Climate deviation model estimation with Kalman filter

Julio Waissman Vilanova

1   The climate deviation model strategy

Let $P_{ij}(t)$ and $L_{ij}(t)$ the observed precipitation and lighting gaussian counts in grid i, j at time t respectively. The aim of the model is to estimate the precipitation $\hat{P}_{ij}(t)$ using the lightning data at the current grid i, j and space and time vicinity. The $\hat{P}_{ij}(t)$ is estimated at two stages: the frist stage using a linear space and time invariant seasonal model (STI) adjusted by a least square methodology and the second, a climate deviation model (CD) adjusted dynamically by means of a Kalman filter. The general scheme is shown in figure 1.
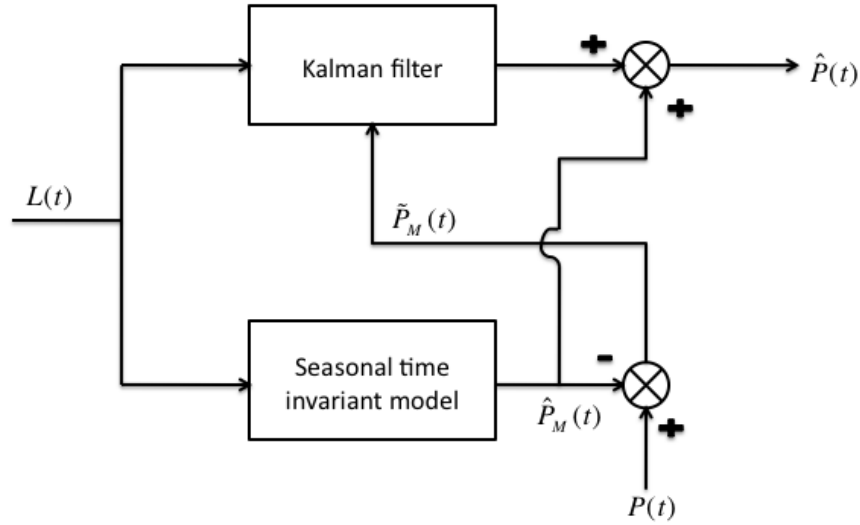

Figure 1: The big picture

1

Let $L(t)$ be a vector of lightning for all the grids of the region of interest in the form

$$L(t) = (L_{11}(t), L_{21}(t), ldots, L_{N1}(t), \ldots L_{ij}(t), \ldots, L_{NM}(t))^T,$$

where M is the number of grid rows and N the number of grid columns. With the STI model, a model precipitation estimation of all the grids is obtained $\hat{P}_M(t)$ like a vector similar to $L(t)$. **In fact, not only $L(t)$ is used, but also $L(t - nl), \ldots, L(t + np)$ but i dont know yet how to express that in a clear form in the scheme**.

The prediction made by the STI model is sustraed to the observed precipitation vector $P(t)$ resulting a error estimation vector $\acute{P}(t)$. The main idea is to develop an estimation system in order to adjust a Climate Deviation Model to estimate the deviation of the seasonal model of the STI estimation. The predicted deviation is then adding to the $P\hat{M}(t)$ in order to obtain a corrected estimated precipitation $\hat{P}(t)$.

Let $nb_V(L_{ij}(t))$ the vector of spatial V neighborhood

$$nb_V(L_{ij}(t)) = (L_{i-V,j-V}(t), \ldots, L_{i,j-V}(t), \ldots, L_{i,j}(t), \ldots, L_{i+V,j+V}(t))^T)^T,$$

and let $\Lambda_{i,j}(t)$ the vector

$$\Lambda_{ij}(t) = (nb_V(L_{ij}(t - nl)^T, \ldots, nb_V(L_{ij}(t + pl)^T, 1)^T,$$

where nl are de number of negative lags and pl the number of positive lags considered in the STI model.

Then, the $\hat{P}M_{ij}(t)$ is obtained as

$$\hat{P}M_{ij}(t) = \Lambda_{ij}(t)^T \Theta$$

where $\Theta = (\theta_1, \ldots, \theta_k, b)$ is the parameter vector for linear regression. In order to compute Θ all the grids of the season marked as convective grid are used in a classical least squares algorithm.

The STI model is static and reflect the general tendency of the season. Thus, a Climate Deviation dynamical model can take in account the diferences along the season. It is important to notice that the CD model only take in account the variation in time of the STI model, because we use one set of parameters for all the grids of the region at time t.

The CD model is assumed with the same neighborhood V and the same positive and negative lags of the STI model. **I don't know if it is a good idea, because if we consider positive lags in the dynamical model, then we assume that we have the $L(t + 1), \ldots L(t + pl)$ data**

available, and then we can not argue that the model is useful to doing nowcasting.

The CD model is then a model of the diferences between the estimation of the STI model respect the observed data at time t,

$$\hat{D}_{ij}(t) = \Lambda_{ij}(t)^T \Phi(t)$$

where $\Phi(t) = (\phi_1(t), \ldots, \phi_k(t), \beta(t))^T$ is the vector of dynamical parmeters of the CD model and $\hat{D}_{ij}(t)$ is the precipitation deviation at time t for the grid i, j.. The model can be represented with a matrix notation in order to obtain the vector of the deviation of entry region at time t as

$$D(t) = \Lambda(t)\Phi(t)$$

where

$$\Lambda(t) = \begin{bmatrix} \Lambda_{11}(t)^T \\ \vdots \\ \Lambda_{ij}(t)^T \\ \vdots \Lambda_{NM}^T \end{bmatrix}$$

The Kalman algoritm used in this work is

Algorithm 1 General CD Kalman filter procedure

Require: Θ

Require: H matrix of coverture (*not explained yet*)

Require: Q covariance matrix of the CD model parmeters uncertainty

Require: R covariance matrix of the gaussian white noise in measurements

Require: τ parameters' decay time constant

Require: h an int to decide that Precip with no Lightning is far enough

Require: $A(0)$ *a posteriori* error covariance matrix

1: $\alpha \leftarrow 0$
2: $\Phi(0) \leftarrow (0,0,\ldots,0)^T$
3: **for** t from $ln+1$ to $t_f - lp$ **do**
4:     Read Lightning data to form $\Lambda(t)$
5:     **if** all entries of $\Lambda(t)$ are zero **then**
6:         $\hat{P}M(t) \leftarrow [0,\ldots,0]]$
7:     **else**
8:         $\hat{P}M(t) \leftarrow \Lambda(t)\Theta$
9:     **end if**
10:     **if** (zero convective grid in all the region at times $t-nl,\ldots,t+pl$ **and** $P(t)$ have at least one entry different from zero) **or** $(P(t-h),\ldots,P(t)$ dont have any entry different from zero) **then**
11:         $\Phi(t) \leftarrow \Phi(t-1)\exp(-\alpha\tau)$
12:         $\alpha \leftarrow \alpha + 1$
13:         **if** $\alpha \geq \frac{3}{\tau}$ **then**
14:             $A(t) \leftarrow A(0)$
15:         **end if**
16:     **else if** $P(t)$ have at least one entry different from zero **then**
17:         $\alpha \leftarrow 0$
18:         $A(t) \leftarrow A(t-1) + Q$
19:         $C \leftarrow H\Lambda(t)$
20:         $K \leftarrow A(t)C^T(CA(t)C + R)^{-1}$
21:         $E \leftarrow HP(t) - C\Phi(t)$
22:         $A(t) \leftarrow (I - KC)A(t)$
23:         $\Phi(t) \leftarrow K\Phi(t-1)$
24:     **end if**
25:     $D(t) \leftarrow \Lambda(t)\Phi(t)$
26:     $\hat{P}(t) \leftarrow \hat{P}M(t) + D(t)$
27: **end for**

Ensure: $\hat{P}(t)$