Subject: Re: cool way to determine durations in time series
Posted by btt on Fri, 20 Jan 2006 21:50:06 GMT
View Forum Message <> Reply to Message

Thomas Pfaff wrote:
> Hello everyone,
>
> I'm doing some IDL-abuse in hydrology, so my question might seem a bit
> odd, maybe.
>
> A task that is occurring quite regularly is to determine the duration of
> certain events. For example "What is the longest contiguous duration of
> stream flow below/above a certain discharge"
>
> Assuming I have an equidistant time series (e.g. one value each day)
> this basically reduces to the question of how can I transform an array
> like this
>
> series = [1,1,0,0,0,0,1,0,1,1,1,0,0,1,1]
>
> into something like this
>
> durations = [2,1,3,2]
>
> which is I want to count all contiguous fields of '1's in an array.
>
> Somehow my brain wants to use HISTOGRAM for this, but I just can't see
> how to do it.
> At the moment I'm helping myself by using CONVOL(to highlight the edges)
> and WHERE(to get the differences between two adjacent edge indices) but
> as the data gets more, this becomes extremely tedious as well as memory
> consuming (see the example below). Besides, CONVOL wouldn't work if a
> series started or ended with '1's as it can't correctly apply the kernel
> to those elements.
>

I think I would use a combination of LABEL_REGION and HISTOGRAM.

```
****START
series = [1,1,0,0,0,0,1,0,1,1,1,0,0,1,1]

nSeries= n_elements(Series)

buffered = [0,series,0]
dummy = FIX(LABEL_REGION(buffered))
label = dummy[1:nSeries]
```

```
H = HISTOGRAM(label, MIN = 1S)

print, series
print, label
print, H
*****END
```

Note the you must pad series with "background" values at the endpoints.

Cheers,
ben
