Planet Hunters Talk

How to fold a light curve using Excel
Started by kianjin
Sign in to watch

Science
discussion

---

kianjin   over 2 years ago

I've been asked this a few times, how to generate good phased curves, and the usual answer is to go to NEA's excellent periodogram service which will find the periods and plot those phased (or folded) curves that you all know and love.

However I've realized that the service isn't perfect. It is sometimes hard to get the precise fold at the correct period. Even trying out various algorithms such as the BLS or the Plavchan, you may still end up with an imperfect period. It will generally get you close to the actual period, but to do better, you can actually use a spreadsheet like Excel to improve on it.

I will use this light curve, SPH101445580 or kic 4678171, which is a rather nice eclipsing binary, to illustrate this.

To start with, you will need the light curve in a format that Excel can understand, this is the CSV format. It simply consists of two columns, the time and the flux.

To get this, you do one of several things.

1. Download it from the PlanetHunter Examine Star page. The CSV by this method needs to be processed further as it contains repeats and some quarters may be missing.
2. Or you can ftp the FITS file from MAST and use a tool like TOPCAT to read it and convert to CSV, and strip off the unneeded columns.
3. The way I do it is to use my own Python scripts getfits and kfits, found here – regular Python for Linux and Mac and Windows formats) which will automatically pull down the FITS file from MAST and prepare the CSV file. With these scripts, you just need to issue the commands, getfits 4678171, followed by kfits -mr kp*4678171*lic.fits >4678171.csv

Note to PlanetHunters - it might be good to reformat your CSV files so that the bulk of it is just an uninterrupted series of TIME and FLUX data. Metadata and how quarters are segmented can be put in a header. This way the meta information can be deleted, leaving the data required for further processing. It is actually not helpful to have the quarter breaks and duplications in the CSV because they need to be reconciled and deleted, adding more complexity.

---

kianjin   over 2 years ago

Once you have a CSV file with 2 columns, Time and Flux, it's very easy to fold this light curve. The secret is just one formula: MOD((A2-start)/period),1)

Open the CSV file in Excel (or some other spreadsheet if you have ideological differences with Microsoft).

Insert a column in between Time and Flux and call it Phase.

Copy the 1st value of Time (in cell A2) and paste it in cell E1. You can also name this cell, 'start'. In cell F1, paste the period you want to refine. If you used the NEA service, the Plavchan approximation for the period is 15.255. Name this cell, 'period'.

In cell B2, the first cell of the Phase column, enter the formula, =MOD(((A2-E1)/F1),1). Now select this entire column from B2 right to the end and fill down, populating the entire column with this formula.

Now select the Phase and Flux columns and Insert a chart (marked scatter) with Phase in the X column and Flux in the Y column. This chart twice, and change the scales so that these two charts focus on the primary and secondary eclipses. This screen shot is what you should have.
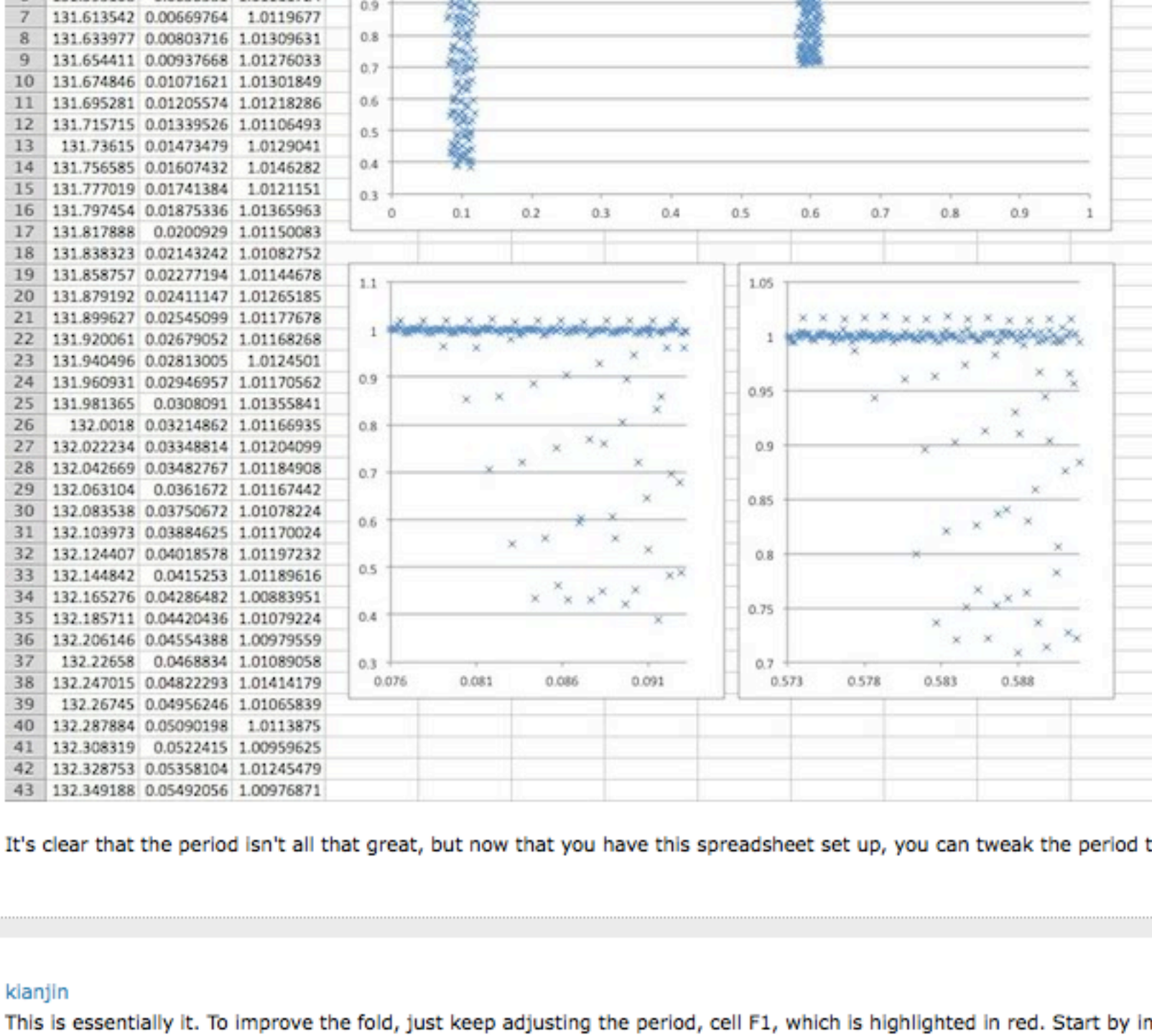

It's clear that the period isn't all that great, but now that you have this spreadsheet set up, you can tweak the period to get the best fold by eye.

---

kianjin   over 2 years ago

This is essentially it. To improve the fold, just keep adjusting the period, cell F1, which is highlighted in red. Start by incrementing and decrementing it by small steps. You should be able to see the graph of the fold get sharper until you get something like this:

---
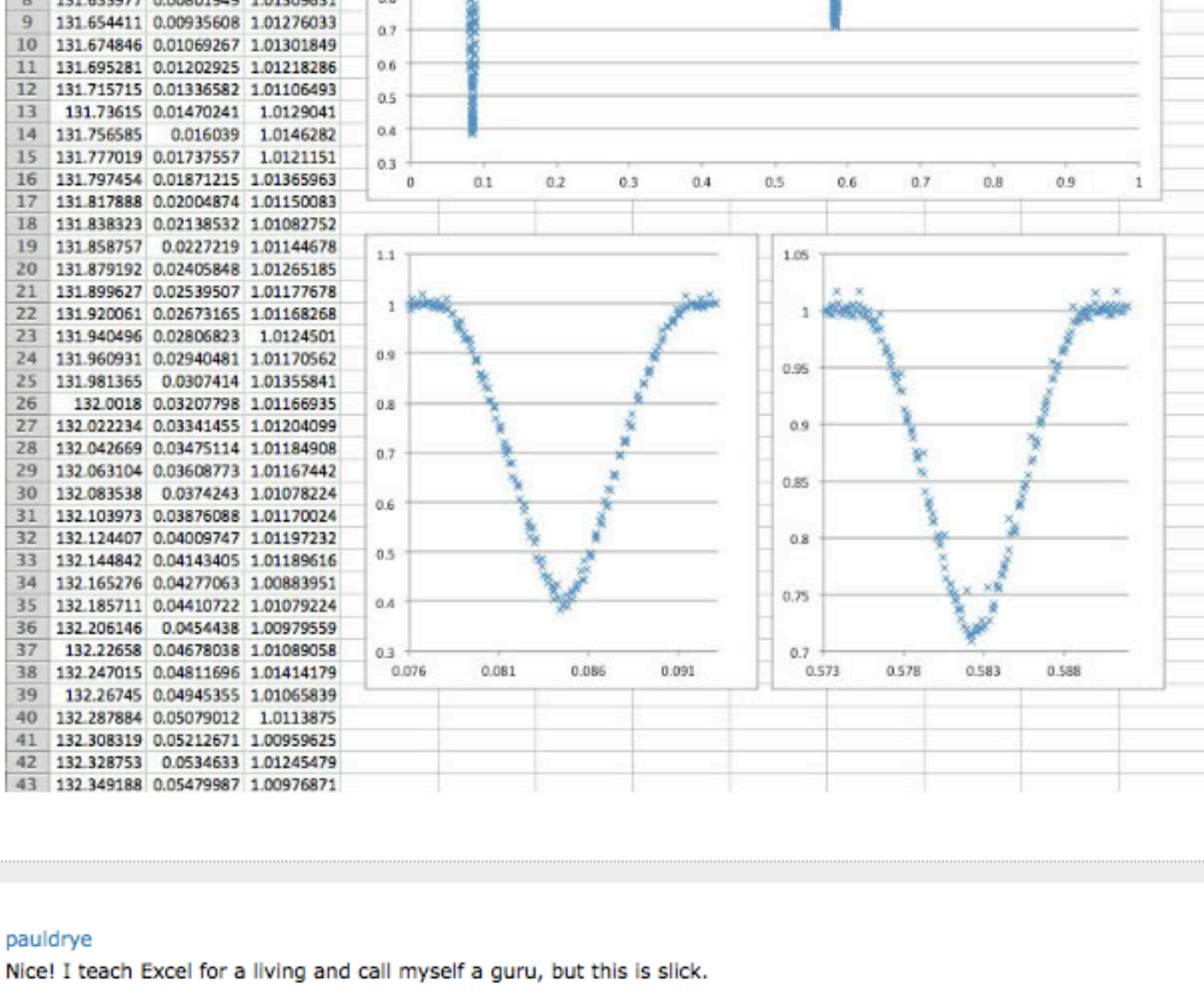
pauldrye   over 2 years ago

Nice! I teach Excel for a living and call myself a guru, but this is slick.

---

kianjin in response to pauldrye   over 2 years ago

Like they say, when it's time to meet your guru, he will appear!

Paul, I was wondering if you know how to add little buttons to the spreadsheet so that you can click to increment or decrement the period by a small amount, say, 0.001, which can be entered in another cell. This would be really cool.

---

pauldrye   over 2 years ago

Yup, depends on which version of Excel you're using. If it's 2007/10:

- Unhide your Developer ribbon tab (let me know if you're not sure how to do that)
- Use the Insert button on it.
- From the ActiveX Controls it shows you (not the similar-looking Form controls) select the Spin Button choice
- Move on to the spreadsheet where you want the spinner to appear and click and drag to draw it. I often put them out in the open somewhere and make them bigger than I need just because that's easier to edit first, then resize and move into place once I'm done.
- Check that your Design Mode button on the Developer tab has turned on (it should have, but turn it on yourself if it hasn't).
- Right-click your new spinner and select Properties. This will display the Properties sheet for the spinner.
- Enter an initial value (usually zero) for the Value field
- Enter an increment for each click on the spinner. Unfortunately you have to use an integer, so if you want 0.001 the best bet is to use 1 and then post process it in your formula later by dividing by 1000. If you use that trick, you need to increase your remaining parameters by the same factor. I'm assuming 1000 for the rest of this.....
- Upper boundary goes in the Max field, say 500,000 so you can go to +500 days. No commas when you type in this field, please!
- Lower boundary goes in the Min field. Thankfully, unlike older versions of Excel you can put a negative in here. Say -500,000 (again, no commas when you enter it).
- The last thing to do is get the spinner to drop its value in a cell somewhere (formulas can't access controls directly, they have to read from a cell). In the Linked Cell field, type the cell reference where you want the number to go. Using your screen shots above, this would be F1. You may actually want to drop it someplace else (say G1), then change F1's formula to read =G1/1000 to compensate for the "integer problem" above. Easier than fixing it in your slick little MOD formula.
- Turn off the Design Mode button on the Developer toolbar and the spinner will go live, at which point you can click the up and down arrows, and the number should change.

The really cool thing here (for Excel nerd values of cool) is that the spinner will pick up on your "initial guess". Type your value into G1 (don't forget to multiply it by 1000 first!) and the spinner will overwrite the last number it was remembering in its internal register with what it sees in G1 and start incrementing from there. I had assumed it wasn't going to do that and was all ready to drop a cartload of Visual Basic macro on you to fix it -- moral of the story: Excel's programmers "have usually thought of that", whatever "that" is :)

Let me know how that works for you and I can walk you through any tricky bits, or refine it to act differently if needed. Or if you're using Excel 2003 or earlier, let me know and I'll post the different instructions (it's mostly the same once you've figured out where the Active X controls are hiding in the older versions and get it drawn onto the sheet).

---

arvintan   over 2 years ago

@Kian

Does this also work with non-EB configs? Say for relatively smaller planets?

Also, I already downloaded the zip file for Windows but haven't a clue on what to do with it. Do I just extract it and let it work itself or what?

Thanks.

---

pauldrye in response to arvintan   over 2 years ago

> arvintan:
> @Kian
> Does this also work with non-EB configs? Say for relatively smaller planets?
> Also, I already downloaded the zip file for Windows but haven't a clue on what to do with it. Do I just extract it and let it work itself or what?
> Thanks.

Do you have Python installed on your computer? You need that language to read and run the script. It's free:

http://python.org/getit/

---

kianjin in response to arvintan   over 2 years ago

> arvintan:
> @Kian
> Does this also work with non-EB configs? Say for relatively smaller planets?
> Also, I already downloaded the zip file for Windows but haven't a clue on what to do with it. Do I just extract it and let it work itself or what?
> Thanks.

Yes, it definitely works on anything that's periodic and you want to fold the curve to see this periodicity.

As Paul said, you'll need python. There are instructions on the README.txt on this too. There are 2 main versions of Python - 3 and 2 (latest is 2.7). Make sure you have 2.5 to 2.7.

There is a bug that I overlooked in the version that is on the server. The files kfits, detrend, getfits all need to have a .py added to the end so that the Python interpreter knows they are python scripts.

I've uploaded the fixed version to the server, so save you the hassle, just download it again: http://www.kianjin.com/kepler/detrend.zip

---

Greg*   over 2 years ago

Hi,

getfits works well but then I've just a little problem with kfit :

kfits -mr kp*4678171*lic.fits >4678171.csv

IOError: [Errno 22] La syntaxe du nom de fichier, de répertoire ou de volume est incorrecte: 'kp*4678171*lic.fits'

meaning in english : incorrect syntax on the file name or directory.

It does not like the "*"

If I give the full name (only one file) it works. Did you have this problem ?

Greg

Please Log in to make comments.