INF 550 Section 3.11 Intro to USA-NPN Culmination Activity

Natasha Wesely

2022-09-28

3.11 Intro to USA-NPN Culmination Activity

Write up a 1-page derived data product or research pipeline proposal summary of a project that you might want to explore using USA-NPN data. Include the types of USA-NPN (and other data) that you will need to implement this project. Save this summary as you will be refining and adding to your ideas over the course of the semester.

For my PhD research, I am trying to constrain a process-based model representing the major ecological processes in four different ecosystem types in the Sevilleta Wildlife Refugre area: a desert grassland, a plains grassland, a desert shrubland, and a juniper savanna. In these four different ecosystem types, LTER experiments have been going on for several years. There are drought experiments (i.e., reducing the total amount of precipitation entering the ecosystem), precipitation variability experiments (i.e., increasing the sporadicity of precipitation entering the ecosystem), and combination experiments with both drought and precipitation variability treatments. These drought and precipitation variability treatments reflect what climatologists anticipate in our future under the new climate regime.

I am planning to use the dryland version of TECO (Hou et al. 2021) as my foundational process-based model, which I will modify to (a) reflect the vegetation structure for each ecosystem type and (b) integrate more data streams.

Once we have successfully constrained the model using data, we can project into the future to make useful inferences about how these four different ecosystem types will be impacted by different potential climate regimes (e.g., more drought and more sporadic precipitation).

For a data product or research pipeline proposal related to NPN, I would like to add in NPN data to help constrain some of the carbon pool parameters in my process-based model. Incorporating observational NPN data could allow us to estimate phenolgical shifts, such as green-up and brown-down, which have major implications for the amount of photosynthesis going on and therefore the amount of carbon being sequestered. Once carbon "enters" the process-based model, it moves through the different plant issue and ultimately decomposes in the soil.

I propose to take NPN observations about spring and fall phenological shifts on plants that match our four ecosystem types of interest from the same around as the Sevilleta Wildlife Refuge, curate the data, and feed in into the model-data fusion algorithms I am using to constrain the parameters in our process-based model.

Question 1

Tables or bullet lists of specific data products

1) USA-NPN observational data products for black gramma grass (*Bouteloua eriopoda*), the dominant species of the desert grassland ecosystem type at the Sevilleta Wildlife Refuge

2) USA-NPN observational data products for blue gramma grass (*Bouteloua gracilis*), the dominant species of the plains grassland ecosystem type at the Sevilleta Wildlife Refuge

3) USA-NPN observational data products for creosote bush (*Larrea tridentata*), the dominant species of the desert shrubland ecosystem type at the Sevilleta Wildlife Refuge

Question 2

An overarching high-quality figure to show how these data align

```
npn.df = npn_download_status_data(
  request_source = 'NAU',
  years = c('2021'),
  states = c("NM")
)
```

```
## using a custom handler function.
```

```
## opening curl input connection.
```

```
##  Found 5000 records... Found 10000 records... Found 15000 records... Found 20000 records... Found 25
```

```
## closing curl input connection.
```

```
desertGrassland.df = npn.df %>%
  filter(genus == "Bouteloua",
         species == "eriopoda",
         phenophase_description == "Initial growth (grasses/sedges)") %>%
  group_by(day_of_year) %>%
  summarise(n = n()) %>%
  mutate(cumSum = cumsum(n))

ggplot(data = desertGrassland.df) +
  # geom_line(aes(x = day_of_year, y = cumSum)) +
  geom_smooth(aes(x = day_of_year, y = cumSum), method = 'loess', color = "forestgreen")+
  labs(y = "Cumulative Count of 'Initial growth' Observations",
       x = "Day of Year (2021)",
       title = "Observations of Initial Growth for B. eriopoda in 2021")
```

```
## `geom_smooth()` using formula 'y ~ x'
```

Observations of Initial Growth for B. eriopoda in 2021
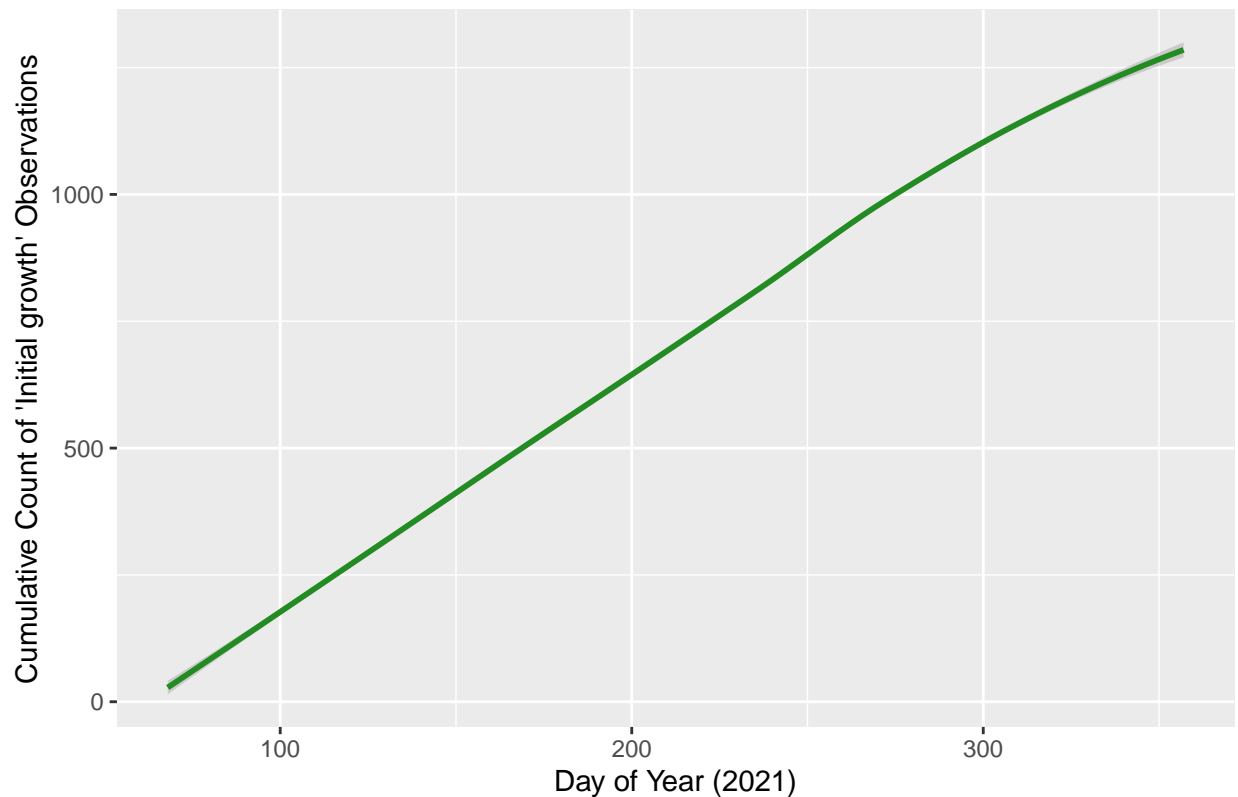
Question 3

One paragraph summarizing how this data or analysis is useful to you and/or the infrastructure.

The NPN observational data about the dominant species in each of the four ecosystem types of interest at the Sevilleta Wildlife Refuge will be incredibly useful for constraining model parameters related to multiple carbon pools in my process-based model of the major ecological processes at play. Data about spring and fall phenological shifts for the dominant plants defining each ecosystem type enable us to make inferences about how much photosynthesis is occurring at different times of the year, and therefore how much carbon is being uptaken by plants at different times of the year.