
Subject: Re: comparing and concatenating arrays...please help!!

Posted by [btt](#) on Fri, 09 Jan 2004 15:30:23 GMT

[View Forum Message](#) <> [Reply to Message](#)

Martin Doyle wrote:

```
> Hello all,
>
> I really hope someone out there can help me with this....I am tearing
> my hair out as my code is so slow!
>
> I have 2 files of data (hourly met data) with one file containing one
> set of parameters, and the other file containing another set of
> parameters. What I am trying to do, is to match the data based on the
> YY, MM, DD and HH values and then write BOTH sets of parameters to a
> seperate file. For example;
>
> file1:
> 1954 12 31 23  90  11  4 366  0.00
>
> file2:
> 1954 12 31 23  2.80  2.10  2.20  95.21
>
> intended result:
> 1954 12 31 23  90  11  4 366  0.00  2.80  2.10  2.20
> 95.21
>
> NOTE: Both files have no order to them, so a simple concatenation
> won't work
>
> I have written some code, but it is wrist slashing-ly slow!;
>
> I read in each variable as a seperate array...
>
> b=0L
> REPEAT BEGIN
> c=0L
> REPEAT BEGIN
> If (year(b) EQ year2(c)) AND (month(b) EQ month2(c)) AND (day(b) EQ
> day2(c)) AND (hour(b) EQ hour2(c)) THEN BEGIN
>
> printf, 3, year(b), month(b), day(b), hour(b), winddir(b), windsp(b),$
> present(b),visib(b), mslpres(b), airt(c), dewt(c), wett(c), relh(c),$
> format = finalformat
> endif
>
> c=c+1
>
> ENDREP UNTIL c EQ lines2-1
```

>
> b=b+1
>
> ENDREP UNTIL b EQ lines1-1
>
> I'm sure there must be a better way than this.
>
> Please help me!
>
> Many thanks in advance, Martin..

Hello,

You have quite a few non-IDL options presented already. But if you're stuck in IDL as I am then I have a suggestion... at least a tempting place to start.

I would collapse the date/time information into a julian day number. Then I would histogram the julian day numbers which is the same as sorting.

```
;merge the year/month/day data of the files  
month = MonthsFile1 + MonthsFile2  
day = DayFile1 + DayFile2  
year = YearFile1 + YearFile2
```

```
;convert to long integer Julian day number  
jul = JULDAY(month, day, year) ; note leave the hour out for now!
```

```
;compute histogram - reverse the reverse indices  
H = Histogram(jul, reverse_indices = r)
```

Now each occupied bin contains the data for a single day (all 24 hours) and I'd bet they are listed in chronological order, too. You'll have to sort through the hours - but you can keep track of the two different data sources since any reverse_index GT n_elements(file1)-1 must have come from file2.

Honestly, I don't know if this would be a savings for you or not since I would expect a pretty flat histogram. JD Smith has written up a tutorial on this very subject and it can be found at the fountain of youth and tennis (www.dfanning.com).

Come to think of it, you might be able to modify the Julian Day number to include the hours.... $jul = (jul * 100) + hours$. But then again, it might squander the efficiency gain using Histogram.

Cheers,
Ben
