

VI Congreso Nacional del Agua - CONA 2023
Curso Pre-Congreso

Ciencia de datos en hidrología

1. Resumen:

La tecnología ha permitido la generación abundante de datos e información del ciclo hidrológico. Desde las clásicas mediciones en campo alimentadas por sensores de alta frecuencia hasta la asimilación de modelos de reanálisis climáticos e imágenes satelitales. El gran tamaño de esta información hace necesario el uso de la programación digital enfocado a la exploración, tratamiento, análisis estadístico-probabilístico y geoespacial en las investigaciones y diversos proyectos de ingeniería relacionados al agua.

De esta forma surgen nuevas propuestas como el empleo del software libre el cual es adaptable a las necesidades de estudio para aplicaciones ilimitadas. El empleo de los lenguajes R y Python ya se encuentran bien establecidos dentro de la comunidad, por ello, se hace necesario su aprendizaje. El presente curso, brinda un primer acercamiento a la ciencia de datos a través del manejo de estos dos lenguajes. Se desarrollan a través de talleres con enfoque práctico y aplicativo.

2. Objetivos del curso:

- Reforzar los conocimientos en estadística y geoestadística con herramientas de la ciencia de datos.
- Introducir al lenguaje R y Python para el estudio de series de tiempo en hidrología.

3. Contenido:

- a. Fundamentos de R y Python (ciencia de datos y tipos de datos)
- b. Procesamiento de datos hidrometeorológicos (estructuras y funciones)
- c. Visualización de datos hidrometeorológicos (matplotlib y ggplot2)
- d. Análisis de series de tiempo (conversiones y AED)
- e. Introducción a datos geoespaciales (fundamentos de vector y raster)
- f. Aplicaciones hidrológicas basadas en productos grillados (PISCO)

4. Audiencia a quién va dirigido:

Profesionales, ingenieros, investigadores y estudiantes interesados en los recursos hídricos, bases de datos en hidrología, operación de sistemas hidrológicos.

5. Fecha y horarios:

Lunes 6 de noviembre de 2023 de 9 am a 6 pm.

6. Lugar

Consejo Departamental de Lima (CD Lima – CIP). Calle Barcelona 240. San Isidro – Lima (Perú).

Aula: a confirmar.

7. Instructores:

Pedro Rau, PhD

Profesor asociado del Departamento de Ingeniería Civil y Ambiental. Universidad de Ingeniería y Tecnología (UTEC)

Investigador principal del Centro de Investigación y Tecnología del Agua (UTEC-CITA). Ingeniero Civil con maestría en medio ambiente. Doctor en Hidrología en la Universidad de Toulouse (Francia) con investigaciones en las ramas de la hidroclimatología y la hidrología de cuencas. Autor de reconocidas publicaciones en revistas científicas y técnicas. Autor de aplicaciones en R, capacitador internacional en R y data science aplicado a la hidrología.

Leonardo Gutierrez, Ing

Ingeniero investigador del Servicio Nacional de Meteorología e Hidrología del Perú (SENAMHI)

Ingeniero Agrícola de la Universidad Agraria la Molina (UNALM). Especialista en data science aplicado a los recursos hídricos. Desarrollador de aplicaciones en lenguaje Python y R. Participación en el diseño del bigdata de productos operacionales a nivel nacional como el producto PISCO-Evapotranspiración de referencia. Experiencia como instructor académico en ciencia de datos. Asistente de investigación UTEC-CITA.

Fiorela Castellón, MSc (c)

Ingeniera investigadora del Instituto Geofísico del Perú (IGP)

Ingeniera Ambiental de la Universidad Agraria la Molina (UNALM). Especialista en data science aplicado a la hidroclimatología e hidrometeorología. Experiencia como capacitadora en ciencia de datos para agencias gubernamentales y académicas. Pasante de investigación en la Universidad de Harvard (EEUU) en proyectos de aplicación del bigdata en climatología. Asistente de investigación UTEC-CITA.

Lorena Dueñas, MSc (c)

Jefe de proyectos de Recursos hídricos en AMPHOS21 RSK Company

Ingeniera Civil de la Universidad Nacional Hermilio Valdizan (UNHEVAL). Especialista en proyectos de recursos hídricos orientado a hidrología y balance de aguas. Experiencia como investigadora en el proyecto PEGASUS (CITA-UTEC) así como especialista de recursos hídricos para minería en AMPHOS21.

Programa del curso

Ciencia de datos en hidrología

Lunes 6-nov 2023

Hora	Tema	Expositor	Asistentes
9:00	Introducción a la ciencia de datos en hidrología. Tratamiento de datos. R y Python	Pedro Rau	Fiorela Castellón Lorena Dueñas Andres Figueroa
10:00 - 12:00	Procesamiento de datos hidrometeorológicos en R. Estructuras en tidyverse	Fiorela Castellón	Lorena Dueñas Andres Figueroa Pedro Rau
12:00	Introducción al entorno geoespacial, netcdf y graficación	Fiorela Castellón	Lorena Dueñas Andres Figueroa Pedro Rau
13:00	Almuerzo libre		
14:00 - 16:00	Procesamiento de datos hidrometeorológicos en Python (matplotlib)	Leonardo Gutierrez	Fiorela Castellón Lorena Dueñas
16:00 - 18:00	Introducción al entorno geoespacial, netcdf y graficación (xarray, geopandas)	Leonardo Gutierrez	Fiorela Castellón Lorena Dueñas Pedro Rau

- a) **Repositorio del curso:** El material del curso se encuentra en el siguiente enlace Github

<https://github.com/hydrocodes/datascience>

- b) **Plataformas online de uso en el curso:**

Rstudio cloud (posit cloud): Crearse una cuenta en el siguiente enlace.

<https://posit.cloud/>

Google Colab: Acceder con una cuenta gmail, google en el siguiente enlace.

<https://colab.research.google.com/>

- c) **Softwares a instalar de acuerdo a la necesidad del alumno:**

Rstudio <https://posit.co/download/rstudio-desktop/>

R <https://cran.r-project.org/>

Jupyter Lab <https://jupyter.org/>

- d) **Paquetes a instalar en R y Python:**

R: *tidyverse*

Python: *pandas, matplotlib, numpy, geopandas*