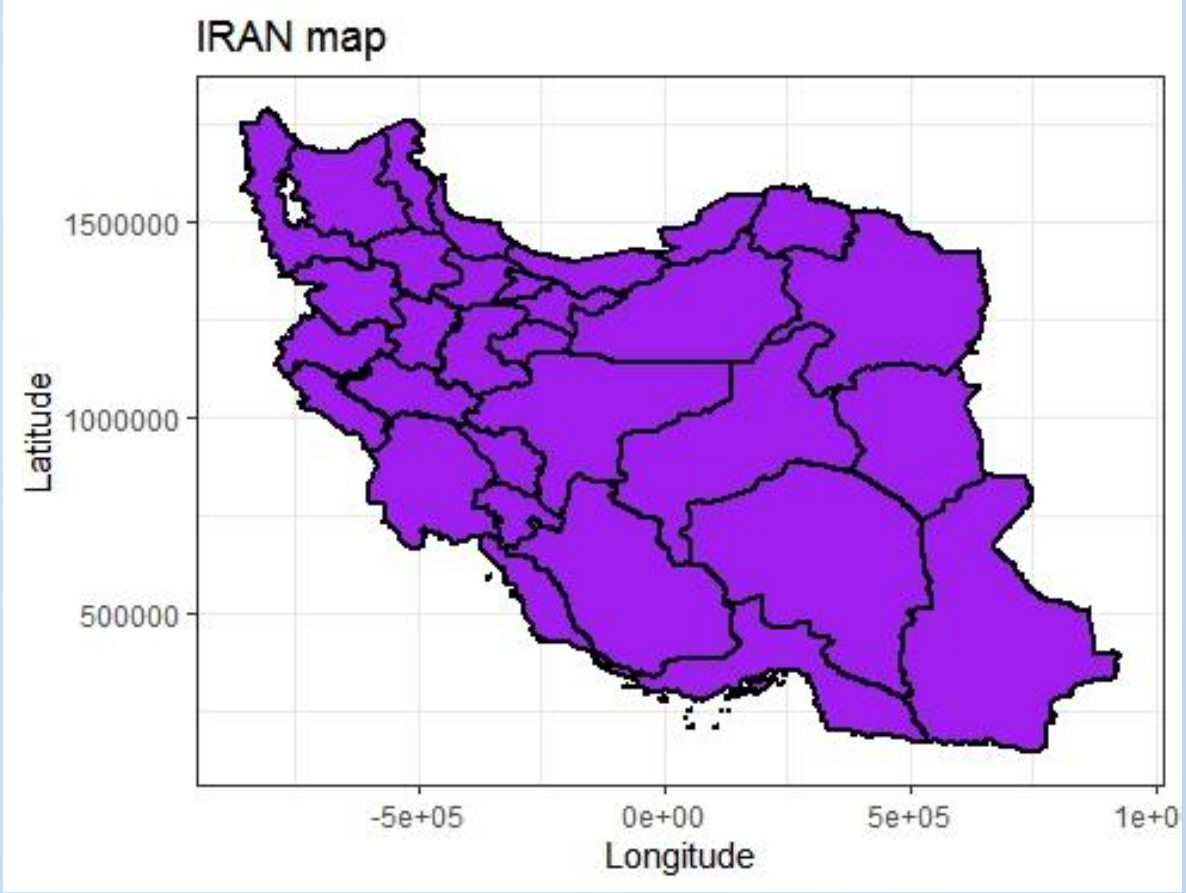
Package name	Usage	Details
fable	tools for fitting univariate time series models	ETS, ARIMA, TSLM, and other models
forecast	provides time series forecasting tools	functions for computing and analyzing forecasts
prophet	Analyzes time series using an additive model that fits yearly and weekly seasonality to nonlinear trends	works best with daily data
tseries	Time series analysis and computational	fits basic GARCH models
tsDyn	Nonlinear Time Series Models with Regime Switching	Implements nonlinear autoregressive (AR) time series models
MTS	all-purpose toolkit for analyzing multivariate time series	VAR, VARMA, seasonal VARMA, VAR models with exogenous variables, and more
TSrepr	methods for representing time series	using dimension reduction and feature extraction
wavelets	Wavelet methods	computing wavelet filters, wavelet transforms, and multiresolution analyses

❑ Hydrological Machine Learning application tools in R

Package name	Usage	Details
nnet	Neural Networks package	ingle-hidden-layer neural network are implemented
deepnet	Deep Learning package	feed-forward neural network, restricted Boltzmann machine, deep belief network, stacked autoencoders
rpart	Tree structure CART model	classification and survival analysis
RWeka	Tree structure models	J4.8-variant of C4.5 and M5 implementation
randomForest	regression and classification	implementation of the random forest algorithm
xgboost	Boosting	tree-based boosting using efficient trees as base learners for several and also user-defined objective functions
rgenoud	Optimization using Genetic Algorithms	offers optimization routines based on genetic algorithms
frbs	Fuzzy Rule-based System	regression and classification using Fuzzy technique

❑GIS IN R Packages






```
ggplot()+
  geom_polygon(data = df1, aes(x = long, y = lat, group = group),
               fill = "purple", color = "black", lwd = 1)+
  xlab("Longitude") + ylab("Latitude") +
  ggtitle("IRAN map") +
  theme_bw()
```

Spatial Analysis Projects and packages (GIS in R)

- sp package=plot points, lines, polygons or grids in R
- maptools and shapefiles packages =serve for accessing vector data
- rgdal package= provides functions to read and write a lot of grid and vector formats, and it provide
- raster package =provides access to data in raster formats and includes analytical tools for this type of spatial data. The raster package provides, among other things, the creation of raster objects from scratch or from a file, the handling of extremely large raster files, raster algebra and overlay functions, distance functions, polygons, lines and points to raster conversion, summarizing raster values, easy access to raster cell values, plotting, reading and writing various raster file types.

Spatial Analysis Projects and packages (GIS in R)

- rasterVis package= complements the raster package, which provides a set of methods for enhanced visualization and interaction
- rgeos package= Vector data manipulation, e.g., topology operations on geometries, are accessible
- gstat package= offers a wide range of univariable and multivariable geostatistical modelling methodologies, prediction and simulation functions, variogram modelling, variogram map plotting, everything from simple global kriging to local universal cokriging, multivariate geostatistics, block kriging, etc
- RgoogleMaps package= package provides tools to access Google Maps data in an image form using the Google Static Maps API, in order to permit background maps to be used in R.

Spatial Analysis Projects and packages (GIS in R)

- ggmap package= allows for the easy visualization of spatial data on top of Google Maps, Open Street Maps, Stamen Maps, or Cloud Made Maps using ggplot2. In ggmap usage, a basic layer, e.g., from Google Maps, is firstly downloaded, and then its object is created in R. Then other layers with lines, points, polygons, texts and other features from various sources are added to it according to its rules of syntax.
- osmar package=Vector data from Open Street Map is also available for downloading

- R is a rapidly growing software with many libraries and packages.

- Recently many scholars in different fields have used R for their daily use and producing publication-quality graphics. Hydrologists also use this software to tidy, transform, visualize, model their hydrometric data.

- R programming language holds interest among the hydrologist's community, and this causes them to develop useful and efficient hydrology packages, workflows, hydrology R-related events, conferences, holding training courses such as the application of R in Hydroinformatics and so on.

Thanks for your Attention