National Taiwan University **CompStats4DataAnalytics**
Department of Civil Engineering Instructor: L.-P. WANG

Final Assignment

Due 23:59, Thursday, December 29, 2022

No Late submission

Final Assignment Submission Procedure (請仔細閱讀)

1. You should submit your Jupyter notebook files.
2. Name a folder using your student id and Final (e.g., b88501025-Final), put all the scripts and the associated files into the folder and zip the folder (e.g., b88501025- Final.zip).
3. Submit your Assignment directly through the NTU COOL course website.

Total 105% + 5% bonus

1. **(30%+5% bonus)** Work with Jupyter notebook `CompStats-Final_Q1_SpatialAnalysis-Morakot`. Spatial structural analysis is widely used to analyse the influence range of a given phenomenon. In the attached `IMERG_HHL` folder, you can find 30-min IMERG precipitation data for period of 2009/08/07 00:00 – 2009/08/09 00:00. Your main task here is to calculate the characteristic range of Typhoon Morakot for the period it passed Taiwan (i.e. 2009/08/07 17:00 – 2009/08/08 08:00 UTC time) based upon the algorithm proposed in Ochoa-Rodriguez et al. (2015) (see Section 3.2.1 in the paper).

   A summary of this algorithm is given as follow:

   1) An empirical (or sample) variogram is computed at each time step as:

   $$\gamma(h) = \frac{1}{2N}\sum_{i=1}^{N}\left[\left(Z(\mathbf{x}) - Z(\mathbf{x}+h)\right)^2\right]$$

   where N is the number of all pairs of satellite pixels separated by a distance h, Z are the precipitation values at the respective pixels and $\mathbf{x}$ corresponds to the centre of a given pixel.

   2) Each empirical variogram is normalised by dividing it by the sample variance.

   3) The normalised variograms obtained for each time step are averaged over the time period of analysis – This yields a climatological sample variogram.

   4) An exponential variogram model is fitted to this climatological sample variogram. The exponential variogram model is termed:

   $$\gamma(h) = C_0 + C\left[1 - \exp(-\frac{h}{a})\right]$$

   where C_0 is the nugget, C is the sill and a is the range. However, please note that the practical (or effective) range of the exponential model, i.e., $r = 3a$, shall be used here.

   5) The integral range measure (A) is then estimated here. This measure summarises the structural information of the variogram provided by the range and the fraction of total variance. For an exponential variogram model, this measure is termed (Lantuéjoul, 1991):

$$A = \frac{2\pi r^2}{9}$$

6) The characteristic length scale (r_c), which represents the mean extent of the spatial structure captured by the data (Garrigues et al., 2008), can be estimated as the square root of A. That is,

$$r_c = \left(\frac{\sqrt{2\pi}}{3}\right) r \approx 0.836r$$

Your specific tasks are (**Note that please use 'degree' directly as distance throughout the following tasks**):

1.1. (15%) Derive climatological sample variogram using attached data files for the period of 2009/08/07 15:00 – 2009/08/08 08:00 (including image at 2009/08/08 08:00). (if you use gstools, please specify 'latlon=False' in vario_estimate).

1.2. (5%) Fit this sample variogram with an exponential variogram model.

1.3. (5%) Plot the sample and fitted variograms in a single chart.

1.4. (5%) Derive the characteristic length (in degree).

1.5. (**5% bonus**) You can get this bonus if you demonstrate that you can download IMERG HHL data files by yourself.

2. **(75%)** Work with Jupyter notebook `CompStats-Final_Q2_TC-NetworkAnalysis`. A tropical cyclone is a rotating, organized system of clouds and thunderstorms that originates over tropical or subtropical waters and has a closed low-level circulation. As quoted by WMO (World Meteorological Organization):

   Tropical Cyclone (TC) are one of the biggest threats to life and property even in the formative stages of their development. They include a number of different hazards that can individually cause significant impacts on life and property, such as storm surge, flooding, extreme winds, tornadoes and lighting.

   Mexico (MX) is located between Atlantics and East Pacific. Every year, MX suffers from several TCs and the financial and societal losses resulting from them.

   Your company is hired by the United Nations to undertake a project quantifying the potential impact TCs to MX. In this project, your tasks are to analyse TCs and MX's rainfall monitoring network. Specifically,

   2.1. (40%) Characterising the rain gauge network covered the entire MX.

   2.2. (5%) Analysing historical records of TC tracks for the period 1979 - 2019

   2.3. (30%) Identifying those *hotspot* states prone to be impacted by TCs passing by 'East Pacific' Ocean, and calculate rainfall sums of a given state for a given TC event.

References

Garrigues et al., 2008: Quantifying spatial heterogeneity at the landscape scale using variogram models, Remote Sensing of Environment, 103(1), 81-96.

Lantuéjoul, C., 1991: Ergodicity and integral range, Journal of Microscopy, 161(3), 387-403.

Ochoa-Rodriguez et al., 2015: Impact of spatial and temporal resolution of rainfall inputs on urban hydrodynamic modelling outputs: A multi-catchment investigation, Journal of Hydrology, 531, 389-407.