

# Sistema experto basado en reglas, para determinar la calidad de las aguas superficiales del litoral atlántico andaluz

Plan de Estudios : Grado en Informática

Área del trabajo final: 75.629 TFG – Inteligencia artificial

Profesor responsable Dr. Carles Ventura Royo

Consultor: Dr. David Isern Alarcón

**Daniel J. Torres Rodríguez**

## **Introducción**

Justificación de la metodología del proyecto  
Justificación del producto. Control de Calidad del Agua  
Problemática a resolver y alcance del proyecto  
  
Objetivos general y específicos  
Breve descripción del producto obtenido

## **Metodología y desarrollo del producto**

## **Conclusiones**

## Contexto y justificación del Trabajo

### Justificación de la metodología del proyecto

#### Sistemas Expertos



- Conocimiento del dominio de un problema específico
- Emulan el razonamiento humano
- Capacidad para incorporar conocimiento
- Adaptables

#### Beneficios



- Capacidad de gestión de datos complejos
- Reducción de errores en análisis
- Reducción de tiempo
- Herramienta de soporte al experto

### Justificación del producto. Control de Calidad del Agua

Destrucción de los limitados recursos hídricos y su ecosistema.

Disminución de la calidad del agua

**Peligro potencial que afecta directamente a la salud pública**

## Contexto y justificación del Trabajo

### Problemática a resolver

Gestionar grandes volúmenes de datos complejos

Legislación variable según el parámetro y la situación geográfica.

Interpretación de datos limitada a experto humano.

### Sistema Experto

Alta capacidad de procesamiento de datos de forma rápida y precisa.

Particularidades definidas en las reglas.

Centrados en la evaluación de resultados

### Alcance del proyecto

Estatal en la evaluación del estado químico.

Autonómica para todos los parámetros presentados

# Objetivos del Trabajo

## Objetivo general

Crear un SE basado en reglas para la evaluación de la calidad de las aguas superficiales litorales de la vertiente atlántica andaluza.

## Objetivos específicos

- Seleccionar la bibliografía.
- Estudiar y evaluar la necesidad del proyecto.
- Crear base de conocimiento.
- Seleccionar las variables.
- Crear el conjunto de reglas.
- Integrar motor de inferencia de la librería Pyknow.
- Implementar el Sistema Experto.
- Generar informes de resultados.
- Validar el Sistema Experto.

## Descripción del producto obtenido

### SeaQ:

Aplicación para evaluar la calidad de las aguas del litoral de la vertiente atlántica andaluza.

**Datos:** medidas de los distintos parámetros Químicos y físico-químicos *in situ*

**Reglas:** Limitadas por la legislación

**Método:** IA. Python



### Resultados de la ejecución de SeaQ:

Evaluación de parámetros de cada estación.

Resultados de puntuación de las estaciones

Evaluación ordenada de calidad de las estaciones

Resultados de calidad por demarcación hidrográfica

Permite procesar una gran cantidad de datos

Facilita la extracción de conclusiones derivadas de los informes de resultados.

## Introducción

Justificación de la metodología del proyecto  
Justificación del producto. Control de Calidad del Agua  
Problemática a resolver y alcance del proyecto

Objetivos general y específicos  
Breve descripción del producto obtenido

## Metodología y desarrollo del producto

Selección de variables  
Base de conocimiento, conjunto de reglas  
Motor de inferencia  
Codificación del Sistema Experto  
Validación.

## Conclusiones

## Selección de variables

### Determinar la tipología de la estación

Demarcación hidrográfica	Masa de Agua	Tipología de la estación	Limitada o Normal
Guadiana	Desembocadura Guadiana (Ayamonte)	AT-T12	L
	Marismas de Isla Cristina	AT-T12	L
	Pluma del Guadiana	AC-T15/18	L

**AT-T12: Estuario atlántico mesomareal con descargas irregulares de río**

**AC-T15/18: Aguas costeras atlánticas expuestas o semi-expuestas con afloramiento medio o intenso, protegidas y no protegidas**



## Selección de variables

## Normas de calidad ambiental del agua

Parámetro	U	AT-T12	AT-T13	AC-T13	AC-T20	AMP-T01	AMP-T04	NCA-CMA	NCA-MA	VPV
Níquel	µg/l							NA [34]	20 [8,6]	720
Terbutilazina	µg/l								1	36
HT MPE	mg/l						0,3			
HT LBSM	mg/l						1			
HT LV	mg/l									25
OD CR	%						90			
OD LBSM	%						40			
OD LV	%									VIL:60 VIN:70
Nitratos LBM	mg/l	11	7,42	0,380 06	0,378 2					
Nitratos LV	mg/l									150
pH										5.5-9.5

RD 817/ 2015

Decreto 109/2015

Fragmento de tabla de parámetros evaluados

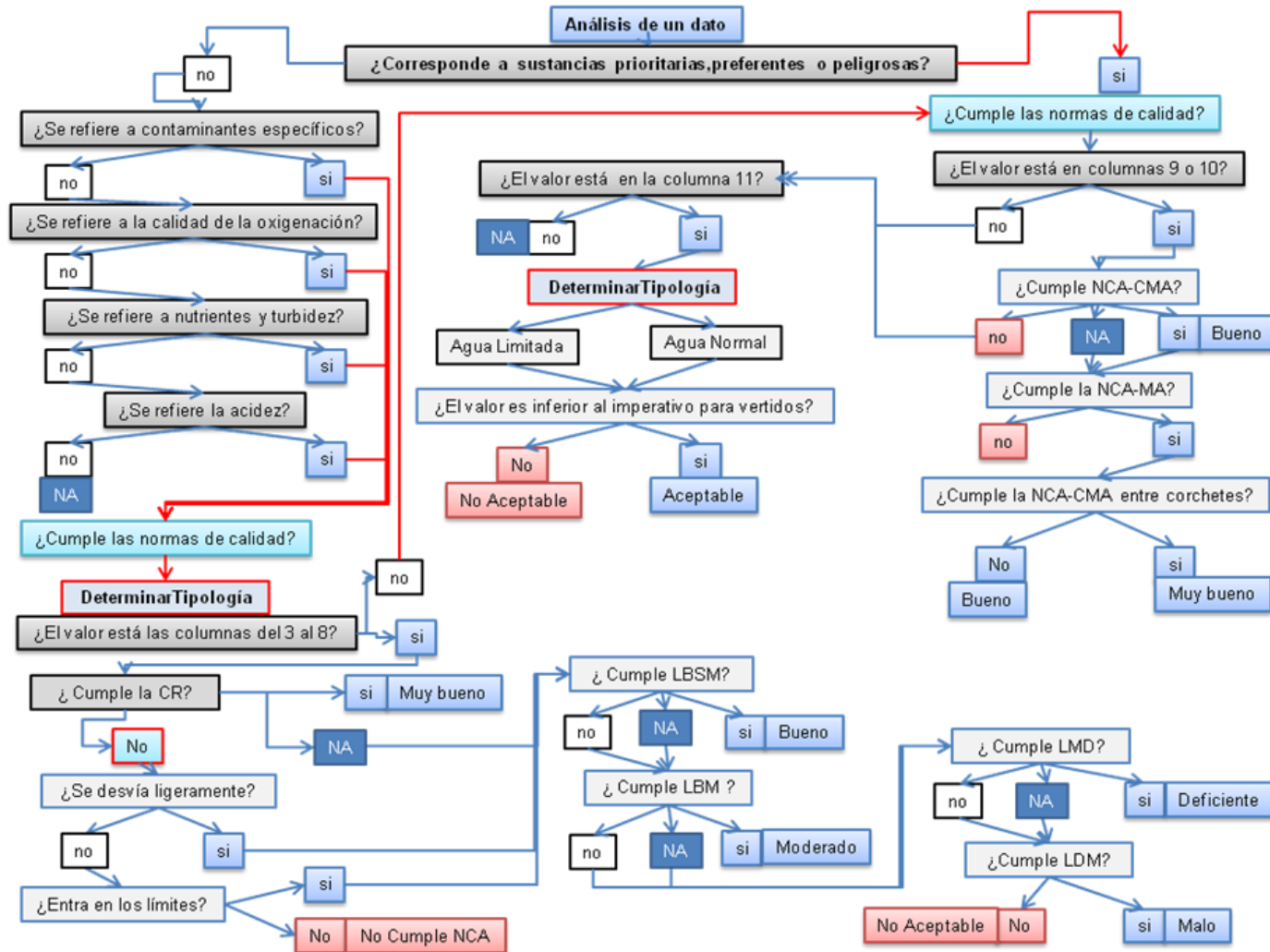
## Selección de variables

Parámetros Evaluados		
Níquel Disuelto	BDE-100	Oxígeno Disuelto( in situ)
Plomo Disuelto	Cobre Disuelto	Nitratos
Clorpirifos	Cinc Disuelto	Nitritos
Diurón	Terbutilazina	Nitrógeno Total
Cadmio Disuelto	Hidrocarburos Totales	Amonio
Mercurio Disuelto	Nitrógeno Kjeldahl	Fosfatos
Ftalato de bis(2 etilhexilo)	Fósforo Total	Turbidez
Benzo[g,h,i]perileno	Carbono orgánico Total	pH (superficie)
Indeno [1,2,3-c,d] pireno	Oxígeno Disuelto (superficie)	pH (secchi)
Tributilestaño	Oxígeno Disuelto (secchi)	pH (fondo)
Endosulfán	Oxígeno Disuelto (fondo)	pH (in situ)

En total se han evaluado 33 parámetros

## Selección de variables

## Proceso de valoración de datos



## Base de conocimiento, conjunto de reglas

## Introducción

class Motor(pk.KnowledgeEngine)

321 Reglas para 33 parámetros

Particularidades respecto al cumplimiento de la legislación

Distintas estructuras de reglas:

Parámetro de estudio

Situación geográfica de la estación de muestreo

Puntuación en base al cumplimiento de la legislación.

Tipos de Reglas		Límites variables	Límites Comunes
1	Hidrocarburos totales	3 para una tipología	1 para el resto
2	Níquel	x	4 para todas
3	Carbono orgánico total	x	3 para dos tipos
4	Oxígeno disuelto	3 para una, 1 de ellos variable	1 para el resto
5	Amonio	2 distintos para cuatro tipos de masas de agua	1 para el resto

## Motor de inferencia

Librería Pyknow

Clase KnowledgeEngine

Ejemplo de ejecución:

```
class Motor(pk.KnowledgeEngine):
```

```
...
```

```
# Ejecución:
```

```
engine = Motor()
```

```
engine.reset()
```

```
engine.run()
```

## Estructura de una regla

Componente de lado izquierdo(LHS)

```
@pk.Rule(Hechos(St_Name=pk.MATCH.St_Name, R_Name='Hidrocarburos_Totales',  
R_Value=pk.P(lambda p: 0<= p <= 0.3), R_Type=pk.MATCH.R_Type,R_Tipo='AMP-T04'))
```

```
def alertaHidrocarburos_TotalesAMP_T04_1 (self):
```

```
    aux.texto.write("\n"+"***Regla activada: Hidrocarburos_Totales*** \nAMP_T04\n"+aux.max_pot_eco  
+" \nEstación:\n"+rules_station_name)
```

```
    aux.texto.write("\n"+"Valor del parámetro:"+str(rules_param_value)+"\n"+"Puntuación obtenida: 4"+" \n")
```

```
    aux.sumaPuntos(rules_station_name, 4)
```

Componente de lado derecho(RHS)

## Conjunto de Reglas

### Ejemplo de Reglas Tipo 4

Parámetro	U	AT-T12	AT-T13	AC-T13	AC-T20	AMP-T01	AMP-T04	NCA-CMA	NCA-MA	VPV
OD CR	%	NA					90	NA		
OD LBSM	%	NA					40	NA		
OD LV	%	NA						NA		VIL:60 VIN:70

```
@pk.Rule(Hechos(St_Name=pk.MATCH.St_Name, R_Name='Oxigeno Disuelto Superficie',
R_Value=pk.P(lambda p: p >= 90), R_Type='L', R_Tipo='AMP-T04'))
def alertaOxigeno_Disuelto_SuperficieAMP_T04_1_L (self):
    aux.texto.write('\n'+***Regla activada: Oxigeno_Disuelto_Superficie***\nAMP_T04\n"+aux.cond_ref
+"Estación:\n"+rules_station_name)
    aux.texto.write("\n"+"Valor del parámetro:"+str(rules_param_value)+"\n"+"Puntuación obtenida: 4"+"")
    aux.sumaPuntos(rules_station_name, 4)
```

Se concreta el parámetro

Se especifica el límite

Se especifica limitada o normal

Se especifica el tipo de estación

Se valora el parámetro

## Conjunto de Reglas

### Ejemplo de Reglas: Tipo 5

Parámetro	U	AT-T12	AT-T13	AC-T13	AC-T20	AMP-T01	AMP-T04	NCA-CMA	NCA-MA	VPV
Amonio LBM	mg/l	0,29	0,15	0,07002	0,06606	NA		NA		
Amonio LV	mg/l									100

```
@pk.Rule(Hechos(St_Name=pk.MATCH.St_Name,R_Name='Amonio',
R_Value=pk.P(lambda p: p <= 0.29),R_Type=pk.MATCH.R_Type, R_Tipo='AT-T12'))
def alertaAmonioAT_T12_1 (self):
    aux.texto.write("\n"+"***Regla activada: Amonio***\nAT_T12\n"+aux.lbm+"\nEstación:\n"
+rules_station_name)
    aux.texto.write("\n"+"Valor del parámetro:"+str(rules_param_value)+"\n"+"Puntuación obtenida: 4"+"")
    aux.sumaPuntos(rules_station_name, 4)
```

Se concreta el parámetro

Se especifica el límite

Se valora el parámetro

Se especifica masa de agua

## Codificación del Sistema Experto

### Estructura de la aplicación SeaQ

#### Módulo principal → SeaQ.py

Implementación de Reglas, Hechos,  
y ejecución del motor de inferencia

#### Módulo secundario → auxiliar.py

Implementación de métodos de lectura de ficheros de entrada  
Estructura de las estaciones de muestreo  
Método de asignación de puntuaciones  
Método para generación de ficheros de salida

#### Ficheros de entrada

Datos de las estaciones de muestreo  
Datos de las tipologías de las estaciones

#### Ficheros de salida

Valoración de los parámetros  
Puntuación de las estaciones  
Calidad de las estaciones  
Calidad de demarcaciones hidrográficas



## Codificación del Sistema Experto

Lectura de datos

Definición de reglas

Declaración de hechos

Generación de resultados

1	Provincia	Código	Estación de muestreo	Naturaleza	Fecha de toma	Parámetro	Unidades	Resultado	Masa de a Demarcación hidrográfica	XUTM	YUTM
2	CADIZ	62T6010	MARISMAS DE CADIZ Y SAN FERNANDO	Aguas	16/01/2018	Cadmio Disuelto	µg/L	0.50	MARISMA GUADALETE Y BARBATE	215538	4038352
3	CADIZ	62T6025	MARISMAS DE CADIZ Y SAN FERNANDO	Aguas	16/01/2018	Mercurio Disuelto	µg/L	<0.01	MARISMA GUADALETE Y BARBATE	214996	4030677
4	CADIZ	62T4005	CURSO FLUVIAL DEL GUADALETE 2 (LA C)	Aguas	18/01/2018	Endosulfan	µg/L	0	CURSO FI GUADALETE Y BARBATE	222029	4060619
5	CADIZ	62T4005	CURSO FLUVIAL DEL GUADALETE 2 (LA C)	Aguas	18/01/2018	Clorpirifos	µg/L	0.0109	CURSO FI GUADALETE Y BARBATE	222029	4060619
6	CADIZ	62T4005	CURSO FLUVIAL DEL GUADALETE 2 (LA C)	Aguas	18/01/2018	Endosulfan alfa	µg/L	<0.00015	CURSO FI GUADALETE Y BARBATE	222029	4060619
7	CADIZ	62T4005	CURSO FLUVIAL DEL GUADALETE 2 (LA C)	Aguas	18/01/2018	Endosulfan beta	µg/L	<0.00015	CURSO FI GUADALETE Y BARBATE	222029	4060619
8	CADIZ	62T4010	CURSO FLUVIAL DEL GUADALETE 1 (1)	Aguas	18/01/2018	Cadmio Disuelto	µg/L	<0.05	CURSO FI GUADALETE Y BARBATE	219753	4058093
9	CADIZ	62T4030	DESEMBOCADURA DEL GUADALETE 2	Aguas	18/01/2018	Cadmio Disuelto	µg/L	0.74	DESEMBOCADURA DEL GUADALETE Y BARBATE	216669	4055171
10	CADIZ	62C1147	PUERTO DE TARIFA (BOCANA EXTERIOR)	Aguas	08/02/2018	Cadmio Disuelto	µg/L	0.087	PUERTO I GUADALETE Y BARBATE	265441	3987958
11	CADIZ	62C1147	PUERTO DE TARIFA (BOCANA EXTERIOR)	Aguas	08/02/2018	Tributilestaño	µg/L	0.00112	PUERTO I GUADALETE Y BARBATE	265441	3987958
12	CADIZ	62C1147	PUERTO DE TARIFA (BOCANA EXTERIOR)	Aguas	08/02/2018	Amonio	mg/L	0.113	PUERTO I GUADALETE Y BARBATE	265441	3987958
13	CADIZ	62C1147	PUERTO DE TARIFA (BOCANA EXTERIOR)	Aguas	08/02/2018	Nitritos	mg/L	0.0138	PUERTO I GUADALETE Y BARBATE	265441	3987958
14	CADIZ	62C1147	PUERTO DE TARIFA (BOCANA EXTERIOR)	Aguas	08/02/2018	Nitratos	mg/L	0.298	PUERTO I GUADALETE Y BARBATE	265441	3987958
15	CADIZ	62C1147	PUERTO DE TARIFA (BOCANA EXTERIOR)	Aguas	08/02/2018	Nitrógeno Total	mg/L	1.04	PUERTO I GUADALETE Y BARBATE	265441	3987958
16	CADIZ	62C1147	PUERTO DE TARIFA (BOCANA EXTERIOR)	Aguas	08/02/2018	Fosfatos	mg P/L	0.034	PUERTO I GUADALETE Y BARBATE	265441	3987958
17	CADIZ	62C1147	PUERTO DE TARIFA (BOCANA EXTERIOR)	Aguas	08/02/2018	Turbidez	NTU	<4	PUERTO I GUADALETE Y BARBATE	265441	3987958
18	CADIZ	62C1147	PUERTO DE TARIFA (BOCANA EXTERIOR)	Aguas	08/02/2018	pH Fondo	Unid. pH	8.14	PUERTO I GUADALETE Y BARBATE	265441	3987958
19	CADIZ	62C1147	PUERTO DE TARIFA (BOCANA EXTERIOR)	Aguas	08/02/2018	pH Secchi	Unid. pH	8.14	PUERTO I GUADALETE Y BARBATE	265441	3987958

The screenshot shows a software window titled "5 - \_main\_station" with a table of station attributes and their values. Below the table are "Aceptar" and "Cancelar" buttons. A second window titled "valores - Diccionario (26 elementos)" is open, displaying a list of parameters and their corresponding values, also with "Aceptar" and "Cancelar" buttons.

Atributo	Tipo	Tamaño	Valor
date	str	1	2018-02-08 00:00:00
demarc	str	1	GUADALETE Y BARBATE
l_n	str	1	N
mda	str	1	PUERTO DE TARIFA
medias	float	1	3.4615384615384617
name	str	1	PUERTO DE TARIFA (BOCANA EXTERIOR PUERTO)
points	int	1	90
tipo	str	1	AMP-T04
valores	dict	26	{'Tributilestaño':0.00112, 'Amonio':0.113, 'Nitritos':0.0138, 'Nitrato ...

Clave/Tecla	Tipo	Tamaño	Valor
Amonio	float64	1	0.113
Cadmio_Disuelto	float64	1	0.11
Carbono_Organico_Total_(COT)	float64	1	2.25
Clorpirifos	float64	1	0.01
Endosulfan	float64	1	0.0
Endosulfan_alfa	float64	1	0.00015
Endosulfan_beta	float64	1	0.00015
Fosfatos	float64	1	0.034
Fosforo_Total	float64	1	304.0
Hidrocarburos_Totales	float64	1	0.1
Mercurio_Disuelto	float64	1	0.01

## Codificación del Sistema Experto

### Lectura de datos

### Definición de reglas

### Declaración de hechos

### Generación de resultados

```
@pk.Rule(Hechos(St_Name=pk.MATCH.St_Name,
R_Name='Hidrocarburos_Totales',R_Value=pk.P(lambda p: 0<= p <= 0.3),
R_Type=pk.MATCH.R_Type,R_Tipo='AMP-T04'))
def alertaHidrocarburos_TotalesAMP_T04_1 (self):
...
engine.declare(Hechos(StationName=S_Name,RuleName=R_Name,
RuleValue=elementos[o][1],RuleType=S_Type, RuleTipo=S_Tipo ))
```

Nombre de la Estación:

PUERTO DE CADIZ - BAHIA INTERNA DE CADIZ (3)

Puntuación de la estación:97

Número de Parámetros:29

```
Lista de parámetros:{'Nitritos': 0.022, 'Tributilestaño': 0.001,
'Benzo_g_h_i_perileno': 0.01, 'Indeno_1_2_3_c_d_pireno': 0.01,
'Fosfatos': 0.084, 'Oxigeno_Disuelto_Superficie': 90.6,
'Oxigeno_Disuelto_Fondo': 90.52, 'pH_Superficie': 8.03, 'Nitratos':
0.57, 'Amonio': 0.12300000000000001, 'Nitrogeno_Total': 0.74,
'Cadmio_Disuelto': 0.05, 'Mercurio_Disuelto': 0.01, 'Endosulfan': 0.0,
'Endosulfan_beta': 0.00015, 'Endosulfan_alfa': 0.00015, 'Clorpirifos':
0.0104, 'Carbono_Organico_Total_(COT)': 1.0, 'Fosforo_Total': 270.0,
'Nitrogeno_Kjeldahl': 304.0, 'Oxigeno_Disuelto_(in_situ)': 101.0,
'pH_(in_situ)': 8.14, 'Turbidez': 4.0, 'Hidrocarburos_Totales': 0.2,
'Plomo_Disuelto': 1.0, 'Terbutilazina': 0.05,
'Oxigeno_Disuelto_Secchi': 100.5, 'pH_Fondo': 7.9, 'pH_Secchi': 7.91}
```

Puntuacion/Numero de Parámetros:

3.3448275862068964

## Validación del Sistema Experto

### Informe 1

Correcta asociación del parámetro con su tipología



Puntuación de cada parámetro conforme a la legislación



Informe del cumplimiento de la legislación de los 33 parámetros



Las reglas están creadas adecuadamente

### Resto de informes

Información derivada de los resultados del informe 1

Ordenación numérica de SeaQ no alcanzable por métodos humanos de forma rápida

Conforme con los resultados



El experto ha extraído conclusiones adicionales de los informes generados



## **Introducción**

Justificación de la metodología del proyecto  
Justificación del producto. Control de Calidad del Agua  
Problemática a resolver y alcance del proyecto  
Objetivos general y específicos  
Breve descripción del producto obtenido

## **Metodología y desarrollo del producto**

Selección de variables  
Base de conocimiento, conjunto de reglas  
Motor de inferencia  
Codificación del Sistema Experto  
Validación.

## **Conclusiones**

Conclusiones de la metodología y del producto obtenido

## Conclusiones de la metodología y del producto obtenido

- Para evaluar el estado ambiental de las aguas del litoral atlántico andaluz, se ha propuesto un SE basado en reglas acorde a la legislación aplicable en materia de seguimiento ambiental, evaluación, normas de calidad y vertidos.
- Se ha creado la base de conocimiento, seleccionado las 33 variables y se ha creado un conjunto de 321 reglas codificadas en lenguaje Python para evaluar aproximadamente 2000 datos complejos.
- Informes de evaluación de la calidad obtenidos:
  - Informe 1: Evaluación de cada parámetro en cada estación de manera individual.
  - Informe 2: Evaluación de calidad de las estaciones.
  - Informe 3: Evaluación ordenada de calidad de las estaciones.
  - Informe 4: Evaluación de la calidad por demarcación hidrográfica.
- Resultados validados por un experto.
- Comparado con el método convencional, SeaQ permite una evaluación más rápida, precisa y permite obtener conclusiones adicionales.

**Gracias por su atención**