**Drought diagnosis based on VHI cumulated values (monotonously increasing)**

In Rojas et al. (2011) “Drought“ flag was raised if

In particular *Threshold* = 35. The mean of VHI was used instead of the sum.

So, mean or integral?

Integral seems more physiologically meaningful. For example, a short peak of FAPAR may have a greater mean than a longer, well-formed season. However, working with integrals we cannot use the “threshold” (the value 35 from Kogan). Other threshold could be defined..



**Proposal:**

DOCUMENTATION OF THE PAST EVENTS

Avoid using the threshold for a hard-classification of drought and proceed in the following way. At the regional level (so assuming that meteo conditions and crop management practices can be considered similar in this region\*), compute VHI’*cum(EOS,y)* for every year y.

VHI’*cum(EOS,y*) is computed from SOS’, the earliest onset in the regional time series, according to

In this way we obtain a distribution of VHI’*cum*. Then we compute percentiles (e.g. every 10%) and for every year we classify each pixel according to these percentiles. So, the lower the rank of each pixel, the higher the probability of “drought”.

\*it’s very likely that this region has to be quite small, as a first trial we may consider to work in 100x100 km grid for example.

NRT-MONITORING

For a given region we can build the scatterplot of VHI’*cum(EOS,y, i, j)* Vs. VHI’*cum(t,y,i,j)*, for t = (SOS’, EOS(y,i,j), where (i,j) are pixel coordinates. This scatterplot enable the forecast of VHI*cum(EOS,y)* from VHI*cum(t,y)*. Consider the following example with fully fantasy data:



With a GSL (Growing season Length) of 35 decades, at decade 5 I have a weak prediction of the final value.



When I am at decade 30 (out of 35), my prediction is better.

So, for any time t during the current growing season I can provide a forecast of the final cumulative VHI, and very important, with an associated uncertainty. Summing VHI from SOS’ (the earliest SOS found in the past) implies:

1. A delay in the onset of the current GS is taken into account (if the current GS does not start, the predicted final value will be low).
2. The monitoring of the current GS starts at the time of SOS’. This means that if the current GS starts earlier I won’t be able to say nothing until SOS’. This should not be a problem in terms of operations. In addition at SOS’ I may think of integtarting, for the current GS, from SOS instead of SOS’ so that the cumulative value of current SOS’ will start higher (and higher predicted final value).

Afterwards, we can rank the forecasted value according to the percentiles defined in documentation phase. So the rank value, can be seen as inversely proportional to the likelihood of “drought”. For example, if

The forecasted value is very low and very likely (uncertainty is associated) to actually end up in a low value.

The main problem seems to connected to the definition of the regional extent of this forecast. In fact if the region is to big we mix up different crops characterized by different levels of VHIcum and different relationships between VHIcum(t) and VHIcum(EOS).

Eduardo suggests two approaches for the definition of the realtionship, after testing the method as it is.

1. Use panel regression Y=b0+b1X, where Y = VHIcum(EOS) – MEAN[VHIcum(EOS)]; and X= VHIcum(t) – MEAN[VHIcum(t)]. Additional normalization for SD can be envisaged?
2. Use polynomial regression including lat and lon as variables, to be modelled with quadratic – cubic polynomials, to model the geographical variations.