

# BRIEF INTRODUCTION TO THE AUTOMATED STATISTICAL DOWNSCALING (ASD) MODEL

The Automated Statistical Downscaling (ASD) tool is an easy to use graphical user interface for the statistical downscaling of GCM outputs to regional or local variables. ASD runs on all platforms that support MATLAB. This short introduction assumes a working knowledge of MATLAB. For a more detailed introduction to the software, see the ASD User Guide (Version 1.1).

## Before You Start

In order to run ASD you will need to provide daily observed local data (the predictand) and daily observed and GCM large-scale atmospheric data (the predictors). In Canada, daily observed local data can be obtained through Environment Canada. Predictor data sets<sup>1</sup>, including both observed and GCM outputs, can be downloaded from the Canadian Climate Change Scenarios Network ([www.cccsn.ca](http://www.cccsn.ca), **statistical downscaling input** section). The ASD naming convention and file formats follow that of SDSM and predictor and predictand files prepared for use in ASD can also be used in SDSM. All input data files should be a single stream of data with each row representing a single daily value. The predictor data sets are already in this format, but you will probably need to reformat observed daily data files.

There are two other files which ideally need preparing before starting the ASD tool – `ncepdata.txt` and `gcmdata.txt`, although they may be created and/or edited whilst within the ASD environment. These files contain a list of the predictor files available for the NCEP data (observed large-scale atmospheric variables) and for the GCM in question. Examples of these files are provided with ASD (look in the `c:\Program Files\ASD` directory) and can be edited if necessary. The order of the file names contained in these two files must be exactly the same and for ASD to work correctly the same predictor variables should be available for both the NCEP and GCM data sets. There is little point in calibrating a statistical model containing a predictor variable which is not available in the GCM data set since, if it is identified as a significant variable and used in the statistical model, it will not then be possible to downscale future climate scenarios.

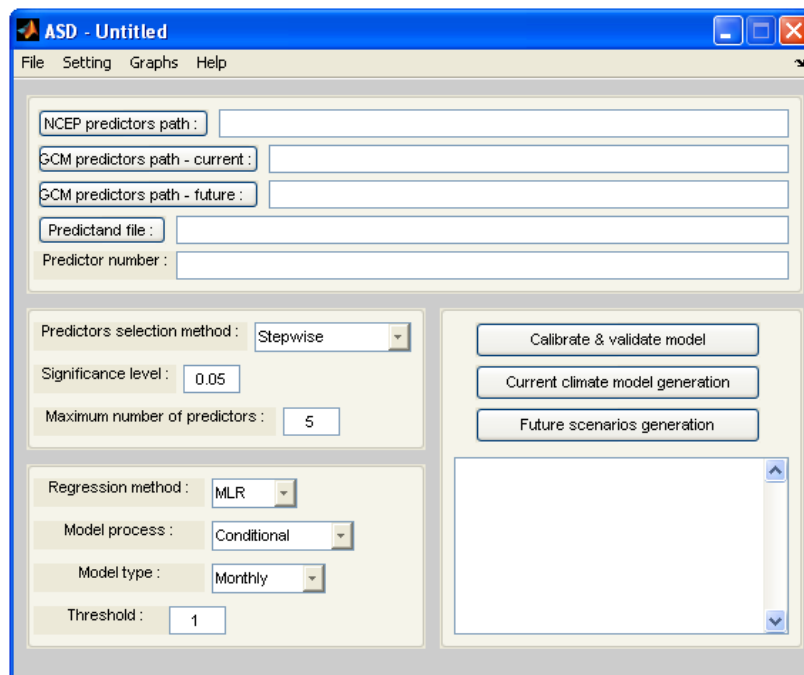
## Starting ASD

Start MATLAB. At the top of the MATLAB main window you will see a Current Directory window containing the directory path of the MATLAB work directory (e.g., `c:\MATLAB7\work`). Change this directory path so that it shows the location of the ASD programs (e.g., `c:\Program Files\ASD`). To start ASD, type “asd” at the MATLAB prompt. The ASD main window opens (Figure 1).

1. Click on the **Setting** menu at the top of the ASD main window and select the **NCEP predictors** or **GCM predictors** option from the drop-down menu to open **ncepdata.txt** or **gcmdata.txt** file. Clicking on the **NCEP predictors** button or **GCM predictors** button will allow you to browse through the directory tree to locate the appropriate directory and files. Both these files contain a list of the available predictor file names – the order of the list in each file should be identical.

---

<sup>1</sup> Note: The GCM predictor files provided by the CCCSN are for the period 1961-2100. If you wish to downscale for particular time periods, e.g., the 2050s (2040-2069) you will need to extract the relevant period from the complete data set before using the ASD tool.

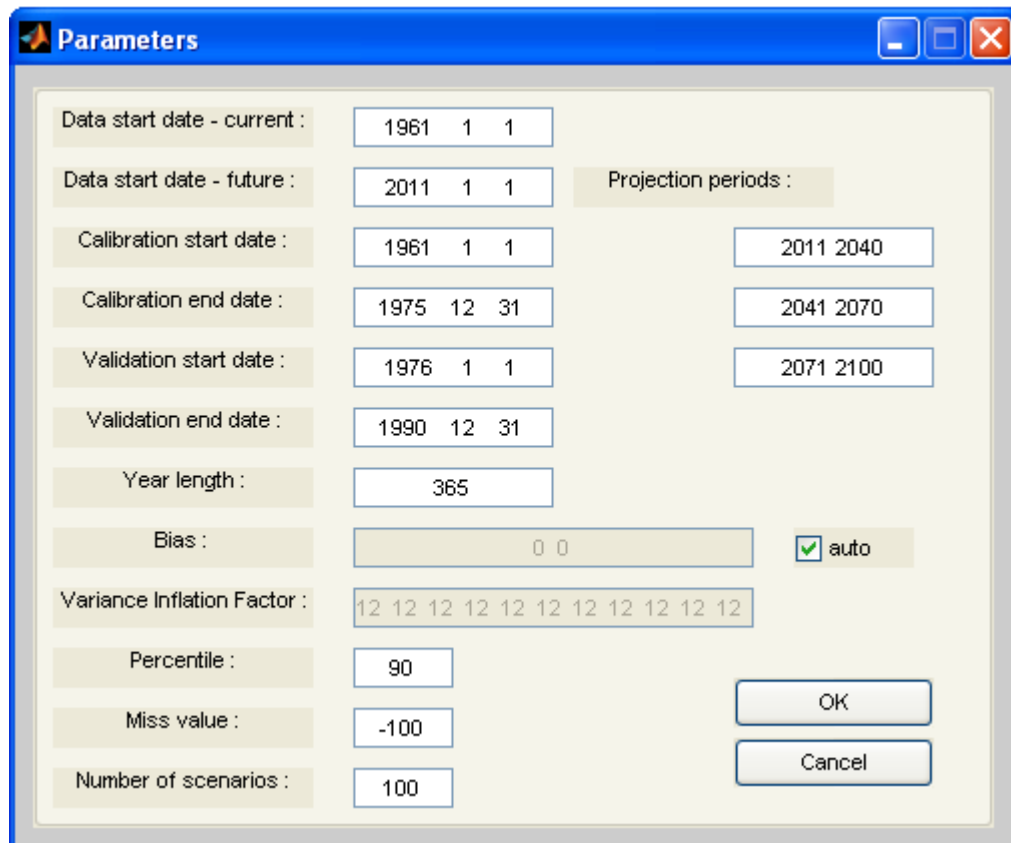


**Figure 1:** The main menu of the ASD (Automated Statistical Downscaling) tool

2. Click on the **Setting** menu at the top of the ASD main window and select the **Parameters** option from the drop-down menu. Enter the **start date** of **current** data, the **start date** of **future** data, the **calibration start and end dates**, the **validation start and end dates** and the **year length** for GCM data<sup>2</sup> (Figure 2). Enter the **Bias** and **VIF (Variance Inflation Factor)** values manually or check the appropriate box to set them automatically (i.e. auto, as in Figure 2). These values are used only when generating scenarios using GCM predictor data (or through **current climate model generation** using GCM predictors). Enter the **Percentile** value that you would like to compute for the **Comparing Statistics** menu. Enter a **threshold** value for the identification of missing values (i.e. **miss value**, any values less than the threshold value will be treated as missing) and finally enter the **number of scenarios** you wish to generate (ex. 100 generated synthetic simulated series, as suggested in Figure 2). Finally, enter the **projection periods** (in Figure 2), the 3 periods of 3 decades, as for ex. 2020s, 2050s and 2080s if GCMs predictors are available over the entire 21<sup>st</sup> century.
3. Return to the ASD Main Menu and click on the **NCEP predictors**, **GCM predictors** and **Predictand** buttons to browse through the directory tree to find the directory paths to the predictor and predictand data files (Figure 3). In the **Predictor number**, enter the number of the predictors which you wish to be considered in the calibration process (e.g., 1:26 if the complete predictor data set is available, or, for example, 1:21 23 24, if only a subset of the complete suite is available). The **predictor selection method** gives three choices: stepwise, partial correlation or none (in this latter case, the process will be based on all predictors entered in the “predictor number”). Chose the **significance level** and the **maximum number of predictors** that you want to select. The choice of two **regression methods** is allowed, i.e. MLR (i.e. multiple linear regression) and the ridge regression (i.e. to alleviate the effect of colinearity between predictor). Select whether the **model process** is conditional (e.g., if downscaling precipitation) and, if so, enter a **threshold** value for the

<sup>2</sup> Some GCMs do not have a standard year length and most do not include leap years. You will need to check the number of days in a year for the GCM you are using. GCMs from the Canadian Centre for Climate Modelling and Analysis (e.g., CGCM1, CGCM2, CGCM3) have a 365-day year, whilst those from the Hadley Centre for Climate Prediction and Research (e.g., HadCM2, HadCM3) have a 360-day year.

boundary between dry and wet days (e.g. 1 mm/day or other). By default, a transformation function<sup>3</sup> (ex. fourth root per default) will be applied only to the predictand in the conditional model. The model process can be conditional on the occurrence of an event (i.e. for precipitation) or unconditional (i.e. for temperature). Also, the model type can be based on monthly downscaling process or annual one (the two choices given in **Model Type** menu). To calibrate the model and generate simulations using NCEP data, click on the **Calibrate & Validate model** button. To generate scenarios using current and future GCM data, click on the **Current climate model generation** and **Future scenarios generation** buttons, respectively.



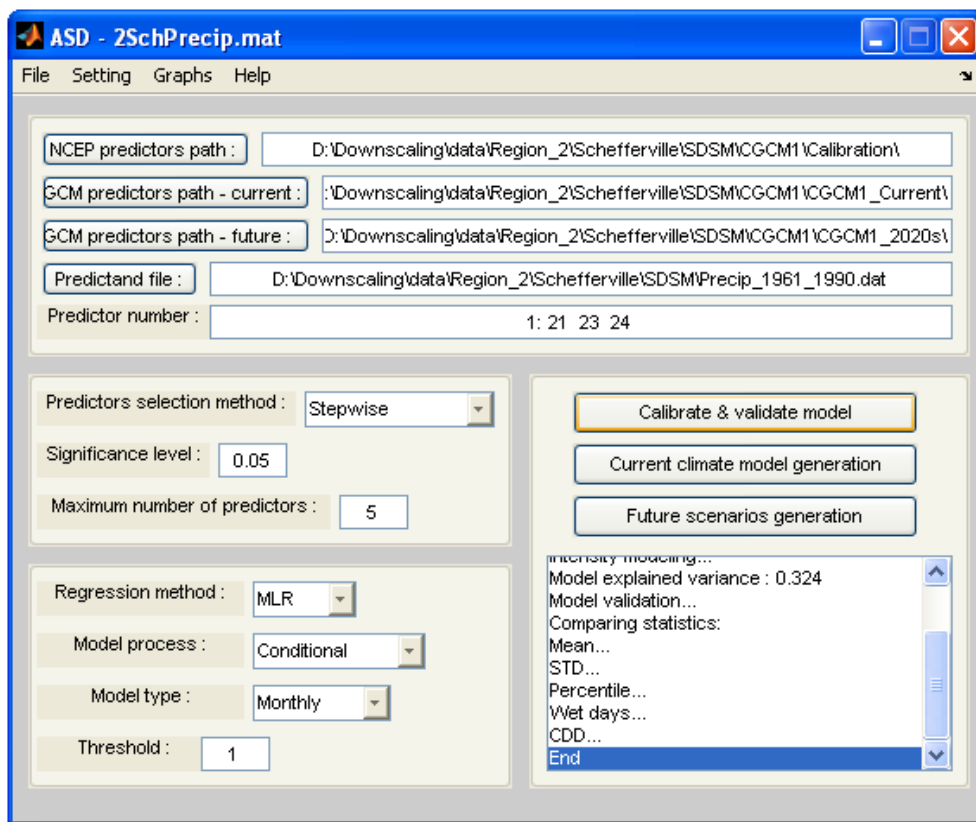
**Figure 2:** The ASD Parameters menu.

4. Click on the **Graphs** menu at the top of the ASD Main menu window and select **Correlation coefficients** from the drop-down menu. The correlations between the predictand and the NCEP predictors will be displayed.
5. Click on the **Graphs** menu at the top of the ASD Main menu window and select **Regression coefficients** from the drop-down menu. The model regression coefficients for the NCEP predictors identified as explaining the most variance in the predictand (up to the maximum number of predictors specified in the ASD Main menu) will be displayed.
6. Click on the **Graphs** menu at the top of the ASD Main menu window and select **Boxplot** or **Barplot** from the drop-down menu. Then click on **NCEP**, **GCM-Current** and **GCM-Future** to compare the statistics for observations/NCEP data, observations/current GCM data and future GCM

<sup>3</sup> For precipitation it is common to apply a transformation to the data to improve the performance of the multiple linear regression since these data tend not to be normally distributed. For temperature, this is not the case and the data are generally used as is. However, if a model is poorly calibrated, examination of the predictand data and of the behaviour of the residuals may suggest the use of a model transformation.

data respectively. Figure 4 shows the observed and simulated mean monthly values of precipitation obtained when using NCEP data. (Other statistical values which can be viewed are listed at the top of this window: standard deviation (STD), values for the selected percentile, percentage of wet days (Wet-day) and maximum number of consecutive dry days (CDD). For temperature variables, the mean, standard deviation, percentile, maximum and minimum values can be viewed.) Figure 5 shows the observed data over the current period (1961-1990) and simulated mean monthly values of precipitation obtained when using GCM predictors and downscaled over a future period (ex. 2011-2040).

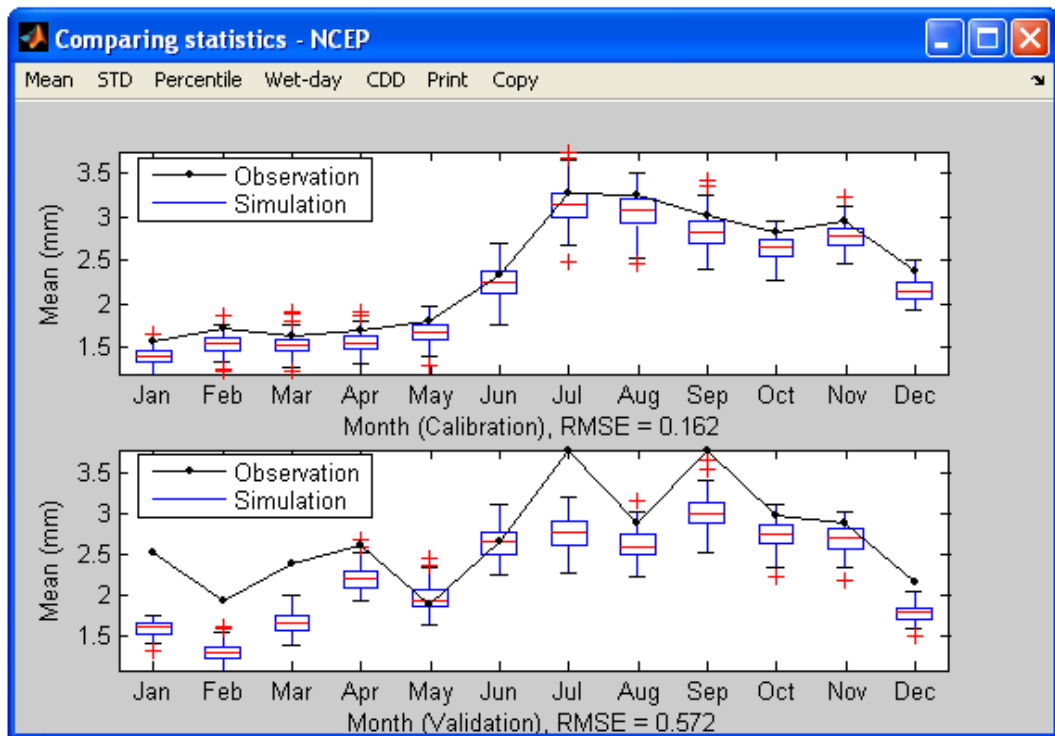
7. Click on the **Graphs** menu at the top of the ASD Main menu window and select **QQ-plot**<sup>4</sup> from the drop-down menu. Then click on **Calibration-NCEP**, **Validation-NCEP**, **Calibration-GCM** and **Validation-GCM** to see the Q-Q-plots between the observations and NCEP data and the Q-Q-plots between observations and GCM data for the calibration and validation periods.



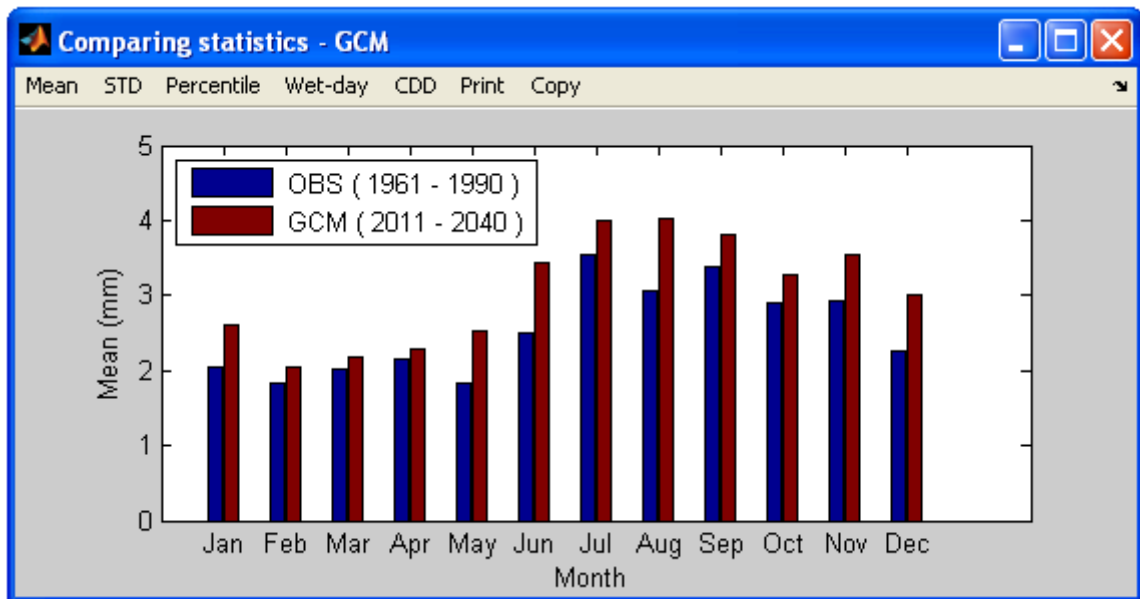
**Figure 3:** Main menu of the ASD tool

---

<sup>4</sup> A quantile-quantile (q-q) plot can be used to determine if two datasets come from populations with a common distribution. In this type of plot, points are formed from the quantiles of the data and if the resulting points lie roughly on a line with a slope of 1, then the distributions are the same.



**Figure 4:** Comparing statistics menu of ASD tool



**Figure 5:** Comparing statistics menu of ASD tool.

### Saving and Retrieving Parameters in ASD

To save the input and output parameters as well as the scenario data generated by the ASD tool, click on the **Save** option in the **File** menu on the ASD Main menu, browse to the desired directory and then type an appropriate filename. The output file extension will be \*.mat. Next time, when ASD starts, you can easily retrieve the selected model parameters by clicking on **Open** in the **File** menu and selecting the file, or by typing **load \*.mat** in the Command Window.

Examination of this \*.mat file also provides the opportunity to view (and save or paste into other software packages) the values of the variables and data generated by ASD. In order to view the values of the outputs, you need to have the MATLAB **Workspace** window displayed – if this is not already visible click on the **Window** menu and select **Workspace** from the drop-down menu, or press CNTL+3. Four variables will be displayed in this window: asdin (input variables), asdout (output variables), asdoutCGCM (output variables using the current period GCM data) and asdoutFGCM (output variables using the future period GCM data). (You can also see this list of variables by typing “who” at the command window prompt.) If you double-click on any of these variables listed in the Workspace window, then the Array Editor opens with a list of the fields associated with each variable. Double-click on the desired field to display the values – these can be printed or you can select and copy them so that they can be pasted into another environment (e.g., Microsoft Excel).

In the event that MATLAB cannot display the values for a particular field (there is a maximum number of elements that MATLAB can display) you can still save this information into an ASCII file. For example, if you wish to save the future scenario output, type asdoutFGCM.Scen at the command window prompt. The data for the number of scenarios generated will be displayed. If you type “who” again, a new variable is displayed – ans. To save this as an ASCII file, type (for example): save scenarios.dat ans –ascii. This file will be saved in the current directory (e.g., c:\Program Files\ASD) unless you specify a different directory path. More information about importing and exporting files can be found in the MATLAB Help facility.

**To quit the ASD tool, select Exit from the File menu, or click the close box.**