

## Part 1: File Identities

### Stephen McIntyre and Ross McKittrick

**Reply to comments made by Professor Mann in response to our paper “Corrections to the Mann et. al. (1998) Proxy Data Base and Northern Hemispheric Average Temperature Series” *Energy and Environment* 14(6).**

November 11 2003

## **PART 1: The Identity of Some Disputed Files**

**SUMMARY:** We refute suggestions by Professor Mann that collation and splicing errors in the proxy data set criticized in McIntyre and McKittrick (2003a) (herein “MM”) were introduced in a special purpose Excel spreadsheet prepared for McIntyre in April 2003. For the 81 out of 112 proxies which were not subjected to principal components calculations by MBH, we have determined that the uncollated series at the Mann FTP site are identical to the versions in the data set we examined and criticized in MM. Accordingly, the criticisms of MM in respect to unjustified truncation and fills, use of obsolete data and geographical mislocations apply equally to the acknowledged MBH98 data archive. The remaining series are unpublished and non-peer reviewed principal components series calculated by MBH. We show that mere identification of the FTP address of the uncollated data is inadequate disclosure, as there is still no way to identify which principal component series were used.

### **1. Introduction**

Following the publication of MM, on or about Oct. 29, 2003, Professor Mann stated that our results, especially before 1600, were entirely “spurious” for three reasons: we used an incorrect collation of data supplied to us by one of Mann’s associates in order to accommodate our alleged request for data in an Excel spreadsheet; we then failed to check this data against the correct data at Mann’s FTP site; and we failed to use a correct dataset consisting of 159 columns of data instead of 112 columns of data.

We replied previously to these comments (see <http://www.uoguelph.ca/~rmckitri/research/Response.Oct29.pdf>) by reviewing the correspondence record. On October 29, 2003, we learned of the existence and location of the anonymous FTP site in question (<ftp://holocene.evsc.virginia.edu/pub/>) and we can now use the information from this FTP site to completely refute the above claims. For ease of reference, we call the FTP site “UVA” hereinafter.

We also report on the incoherent structure of the UVA site and the impossibility of identifying the alleged 159 columns of data within this site under existing disclosure.

We will occasionally make reference to the reply written by Professors Mann, Bradley and Hughes and released on November 3; denoting it “MBH-r.”

## Part 1: File Identities

### 2. The Data Files *pcproxy.txt* and *pcproxy.mat*

As of April 2003, the collated proxy dataset used in MBH98 was not included in the Supplementary Information to MBH98 at *Nature*, was not archived at the World Data Center for Paleoclimatology or similar public archive, was not privately archived at the University of Massachusetts and was not linked in any of the University of Virginia website pages referring to MBH98. In April 2003, in response to a request from one of us (McIntyre) for the FTP location of the proxy data used in MBH98, both Professor Mann and Scott Rutherford, Mann's associate, either refused to provide or were unable to provide the requested FTP location. Subsequent to MM, Professor Mann has identified the UVA site. Even if the archive was located on an anonymous FTP site, such location falls well short of constituting *public* archiving, especially since a specific request for the location of the site went unanswered.

Instead of providing the FTP address, in April 2003, Rutherford emailed to McIntyre the file **pcproxy.txt**, consisting of 112 proxies, the number referred to in MBH98, together with a file listing proxy names and weights. The dataset **pcproxy.txt** consisted of 112 proxies - 31 principal component series and 81 proxy series. We provide all correspondence at the site <http://www.uoguelph.ca/~rmckitri/research/trc.html>, hereinafter denoted \*trc.html.

Following the publication of MM, Mann identified the UVA FTP site to David Appell, a freelance journalist. Professor Mann faults us for not getting the raw data from UVA in the first place, and he faults us for not going to the UVA site when we noticed problems in the file.

When we noticed apparent collation errors in the principal components series (not present in the non-PC series), we sought specific clarification from Scott Rutherford and we asked for confirmation of the validity of the dataset from Mann. Neither Mann nor Rutherford referred us to the UVA site or suggested the possibility that collation errors had been introduced in **pcproxy.txt**. Instead Professor Mann told us he would not respond to further inquiries on the matter.

Nevertheless had we found out about the UVA site, we would have found that the only two collated data sets there contained exactly identical data as the file (**pcproxy.txt**) we had been sent. One was a text file (**pcproxy.txt**) identical to the text file sent to McIntyre (see <ftp://holocene.evsc.virginia.edu/pub/sdr/pcproxy.txt>). In addition to the text file we found a matlab file **pcproxy.mat**, the header to which reads: "MATLAB 5.0 MAT-file, Platform: SOL2, Created on: Thu Aug 8 10:18:19 2002." Likewise, the date of creation of the **pcproxy.txt** file at UVA was August 8 2002. Both files contain the identical data to the file **pcproxy.txt** emailed to McIntyre in April 2003, including all collation errors, fills and other problems identified in MM. Furthermore the data in the file of weights emailed to McIntyre at that time is identical to the data in the file **multiproxy.inf** at the UVA site.

We now wish to examine the veracity of Professor's Mann's claims in light of the above. Mann's comments are reported in three segments at [www.davidappell.com](http://www.davidappell.com) and we refer to them collectively as "MC." The first comment is:

In short, here's what happened: M&M asked an associate of Mann to supply them with the Mann et. al. proxy data in an Excel spreadsheet, even though the raw data is available [here](#) [UVA site link]. An error was made in preparing this Excel file, in which the early series were successively overprinted by later and later series, and this is the data M&M used. Mann says:

"...the authors results are entirely spurious. The mistake made insures that the estimates, in particular prior to 1600-1700, are meaningless."

## Part 1: File Identities

This leads, Mann says, to "the use of series that are artificial combinations of early [e.g. 15th-16th century] and late [e.g. 19th-20th] information accidentally spliced together" with "no relation" to the proxy data used by Mann et. al. in their 1998 (and subsequent) work.

**10/29/2003: "E&E paper is 'wrong'"**

The second comment is:

Mann adds:

The spreadsheet file they used was a complete distortion of the actual Mann et. al. proxy data set, and was essentially useless, particularly in the earlier centuries. The authors had access to the full data, which has been available on a public ftp site for nearly two years. When they noticed, as described in their paper, some signs of problems with the Excel spreadsheet version of the data, one might think that they would have bothered to check the data available on our public ftp site.

Mann says that the crux of M&M's error is their use of a Excel dataset with only 112 columns (where each column represents one set of proxy data--tree rings, ice cores, historical temperature data, etc.), when in fact the full paleoclimatic data series requires 159 to be used properly in the analysis behind the [Mann, Bradley, and Hughes 1998 paper](#) whose results they're trying to check.

This email indicates that M&M were at least initially aware of this, and could have used raw data if they'd chosen. **10/29/2003: "M&M: The Details"**

The third comment is:

M&M apparently did not start from MBH's raw data--which consists of 2,077 different files--but instead asked Michael Mann to provide them with an Excel spreadsheet of summarized and collated data. Mann and a colleague who was assisting him indicated to M&M that they could obtain their raw data via FTP on the Web, and also did their best to provide the data in the format requested.

However, Mann said today, a transcription error was inadvertently made in preparation of the spreadsheet, in which some multiple data that should have appeared in multiple columns was mistakenly overwritten into some single columns. A dataset that should have contained 159 columns of data in fact only contained 112 columns. So when M&M slid this dataset into their calculations, the results that came out were naturally in error. **10/29/2003: "M&M (&M), Take 2**

It is self-evident that Mann's comments are a pastiche of false statements. (1) We requested an FTP location, not an Excel spreadsheet. (2) The FTP site location is anonymous. At no point prior to the identification of the FTP site on Oct. 29, 2003 by David Appell, was the site identified for us nor did we have access to the full data. (3) When we noticed problems, we brought it to the attention of both Rutherford and Mann and were refused. (4) More significantly, the file in question was prepared on or before August 2002 and contained collation errors at that time; it was not a special purpose collation for us. (5) The original form of the file was Matlab, rather than Excel, indicating that it derived from Mann's own calculations. (6) We did not "slide" the defective error set into our calculations. The paper MM stands as copious evidence that we noticed the collation errors in **pcproxy.txt** and firewalled ourselves against them. We posted our own re-collation **proxy4.txt**, which does not contain the "sliding" of incorrectly collated data, falsely alleged by Mann.

## Part 1: File Identities

More importantly for efforts to verify MBH98, the files **pcproxy.txt** and **pcproxy.mat** show that collation errors were embedded at Professor Mann's UVA site long before our request for data.

### *Events Since MM*

Following the identification of UVA as the location of MBH98 data, we downloaded the files **pcproxy.txt** and **pcproxy.mat** from UVA, together with the entire contents of <ftp://holocene.evsc.virginia.edu/pub/MBH98/>. On November 8, 2003, we re-visited this site and discovered the following changes: (1) the file **pcproxy.mat** had been deleted from Mann's FTP site; (2) the file **pcproxy.txt** no longer was displayed under the directory, although it could still be retrieved through an exact call if one knew the exact file name; (3) without any notice, a new file named "**mbhfilled.mat**" was inserted into the directory. The header in this file reads "Created on: Tue Nov 4 23:37:07 2003." Interested readers may partly verify (1) and (2) through the following FTP calls: <ftp://holocene.evsc.virginia.edu/pub/sdr/> - showing that neither file is listed, but listing the new file; <ftp://holocene.evsc.virginia.edu/pub/sdr/pcproxy.txt> - retrieving pcproxy.txt; <ftp://holocene.evsc.virginia.edu/pub/sdr/pcproxy.mat> - which does not retrieve pcproxy.mat, formerly located in this directory. We have just reconfirmed these matters (November 11, 2003, 1:46 PM).

Given the significance which Mann placed on our alleged use of incorrect data and his insistence that we ought to have relied upon the contents of the UVA site, we are concerned about the deletion from this archive of **pcproxy.mat**, a file which is right at the center of the current dispute. This is the file which provided evidence of the prior existence of the contested data in the form sent to us, and showed it to be a Matlab product rather than an Excel product. Under the circumstances the deletion of **pcproxy.mat**, and the removal of **pcproxy.txt** from the folder directory seems ill-advised. We have posted up the deleted file at [www.climate2003.com/data/pcproxy.mat](http://www.climate2003.com/data/pcproxy.mat).

We are currently analyzing the newly archived file **mbhfilled.mat**. This is a file which has dimensions 381x2016. It lacks any descriptive information whatever. Within the matrix, 2006 columns have values for all years and 10 columns (28, 55, 64, 65, 611, 683, 718, 1105, 1106, 1936) have no values. It is possible that this dataset may have nothing whatever to do with MBH98, and its appearance within the folder that until recently held **pcproxy.mat** may only be a coincidence. It remains enigmatic for now.

### **3. Uncollated Data at the FTP Site**

Professor Mann has seemingly taken the position that the mere identification of the FTP site is adequate disclosure. We show here that this is not the case.

The explanation of the principal component information at the UVA site consists of the following text from MBH98: "Certain densely sampled regional dendroclimatic data sets have been represented in the network by a smaller number of leading principal components (typically 3–11 depending on the spatial extent and size of the data set)." There is no other disclosure of the methodologies used for proxy principal components calculations or stepwise rosters in MBH98 proper or in any Supplementary Information and an email request on September 25 specifically asking for more information on this was refused.

## Part 1: File Identities

Judging from Professor Mann's comments, anyone carrying out due diligence on this data would be required, first, to guess that 159 proxy series are needed (rather than 112 as stated in MBH98), and, second, to pick them out of the uncollated data at the UVA site.

Within Professor Mann's UVA site, there is a separate directory for each of the six regions identified in MBH98 for PC analysis. These six directories contain a total of 29 subdirectories (each identified with a start date through a name in the format BACKTO\_1700 etc.) for the periods after 1400; each subdirectory contains between 10-15 principal components. (Additionally, three of the directories also contain principal components for a period different than any of the subdirectories.) In total, the UVA site contains 430 different principal component series for periods after 1400. Since there had been no previous public disclosure that there were 159 separate series behind MBH98 rather than 112 as stated in *Nature*, it would have been impossible to guess that that was the number sought. Subtracting the 81 non-PC series, this requires the identification of 78 PC series from these 430 series in order to successfully replicate the procedures carried out in MBH98.

The roster files located at the FTP site do not explain the formation of these 78 series. The file **multiproxy.inf**, which was sent to us in April 2003, contains a roster description, but not of 78 series. Even following the subdirectory information in **multiproxy.inf** would not help. The subdirectories and file names in **multiproxy.inf** are, in every case, accurate renderings of the locations of the non-PC proxies. Accordingly, a reader might assume that the subdirectory information in **multiproxy.inf** also located the PC proxies. If so, he would collate the following proxies: Stahle/OK – BACKTO\_1700; Stahle/SWM – BACKTO\_1700; ITRDB/NOAMER – BACKTO\_1750; ITRDB/SOAMER – BACKTO\_1600; ITRDB/AUSTRAL – BACKTO\_1750 and VAGANOV – BACKTO\_1750. But the results of this roster selection would conflict with the requirement for 78 series.

We presume that the above elucidates the problems in replicating MBH principal component rosters. There are more problems, which are discussed in the Appendix. We deliberately sought clarification on these matters from Professor Mann or a reference to a more expanded description than MBH98 and were refused. The existing disclosure does not even begin to satisfy even basic standards of data disclosure. In this context, any question of whether the uncollated data was available online as of April 2003 or earlier (the present FTP site being established only on July 30, 2002) is moot, since even if it was, the series to be collated could not have been identified and the online collations are now stated by Professor Mann to be incorrect.

As an aside, one of the subdirectories in the ITRDB/NOAMER has the offputting title "BACKTO\_1400 – CENSORED". In light of the identified sensitivity of early 15<sup>th</sup> century values to very slight variations in proxy indicators and the evidence elsewhere of truncation (censoring?) of important temperature series, we believe that disclosure of the censoring process would be helpful.

### 4. Conclusions

Professor Mann's public comments regarding MM contain many provably false statements, including statements regarding the preparation of **pcproxy.txt** which are refuted by available correspondence and by evidence that this same file was located at Mann's FTP site, dated long before April 2003. For the 81 out of 112 proxies which were not subjected to principal components calculations by MBH, we have determined that the uncollated series at the Mann FTP site are identical to the versions in the data set we examined and criticized in MM. Accordingly, the criticisms of MM in respect to unjustified truncation and fills, use of obsolete

## Part 1: File Identities

data and geographical mislocations apply equally to the acknowledged MBH98 data archive. Professor Mann has provided no evidence for his claim that the collation and other errors latent at his FTP site did not affect the MBH98 calculations.

Professor Mann has stated that, if **pcproxy.txt** as sent to us were used for computation, “the estimates, in particular prior to 1600-1700, are meaningless.” Since this is true of his data collations **pcproxy.txt** and **pcproxy.mat** on his own FTP site, Professor Mann himself has made a prima facie case for a new refereeing of MBH98.

Existing disclosure by MBH98 does not enable the identification of the “159 series” allegedly used in MBH98 within the population of uncollated proxy and principal component series located at the UVA FTP site. Adequate disclosure would include the following: the collated data in the 159-column version cited by Mann, but conspicuously absent from his FTP site, or a computer program evidencing the correct use of uncollated data, again notably absent from his FTP site.

## Part 1: File Identities

### APPENDIX:

While our main concerns with the newly disclosed FTP information pertains to the above, we also note that the existing disclosure by MBH relating to principal component selections is insufficient to permit a coherent interpretation of the data located at the FTP site, even with the recent additional information on MBH98 principal components methodology.

MBH-r states and MC refers to 159 proxies, although no more than 112 proxies were ever disclosed previously. MBH-r disclosed for the first time that MBH98 did not use “conventional” principal components methods, but instead:

“the procedure of MBH98 of calculating the PC series separately for all intervals used in their stepwise reconstruction approach.”

#### *(1) Disclosure*

The only disclosure in MBH98 specific to principal components methodology are the following comments with respect to the calculation of temperature principal components:

- (1) that “a conventional Principal Components Analysis” was performed (see top left column, page 786);
- (2) that Preisendorfer’s Rule N was used to select a subset of principal components used for downstream analysis.

MBH98 uses principal components methods for 6 tree-ring regions, the resulting indicators making up (as shown below) 31 of the 112 proxies in the roster of proxies after 1820. These calculations have never been separately published or, to our knowledge, been peer reviewed. There is no disclosure in MBH98 on the principal components methodology used for proxies other than the following:

“Certain densely sampled regional dendroclimatic data sets have been represented in the network by a smaller number of leading principal components (typically 3–11 depending on the spatial extent and size of the data set).”

In passing, we note that we are unaware of any rosters in which 11 principal components are used. MBH98 Supplementary Information identifies the following “regions”, as being used for the calculation of principal components, and, in all cases but Vaganov, provides a roster of series (some sites having more than one series) used in the regional calculations. In the Vaganov case, no roster is given, but the total number of sites is listed. The number of PCs used and the number of sites is shown below.

Region	# of PCs	# of Series
Stahle/OK	3	14
Stahle/SWM	9	20
ITRDB/NOAMER	9	232 <sup>1</sup>
SOAMER	3	18
AUSTRAL	4	16
Vaganov	3	61

---

<sup>1</sup> Of which source data has been located at the World Data Center for Paleoclimatology for 231 sites.

## Part 1: File Identities

Figure 1a of MBH98 shows the geographical distribution of indicators, with the number of “+” symbols “indicating the number of retained principal components.”

In an email of September 25 2003, we specifically asked Professor Mann how they dealt with missing values, and he refused to answer. Accordingly, in MM, we stated the following:

The start dates of the MBH98 PC’s are not consistent with those of available data. In 12 cases, MBH98 commenced their calculation after the date in which all records were available (e.g. Australia-New Zealand region where MBH98 commenced in 1750, although a start date of 1625 was possible.) In 16 cases, MBH98 commenced their PC index in a period *prior* to that available in the data (e.g. Texas-Mexico). Because standard PC algorithms fail in the presence of missing data, an important part of the methodology—namely how missing data were treated in the PC calculation—remains unexplained in MBH98. We computed all 28 PCs, together with their explained variances, using a standard principal component algorithm for the maximum period in which all records were available within each region.

In Table 8 of MM, we listed the maximal available period according to our calculations, and in the SI to MM we provided both the collations and the algorithms used to carry out these calculations. The complaint in MBH-r that detective work on their part was required is laughable.

We acknowledge that a “stepwise” principal components approach may result in different indicator availabilities than a “conventional” principal components algorithm, the methodology implied by the language of MBH98. In such a case, we believe that MBH had an obligation to fully disclose both that they used a “stepwise” principal components approach and all the parameters involved in this “approach”, including the number of principal components used in each region on a period-by-period basis.

We will show first that the information at the newly disclosed FTP site does not clarify these questions. We will then outline some of the puzzles governing MBH procedures for making proxy rosters and culling the number of principal components used in certain regions over time.

### *Principal Components Subdirectories at UVA*

Within Professor Mann’s UVA site, there is a separate directory for each of the six regions identified in MBH98 for PC analysis. These six directories contain a total of 29 subdirectories (each identified with a start date through a name in the format BACKTO\_1700 etc.) as shown below (with the “\_” deleted in each subdirectory name for reasons of space):

Subdirectory at UVA						
Period Start	Stahle/OK	Stahle/SWM	NOAMER	SOAMER	AUS	VAG
1400		BACKTO1400	BACKTO1400			Main
1450		BACKTO1450	BACKTO1450	BACKTO1450		BACKTO1450
1500		BACKTO1500	BACKTO1500			
1600		BACKTO1600	BACKTO1600	BACKTO1600	BACKTO1600	BACKTO1600
1700	BACKTO1700	BACKTO1700		Main	BACKTO1700	
1730			BACKTO1750		BACKTO1750	
1750						BACKTO1750
1760						
1780						
1800						
1820						

**List of subdirectories containing principal components at Mann FTP site**

## Part 1: File Identities

Each subdirectory contains between 10-15 principal components. (Additionally, three of the directories also contain principal components for a period different than any of the subdirectories.) In total, the FTP site contains 430 different principal component series pertaining to the period in question. The FTP site also contains numerous principal component series relating to earlier periods for the ITRDB/NOAMER region.

MBH-r stated that a collation of 159 series was required to effect the calculations carried out in MBH98. Since this is the first public disclosure that there were 159 separate series behind MBH98 rather than 112 as stated in *Nature*, it would have been impossible to guess that that was the number needed. Subtracting the 81 non-PC series, this requires the identification of 78 PC series from these 430 series in order to successfully replicate the procedures carried out in MBH98 – an obvious impossibility.

### *Principal component rosters*

There is no direct statement in MBH98 or elsewhere of the number of principal components (and thus the total number of proxies) for each region used in the MBH98 reconstruction. However, the following information can be deduced.

The actual number of principal components used (after 1820) for each region are: Stahle/OK – 3; Stahle/SWM –9; ITRDB/NOAMER – 9; SOAMER – 3 and AUSTRAL –4 for a total of 31. Based on this identification, we note the vast and inexplicable loading of North American tree ring sites relative to other regions, as summarized below (for 1820 on):

Continent	Maximum PC Proxies	Individual Proxies	Total
North America	21	17	38
South America	3	2	5
Europe and Morocco	0	10	10
Asia	3	7	10
Australasia	4	3	7
TOTAL	31	39	70

### *Number of PC proxies by MBH Period*

The total number of proxies reported in MBH98 can be compared with the number of non-PC proxies by period in pcproxy.txt to obtain the number of PC proxies by MBH period as shown below. There are only 3 PC proxies from 1400 to 1600, rising to a full roster of 31 PC proxies by 1750.

Period Start	Total Proxies	Non-PC Proxies	PC Proxies (Difference)
1400	22	19	3
1450	24	21	3
1500	28	25	3
1600	57	39	18
1700	74	50	24
1730	79	55	24
1750	89	58	31
1760	96	62	31
1780	97	66	31

## Part 1: File Identities

1800	102	71	31
1820	112	81	31

### Number of Proxies by MBH Period

In trying to identify the PC proxies used before 1600 we note the following.

a) The Australian proxies seem to start in 1750 due to the increase of 7 proxies from 1730 to 1750 (most of the proxy totals are divisible by 3; only Australia has 4) and the lack of any collation errors in **pcproxy.txt**. Strangely all Australian sites are available as early as 1625, yet the PC does not begin until 1750.

b) the South American proxies collated into **pcproxy.txt** are all coherent and correspond to BT1600, by which time there was a full roster.

c) the Stahle/OK roster reasonably seems to be from BT1700, based on **pcproxy.txt** and proxy availability.

This suggests that there were 3 PC proxies used between 1400 and 1600. One was from NOAMER and one from Stahle/SWM, but since the Vaganov PC1 is reported to commence in 1450, there is one unaccounted for in the 1400-1450 interval. Consequently, this schedule is impossible to reconcile with disclosure in the SI on Vaganov PC rosters.

### *Number of PC Proxies by Region and Period*

information on the number of principal components by region and period, critical to any replication of MBH98, remains undisclosed to this late date. Some evidence on this comes from **pcproxy.txt**. In MC, Professor Mann stated that **pcproxy.txt** contains improper splices of various principal component series. With some patience, one can identify the source series for the splices – an exercise not dissimilar to trying to interpret Manetho’s rendering of ancient Egyptian dynasties – and thereby obtain some evidence on the supposed rosters of principal components.

For example, the ITRDB/NOAMER PC1 in **pcproxy.txt** can be seen to be a splice of PC1 versions in the 5 different subdirectories listed above (with a one-year collation error in all values after 1449). For some reason, the 1980 value is taken from the BACKTO\_1450 file. Higher PCs splice fewer versions, but all contain collation errors, including the colourful use of the 1980 value of the PC1 from the BACKTO\_1450 file for PCs 1 through 9. The versions used in the splicing of the higher PCs provide some indirect information on the series originally used. Similar splicing patterns can be seen in the Stahle/SWM region and the Vaganov region.

Strangely, the **pcproxy.txt** versions of the Stahle/OK (1700), the AUSTRAL (1750) and SOAMER (1600) regions are all taken from one subdirectory as shown in brackets and collated correctly – suggesting that, in these cases, contrary to MBH-r, the principal components were not re-calculated for each period.

By 1750, all PC proxies are presumably present. Prior to then establishing an accurate record of proxy initiation is virtually impossible with current information. The Table below shows the best we were able to come up with. This illustrates the pervasive difficulties involved in attempting to replicate MBH98, even with the recent disclosure that stepwise principal component calculations were used.

Period Start	Stahle/OK	SOAMER	AUS	VAG	Stahle/SWM	NOAMER	TOTAL
1400	0	0	0	10-1	1	1-2 BT1400	3
1450	0	0	0	0-1	1	1-2 BT1450	3

## Part 1: File Identities

1500	0	0	0	0-11	1	1-2 BT1500	3
1600	0	3-BT1600	0	0-2	3-6 BT1600	7 to 9 BT1600	18
1700	3-BT1700	3-BT1600	0	0-2	9-BT1700	7 to 9 BT1600	24
1730	3-BT1700	3-BT1600	0	0-2	9-BT1700	7 to 9 BT1600	24
1750	3-BT1700	3-BT1600	4- BT1750	3-BT1750	9-BT1700	9-BT1750	31
1760	3-BT1700	3-BT1600	4- BT1750	3-BT1750	9-BT1700	9-BT1750	31
1780	3-BT1700	3-BT1600	4- BT1750	3-BT1750	9-BT1700	9-BT1750	31
1800	3-BT1700	3-BT1600	4- BT1750	3-BT1750	9-BT1700	9-BT1750	31
1820	3-BT1700	3-BT1600	4- BT1750	3-BT1750	9-BT1700	9-BT1750	31

However, the ludicrousness of being obliged to engage in this type of cryptology merely illustrates the inadequate disclosure of MBH98 methodology.

### *Site Availability*

For most regions, the roster of sites declines before 1700 and is fairly constant thereafter until about 1970, as shown below. It is hard to rationalize the precise MBH rosters: for example, why does the Stahle/SWM PC1 extend back to 1400, but not the South American PC1 or Australian PC1? Why do the Australian PCs apparently only begin in 1750?

Period Start	Stahle/OK	Stahle/SWM <sup>2</sup>	NOAMER	SOAMER	AUS	VAG
1400	0	2	75	9	6	3
1450	0	2	92	10	7	12
1500	0	8	118	14	8	16
1600	0	14	210	18	15	40
1700	14	18	231	18	16	54
1730	14	18	231	18	16	57
1750	14	18	231	18	16	58
1760	14	18	231	18	16	58
1780	14	18	231	18	16	59
1800	14	18	231	18	16	60
1820	14	18	231	18	16	60

**Available sites in each MBH period by region.**

**To summarize:** regardless of the merits of their methodology, until MBH provide the long overdue public disclosure of their PC rosters, one is still involved in an ongoing guessing game, which is completely unedifying for a paper on which there is considerable public reliance. The disclosure within MBH98 and the Supplementary Information to MBH98 is inadequate and further disclosure was not given upon private request. Material differences may result from a reconstruction using stepwise PC calculation rather than conventional PC calculation. The non-disclosure in MBH98 of the use of stepwise PC methods is accordingly a material non-disclosure. More adequate disclosure by MBH in MBH98 may well have resulted in a more searching examination of their methodology by statistical specialists long before now.

<sup>2</sup> Excluding Spruce Canyon CO