ESS 224/CEE260D: Remote Sensing of Hydrology

Problem Set 3, due Thursday May 13th, 2021 at 10:30 AM PDT

**Question 1 (3 points, plus 2 points extra credit)**

In this problem, we will try to better understand the relationship between reflectivity and rainrate. This relationship depends on the drop-size distribution of the specific storm, so this problem considers an example drop-size distribution.

(*Hint: If you don't have an easy capacity to calculate analytical integrals, Wolfram Alpha online is a good quick and dirty choice)*

a) (2 points) In class, we discussed how the radar reflectivity *Z* can be related to the drop-size distribution. In this problem, we'll assume the Marshall-Palmer drop-size distribution applies. That is,

$$N(D) = N_0 \exp(-cD)$$

Integrate this expression to derive an equation expressing *Z* as a function of the Marshall-Palmer parameters *c* and $N_0$.

b) (1 point) We also derived an equation for precipitation as a function of the drop-size distribution. Assuming the terminal velocity is much greater than the updraft velocity and is directly proportional to drop-size diameter ($V_t = \alpha D$), and assuming the minimum drop size is effectively zero, derive an equation for the precipitation rate P in terms of α and the drop-size distribution N(D).

c) (1 point extra credit) Now assume once more that the drop-size distribution of a given rainstorm follows the Marshall-Palmer distribution, as in part a. Derive an equation (including solving the integral) for the precipitation rate P in terms of *c, α,* and $N_0$.

d) (1 point extra credit) Derive an equation for the relationship between observed radar reflectivity and precipitation rate (in terms of α and $N_0$) under the above assumptions. What are the coefficients a, and b in the equation $Z = a P^b$? How do *a* and *b* depend on $N_0$? Draw a schematic of these functional forms (i.e. sketch what a plot of a vs. $N_0$ and a plot of b vs. $N_0$ might look like)

**Question 2 (7 points + 1 point extra credit)**

You've been hired as a consultant by an exceptionally wealthy and somewhat eccentric farmer (think Willy Wonka but with corn and soybeans instead of candy). The farmer has asked you to develop an ET model using remote sensing, publicly available climate data, and ET measurements from an eddy covariance tower for his fields. You are able to train the model using data from one of the irrigated fields (‘Ne1’) and also are provided with measurements from another irrigated field (‘Ne2’) and rainfed field (‘Ne3’) a little further away. All fields are nearby each other. Flux tower ET measurements have the units $W/m^2$. The farmer needs the model RIGHT AWAY, so you don’t have time to do anything too sophisticated.

Given the time constraints, we suggest you elect to build a very simple empirical model, using an approach similar to how some “real” reflectance-based ET remote sensing approaches work). This approach is based on using established equations that account for the effect of meteorological forcing (e.g. radiation, wind speed, temperature) on ET, but cannot account for the effect of the specific land surface conditions. The meteorological forcing is accounted for in a ‘reference ET’ (ET_ref), which is similar to ‘potential ET’ but calibrated to a specific reference crop that is ‘well-watered’ (no water limitations on ET). The model will estimate actual ET by scaling a reference ET value using information from remotely sensed measurements of vegetation greenness (as approximated using the Normalized Difference Vegetation Index, or NDVI). NDVI is useful for predicting ET since it is correlated with the leaf surface area.

a) (1 point) Generate an NDVI time series using the data provided (in the files `ne[1-3]_ndvi.csv). This data is extracted from [MOD13 data products]() using the [AppEARS]() point query tool. The data includes QA flags (QA stands for quality assurance), which indicate whether there may have been any contamination or data processing issues with the pixel in question. Filter out any observations where there is snow/ice or clouds in the pixel using the QA flags provided.

b) (2 points) In order to control for meteorological forcing variables, calculate the daily ‘evaporative fraction’ ($ET_oF$) by dividing daily flux tower ET by daily ET_ref. The ET_ref values we provide are extracted from the [gridMET reference ET data product]() and converted from mm/day to $W/m^2$. $ET_oF$ will be the response variable in your empirical model. Assuming leaf surface area (and thus NDVI) influences the EToF, we can write

$$ET_oF = aNDVI + b$$

To build your model, estimate the coefficients *a* and *b* of the above linear regression using all available Ne1 data (2008-2012). What are their values? Then use the resulting a and b to calculate ET based on NDVI, your coefficients, and ET_ref. Plot the flux tower ET and your model estimates of ET on the same axes.

c) (1 point) Now it is time to test your model at a different location! Use your NDVI-based regression model to estimate the ET for 2011-2012 for a nearby field (‘Ne2’). Plot the flux tower ET estimates and your model ET estimates on the same axis. Compute the RMSE and correlation between flux tower estimates and your model estimate.

d) (1 point) To get a sense of how well this simple model does, compare your model to the [MODIS ET product MOD16](). MOD16 data is provided in the file `ne2_pm_mod.csv` ('pm' stands for Penman-Monteith). These data are in the same format as you would get them if you downloaded them from the AppEEARS tool. First, filter out the observations with cloudy pixels using the provided flags, as these are unreliable. Then plot the MOD16 ET estimates on the same axis as the flux tower observations. Note that MOD16 estimates are averaged over 8-day composites, and the data product does not say which day-of-year the particular pixel observation comes from (unlike MOD13). Therefore you will need to aggregate the flux tower data into 8-day composites that match those of the MOD16 data.

(Note that we ignore representativeness errors in this calculation for simplicity, given the high resolution of MOD16 data. But that doesn't mean they don't exist!)

*Hint: check the units on MOD16 LE versus flux tower LE.*

e) (1 point) Apply your model to a third site ("Ne3"). Compute the RMSE and correlation with the flux tower estimates. Calculate and compare the bias in your model at Ne2 versus Ne3. What might be causing the difference in model accuracy at Ne2 versus Ne3?

*Hint: what do you know about these fields?*

f) (1 point) Now imagine that the farmer wants you to do similar calculations for a field much further away in New Jersey rather than Nebraska. Do you expect qualitatively similar results in your model performance? Why or why not?

g) (1 point extra credit) To compare the MOD16 data to your model more rigorously (and in a fair way), average your NDVI-based model's predicted ET across the same 8-day periods as MOD16. Compute relevant summary statistics that compare the accuracy of your simple model to the accuracy of MOD16 across the three sites. Qualitatively, how do they compare?