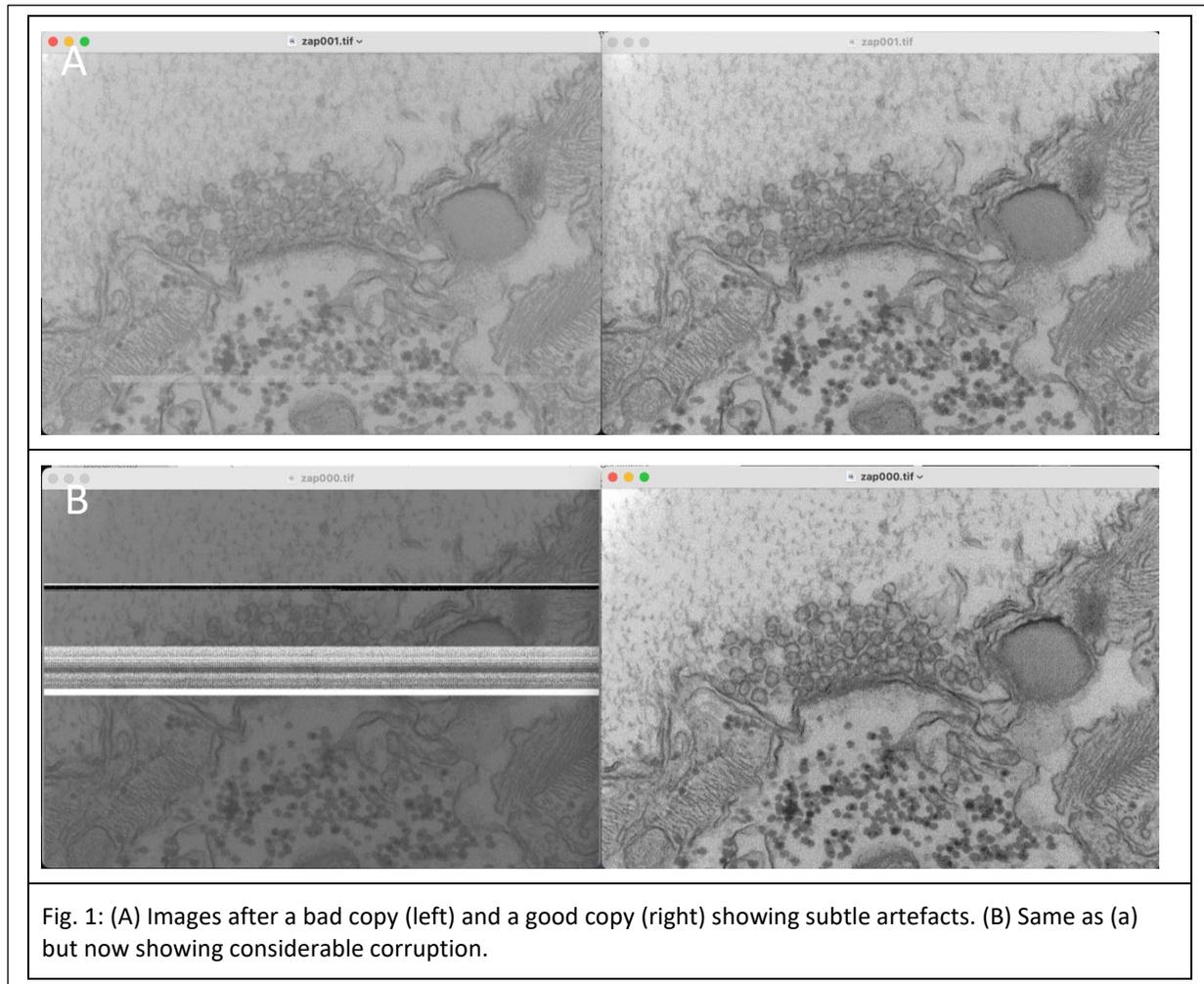


Tech Note: SerialEM, large files and robocopy

Tomographic and microED data sets created by SerialEM are single files with the data stored typically as 16-bit signed or unsigned integers. When these files have to be copied from the computer onto a USB drive, problems can arise. Consider the case below - shown are sections from one file with artefacts arising from a bad file copy alongside uncorrupted images from a good file copy.



This problem was seen after copying a 6.4Gb data file to two different USB drives, both type USB-2 using a CTRL-C, CTRL-V operation. Both copy steps proceeded to full completion without apparent errors until the files were viewed in 3dmod. After consulting with experts on the SerialEM list server the consensus solution is to use 'ROBOCOPY' under Windows. This program executes from the command line prompt and is standard available on any Windows platform, since Windows Vista and Windows server 2008.

Robocopy has many features enabling it to copy accurately large files with checksum capabilities and can take into account network bottlenecks. In a way, it resembles FTP.

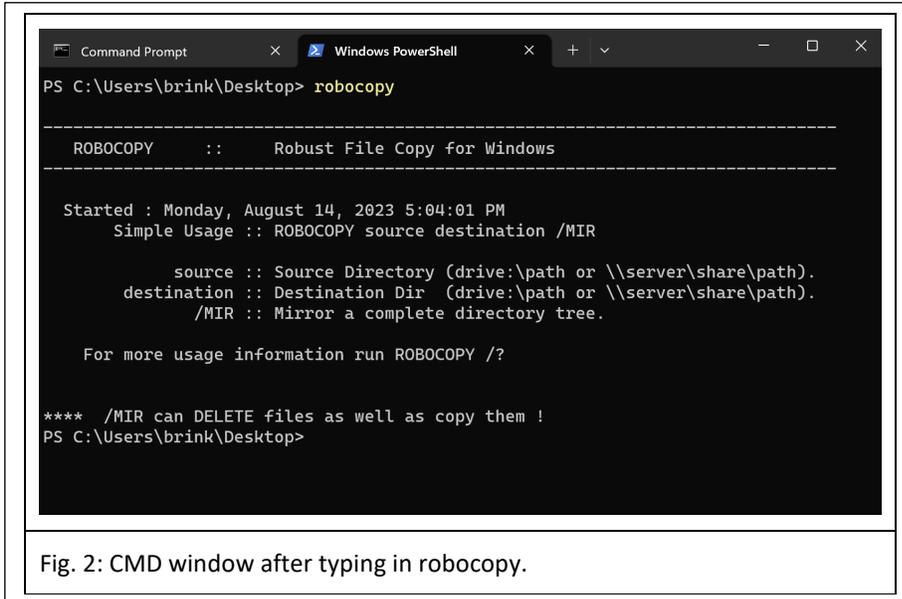


Fig. 2: CMD window after typing in robocopy.

The syntax for robocopy is 'robocopy source-volume destination-volume file'. As an example, the file records-50kx.st was copied from a USB drive (E:) to the desktop of my virtual machine (C:\Users\brink\Desktop\) (Fig. 3). Multiple files can be copied by employing a wildcard in the filename, e.g records-50kx.*.

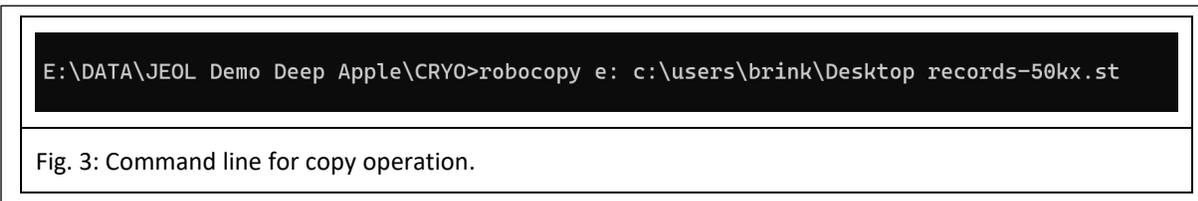


Fig. 3: Command line for copy operation.

Afterwards, robocopy returns with some statistics, e.g. any errors, repeats (Fig. 4). There should be none.

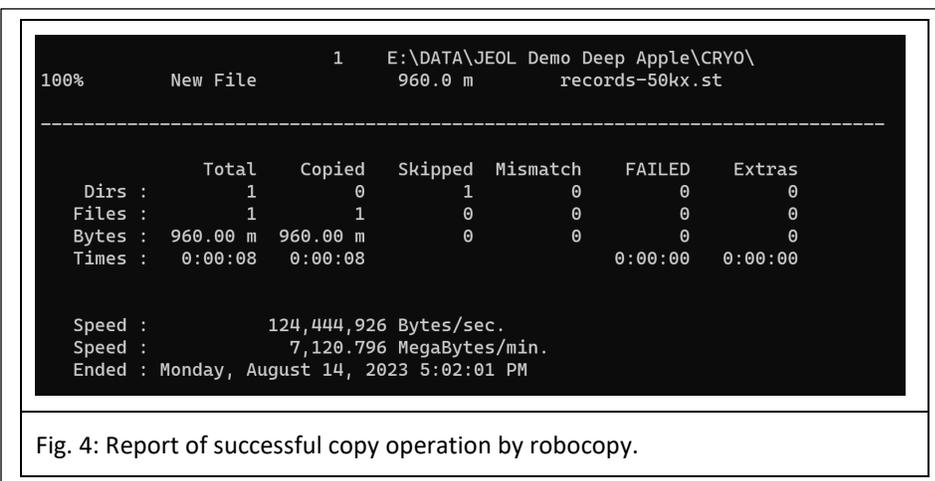


Fig. 4: Report of successful copy operation by robocopy.