



# UNIVERSIDAD DE SONORA

División de Ciencias Exactas y Naturales

Licenciatura en Física

Física Computacional I

---

## Evaluación 1

*"Análisis de las Mareas y Salinidad en el Manglar  
El Sargento"*

---

Michelle Contreras Cossio

Profr. Carlos Lizárraga Celaya

Hermosillo, Sonora

Marzo 8 de 2018

# 1 Introducción

Esta es la primera evaluación que se aplica para la materia de Física Computacional; en este documento se presenta tanto el procedimiento realizado, como los resultados.

Se utilizan datos reales de la estación de monitoreo de variables atmosféricas en el Manglar "El Sargento", una bahía costera frente a la parte norte de la Isla del Tiburón; las variables con las que se trabajó en la evaluación fueron Nivel del Agua y Salinidad.

El objetivo de esta evaluación es dar constancia de lo aprendido en las cinco actividades ya realizadas, haciendo un análisis y lectura de datos pertinentes, para poder graficar y comparar algunas variables como Temperatura, Salinidad, Nivel del agua. Para ello, se hizo uso de pandas en jupyter y de bibliotecas como matplotlib o seaborn.

## 2 Descripción de los conceptos físicos: Salinidad y Nivel del Agua

Como se mencionó, las variables de interés en este caso son la salinidad y el nivel del agua, a continuación se definen:

La salinidad es el contenido de sales minerales disueltas en un cuerpo de agua. Dicho de otra manera, es válida la expresión salinidad para referirse al contenido salino en suelos o en agua.

El nivel del agua, por su parte, es la altura de las aguas del mar cuando está en calma, que sirve de referencia para medir la altura o la profundidad de una montaña, un punto geográfico, etc.

## 3 Descripción de la estructura de los archivos

Los archivos mencionados en esta sección fueron proporcionados por el maestro para realizar la evaluación, descargados del portal de la materia.

El archivo "sargento-201117.csv", contiene los datos del nivel del mar. Primeramente, su nombre fue cambiado a "sargento\_201117.csv", para permitir su lectura en pandas. Este archivo de datos, cuenta con 5 columnas: la primera indica el número del dato, la segunda fecha y hora, la tercera la presión en kPa, la cuarta la temperatura en grados

centígrados y la quinta el nivel del agua en metros; las columnas se encuentran separadas por comas. Este archivo cuenta con 2395 datos, proporciona datos de cada 15 minutos, empezando el 26 de Octubre del 2017 a las 13 horas y terminando el 20 de Noviembre del 2017 a las 11:30 horas. Un segmento del archivo:

```
"Plot Title: sargento-salinidad"
"#","Date Time, GMT-07:00", "Cond High Rng, ?S/cm", "Temp, C", "Specific
Conductance, ?S/cm", "Salinity, ppt"
1,10/26/2017 12:45:00,54525.5,25.21,54301.2,35.9195
2,10/26/2017 13:00:00,54525.5,24.91,54622.1,36.1588
3,10/26/2017 13:15:00,54525.5,24.82,54719.0,36.2311
4,10/26/2017 13:30:00,54525.5,24.76,54783.8,36.2794
5,10/26/2017 13:45:00,54525.5,24.75,54794.6,36.2875
```

El archivo "salinidad\_sargento-201117.csv", contiene los datos de la salinidad, de igual manera sufrió un cambio de nombre, por la misma razón, a "salinidad\_sargento\_201117.csv". Este archivo cuenta con 6 columnas: número del dato, fecha y hora, conductancia en S/cm, temperatura en grados centígrados, conductancia específica en S/cm y salinidad en ppt. El archivo cuenta, de igual manera, con 2395 datos, sin embargo empieza y termina 15 minutos antes que el anterior, por lo que hay un pequeño desfase. A continuación se muestra un segmento del archivo:

```
"Plot Title: sargento-salinidad"
"#","Date Time, GMT-07:00", "Cond High Rng, ?S/cm", "Temp, C", "Specific
Conductance, ?S/cm", "Salinity, ppt"
1,10/26/2017 12:45:00,54525.5,25.21,54301.2,35.9195
2,10/26/2017 13:00:00,54525.5,24.91,54622.1,36.1588
3,10/26/2017 13:15:00,54525.5,24.82,54719.0,36.2311
4,10/26/2017 13:30:00,54525.5,24.76,54783.8,36.2794
5,10/26/2017 13:45:00,54525.5,24.75,54794.6,36.2875
```

## 4 Análisis de datos utilizando Pandas

En esta sección se muestra el procedimiento que se realizó al momento de trabajar con los archivos de datos, para poder crear las gráficas:

1. Como se mencionó en la sección anterior, se descargaron los archivos de la página y el único cambio que se les hizo fue el nombre, para poder leerlos con pandas.
2. Se inicializó Jupyter en la carpeta creada de la Evaluacion1, misma donde se encuentran los archivos descargados. Creando un archivo nuevo "Evaluacion1.ipynb".
3. Se descargaron las bibliotecas necesarias: pandas, numpy, matplotlib y datetime.

```
#Descargamos las bibliotecas necesarias
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
from datetime import datetime
```

4. Se procedió a leer los archivos asignándolos a su data frame correspondiente. Se tomó en cuenta que el archivo con la información de salinidad (df2) debía comenzar un renglón después, para que ambos contaran con datos a partir de las 13:00:00 del 26 de Octubre. Se les dio nombre a las columnas.

```
#Leemos los archivos, asignandoles los nombres a las columnas
df1=pd.read_csv('sargento_201117.csv', skiprows=2, sep=',', header= None ,
               names=['#', 'Fecha', 'AbsPres', 'Temp', 'WaterLevel'] )
df2=pd.read_csv('sargento_salinidad_201117.csv', skiprows=3, sep=',', header= None ,
               names= ['#', 'Fecha', 'CondHighRng', 'Temp', 'SpecificConductance', 'Salinity'])
```

5. Con el fin de que ambos dataframes terminaran a las 11:15:00 del 20 de Noviembre, se eliminó el último renglón del data frame con la información del nivel del agua (df1). Y se borró la columna del número de dato de ambos, ya que aprovecharemos el índice para esto.

```
#Eliminamos la fila extra, para que los dos data frames tengan la misma cantidad de datos
df1=df1[:-1]
df1=df1.drop(['#'],1)
df2=df2.drop(['#'],1)
```

6. Se convirtieron las columnas de fecha en formato de fecha y se creó una nueva columna de mes, que nos indica el número del mes, para poder utilizarla posteriormente.

```
#Convertimos las columnas a formato de fecha
df1['Fecha'] = pd.to_datetime(df1['Fecha'], format='%m/%d/%Y %H:%M:%S')
df1['Mes'] = df1['Fecha'].dt.month
df2['Fecha'] = pd.to_datetime(df2['Fecha'], format='%m/%d/%Y %H:%M:%S')
df2['Mes'] = df2['Fecha'].dt.month
```

7. Se reinició el índice de número de dato, para que el de ambos archivos fuera idéntico. Y se verificó que todo haya funcionado utilizando las funciones head y dtypes.

```
df1.reset_index()
df1.head()
```

	Fecha	AbsPres	Temp	WaterLevel	Mes
0	2017-10-26 13:00:00	105.612	24.448	-0.150	10
1	2017-10-26 13:15:00	105.513	24.351	-0.160	10
2	2017-10-26 13:30:00	105.433	24.351	-0.168	10
3	2017-10-26 13:45:00	105.385	24.351	-0.173	10
4	2017-10-26 14:00:00	105.321	24.351	-0.179	10

```
df1.dtypes
```

```
Fecha          datetime64[ns]
AbsPres        float64
Temp           float64
WaterLevel     float64
Mes            int64
dtype: object
```

8. Comenzando con las gráficas, las primeras fueron los boxplots, las variables WaterLevel, Salinity y Temp fueron graficadas para visualizar su variabilidad en los datos que se tienen de los meses Octubre y Noviembre, para ello se utilizó la columna de mes sobre el eje x. El siguiente segmento de código muestra un ejemplo de como se realizó el boxplot de WaterLevel, las demás se realizaron de manera análoga:

```
#Boxplot de Nivel del Mar por mes
import seaborn as sns
ax = sns.boxplot(x='Mes', y="WaterLevel", data=df1)
plt.show()
```

9. Se utilizó la función describe para poder tener el dato exacto sobre los cuartiles.
10. Posteriormente se realizaron los diagramas de Pearson. El primero pide ver la relación entre el nivel del mar y la salinidad. Para ello se creó un nuevo data frame (df3), que contuviera a estas dos variables, ya que pertenecían al df1 y df2, respectivamente.

```
#Nuevo data frame con WaterLevel y Salinity en uno solo
df3=pd.DataFrame()
df3['WaterLevel']=df1['WaterLevel']
df3['Salinity']=df2['Salinity']
```

11. Una vez que se tuvieron ambas variables juntas, se graficó el Pearson que relacionara ambas.

```
#Gráfica de Pearson de Nivel del Agua y Salinidad
import seaborn as sns
sns.set(style="darkgrid", color_codes=True)

g = sns.jointplot("WaterLevel", "Salinity", data=df3, kind="reg", color="r", size=7)
plt.show(g)
```

12. Para los siguientes dos Pearson no hubo necesidad de crear nuevos data frame, ya que las variables necesarias pertenecían ya a uno mismo. El segundo Pearson fue de nivel de agua y temperatura y el tercero de salinidad y temperatura y se realizaron de manera análoga al paso anterior.
13. Las siguientes gráficas fueron de las variables nivel del agua, salinidad y temperatura, cada una en función del tiempo, para ello se utilizó la columna de fecha sobre el eje x. A continuación se muestra el código para realizar la del nivel del agua, las otras dos se realizaron de manera análoga:

```
#Gráfica de Nivel del Agua en función del tiempo
plt.plot_date(x=df1.Fecha, y=df1.WaterLevel, fmt="b-")
plt.title("Variación del Nivel del Agua")
plt.ylabel("Water Level (m)")
plt.xlabel("Fecha")
plt.grid(True)
plt.show()
```

14. A continuación, se crearon dos gráficas con doble eje y, uno en la izquierda y uno en la derecha, ya que tienen distinta escala o graduación, en la primera, el eje de la izquierda correspondía a la salinidad y el de la derecha al nivel del agua; en la segunda el nivel del agua y la temperatura, respectivamente. Ambas en función del tiempo, tomando al eje x como la fecha. Se muestra el código para la gráfica Salinidad-Nivel del agua:

```
#Grafica de doble eje Salinidad y nivel del mar
fig, ax1 = plt.subplots()
x=df1['Fecha']
y1=df3.Salinity
y2=df3.WaterLevel
ax1.plot(x,y1,'b-', label='Salinity')
plt.legend(loc='upper right')
ax1.set_xlabel('Fecha')
ax1.set_ylabel('Salinity (ppt)')
ax2 = ax1.twinx()
ax2.plot(x, y2 , 'xr-', label='Water Level')
plt.legend(loc='upper left')
ax2.set_ylabel('Water Level(m)')
fig.tight_layout()
plt.show()
```

15. Finalmente, las mismas gráficas realizadas en el punto anterior, se volvieron a crear pero limitando las fechas del eje x, en ambos casos de límite del 1ro de Noviembre del 2017 a las 00:00:00 hasta el 5 de Noviembre a las 23:45:00. Se muestra el código para la gráfica Salinidad-Nivel del agua:

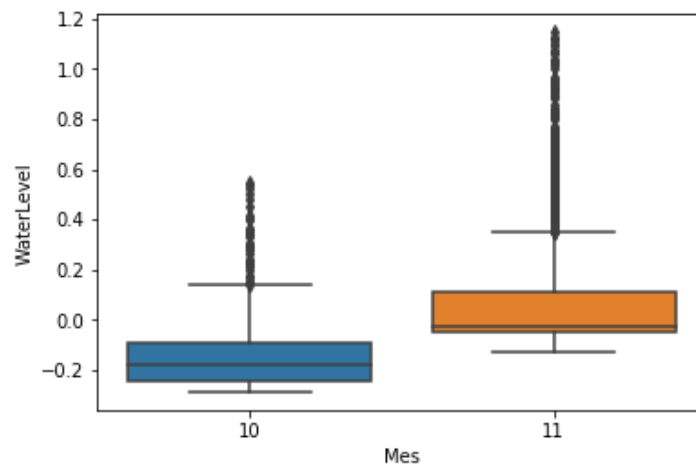
```
#Grafica de doble eje Salinidad y nivel del mar (1 Nov 2017 - 5 Nov 2017)
fig, ax1 = plt.subplots()
x=df1['Fecha']
y1=df3.Salinity
y2=df3.WaterLevel
ax1.plot(x,y1,'b-', label='Salinity')
plt.legend(loc='upper right')
ax1.set_xlabel('Fecha')
ax1.set_ylabel('Salinity (ppt)')
ax2 = ax1.twinx()
ax2.plot(x, y2 , 'xr-', label='Water Level')
plt.legend(loc='upper left')
ax2.set_ylabel('Water Level(m)')
fig.tight_layout()
plt.xlim('2017-11-01 00:00:00', '2017-11-05 23:45:00')
plt.show()
```

## 5 Resultados del análisis

Como bien se mencionó en la sección anterior, el fruto que tuvo este análisis de datos fueron diversas gráficas, a continuación mostradas y discutidas:

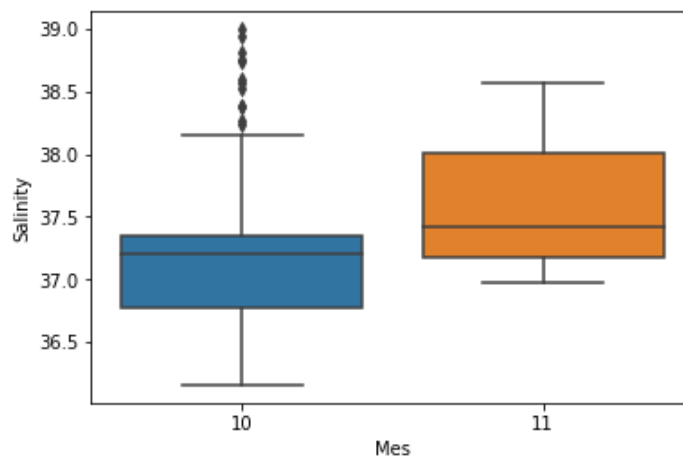
### 5.1 Boxplots

- Nivel del agua:



Se puede observar que la media de nivel del agua de octubre a noviembre aumentó aproximadamente 20 cm. En octubre se encontraba en  $-0.2\text{m}$  y en noviembre alrededor de los  $0\text{ m}$ . Ambas tienen una distribución bastante pequeña, se puede observar que el nivel del agua con aumenta de manera considerable, únicamente centímetros.

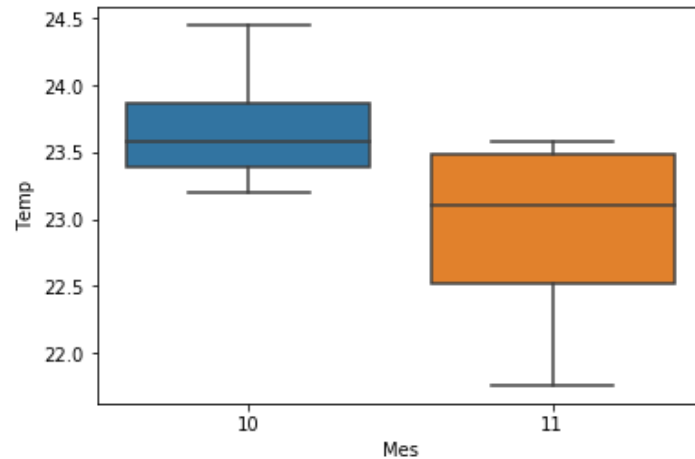
- Salinidad:





En este diagrama de caja se observa que aunque las cajas de octubre y noviembre se encuentran algo separadas, la media sigue apareciendo casi en el mismo valor, entre 37 y 37.5.

- **Temperatura:**



Este diagrama, al igual que los anteriores, no muestra un cambio bastante considerable de octubre a noviembre, a pesar de que si lo hay, lo que se puede observar es que las temperaturas están bastante fijas entre un rango total de 22 a 24.5°C.

- **Describe:**

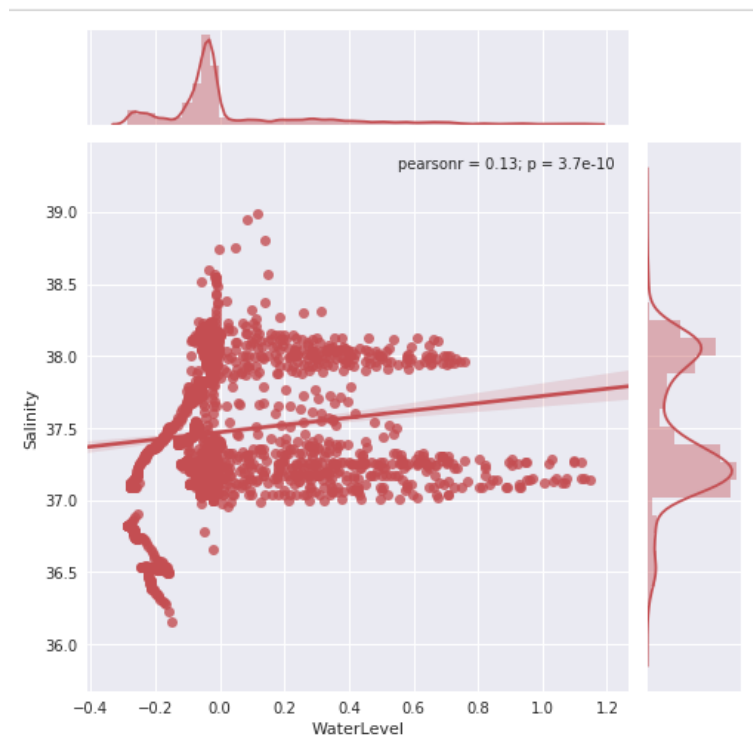
Estos permiten ver un análisis completo de los datos. Sin embargo, no dan la información de donde se encuentran los cuartiles, máximos, mínimos y mediana, necesarias, ya que los datos posteriormente se clasifican por el mes.

	AbsPres	Temp	WaterLevel	Mes
<b>count</b>	2394.000000	2394.000000	2394.000000	2394.000000
<b>mean</b>	107.430007	23.120883	0.030863	10.781119
<b>std</b>	2.371844	0.563555	0.235974	0.413574
<b>min</b>	104.229000	21.760000	-0.288000	10.000000
<b>25%</b>	106.407000	22.525000	-0.071000	11.000000
<b>50%</b>	106.764000	23.388000	-0.035000	11.000000
<b>75%</b>	107.305000	23.484000	0.018750	11.000000
<b>max</b>	118.641000	24.448000	1.146000	11.000000

	CondHighRng	Temp	SpecificConductance	Salinity	Mes
<b>count</b>	2394.000000	2394.000000	2394.000000	2394.000000	2394.000000
<b>mean</b>	54524.972807	23.316646	56386.831662	37.479737	10.781119
<b>std</b>	11.876669	0.547033	619.501987	0.464974	0.413574
<b>min</b>	54105.700000	21.490000	54622.100000	36.158800	10.000000
<b>25%</b>	54525.500000	22.730000	55949.700000	37.151400	11.000000
<b>50%</b>	54525.500000	23.490000	56185.600000	37.328300	11.000000
<b>75%</b>	54525.500000	23.700000	57053.700000	37.980300	11.000000
<b>max</b>	54525.500000	24.910000	58398.700000	38.994200	11.000000

## 5.2 Diagramas de Pearson

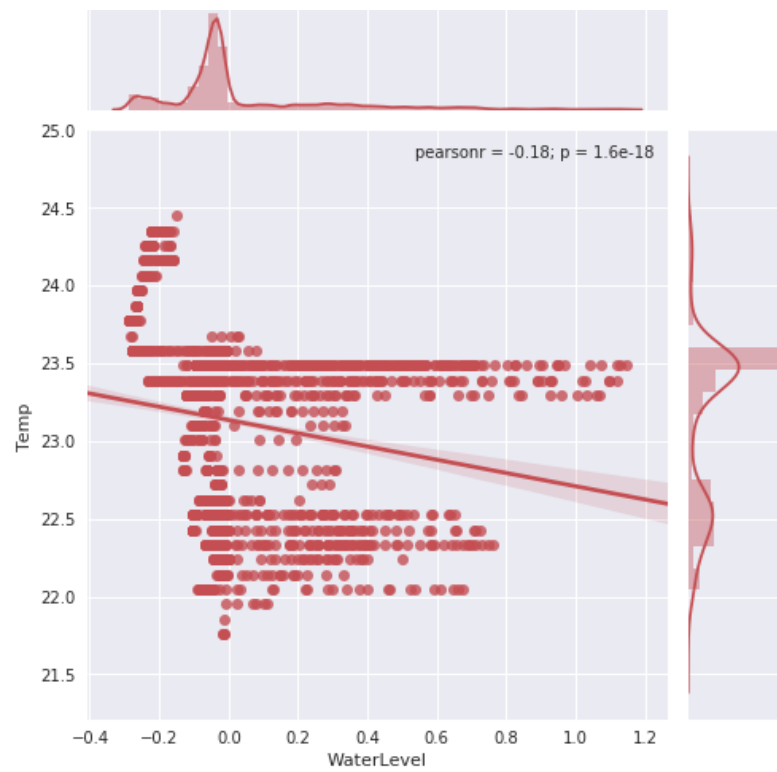
- Salinidad-Nivel del agua:



Con esta gráfica podemos observar que si existe una correlación lineal entre la salinidad y el nivel del agua, aunque no muy fuerte, el  $\text{pearsonr} = 0.13$ . Se observa que la distribución de la salinidad cuenta con dos picos, mientras que la del nivel del agua con uno.

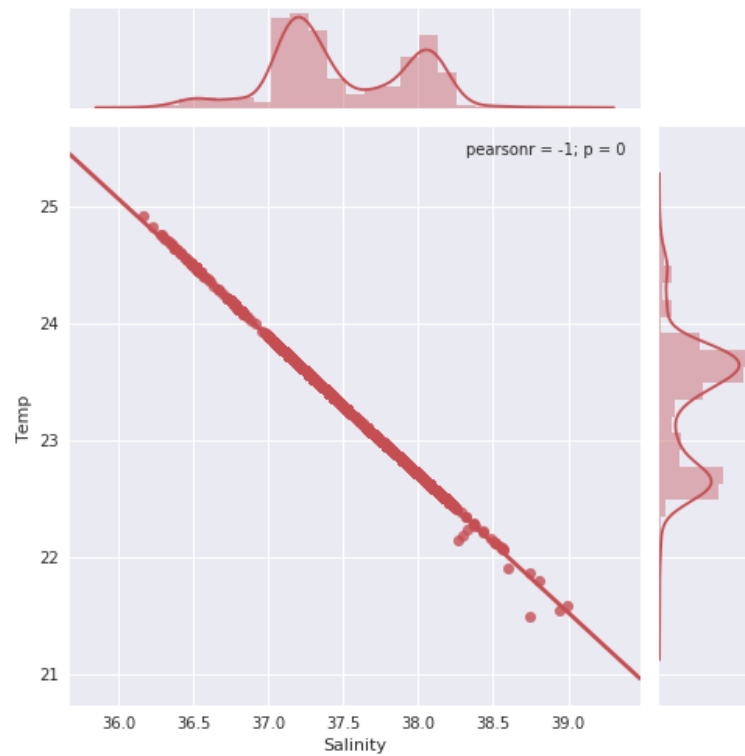
- Temperatura-Nivel del agua:

La relación que existe entre estas dos variables es negativa y tampoco es muy grande,



aunque existe, ya que el pearson es diferente de 0 y es igual a -0.18

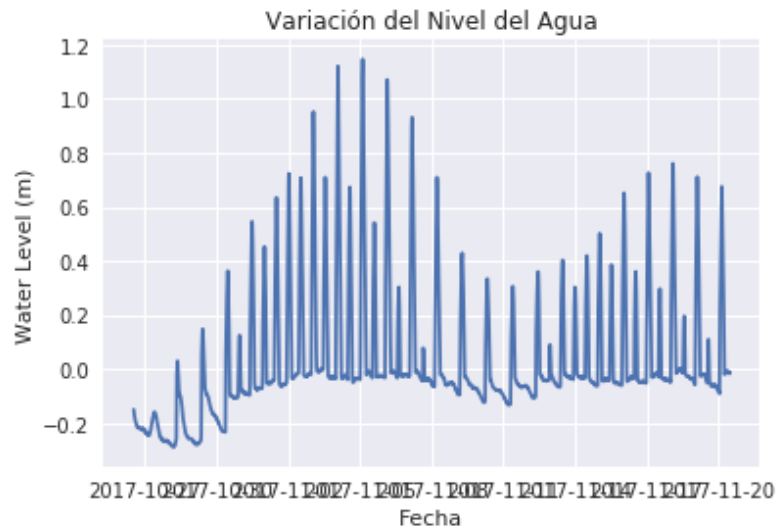
- **Temperatura-Salinidad:**



Estas dos variables se encuentran perfectamente correlacionadas negativamente, el pearson es igual a -1, se puede observar que ambas tienen los mismos picos, sus distribuciones, como es de esperarse, son bastante similares.

### 5.3 Gráficas en función del tiempo

- Nivel del agua:



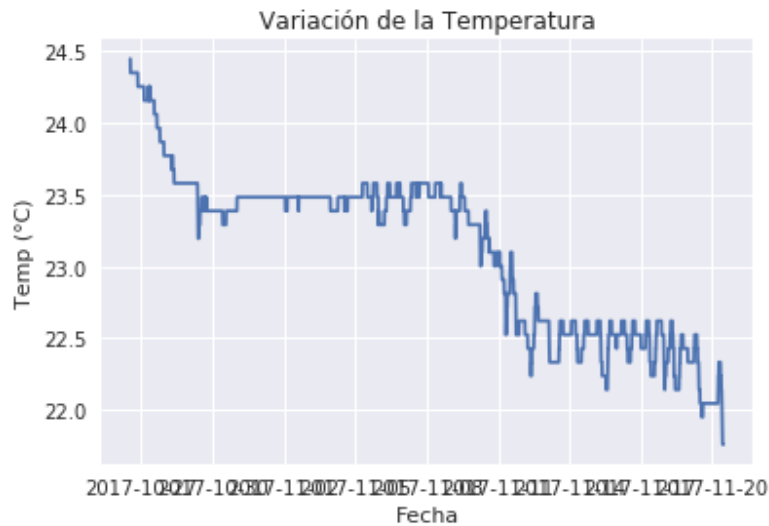
El nivel del agua, conforme el tiempo, empezó a aumentar a principios de noviembre, pero posteriormente disminuyó, para volver a incrementar, teniendo con esto una distribución con dos picos.

- Salinidad:



La salinidad, de manera general y aunque cuenta con dos picos a principios de noviembre, fue aumentando, conforme el tiempo iba pasando.

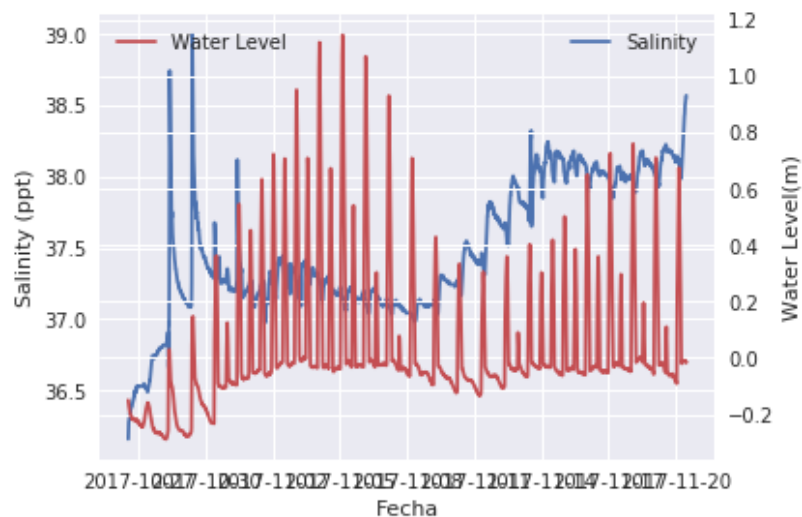
- **Temperatura:**



Como era de esperarse, la temperatura fue disminuyendo al paso del tiempo, esto debido a que se acercó el invierno.

## 5.4 Gráficas de doble eje vertical

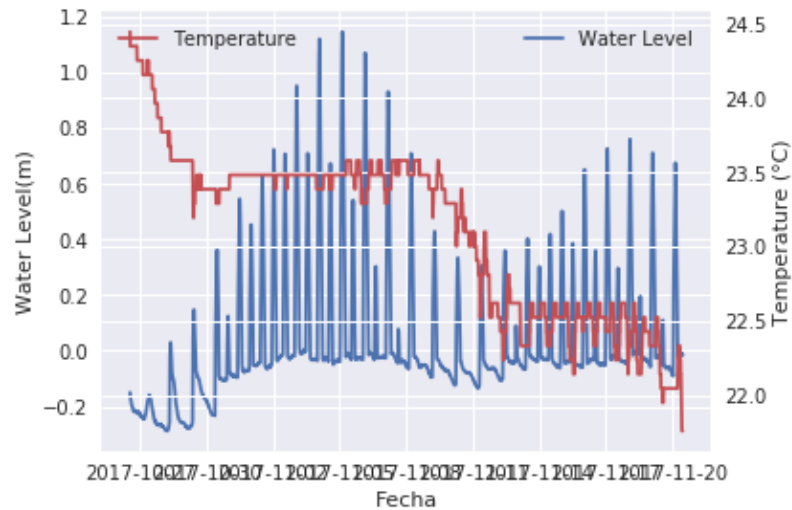
- **Salinidad-Nivel del agua:**



En esta gráfica se puede comparar la distribución que tienen la salinidad y el nivel del mar en función del tiempo. Se puede ver que las distribuciones no son precisamente

parecidas, aunque como se vió con el diagrama de Pearson, podría existir una relación entre ambas.

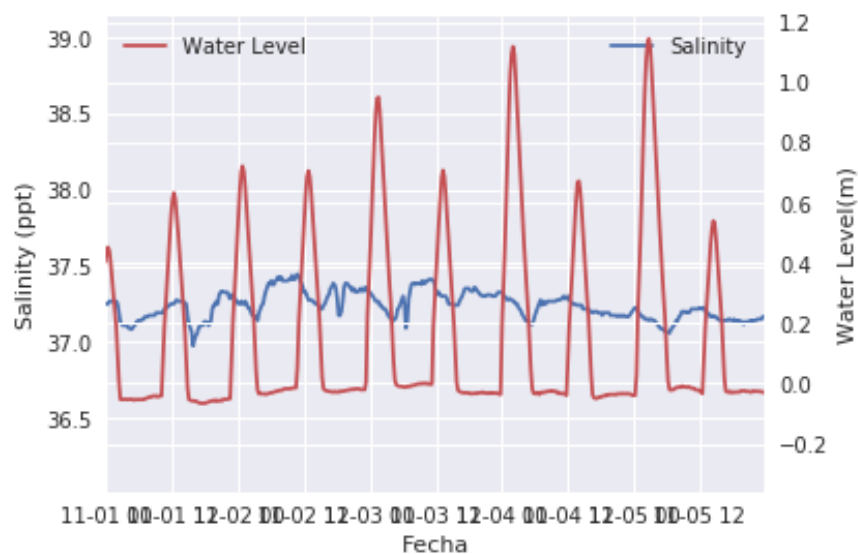
- **Temperatura-Nivel del agua:**



En esta gráfica se puede comparar la distribución que tienen la temperatura y el nivel del mar en función del tiempo. A simple vista, no se puede observar una dependencia o relación entre ambas, aunque por el diagrama de Pearson, no se puede descartar la existencia de una.

## 5.5 Gráficas de doble eje vertical en 5 días

- **Salinidad-Nivel del agua:**



En mi opinión no se observa una dependencia entre ambas, ya que mientras el nivel del mar tiene un comportamiento, casi periódico a lo largo de un día, la salinidad permanece alrededor de una constante, no se ve que el cambio de una afecte a la otra.

The graph displays two data series over time from 11-01 to 12-05. The left y-axis represents Water Level (m) from -0.2 to 1.2. The right y-axis represents Temperature (°C) from 22.0 to 24.5. The Temperature series (red line) is mostly constant at 23.5°C, with some step-like changes. The Water Level series (blue line) shows frequent, sharp fluctuations between approximately -0.1m and 1.1m.

De igual manera, el nivel del mar tiene un comportamiento que se repite en un mismo día donde a ciertas horas aumenta y a ciertas horas disminuye, mientras que la temperatura parece permanecer constante independientemente de que el nivel del agua cambie.

## 6 Conclusiones

Por último y para concluir, podría agregar que la única relación que vi bastante evidente fue la de la temperatura y la salinidad, a medida que la temperatura disminuye, la salinidad aumenta, y viceversa. La demás relaciones de dependencia no las podría defender como ciertas, a pesar del índice de pearson diferente de cero.

Además, como conclusión sobre el examen, me pareció bastante entretenido, aunque un poco largo e imposible de hacer en dos horas, si fue algo agotador, pero valió la pena porque en este me di cuenta de lo mucho que he aprendido y de la facilidad que adquirí para resolver los problemas con los que me encontraba utilizando tutoriales de internet, cosa que se que complicaba al entender la sintaxis con la que se trabajaba. Me gustó y el contenido fue adecuado a lo visto en las actividades anteriores.

## 7 Bibliografía

- Nivel del mar (2018). Consultado: 8 de Marzo del 2018, de Wikipedia. Sitio web: [https://es.wikipedia.org/wiki/Nivel\\_del\\_mar](https://es.wikipedia.org/wiki/Nivel_del_mar)
- Salinidad (2018). Consultado: 8 de Marzo del 2018, de Wikipedia. Sitio web: <https://es.wikipedia.org/wiki/Salinidad>