# Determinants of Temperature and Salinity in the Levantine Sea

Cem Serimozu – cem.serimozu@metu.edu.tr

Middle East Technical University – Institute of Marine Sciences – TURKEY

## Abstract

The Levantine Sea (LS), part of the Eastern Mediterranean, and four of its sub-regions, Cilician (CB) and Levantine (LB) Basins, Coastal Nile Delta (CND) and Rhodes Gyre (RG) is the focus of this study. The aim is to contribute to the investigation of Sea Water Temperature (SWT) and Salinity (SWS), two critical oceanographic parameters in the context of climate change. An aggregated in-situ historical dataset, covering the area between 1960 and 2017, is retrieved to conduct an in-depth descriptive analysis for multiple depth layers and seasons. Empirical and theoretical cumulative distribution functions of the data supported by classical statistical methods are used to establish optimal theoretic models for modelling purposes showcasing that SWT follows a lognormal distribution. On the other hand, SWS is compliant with beta distribution in shallow waters and non-parametric at greater depths. Regression analyses are performed to measure spatiotemporal parameters' significance using predicted values showcasing that in-situ data suffers from sampling bias and is ill-suited for giving a genuinely unbiased overview of regional seawater properties. Comparison between empirical and predicted value show discrepancies which indicate a need to use determinants to mitigate endogeneity and perform exogenous linear regression when models are developed combined with careful data selection, especially at sub-regional level.

## Methodology

Climatological and isothermal maps are used to sketch the boundaries of the sub-regions. Meanwhile, Cluster analysis using Clustering Large Application (CLARA) [1] algorithm is employed to define the limits and further identify distinct regional properties. Moreover, absent observations are replaced with approximations using chained equation method to increase the robustness of the analysis. The resulting datasets are analysed on a yearly and monthly basis to determine variability.

To determine the distribution of the data, empirical and theoretical distribution of the seasonal observations at each depth level is conducted and tested using Cullen & Frey graphs (ex: Fig. 1) and Cumulative Distribution Functions. Anderson Darling (AD) goodness of fit test is used to ascertain which of the theoretical models correspond to our data statistically and which model is the most suitable at a certain depth level and season. The results for each season and depth level are further controlled using an Akaike's goodness of fit information criterion (AIC) measurement. Additionally, regression analysis according to the distribution patterns are conducted on the data to determine the significance of different geospatial factors on SWT and SWS and their interaction with each other.

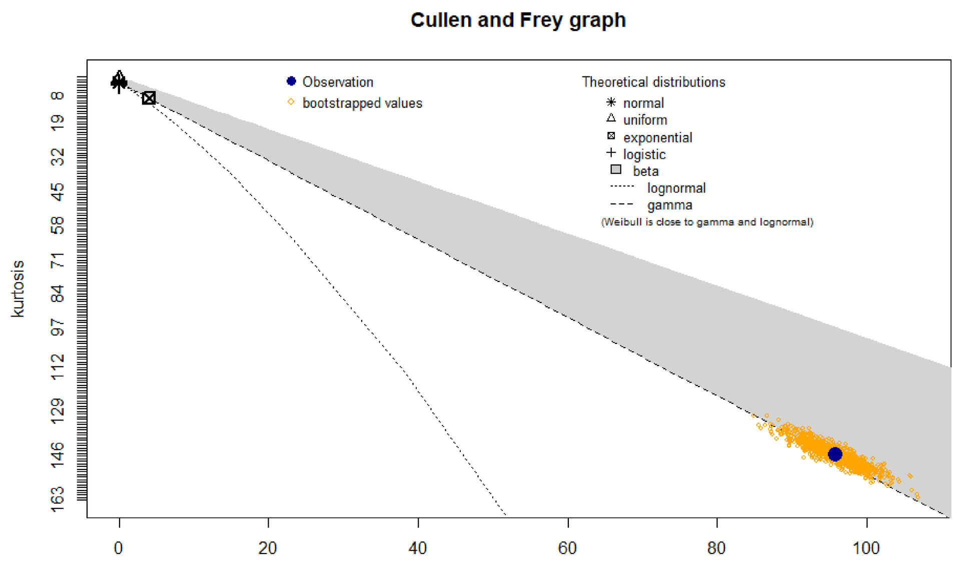



Fig. A1 Cullen and Frey graph for the 500m layer in summer of the LS, 1960-2017 period

## Data

The aggregated in-situ dataset, containing 81,317 stations and 10,590,891 individual entries (Fig. 2), "Mediterranean Sea-Temperature and Salinity Historical Data Collection SeaDataCloud V1" [2], is the most expansive collection of regional measurement currently available and is used in this work

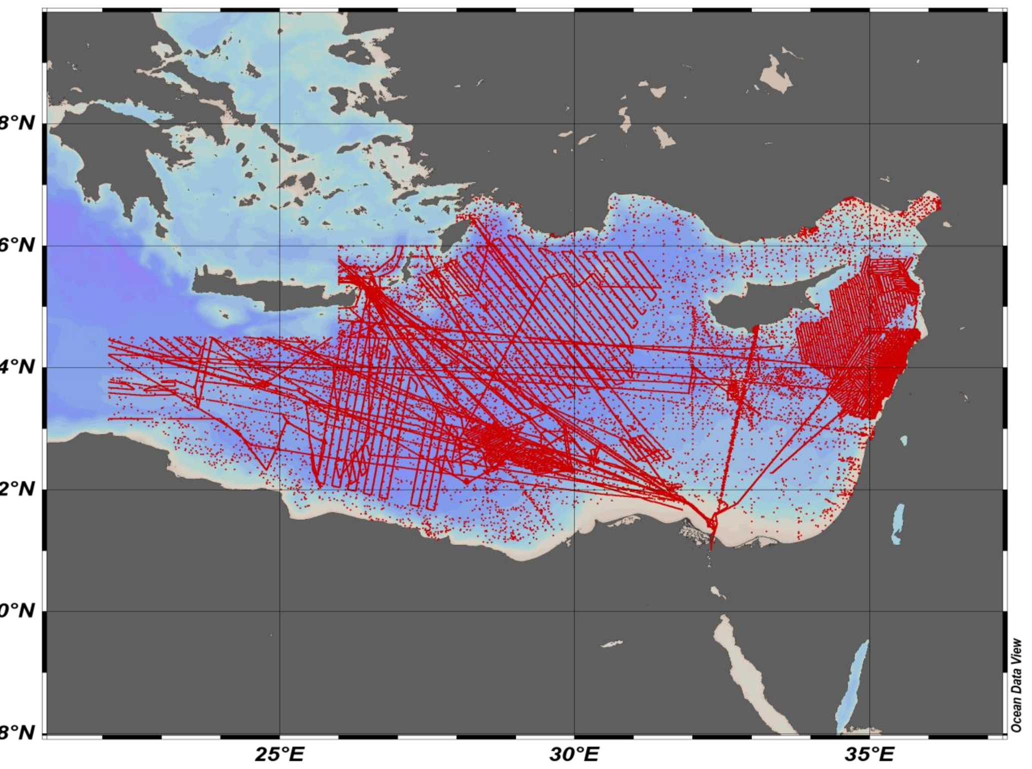
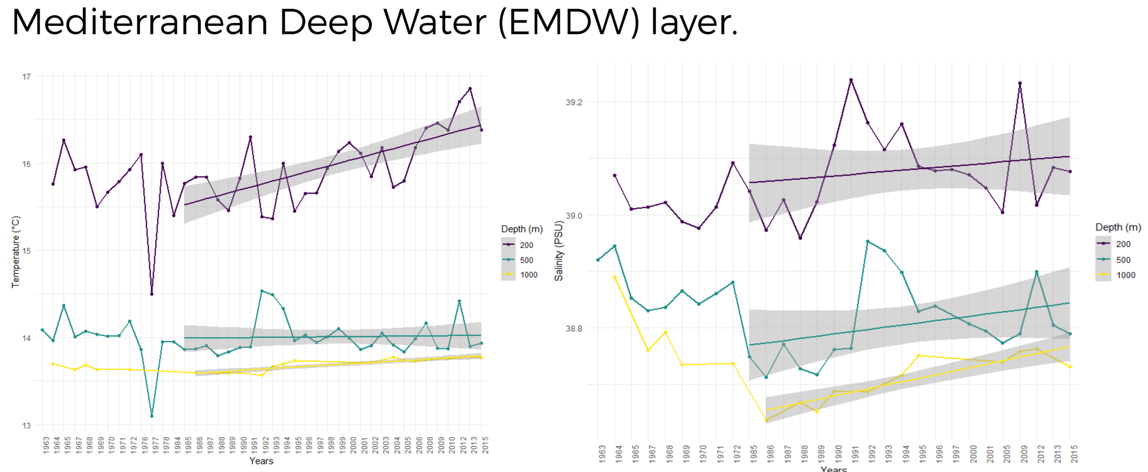


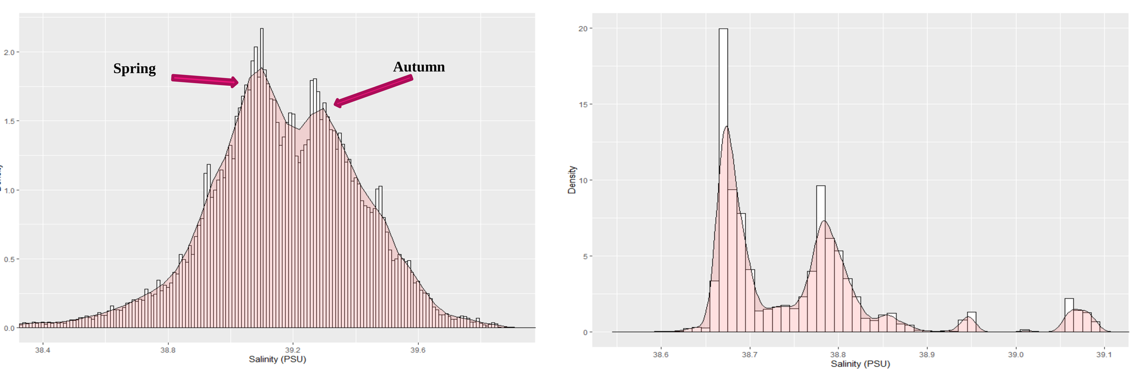

Fig. 2: Map of the region showing spatial distribution of the historical data

## Descriptive Analysis

There is a significant change in data availability post-1979 and in the parameter absence ratio. The LS's most studied region is LB, followed by RG, while the least explored areas are the CB and CND. Yearly plots have clear positive trends for SWT and SWS in the LS at the surface and intermediate waters for all sub-regions with a noticeable signature through the Levantine Intermediate Water (LIW) (Fig.3 & 4), particularly in the CB in the past 20 years. The data density plots for SWT and SWS are bimodal at the surface in accordance to seasonal atmospheric temperature differences (Fig. 5) before taking a mounded appearance around 200m to 500m in the LIW interval. SWS becomes bimodal at 2000m (Fig. 6), in the Eastern Mediterranean Deep Water (EMDW) layer.



Figures 3 & 4: SWT (Left) and SWS (Right) weighted average plots for the CB 1963-2015 between 200m-1000m




Figures 5 & 6: Distribution and Density plot of SWS values at: 10m (Left); 2000m – EMDW (Right)

## Distribution and Empirical Regression

From the preliminary density plots, it was expected that the potential distribution of SWT and SWS to be a lognormal, normal, beta or gamma trend. Calculations indicate that approximatively 25% of the layers have non-parametric distributions for SWT and SWS (ex. Fig. 7).

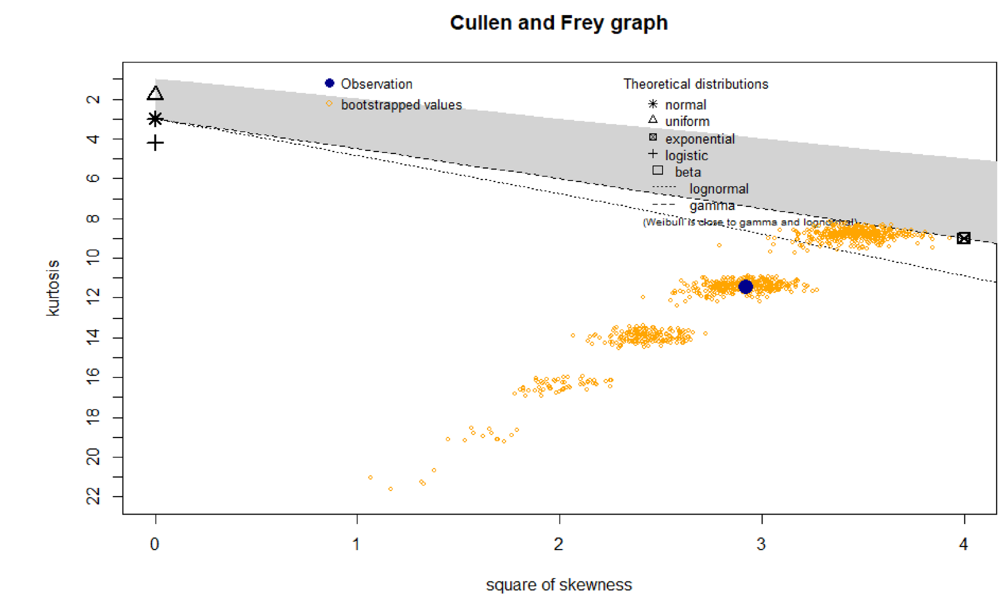



Fig. 7: Cullen and Frey graph for the 500m layer in winter of the LS, 1960-2017 period

Theoretical lognormal distribution has the highest ration of the goodness of fit according to the AIC, where 12 out of 13 results are confirmed. Lognormal distribution is followed closely by the beta distribution with a ratio of 11 to 15. Weibull is only ascertained for 5 out of 16 AD outcomes. Meanwhile, gamma distribution is not established for any layer. SWT seems to conform majorly to a lognormal, secondly Weibull distribution and few instances of beta distribution. On the other hand, SWS is mostly compliant with beta distribution, followed by lognormal and few instances of Weibull.

The actual distribution pattern of the data make it impossible to do regression analysis using exogenous (i.e. predicted) variables for SWT and SWS with the established distribution functions. The way the data is sampled creates clusters that are densely condensed both in time and space compared to the total area we are working with, resulting in very similar groups of values and preventing convergence, especially for SWS. The aforementioned issues and many other indicators showcase that the data that ill-suited, meaning that the sampling process is exceptionally uneven in the region. Fortunately, due to the size of the dataset, the distributions of both SWT (Tbl. 1) and SWS (Tbl. 2) be taken as normal (Assumption of Normality), meaning that a linear regression is possible.

| Seasonal Temperature | Latitude | Longitude | 10m | 25m | 50m | 100m | 200m | 500m | 1000m | 2000m |
|---|---|---|---|---|---|---|---|---|---|---|
| Winter | ✓ | ✓ | ✗ | ✓ | ✓ | ✗ | ✗ | ✗ | ✗ | ✗ |
| Spring | ✓ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ |
| Summer | ✓ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ |
| Autumn | ✓ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ |

Tbl. 1: Linear regression for SWT using empirical data in LS from 1960 to 2017

| Seasonal Salinity | Latitude | Longitude | 10m | 25m | 50m | 100m | 200m | 500m | 1000m | 2000m |
|---|---|---|---|---|---|---|---|---|---|---|
| Winter | ✓ | ✓ | ✓ | ✓ | ✓ | ✗ | ✗ | ✗ | ✗ | ✗ |
| Spring | ✗ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✗ | ✗ | ✗ |
| Summer | ✓ | ✓ | ✓ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ |
| Autumn | ✓ | ✓ | ✓ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ |

Tbl. 2: Linear regression for SWS using empirical data in LS from 1960 to 2017

The ✓ symbol indicate an increase in the value of SWT or SWS as the paired value increases or as we approach the defined depth layer while ✗ denotes a decrease.

## Predicted Value Regression

Using empirical SWT and SWS values combined with ordinary least squared method, predicted SWT and SWS, are obtained to mitigate the endogeneity. This allows for exogenous linear regression by replacing empirical SWT (Tbl. 3) and SWS (Tbl. 4) values with the predicted ones and define determinants.

| Seasonal Temperature | Latitude | Longitude | Predicted SWS |
|---|---|---|---|
| Winter | ✓ | ✗ | ✓ |
| Spring | ✓ | ✗ | ✓ |
| Summer | ✓ | ✗ | ✓ |
| Autumn | ✗ | ✗ | ✗ |

Tbl. 3: Exogenous seasonal linear regression for SWT using predicted SWS in LS from 1960 to 2017

| Seasonal Salinity | Latitude | Longitude | Predicted SWT |
|---|---|---|---|
| Winter | ✗ | ✓ | ✓ |
| Spring | ✗ | ✓ | ✓ |
| Summer | ✓ | ✓ | ✓ |
| Autumn | ✓ | ✓ | ✓ |

Tbl. 4: Exogenous seasonal linear regression for SWS using predicted SWT in LS from 1960 to 2017

## Conclusions

SWT and SWS show a positive trend in the region after 1985 with different magnitudes across the study period. The easternmost areas such as CB and LB have more significant positive trends compared to the whole LS. LIW properties tend to be consistent throughout the LS, although positive changes to the SWT and SWS values in the intermediate layer are perceptible. Density plots of SWS potentially indicate two slightly different water masses displaying different SWS properties instead of the traditionally accepted homogenous distribution of the EMDW in the oceanographic literature centred on the RG deep water formation.

Results of the empirical regression analysis were within expectation except for SWT decreasing the further east we moved, potentially an outcome of the fact that the analysis does not account for the depth indicating. The LIW and EMDW might have variable SWT at deeper layers compared to the assumed more homogenous distribution and require further study, as shown by the descriptive analysis and the clear depth signal shown in the SWS regression. Exogenous regression shows that latitudinal results using predicted values are similar to those found in the second-stage linear regression with an increase in SWT when moving eastward which was not expected. Northward SWS increases which might be due to the properties of the LIW and EMDW and further analysis is required.

This discrepancy between predicted and empirical values, originating from data collection patterns in the region, as sources for models demonstrate that a raw coverage of the area might create conflicting results and great care should be given to data collection patterns spatiotemporally to monitor the changing water conditions in the region.

## References

[1] Mediterranean Sea - Temperature and Salinity Historical Data Collection SeaDataCloud V1 (Last Accessed 29/06/2021)

[2] Rousseeuw, P. J., & Kaufman, L. (1990). Finding groups in data. Hoboken: Wiley Online Library. DOI: 10.1002/9780470316801