# Dependent Indicators for Environmental Evaluation of Desalination Plants

Ghanima Al-Sharrah and Haitham M.S. Lababidi

Chemical Engineering Department
College of Engineering & Petroleum
Kuwait University

Next Generation Brine Desalination and Management for Efficiency, Reliability, and Sustainability

Kuwait Foundation for the Advancement of Sciences
Massachusetts Institute of Technology

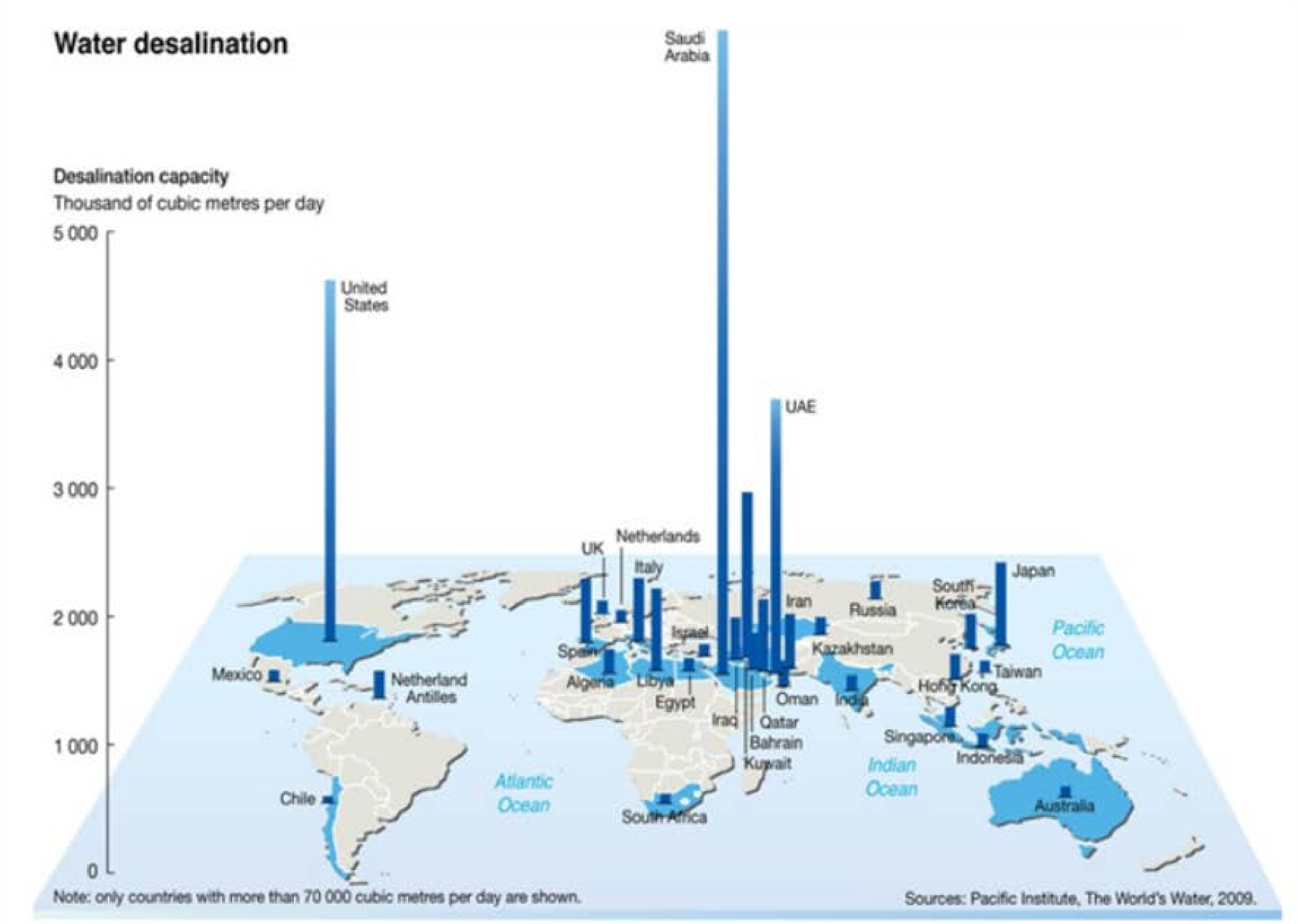
27th October 2018

12th Partial Order Workshop
International Workshop on Partial Orders in Applied Sciences

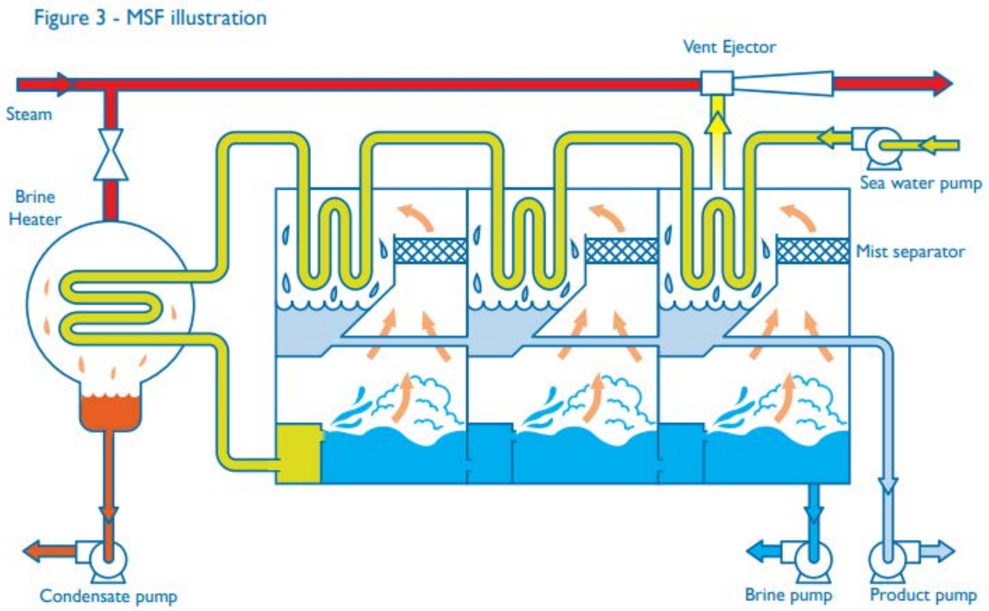
# Outline

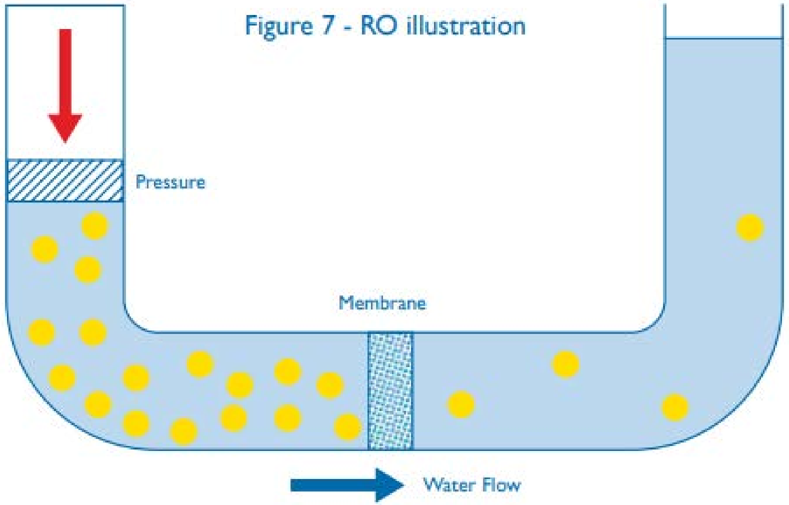
- Background
- Dependent data for brine and sea water
- Ranking with correlated data
- Case study
- Conclusions

# Background: Desalination




(Advision, 2018)



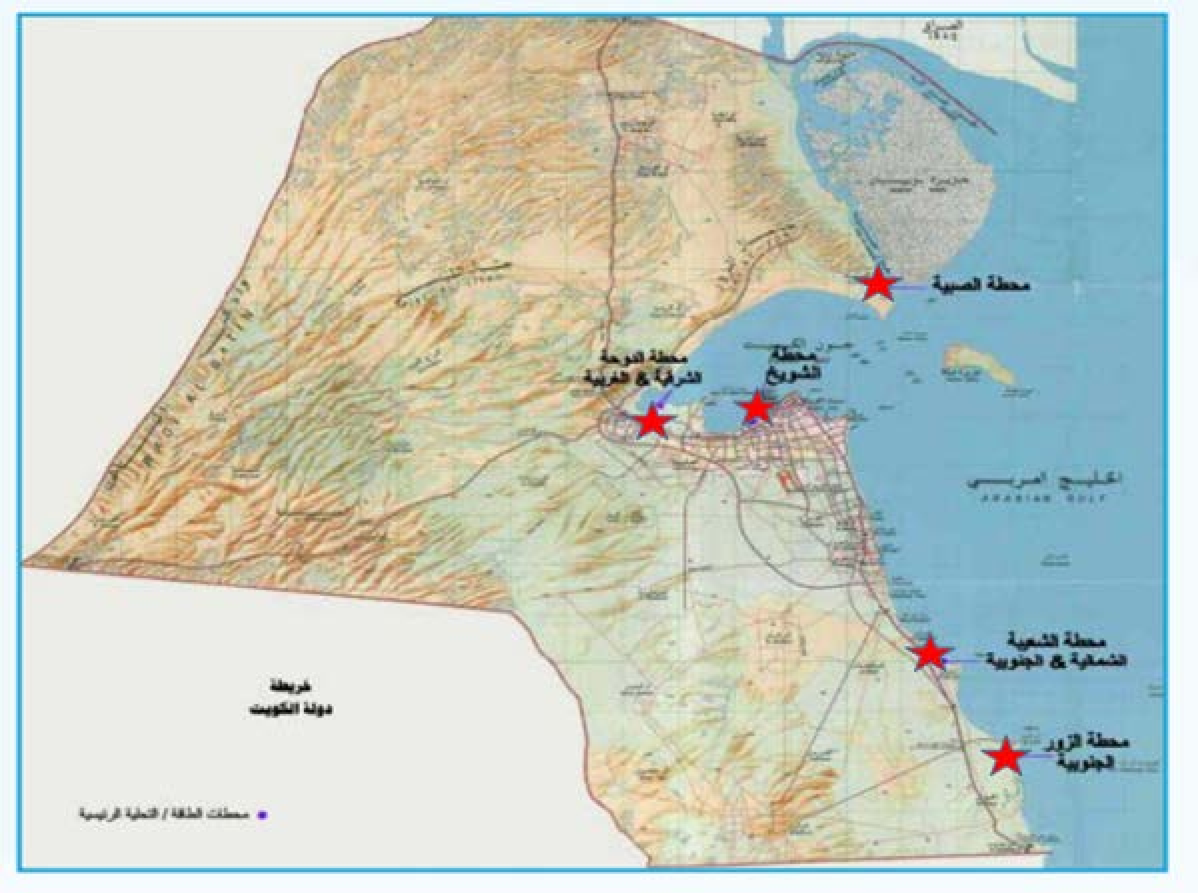

# Desalination

- MSF and RO are the main seawater desalination technologies in the GCC countries

(Desalination in the GCC, 2014)

# Desalination Plants in Kuwait

| Station | Technology | Capacity (MIG/d) |
|---|---|---|
| Shuwaikh | MSF | 19.5 |
| | RO | 30 |
| Shuaiba South | MSF | 36 |
| North | MSF | 45 |
| Doha East | MSF | 42 |
| West | MSF | 110 |
| Azzour | RO | 30 |
| | MSF | 115 |
| Sabiya | MSF | 100 |



# Dependent indicators

- We need indicators to study desalination activities.
- Limited data available.
- Some of the indicators are dependent on each other.
  - Example: energy consumption and $CO_2$ emission.
- They are usually correlated using mathematical models.
- Dependencies between indicators can be identified through:
  - Expert knowledge, or
  - Mathematically by using statistical analysis of collected data.

# Dependent Indicators for Seawater and Brine

| No | Dependent Indicators | | Note | Reference |
|---|---|---|---|---|
| 1 | Electric conductivity | chlorine ($Cl^-$) and sodium ($Na^+$) ions | linearly dependent | Fondriest (2015)<br>Sharp and Culberson (1982) |
| 2 | Total hardness , expressed as equivalent $CaCO_3$ | $Ca^{2+}$ and $Mg^{2+}$ ions | $[CaCO_3]=2.5[Ca^{2+}]+4.1[Mg^{2+}]$ | Venkateswarlu (1996)<br>LENNTECH 2016 |
| 3 | Langelier Saturation Index | total dissolved solids (TDS), concentrations of calcium ($Ca^{2+}$) and bicarbonates ($HCO_3$), and water temperature | The LSI is expressed as the difference between the actual system pH and the saturation $pH_s$ (LSI = pH – $pH_s$). The saturation $pH_s$ a log function | Alvarez-Bastida et al. (2013) |
| 4 | Density , viscosity | Temperature Salinity | Empirical correlation valid for salinities between 0 to 160 ppt and temperature between 10 to 180ºC at pressure of 1 atm | El-Dessouky and Ettouny (2002) |
| 5 | Dissolved oxygen | pH | Nonlinear | Makkaveev 2009 |
| 6 | Dissolved oxygen | Temperature Salinity | Nonlinear | Lewis (2005) |
| 7 | Total alkalinity (At) | total amount of calcium carbonate | $A_T=[HCO_3^-]+2[CO_3^{-2}]$ (mmol/l)<br>$CO_3^{-2}$ (mg/l)=0.6 At (mg/l)<br>$HCO_3^-$ (mg/l)=1.22 At(mg/l) | Danoun (2007)<br><br>California environmental protection agency (2016) |
| 8 | Carbonate ($CO_3^{-2}$ and $HCO_3^-$ ) | pH | Non-linear | Holmes-Farley (2002) |
| | | | | |

# Dependencies

- Can dependencies in indicators help?
  - How? And to what extent?
- Can correlation help reduce the number of indicators and complete/check collected data?
- Is the exclusion of indicators possible?
- What is the effect on decision making using ‘’Ranking Methods’’?
- We need to pereared before we receive data.

# Dependencies : Approach

- Original Data:
  - Sequences of uncorrelated normal distributed random indicators $X_1, X_2, \ldots, X_n$ for hypothetical objects $Obj_1, Obj_2, \ldots, Obj_m$ will be generated.
- Extended Data:
  - A correlation model is selected and a dependent indicator $Y_1$ is generated for all objects.
- Decision ranking:
  - Ranking is performed using the original and the extended data.
  - Results of ranking will order objects (assigned a numerical rank) from top to bottom to represent the most and the lease important object i.e. decision-making.
- Comparison:
  - Ranks from original and extended data are compared using the Spearman's rank correlation coefficient (SRCC).

# Ranking Method

- Modified Copeland method:
  - Sum of greater than minus less than relations in a pairwise comparison between indicators

Three Indicators ⟷ Two indicators

| | | Indicator | | | (a) | | Indicator | | (b) |
|---|---|---|---|---|---|---|---|---|---|
| | | $I_1$ | $I_2$ | $I_3$ | Copeland rank | | $I_1$ | $I_2$ | Copeland rank |
| Objects | $O_1$ | 1 | 3 | 6 | 1 | | 1 | 3 | 0 |
| | $O_2$ | 0.5 | 2 | 4 | -5 | | 0.5 | 2 | -4 |
| | $O_3$ | 5 | 1 | 2 | -4 | | 5 | 1 | -1 |
| | $O_4$ | 5 | 5 | 10 | 8 | | 5 | 5 | 5 |

# Ranking with Correlation

- Random data generated to represent *n* indicators for *m* objects.
- Different correlation models were used to represent dependencies.
- Original data extended using the correlation model (data extended to *n*+1 indicators for *m* objects).
- Ranks were evaluated for both data sets
- Ranks were compared.
- The methodology was implemented using MATLAB.

| Model for correlation | example |
|---|---|
| Linear | $Y_1 = a_0 + a_1 X_1 + \varepsilon_1$ |
| Multiple linear | $Y_1 = \rho X_1 + \sqrt{(1-\rho^2)} X_2$ |
| Non linear | $Y_1 = a_2 X_1^2 + \varepsilon_1$ |
| Complex | $Y_1 = \log(Y_2 \times Y_3) + Y_4$ |

# Ranking with Correlation : Results

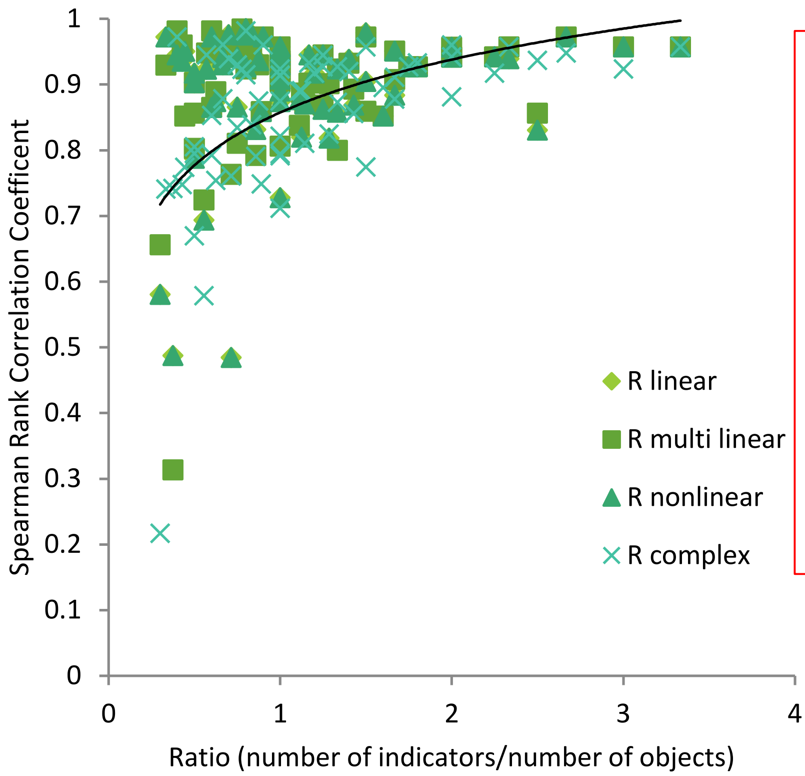
- 64 data sets
- Sizes up to 10 objects with up to 10 indicators
- Each data is ranked and then re-ranked after extension using 4 models.
- SRCC was relatively high. But is it guaranteed for all data sets?

| Model for correlation | example | SRCC |
|---|---|---|
| Linear | $Y_1 = a_0 + a_1 X_1 + \varepsilon_1$ | 0.892 |
| Multiple linear | $Y_1 = \rho X_1 + \sqrt{(1-\rho^2)} X_2$ | 0.894 |
| Non linear | $Y_1 = a_2 X_1^2 + \varepsilon_1$ | 0.892 |
| Complex | $Y_1 = \log(Y_2 \times Y_3) + Y_4$ | 0.857 |

Exclusion of dependent indicators will slightly affect rank of objects .

# Ranking with correlation

- What is the effect of data size?
- We need to know is the pervious result applicable for high/low number of indicator or high/low number of object?
- Use  ratio → Indicators/Objects
- Now test random data then test deal data.




if number of indicators are high relative to objects, excluding an indicators will not affect ranking result and vice versa

# Case Study: Desalination plants in GCC

| No | Parameter | Alssadanat, Oman (a) | Umm Alquain, UAE (b) | Hamriyah, Sharjah, UAE (c) | Saja'a Sharjah UAE (d) | Buwaib Saudi Arabia (e) | Salboukh Saudi Arabia (f) |
|---|---|---|---|---|---|---|---|
| 1 | $Ca^{++}$, mg/l | 923 | 202 | 173 | 188 | 573 | 404 |
| 2 | $Mg^{++}$, mg/l | 413 | 510 | 311 | 207 | 373 | 257 |
| 3 | $Na^{++}$, mg/l | 2780 | 3190 | 1930 | 4800 | 2327 | 1433 |
| 4 | $K^{++}$, mg/l | 81.5 | 84.5 | 50.7 | 60 | NA | NA |
| 5 | $Sr^{++}$, mg/l | 28.2 | 21.1 | 14.2 | 40 | NA | NA |
| 6 | Sum cation, meq/l | 203.06 | 192.98 | 119.48 | NA | NA | NA |
| 7 | pH | 7.21 | 7.54 | 7.66 | 7.95 | 4.1 | 4.5 |
| 8 | Electrical conductivity, mS/cm | 16.8 | 14.96 | 127.41 | NA | NA | NA |
| 9 | TDS, mg/l | 10553 | 10923 | 7350 | 12239 | 10800 | 6920 |
| 10 | $NO_3$, mg/l | 7.2 | 27.4 | 15.9 | NA | 143 | 142 |
| 11 | $F^-$, mg/l | 0 | 1.6 | 1.3 | 8.0 | NA | NA |
| 12 | $Cl^-$, mg/l | 4532 | 4108 | 2933 | 4860 | 2798 | 1457 |
| 13 | $SO_4$, mg/l | 1552 | 2444 | 1537 | 2400 | 4101 | 2840 |
| 14 | $SiO_2$, mg/l | NA | 164.09 | 133.71 | 120 | NA | NA |
| 15 | Carbonate ($CO_3^-$), mg/l | NA | NA | NA | NA | NA | NA |
| 16 | Bicarbonate ($HCO_3^-$), mg/l | 466 | 656 | 753 | NA | NA | NA |
| 17 | $N^-$ | 1.6 | 6.2 | 3.6 | NA | NA | NA |
| 18 | Sum anions, meq/l | 167.88 | 198.05 | 127.41 | NA | NA | NA |
| 19 | Ion balance | 9.48 | 4.02 | -3.21 | NA | NA | NA |
| 20 | SAR | 19.12 | 27.2 | 20.3 | NA | NA | NA |
| 21 | SER | 59.55 | 71.91 | 70.27 | NA | NA | NA |
| 22 | L.I | 1.24 | 1.04 | 1.26 | NA | NA | NA |
| 23 | R.I | 4.73 | 5.46 | 5.14 | NA | NA | NA |
| 24 | Total ion, mg/l | 10781 | 11245 | 7719 | NA | NA | NA |
| 25 | Total alkalinity | 380 | 538 | 617 | NA | NA | NA |
| 26 | Total hardness | 4041 | 2630 | 1730 | NA | 2968 | 2066 |
| 27 | Fe, meq/l | 0.06 | 0.08 | 0.05 | NA | 65.5 | NA |

Some data is missing

# Case study

- For dependencies, the analysis requires  deep analysis and subset data selection.
- The data can be divided into sets to test indicators dependencies with different data sizes.
- Data selected then correlated.

# Case study

| Sub set No | Objects | Original Indicators | Ratio= Indicator to objects | Extended Indicator | Type of Correlation | SRCC |
|---|---|---|---|---|---|---|
| **1** | a,b,c | All indicators (rows) except:<br>• $SiO_2$<br>• Carbonate<br>• Hardness | 24/3 = 8 | Carbonate | Linear | 1 |
| | | | | LSI | Complex | 1 |
| | | | | Hardness | Multi-linear | 1 |
| **2** | a,b,c | • $Ca^{++}$<br>• $Mg^{++}$<br>• pH<br>• Electrical Conductivity<br>• Bicarbonate<br>• Alkalinity | 6/3 = 2 | Carbonate | Linear | 1 |
| | | | | LSI | Complex | 1 |
| | | | | Hardness | Multi-linear | 1 |
| **3** | a,b,c,d,e,f | • $Ca^{++}$<br>• $Mg^{++}$<br>• pH<br>• $Na^{+}$ | 5/6 = 0.83 | Carbonate | Linear | 0.8286 |
| | | | | LSI | Complex | 0.89 |
| | | | | Hardness | Multi-linear | 0.9559 |

# Conclusions

- Exclusion of indicators is possible with the Copeland method.
- Exclusion of dependent indicators is safe if number of indicators to number of objects is relatively high.
- Exclusion of dependencies is safer for simple dependence relationship.
- Environmental indicators are highly dependent
- It is important to study dependencies for assessment or ranking.
- Identifying dependencies will provide a convenient way for completing missing data.