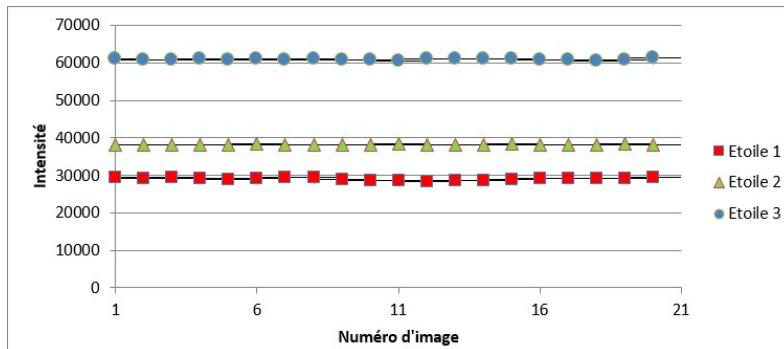


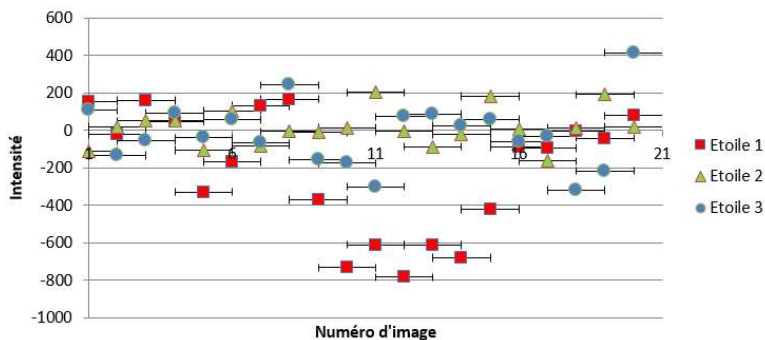
More information on the normalisation process

If you use the spreadsheet provided, which allows you to have a "Raw data" tab to copy the raw data directly measured on the images, then a "plot and analysis" tab on which the spreadsheet will then copy the values and plot the curves of the three stars according to the normalisation method chosen afterwards (cell I30 - the default value being to plot the raw data – the default normalisation must be set to 0 at the beginning of the activity, see below for details).

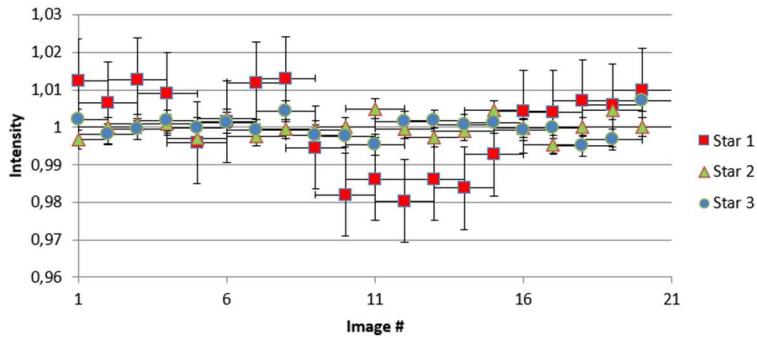
First, after copying the data in the "measurement" tab, you must obtain this curve in the "normalisation" tab:



If one subtracts the average outside transit (cell I30 = -1), the curve shows a "transit" (see below). This is the most "natural" way to get all stars at the same level (which is zero), but this will not be able to interpret it correctly. *It is then important that in the "modélisation", the teacher explains that there are photons received all the time, so the value in the Intensity should never be equal to 0 or be negative...*



as the planet does not align with Earth and star, the average intensity is $I_0 = K \cdot S$; S is the surface of the star and K a factor including the photon flux per unit area and the sensitivity of the telescope. The measured intensity oscillates then around the value of I_0 . When the planet passes in front of the star, during the transit, the surface s of the planet is occulted. The average intensity during the transit is then $I_1 = K \cdot (S - s)$. The K factor is different for each star, and is not of interest to us. In the case of spatial telescope such as Spitzer, K does not change with time. We then cancel this unknown factor by dividing the observed intensity by its average value outside the zone of the transit (cell I30 is set to "1"). Then, the values equal 1 outside the zone of the transit and $1 - s/S$ during the transit.

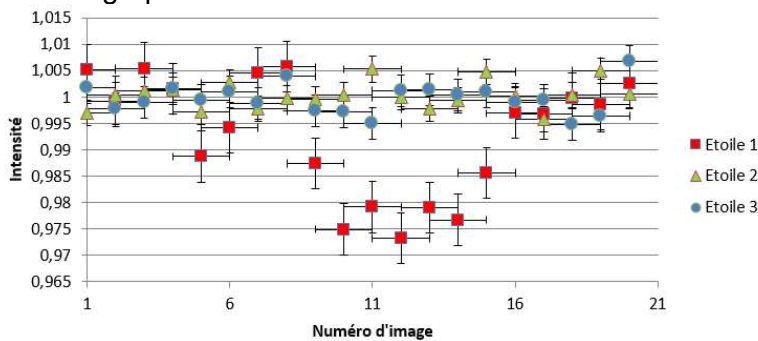


Last, it is necessary to calculate the average excluding transit, when we do not yet know whether there is transit. Therefore, it is necessary to proceed by iteration.

On the spreadsheet provided, column D is used to indicate whether there is a "transit" for each single date. By default, all rows show "0" (no transit). Rows where students write "1" in this column will be considered as dates with a transit and will therefore not be used to calculate the average.

The averages during and out of transit are shown for each star in the first rows. Students should therefore put "1" in each corresponding cell as they find out where the transit is taking place...

You can try and play by considering no cell with transit, then one cell (where the signal is much lower for example), then another one... If no signal is to be observed, the graph should not change too much. If a signal is to be observed, you will see that the "outer zone" starts above a mean value of 1 and ends up around 1 while the transit zone will keep being below 1... and you end up with this graph



Final note : The spreadsheet calculates the dispersion of the measurements and thus plots the error bars. Error bars should decrease as transit dates are correctly identified.

Students can therefore use the spreadsheet provided, and do some reflection on the error bars. For (much) older students they can also understand the (complex!) formulas used in the spreadsheet.