LUCA: A WIZARD-STYLE GUI FOR EASY AND SYSTEMATIC MULTI-OBJECTIVE, STEP-WISE CALIBRATION OF HYDROLOGIC MODELS

Makiko Umemoto, Computer Scientist, U.S.Geological Survey, Denver,CO, makiko@usgs.gov;
Lauren E. Hay, Hydrologist, U.S.Geological Survey, Denver,CO, lhay@usgs.gov;
Steve L. Markstrom, Hydrologist, U.S.Geological Survey, Denver,CO, markstro@usgs.gov;
George H. Leavesley, Hydrologist, U.S.Geological Survey, Denver,CO, george@usgs.gov

Abstract: This paper introduces the graphical user interface (GUI), Luca (Let us calibrate), a tool for building and performing a procedure to calibrate parameters for a hydrologic model. Luca integrates the following components in a user-friendly GUI: (1) multiple-objectives step-wise calibration; (2) Shuffled Complex Evolution (SCE) global-search parameter optimization; and (3) the Modular Modeling System (MMS). This paper presents detailed instructions for using Luca to build a calibration procedure, including all available functionalities.

INTRODUCTION

A multi-objective, step-wise, automated calibration procedure was presented by Hay et al. (2006) and shown to be an effective technique to calibrate a hydrologic model. This approach uses the Shuffled Complex Evolution (SCE) global search algorithm (Duan et al., 1992; 1993; and 1994) to calibrate parameters for the Precipitation Runoff Modeling System (PRMS, Leavesley et al., 1995), a distributed hydrologic model implemented in the Modular Modeling System (MMS) (Leavesley et. al., 1996). This paper presents a wizard-style graphical user interface (GUI) called Luca (Let us calibrate) that allows the user to build a multiple-objective, step-wise calibration procedure for a hydrologic model calibration in an easy and systematic manner.

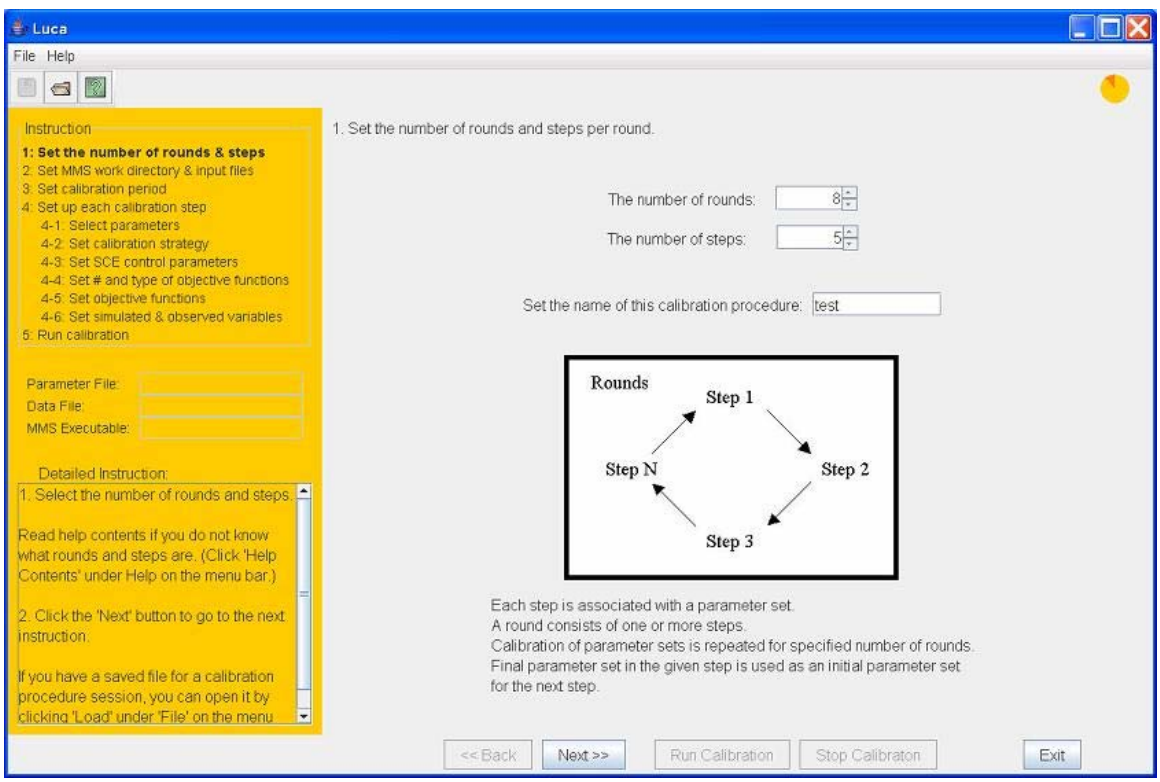

Figure 1: Initial screen of Luca

Luca

The following sections describe each step in Luca – the wizard-style graphical user interface for multi-objective, step-wise, automated calibration. Each screen in this wizard-style GUI consists of three components (see Figure 1):

- the instruction panel, which is the orange colored panel located on the left
- the main panel located in the center
- the buttons at the bottom

The instruction panel provides an overview of all tasks the user has to complete in order to build the calibration procedure. These tasks are called 'instructions' in Luca terminology. This panel also shows the current instruction indicated by the bold text. The instruction panel in Figure 1 shows that there are a total of 5 instructions, and the current instruction is Instruction 1. The user builds the calibration procedure by following the instructions in the main panel. The user can proceed to the next instruction by clicking the *Next* button located on the bottom. The *Back* button allows the user to go back to previous instructions.

Luca provides three useful additional functionalities: save, load, and help buttons (from left to right) are located on the top left of the panel beneath the menu bar (See Figure 1). After the calibration procedure is built, it can be saved by clicking the *Save* button and specifying the file to save. The saved calibration configuration can be loaded anytime by clicking the *Load* button and specifying the file to load. When the *Help* button is clicked, the Help contents, which contain the detailed instructions of how to use Luca including the contents of this paper, are displayed.

INSTRUCTIONS FOR BUILDING A CALIBRATION PROCEDURE

Instruction 1: Set the number of rounds and steps per round: Figure 1 shows the initial Luca screen. In this example, *the number of rounds* and *steps* are set to 8 and 5, respectively. A step is associated with a parameter set to be calibrated. The parameter set can contain one or more parameter values. A round consists of one or more steps. The steps are performed in a sequential order. Also, the name of this calibration procedure, which is used to distinguish this procedure from others, is set to *test*.

In the paper by Hay et al. (2006) a procedure that includes the sequential calibration of simulated: (1) solar radiation, (2) potential evapotranspiration (PET), (3) annual water balance; (4) snow-covered area; and (5) components of daily runoff is presented. Implemented in Luca, and shown in Figure 1, this would be a 5-step procedure with a sufficient number of rounds to reach an objective function minimum (generally less than 8 rounds).

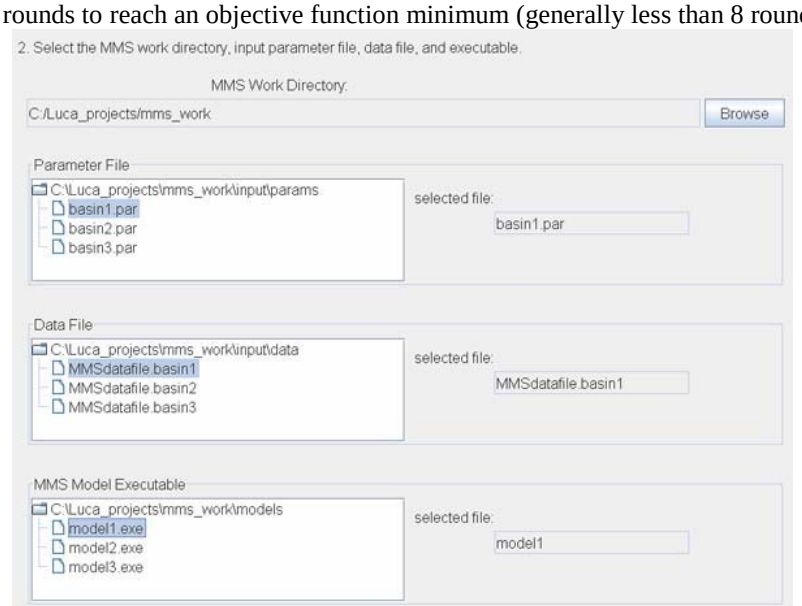

Figure 2: Instruction 2

Instruction 2: Set MMS work directory and input files: Figure 2 shows the Luca screen for MMS work directory selection. When the MMS work directory is selected, the parameter files, data files, and MMS models contained

within this area are shown (see Figure 2). The user then selects an input *Parameter File*, an input *Data File*, and an *MMS Model executable*.

<u>Instruction 3: Set calibration period</u>: This instruction (see Figure 3) prompts the user for the time period in which the MMS executable performs a simulation. Note that the first year of this user-defined time period is devoted to model initialization and is not used for objective function calculation. Hay et al. (2006) used 8 water years (1997-2004) for calibration (Water years run from October through September). Figure 3 shows the *Start Date* and *End Date* of the data to reflect this calibration period, including the initial year (water year 1996) for model initialization.

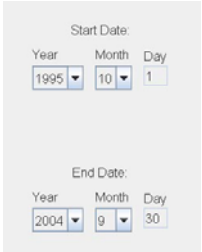

Figure 3: Instruction 3

<u>Instruction 4: Set up each calibration step</u>: This next set of instructions (4-1 to 4-6) are for the MMS executable setup (4-1 to 4-3) and the objective function selection (4-4 to 4-6). In each of these instructions, tabs are displayed on the top of the main panel (see Figure 4). Each tab is associated with each calibration *step*. The five tabs shown on the top of Figure 4 indicate that there are five calibration *steps* to set up. Luca allows the user to establish unique settings for each calibration step by completing the instructions after selecting each tab.

Instruction 4-1: Select parameters: A parameter set for each calibration step is defined by selecting parameters from available parameters (displayed in the list in Figure 4). These parameters are automatically displayed based on the input parameter file selected in Instruction 2. The user is allowed to create a distinct *parameter set* for each calibration step by clicking the appropriate *Step* on the tab and completing this instruction for each step. In the paper by Hay et al. (2006), step 2 of the calibration procedure calibrates the PET parameter jh_coef (see Figure 4 and Hay et al., in press for further details).

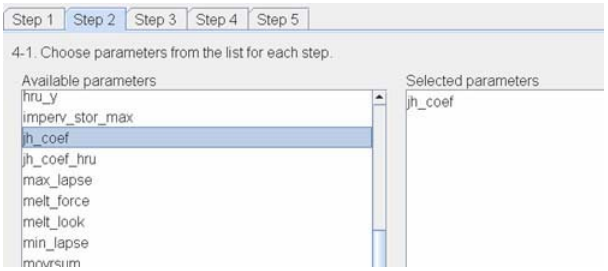

Figure 4: Instruction 4-1. Step 2 is shown.

Instruction 4-2: Set calibration strategy: The initial parameter values from the input parameter file selected in Instruction 2 are shown on the screen. The user has three choices for calibrating the parameter values:

(1) *Use the mean value*: the mean, instead of individual values, is calibrated for a given parameter (see Figure 5).
(2) *Use the individual values*: the individual values of a given parameter are calibrated. The parameter shown in Figure 6 (jh_coef) has a dimension of 12. The user is allowed to select which individual values should be calibrated by checking the boxes as shown in Figure 6. All 12 individual values were calibrated for step 2 in the Hay et al. (2006) paper.
(3) *Parameters are binary (0, 1)*: In the case of a binary parameter, the user needs to choose this option. The parameter values must be either 0 or 1. Calibration will be done in the same way as *Use the individual values*; the difference is that if the calibrated value is greater than or equal to 0.5, the parameter value is set to 1, and set to 0 otherwise.

The lower and upper bounds must be defined for each parameter in instruction 4-2. After entering the values for lower and upper bounds, the actual range is displayed below them. The actual range is exactly the same as the user defined bound if *Use the individual values* or *Parameters are binary* is selected as the calibration type. If *Use the mean* is selected, then the actual range is calculated as follows:

$$\text{Actual lower bound} = \frac{(L_u + C)(M + C)}{(min + C)} - C \tag{1}$$

$$\text{Actual upper bound} = \frac{(U_u + C)(M + C)}{(max + C)} - C \tag{2}$$

where L_u is the user defined lower bound, U_u is the user defined upper bound, M is the mean of initial parameter values, C is the absolute value of (L_u + 10, *max* is the maximum parameter value, and *min* is the minimum parameter value. The derivation and reasoning of these formulas can be found in the Luca Help contents (Figure 1).

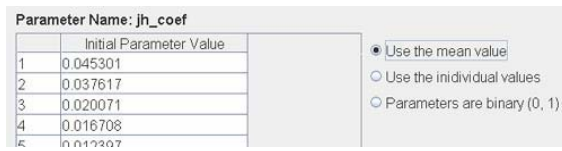

Figure 5: Instruction 4-2. The calibration type *Use the mean value* is selected.

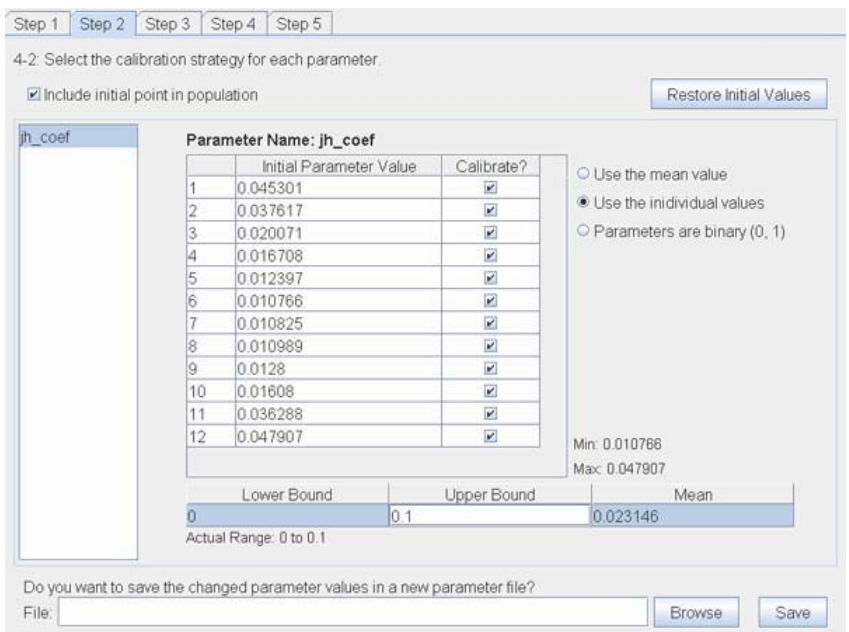

Figure 6: Instruction 4-2. The calibration type *Use the individual values* is selected

Instruction 4-3: Set SCE control parameters: For each Step, SCE control parameters need to be defined (see Figure 7). These values are automatically calculated based on the suggestions in the Duan et al. (1994) paper. For more information on these parameters the user is referred to Duan et al. (1994). For each step and round a parameter file is saved. The user can change the name of this parameter file in the lower panel.

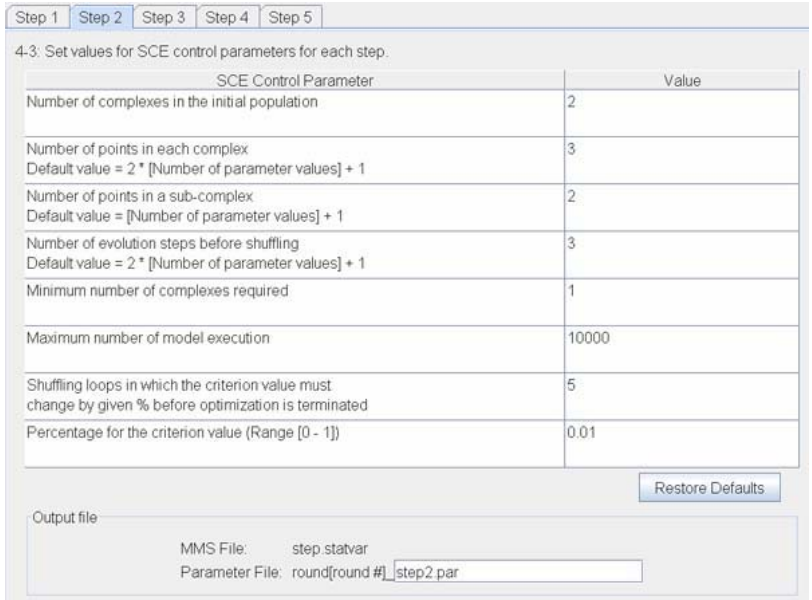

Figure 7: Instruction 4-3.

Instruction 4-4: Set number and type of objective functions: In each Step the user must define *The number of objective functions* and the *Objective Function Type* to use (see Figure 8). For each step, the objective functions must be the same type. Figure 8 shows the selection for Hay et al. (2006), Step 2. In this case, one objective function with an objective function type *Absolute Difference [log]* was chosen.

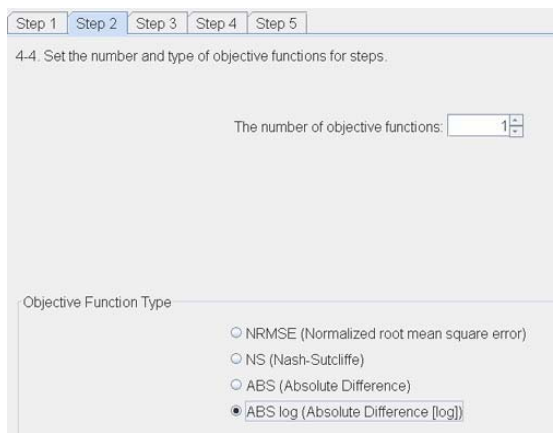

Figure 8: Instruction 4-4.

Instruction 4-5: Set the weight and the time step for each objective function: In Luca, the final objective function value for each step is calculated as follows:

$$OF_f = \sum_{i=1}^{nOF} w_i OF_i \tag{3}$$

where OF_f is the final objective function value, nOF is the number of objective functions for given step, w_i is the ith weight associated with the ith objective function (OF_i). The user is prompted to set the weight for each objective function in every step (Figure 9).

The user also has a choice of using a subset of the observed data for model calibration. If the *Use data_subdivide* is chosen, the user is prompted for an input file that consists of the date and a value. When the data value for a given time step is equal to the value listed in the *data_subdivide value*, then data for that day is used in model calibration.

For example, the user may only want to calibrate runoff on days with irrigation diversions, peak flows, or significant hydrometeorological events. The details are explained in the Luca Help contents (Figure 1).

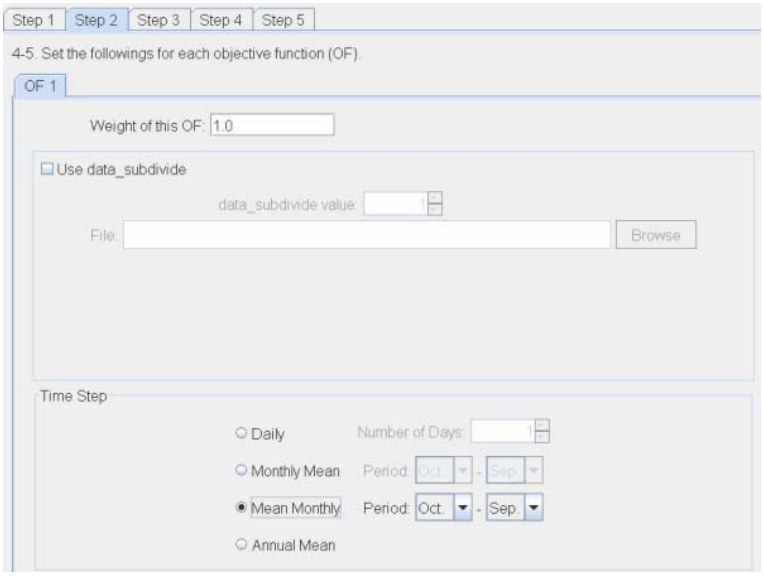

Figure 9: Instruction 4-5

The time step specified in instruction 4-5 is used to calculate an objective function value for simulated and observed values. The available time steps are *Daily*, *Monthly Mean*, *Mean Monthly*, and *Annual Mean*. The monthly mean is the arithmetic mean of individual daily values within a given month. The mean monthly is the arithmetic mean of the monthly means of a given month during a specific period of years. *Annual Mean* is a water year mean. The calibration period must be at least 2 years (including 1 year for model initialization) to select *Annual Mean*. If the user selects *Daily*, then the *Number of Days* must be chosen. A value greater than 1 uses an n-day moving average for objective function calculation. The *Period* field must be specified if the *Monthly Mean* or *Mean Monthly* time step is selected. *Period* allows the user to select a period of months whose monthly mean or mean monthly values are used for objective function calculation. For example, if *Period* is set to Oct. – Mar., then the monthly mean or mean monthly values for the period from October to March of each year are used for objective function calculation, and values from April to September are ignored. If *Period* is set to Oct. – Sept., then all monthly mean or mean monthly values are used.

Hay et al. (2006) calibrated the PET parameter jh_coef using one objective function and mean monthly data. Figure 9, reflects these choices: one tab *OF 1* is seen under the *Step 2* tab, indicating one objective function with a weight of 1.0, no data subdivide, and *Mean Monthly* time step.

Instruction 4-6: Set simulated and observed variables for each objective function and calibration step: In this instruction the user must select the simulated variable from a list of variables that are produced by the model and saved in an MMS file called *statvar* (see Figure 10). The corresponding observed data must be identified and the source must be indicated as either *Statvar File* or *External Source*. If the user chooses *External Source*, the user must supply the observed data file, which must contain the data covering at least the calibration period excluding the first year model initialization period. When this *External Source* is selected, the observed data type (*Value* or *Range*) must be set.

(1) Observed Data File from External Source, format for the *Value* type.
If the observed data file contains an observed value for each date, it is the *Value* type. If the *Value* type is selected, the file format of the observed data must be selected from several choices:
Daily (Year Month Day Value): This format can be selected for any time step used for objective function calculation. These daily values in the file are converted into monthly mean, mean monthly, or annual mean values as necessary.

Monthly Mean (year month value): This format can be selected if the time step for objective function calculation is anything except for *Daily*. These monthly mean values in the file are converted into mean monthly or annual mean value as necessary.
Mean Monthly (month value): This format can be selected only if the time step for objective function calculation is *Mean Monthly*. The file does not have to contain values for all 12 months, but must contain mean monthly values of the period selected in Instruction 4-5.
Annual (year value): This format can be selected only if the time step for objective function calculation is *Annual Mean*.

(2) *Observed Data File from External Source*, format for the *Range* type.
If the observed file contains the lower and upper bounds of the observed value for each date, then it is the *Range* type. The observed values used in the objective function calculations are determined based on the range defined by the lower and upper bounds in the observed data file and the simulated value generated by the MMS executable. If a simulated value is within the range for a given date, the observed value is set equal to the simulated value. If it is less than the lower bound of the range, the observed value is set to the lower bound. If it is greater than the upper bound, then the observed value is set to the upper bound.

If the *Range* type is selected, the format which the file must follow is displayed in the panel:
Daily (Year Month Day Lower_bound Upper_bound)
Monthly Mean (Year Month Lower_bound Upper_bound)
Mean Monthly (Month Lower_bound Upper_bound)
Annual Mean (Year Lower_bound Upper_bound)

Figure 10 shows the Luca setup for Instruction 4-6 used by Hay et al. (2006) for *Step 2* of their calibration procedure. In this case *Simulated Variable* is set to *basin_potet.potet* (basin PET), observed calibration data is from an *External Source*, and *Observed Data File from External Source* is of the *Value* type with a *File Format* of *Mean Monthly*.

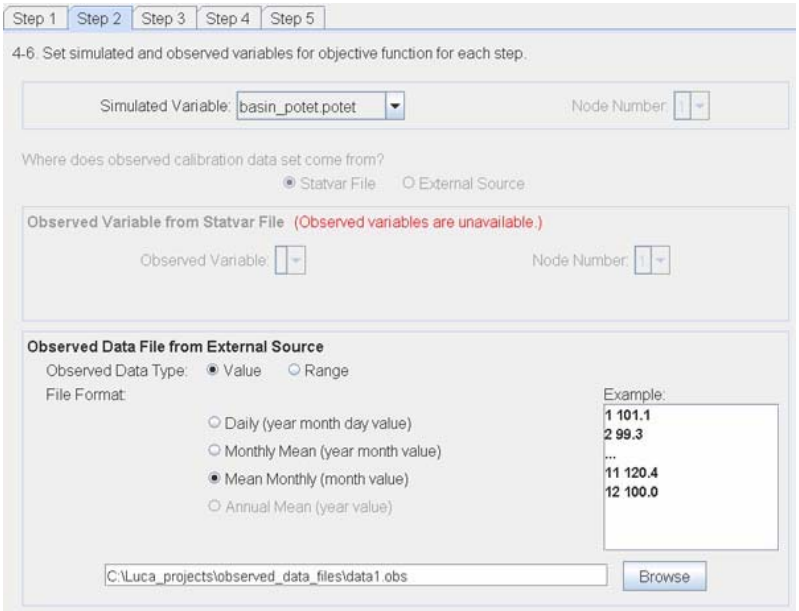

Figure 10: Instruction 4-6.

<u>Instruction 5: Run calibration</u>: In this last instruction the user must click the "Run Calibration" button at the bottom of the panel (see Figure 11), which runs SCE for every step of every round. While it is running, Luca shows the program's current status, and the status of every step of every round including the current best parameter values and the objective function values. After the calibration ends, the end statement is displayed in the box located on the bottom left. The end statement explains why the SCE is terminated for a given step of a round

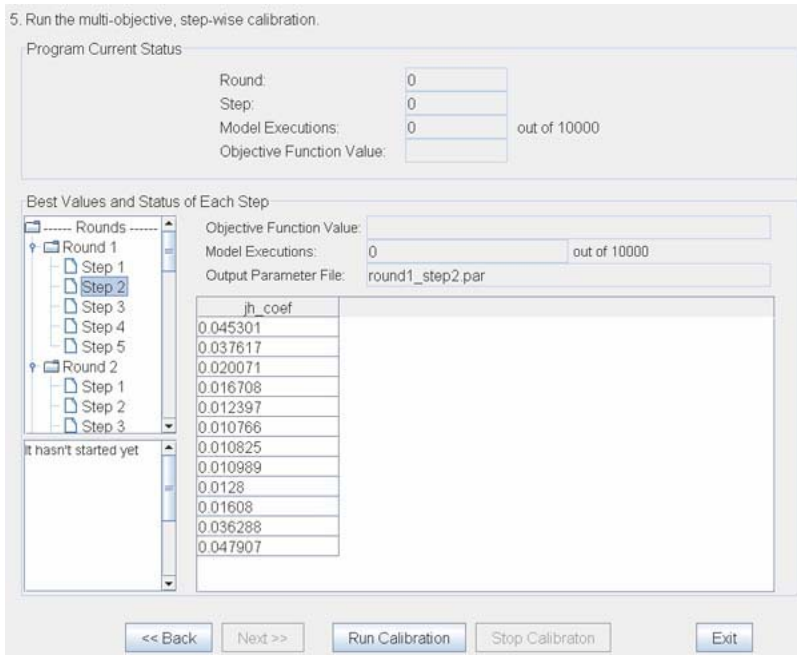

Figure 11: Instruction 5 (before calibration).

CONCLUSION

This paper introduced the wizard-style graphical user interface, Luca, which is a tool for building a multi-objective, step-wise, automated calibration procedure to calibrate a hydrological model using the Modular Modeling System. The detailed instructions of Luca to build the procedure presented in this paper are intended to help the user get started with Luca.

REFERENCES

Duan, Q., Sorooshian, S. and Gupta, V.K., (1992). Effective and efficient global optimization for conceptual rainfall-runoff models. Water Resources Research 28 (4), 1015-1031.

Duan, Q., Sorooshian, S. and Gupta, V.K., (1993). A Shuffled Complex Evolution approach for effective and efficient global minimization. *J. of Optimization Theory and its Applications*, 76 (3), 501-521.

Duan, Q., Sorooshian, S. and Gupta, V.K., (1994). Optimal use of the SCE-UA global optimization method for calibrating watershed models. *Journal of Hydrology*, 158 265-284

Hay, L.E., Leavesley, G.H., Clark, M.P., Markstrom, S.L., Viger, R.J., and Umemoto, M. (in press). Step-wise, multiple-objective calibration of a hydrologic model for a snowmelt-dominated basin. *Journal of the American Water Resources Association*.

Hay, L.E., Leavesley, G.H., and Clark, M.P., (2006). Use of Remotely-Sensed Snow Covered Area in Watershed Model Calibration for the Sprague River, Oregon. Joint 8th Federal Interagency Sedimentation Conference and 3rd Federal Interagency Hydrologic Modeling Conference, Reno, Nevada, April, 2006.

Leavesley, G.H. and L.G. Stannard, (1995). The precipitation-runoff modeling system- PRMS. In: Computer Models of Watershed Hydrology, Water Resources Publications, Highlands Ranch, CO, edited by V.P Singh, Chapter 9, 281-310.

Leavesley, G.H., Restrepo, P.J., Markstrom, S.L., Dixon, M., and Stannard, L.G., (1996). The modular modeling system - MMS: User's manual: U.S. *Geological Survey Open File Report* 96-151, 200 p.