Post Process Cynet Runs

ishanu chattopadhyay

1 Q1 & Q2: PRED. VISITORS & MEETINGS

- 2 scenarios (A. No change, B. Closed sites)
- Need confidence bounds for Q1 (prediction of visits to rec sites)

1a. Scenario A: No Change

Let $\{s^0\}$ be the set of cynet-run sites, and let $s^\star$ be a target site.

Let $\mathcal{N}_k(s^\star) \subset \{s^0\}$ be the set of k closest neighboring sites of $s^\star$ in $\{s^0\}$. Note, "closest" implies we are using a specific distance metric, the choice of which is described later.

Step 1. Compute predictions for all sites $\{s^0\}$, by computing the regressors on FTX files as done before. The predicted time series for site s is denoted as s_t.

Step 2. We then use: (Denoting a neighboring site as s^i)

$$s^\star_t = \sum_{s^i \in \mathcal{N}_k(s^\star)} \alpha_i s^i_t \tag{1}$$

where

$$\alpha_i = \frac{1/\theta(s^i, s^\star)}{\sum\limits_{s^i \in \mathcal{N}_k(s^\star)} 1/\theta(s^i, s^\star)} \tag{2}$$

Note: we still need to choose k. We set it to some small value in the range $[10, 35]$.

1b. Distance Metric

Three different distance metrics may be used. We will try all of them:

1) Euclidean
2) Travel distance
3) Learned from prediction behavior (discussed later)

1c. Scenario B: Site Closing

Let $\mathscr{S}$ be the set of all rec sites.

Let s^c be a closed site, and let $\mathcal{N}^{\mathscr{S}}_k(s^c) \subset \mathscr{S}$ be the set of k neighbors of s^c. Note this is the set of neighbors among **all** rec sites, which is different from $\mathcal{N}_k(s^\star)$.

Step 1. Predict s_t for all $s \in \mathcal{N}^{\mathscr{S}}_k(s^c)$ using the approach in the previous section.

Step 2. Compute:

$$\forall s^i \in \mathcal{N}^{\mathscr{S}}_k(s^c),\ \beta_i = \frac{1/\theta(s^i, s^c)}{\sum\limits_{s^i \in \mathcal{N}^{\mathscr{S}}_k(s^c)} 1/\theta(s^i, s^c)} \tag{3}$$

Step 3. Predict s^c_t using the approach in the previous section

Step 4. Change for sites $s^i \in \mathcal{N}_k(s^c)$ as:

$$\delta s^i_t = \beta_i s^c_t \tag{4}$$

Step 5. Do this for all closed sites (which updates a bunch of neighboring rec sites). Let the set of sites updated be denoted as $U \subset \mathscr{S}$. We know, the total change for each site in U, which is denoted as δs_t^u for $s^u \in U$.

Step 6. Update sites in $\{s^0\}$ as follows: For each $s^u \in U$, compute the α-coefficients as in previous section. In other words,

$$\alpha_i^u = \frac{1/\theta(s^i, s^u)}{\sum\limits_{s^i \in \mathcal{N}_k(s^u)} 1/\theta(s^i, s^u)} \tag{5}$$

Then, for each $s^u \in U$:

$$s_t^i \leftarrow s_t^i + \sum_{s^u \in U} \alpha_i^u \delta s_t^u \tag{6}$$

Step 6. Update target sites using approach in previous section.

The last two sections are sufficient to answer both Q1 and Q2 for both scenarios. Now, we define the computation of confidence bounds.

Confidence Bounds