



Herramienta para Monitorear la Calidad de Datos

Franco Chiappe

Vicente Garay

Ziyu Guo



Ainwater

LA INTELIGENCIA DEL AGUA

- ◆ Quiénes son
- ◆ Qué hacen
- ◆ Contraparte



Ainwater

LA INTELIGENCIA DEL AGUA

◆ **Quiénes son**

Ainwater es una startup chilena de tecnología que desarrolla soluciones de inteligencia artificial para optimizar plantas de tratamiento de agua.

◆ Qué hacen

Opera actualmente en Chile, México y Brasil, están en proceso de expandirse a Colombia y España.

◆ Contraparte



Ainwater

LA INTELIGENCIA DEL AGUA

◆ Quiénes son

◆ Qué hacen

◆ Contraparte

Su solución monitorea en tiempo real el funcionamiento de plantas de aguas (residuales, industriales, etc.) y emite alertas predictivas con recomendaciones de control.



Ainwater

LA INTELIGENCIA DEL AGUA

◆ **Quiénes son**

Constanza Córdova

Líder de Ciencia de Datos

Ingeniera Civil en Química y Analista Data Sciences.

◆ **Qué hacen**

Yaniela Fernández

Data Engineer

◆ **Contraparte**

Licenciada en Ciencias de la Computación y Máster en Bioinformática.

¿Por qué nace este Proyecto?

Actualmente, el control de calidad de los datos que recibe AinWater se realiza de forma manual y reactiva, lo que genera:

- Retrasos en la detección de fallas, drift, valores anómalos y/o descalibraciones.
- Ausencia de indicadores de calidad del dato que entra a la plataforma.
- Impacto directo en la precisión y confiabilidad de los modelos predictivos.



Objetivos

Objetivo General

- ❖ Diseñar e implementar un prototipo funcional que evalúe automáticamente la calidad de datos de sensores, detecte problemas como anomalías, drift y datos faltantes, y entregue indicadores que permitan mejorar la confiabilidad de los modelos predictivos.

Pasos



Diagnóstico inicial
de los datos



Definición de
métricas de calidad



Diseño del sistema
de evaluación

Plan general/ Timeline

1	Diagnóstico de la calidad de datos	Informe con resultados del diagnóstico.
2	Definición de métricas y umbrales	Documento de métricas
3	Desarrollo del pipeline de evaluación	Script de Python (pipeline) + documentación técnica



Métodos para la detección de valores atípicos



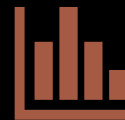
1. Métodos estadísticos
(distancia a la media, IQR, z-score)



2. Utilizando aprendizaje no supervisado (p. ej., PCA).



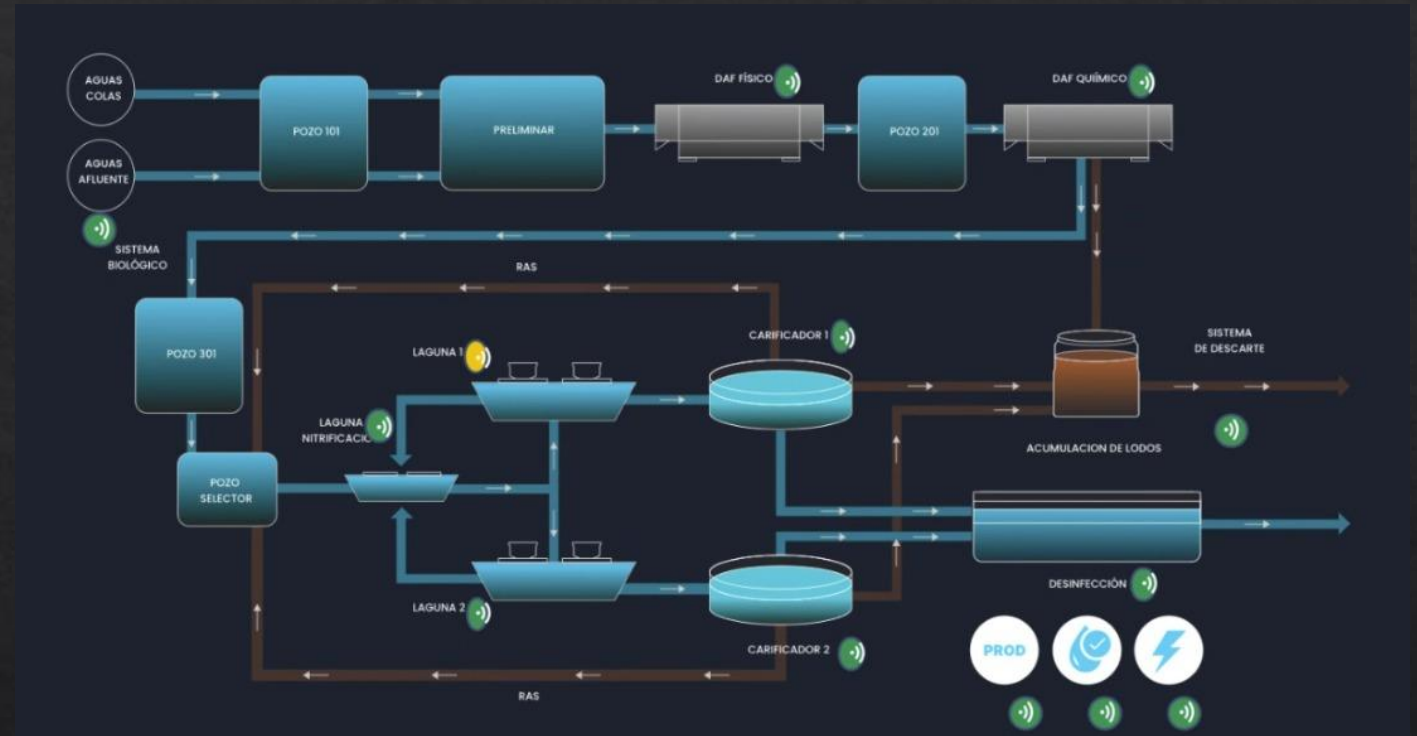
3. Serie de tiempo (calculo residuos)



4. Naturaleza de los datos

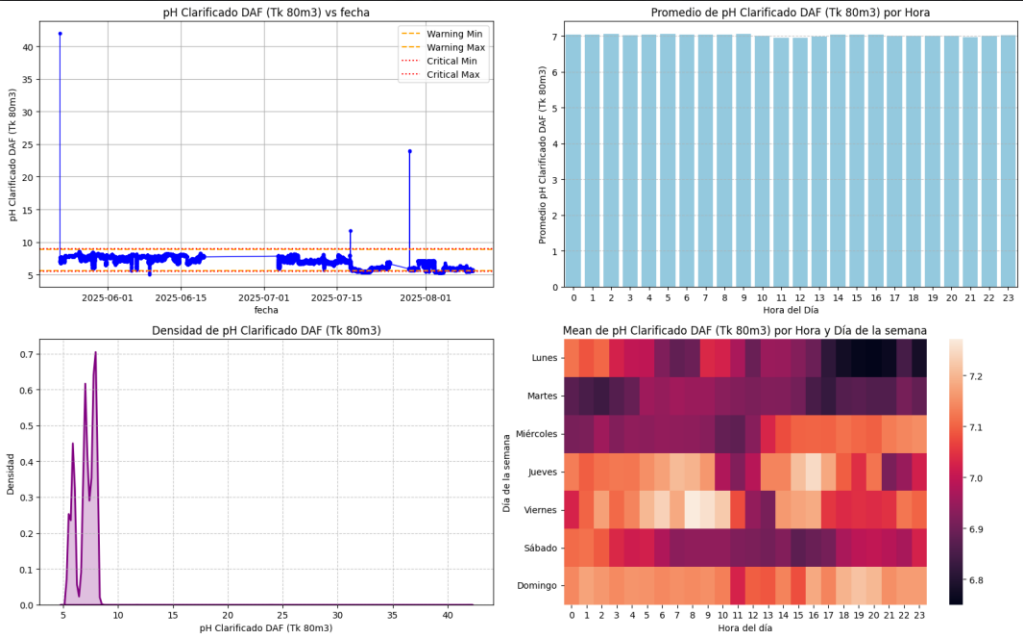
Planta

- ◆ Funcionamiento de una planta de trata de aguas
- ◆ Varios procesos
- ◆ Cada proceso tiene ciertos parámetros que influyen más
- ◆ Sensores que miden estos parámetros

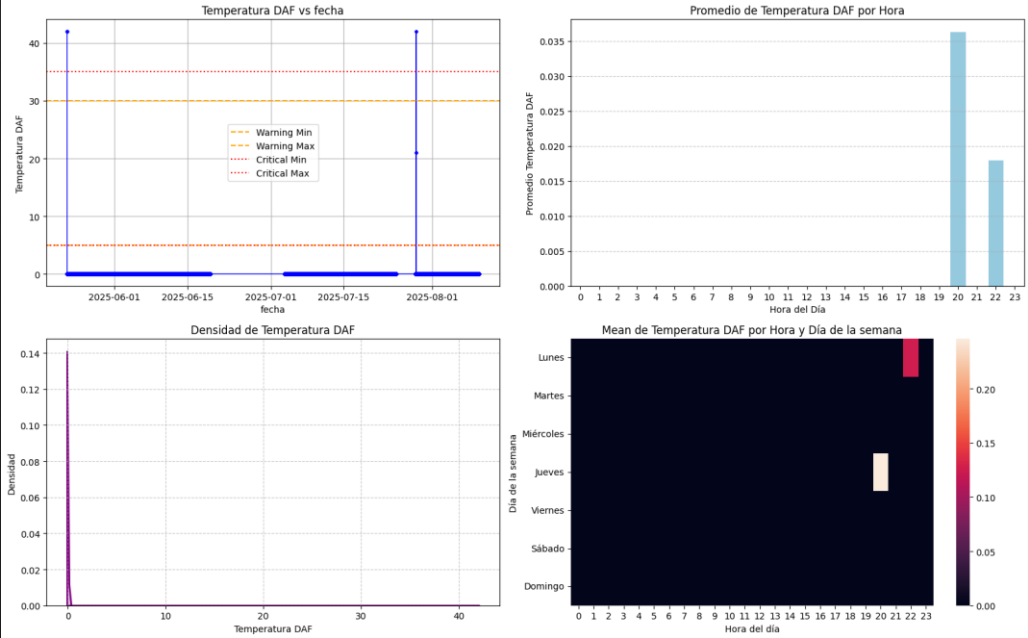


Situación actual

Ph Clarificado

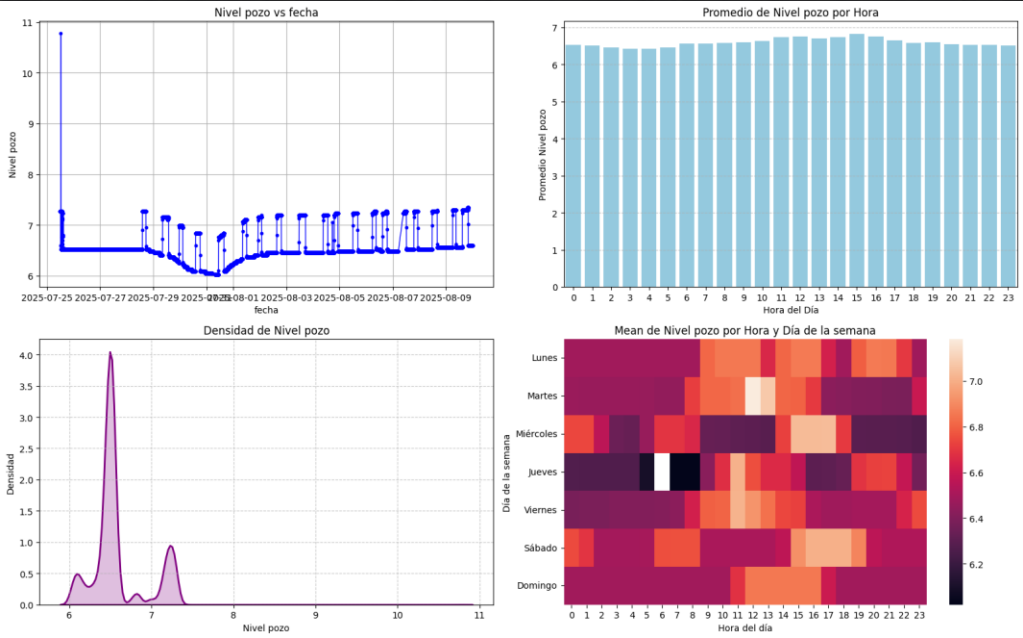


Temperatura DAF

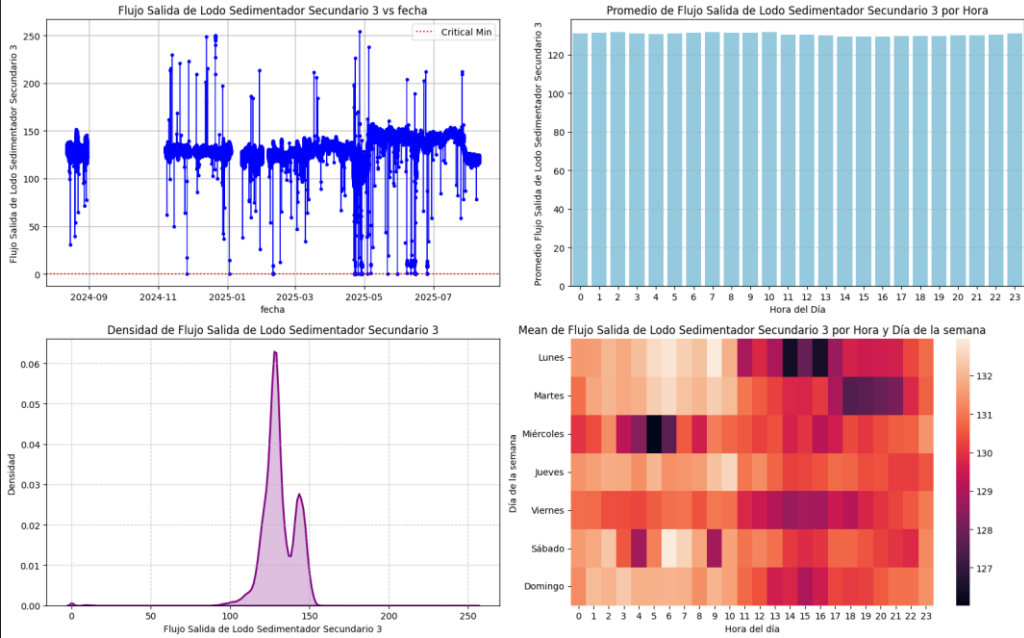


Situación actual

Nivel Pozo



Flujo Salida de Lodo





¿Preguntas o comentarios?

Gracias por su Atención