3. HP Starter File

3.1 The Role of the Starter File

The serial version of standard PEST (including I64PEST, a 64 bit version of PEST) can now be run with a new command line option. This option is "/hpstart". Let us suppose that a PEST control file is named *case.pst*. Then to run PEST with this new option, use the command

```
pest case /hpstart
```

If initiated using this option, PEST runs the model only once using parameter values supplied in the PEST control file; it then ceases execution. As well as its usual set of output files, PEST writes a binary file named *case.hp*; this is referred to herein as a "hp starter file". This file is used by PEST_HP. If PEST_HP is to be run on another machine, or in the cloud, this file should be placed in the folder used by the PEST_HP manager, and in folders used by all run agents.

When execution of the PEST_HP manager is initiated, the following command can be used.

```
pest_hp case /hpstart /h :4004
```

When the PEST_HP manager is started in this way, it reads file *case.hp*, along with its normal set of PEST input files (i.e. the control, template and instruction files). The information contained in file *case.hp* eliminates the requirement for PEST_HP to undertake an initial model run in which it calculates values for model outputs using parameter values supplied in the PEST control file. (These outputs are used as reference values in finite-difference calculation of derivatives that are used to fill the first Jacobian matrix.) Instead, PEST_HP obtains these reference values from the hp starter file. These reference model outputs are also used to calculate the initial objective function, which is reported to the screen and to the run record file.

Because the PEST_HP manager obtains initial reference values from the hp starter file, and because the need for a model run based on initial parameter values is thereby eliminated, PEST_HP can commence finite-difference calculation of the initial Jacobian matrix immediately upon commencement of execution. All run agents can therefore be immediately put to work; there is no "gap time" in which they are waiting for completion of the initial model run (and thereby losing money if cores have been purchased on the cloud). This is a better option for handling of the initial model run than that provided by the "/p1" command line switch which is supported by BEOPEST, but not by PEST_HP. The "/p1" switch bundles the initial model run with the model runs required for filling of the first Jacobian matrix. Hence this first bundle of runs is comprised of N+1 model runs, where N is the number of adjustable parameters. Filling of the second Jacobian matrix requires only N model runs however, as reference observations used in finite-difference approximations to derivatives are calculated during the preceding Marquardt lambda selection and parameter upgrade testing process. Where inversion takes place in the cloud, this disparity between the number of model runs required to fill successive Jacobian matrices can make selection of an optimal number of processors difficult. Ideally N should be an integer multiple of the number of available processors; no processors are therefore idle as model runs are undertaken for filling of the Jacobian matrix.

(Note that the above logic assumes that one model run is required per adjustable parameter. If derivatives with respect to all parameters are calculated using higher order finite-differences, the same logic applies, as the number of model runs required for filling the Jacobian matrix remains an integer multiple of *N*. The argument breaks down, however, where forward differences are used for calculating derivatives with respect to some parameters and higher order differences are used for calculating derivatives with respect to other parameters. It is also invalidated where PEST's observation re-referencing functionality is invoked. This does not erode the value of the hp starter file however.)

The PEST_HP run agent (i.e. AGENT_HP) also reads the hp starter file. However it does not need to be started using the "/hpstart" command line switch; in fact if this switch is provided to AGENT_HP, it is ignored (as are any other command line options). Instead, AGENT_HP is informed of the fact that it should read a hp starter file by the PEST_HP run manager (if its execution is initiated using the "/hpstart" command line option). The contents of the hp starter file include information which can facilitate the reading of lengthy model output files by AGENT_HP. In particular, AGENT_HP is informed of the order in which observations are encountered as model output files are read. Where observations number in the hundreds of thousands, and/or where ordering of model outputs corresponding to observations does not coincide with the ordering of observations in the PEST control file, this can reduce the time to read model output files considerably on the first occasion on which they are read by AGENT_HP. (Note that if a hp starter file is not employed, then AGENT_HP orders observations internally as it reads model output files for the first time; its second reading of these files is therefore fast, regardless of the presence, or otherwise, of a hp starter file.)

If either PEST_HP or AGENT_HP is asked to read a hp starter file, but the file is not actually present in the folder from which it is run, the absence of this file will be tolerated. PEST_HP will commission the first model run using a single agent, just as if it had not been requested to read a hp starter file at all. If AGENT_HP does not find an expected hp starter file in its folder, it will read model output files in the usual way, without the benefit of ordering information contained in the hp starter file. Unless observation numbers are very large, this is not a serious disadvantage. Where observation numbers are very large, the benefits of ordering are available on subsequent occasions that AGENT_HP reads model output files.

A hp starter file must be written by PEST or I64PEST. It cannot be produced by BEOPEST. A serial version of PEST must be used to produce a hp starter file because it stores both reference observations and the ordering of observations in model output files. Where an initial model run is undertaken by BEOPEST, the former information is available to the manager whereas the latter information is obtained by an agent. It is therefore difficult to record both sets of information in the same file. Use of a serial version of PEST to write the hp starter file is not a constraint however, as the concept of "parallel" makes little sense when only one model run is undertaken.

If execution of PEST_HP is re-commenced using the "/s" switch, then the presence of a "/hpstart" switch is ignored – except for the fact that the PEST_HP manager instructs its agents to read this file in order to accrue the efficiency gains discussed above where observation numbers are large.

3.2 The "/i" Command Line Option

Execution of PEST_HP can be initiated using the "/i" command line switch. (This switch is also available for PEST and BEOPEST.) If started in this way, PEST_HP requests the name of a JCO file

which it must read to obtain the initial Jacobian matrix; it does not therefore need to fill this Jacobian matrix using finite-difference-calculated derivatives. See PEST documentation for further details.

If execution of PEST_HP is initiated using both the "/hpstart" and "/i" command line options, then the need to undertake both the initial model run, and any model runs required for filling of the initial Jacobian matrix, are eliminated. PEST_HP can therefore commence calculation of parameter upgrades using different Marquardt lambdas immediately. However it has been programmed to pause for about a minute before commencing this process; this provides time for all run agents to connect to the manager. As has been discussed above, the Marquardt lambda and line search strategy adopted by PEST_HP for calculation of parameter upgrades depends on the number of agents to which it has access.

Note that, as stated above, AGENT_HP ignores all command line options (except the "/h" option through which it learns the contact details of its manager).

3.3 SVD-Assist and the "/hpstart" Option

There is nothing to be gained by PEST_HP reading a hp starter file if SVD-assisted inversion is being undertaken. The Jacobian matrix file produced by PEST (or by BEOPEST or PEST_HP) as a precursor to initiation of SVD-assisted inversion actually includes reference values for observations. Neither the initial model run, nor model runs required for filling of the first Jacobian matrix, need therefore be undertaken by either PEST_HP or by PEST/BEOPEST at the commencement of an SVD-assisted inversion process. Use of the "/hpstart" command line switch therefore has no effect on the operation of PEST_HP, except for the fact that it informs its run agents to expect a hp starter file in their working folders. As is explained above, the agents can then use this file to order observations prior to reading model output files so that the reading of these files can be expedited.