Hydro-extension 6.0 User Guide

1. Introduction

Model calibration and data assimilation are the two major issues for hydrology models to be applied in both research and real-time application. We solved these problems with breakthroughs of model performance in developing a new hydrology model framework “hydro-extension” based on WRF-Hydro/National Water Model (NWM). The calibrated high resolution distributed parameters and data assimilation are applied to improve the model accuracy. The event-based calibration and data assimilation are integrated in the new model framework. The long term calibration method is not needed. The coupled atmosphere-water model is available in the latest release, the first coupling system which can demonstrate the reasonable hydrology information.

The company provides the professional service to install and support the real-time system on the user's platform, as well as WRF-Hydro related development. The real-time system includes automatic parameter adjustment/calibration, data assimilation and forecast. The following features can be available in the same real-time system: river flooding, overland flooding, city flooding, water management and related downstream impact, storm surge related impact, water resource estimation and prediction. Before we solve the copyright issues for the newly developed codes, we will only provide the binary installation for limited compilers. There is a one year expiration code, but it is free to renew for users who are working on pure research (thesis, publications, etc), if there is no conflict with our development. For questions, please send an email to “ramiweather@gmail.com”.

2. Highlight of the new features

WRF-Hydro is an integrated modeling system, including a land surface model, terrain routing and channel routing. There are many limitations and uncertainties in connecting these components. As an example, the WRF and WRF-Hydro fully coupled system failed to provide the appropriate streamflow information compared to the WRF-Hydro offline runs. In order to solve those issues, we developed a new model framework, providing a flexible and robust model performance. The hydrology model can be driven by regular WRF-Hydro forcing data, output from any land models, or atmosphere models. The new model framework can be directly applied to the current atmosphere ensemble forecasting system to provide hydrology forecasts. A simple self learning function has been added into the model system to automatically pick up the best distributed parameters, which can be used to model simulation or real time forecasts.

The data assimilation can be applied to not only the water model real time application, but also improving the soil moisture condition for the atmosphere model based on streamflow observations. The feedback of the subgrid impact of the river to the atmosphere model grid is the new feature.

Functions/Package	**Hydro-extension**
Nesting	Coarse grid domain or upper streamflow can provide the boundary to a fine mesh grid domain.
City/Residential area flooding	The new land routing scheme has the capability to forecast overland flooding for city/residential areas.
Water management function	The model has a dynamic reservoir discharge scheme. The future discharge plan and its impact can be simulated. The lakes/reservoirs can be created either manually or through the ARC-GIS tool.
Data assimilation	A new data assimilation scheme, which modifies the streamflow conditions from both upstream and downstream watersheds.
Storm surge	Add the ocean boundary to allow the water moving from ocean/river to land.
Flooding risk estimation	Capable of estimating the flooding risk based on the hydro dynamic process from either real cases or estimated precipitation.
Sediment	Code developed by LSU as a collaboration.
River geometry	The new functions provide the land and river interaction, including the prediction of overbank flooding.
Channel Routing Schemes	Both the diffusive and muskingum methods are available.
Coupled WRF-water model	The river flow information is consistent between the fully coupled run and offline experiment. The river feedback provides additional boundary layer features to the atmosphere model.

3. Quick Start/Installation

3.1 General steps to run the model

1) Create the domain file: geogrid (from WPS), wrfinput Fulldom files.
2) Create the forcing files and prepare the observation data.
3) Map the observation data to the model domain.

4) Run the model. It includes either parameter adjustment, data assimilation or forecast. They are in the same model framework. Here we do not need the long term calibration. The parameter adjustment function helps to get the best model performance.
5) All exe commands require the existence of a “usage.dat” file. Copy it from the run_template directory.

3.2 Installation

3.2.1 Install WRF-WPS

Please refer to the NCAR WRF model user guide for detailed information. Here we provide a simple procedure for experienced users.
As an example, the WRF model will be installed under the $HOME/hydro_extension/WRF directory.

```
cd hydro_extension/WRF/WRFV3
csh clean -a
./configure → choose 15 for intel, 34 for GNU
Then choose 1 (basic)
./compile em_real
(Note: if the MPI software is openmpi, send email for additional instructions.)

cd WPS
csh clean -a
./configure    ---> choose 17 for intel, 1 for GNU
./compile
```

Note: WPS package will be used to generate the domain file (see 4.1). To build WPS, users need to compile the WRF model first, since WPS uses some functions from WRFV3/ directory.

3.2.2 Install hydro-extension packages:

For example, extract the packages at some place. All should be under the “hydro_extension/” directory. And then link “hydro_extension” to $HOME directory so that we have $HOME/hydro_extension/hydro1 and $HOME/hydro_extension/hydro2
Note: hydro1 (and hydro2) and run_template will both be located under the “hydro_extension/” directory.

```
cd $HOME/hydro_extension/hydro1
mkdir ../exec
Edit the file “install.dhydro.csh”,
        configure 2 → for gfortran
        configure 3 → for intel
./install.dhydro.csh
```

The compiled exe files will be copied to $HOME/hydro_extension/exec and run_template.

Repeat above procedure for hydro2

```
cd $HOME/hydro_extension/hydro2
./install.dhydro.csh
```

With the above steps, the installation finished. The exe files are located at $HOME/hydro_extension/exec

In this release, we provide three model exe files. “wrf_hydro.exe”, “dhydro.exe” and “NoahMP_dhydro.exe”.
“wrf_hydro.exe” can be used as the standalone land model and provide the input for “dhydro.exe”. When running the “wrf_hydro.exe”, set “SUBRTSWCRT = 0, OVRTSWCRT =0, CHANRTSWCRT =0” from “hydro.namelist”. The output of surface runoff can be used to drive “dhydro.exe”.

“dhydro.exe” includes the hydrology part: land routing and channel river part. It has the feature to use the distributed parameters and do the event-based parameter adjustment.
“NoahMP_dhydro.exe” includes the NoahMP land model, land routing and channel routing. It has the feature to use the distributed parameters and do the process based parameter adjustment. It combines “wrf_hydro.exe” and “dhydro.exe”.

3.3 Offline testing case

“run_template/ “ provides a testing case to run the model.
An example of the running command is : `mpirun –np 2 ./dhydro.exe`
or

```
 mpirun -np 2 ./wrf_hydro.exe
```

or

```
mpirun -np 2 ./NoahMP_dhydro.exe
```

 wrf_hydro.exe: uses the namelist files: “hydro.namelist” and “namelist.hrldas”. It is the compatible model framework as WRF-Hydro.
 dhydro.exe: uses the namelist files: “hydro.namelist” and “namelist.dhydro”. It is the new model framework.
NoahMP_dhydro.exe uses the namelist files: “hydro.namelist”, “namelist.dhydro” and “namelist.hrldas”. The start date is from “namelist.hrldas.”
Here “dhydro.exe” is the basic framework with parameter adjustment.
“NoahMP_dhydro.exe” combines the NoahMP land model with the new framework for the parameter adjustment.
Note: When parameter adjustment applied, “NoahMP_dhydro.exe” may have problems on the land model side.

3.4 How to select the binary installation package

Run the command “mpif90 --version” or “gcc --version”, it will tell which GNU version installed on your computer.

3.5 How to generate the model domain

1) Using WPS procedure to create the domain file and name it as “DOMAIN/geo_finemesh.nc” (see section 4.1)
2) using ARC-GIS file to extract the river information and name it as “DOMAIN/Fulldom_finemesh.nc” (see WRF-Hydro related document).
3) The “DOMAIN/wrfinput_d01” will be created by running the command “cwrfInput.exe” (see section 4.1).
   Note: all required land model informations (including the vegetation type information) are included in “geo_finemesh.nc”, “wrfinput_d01” and *.TBL files.

3.6 Mapping the gauge observation to model domain

The script “dhydro_update_Fulldom.csh” provides a quick configuration. The script has a self-described document.

1) Set the correct land routing grid resolution (in meters) and aggregation factor (the resolution ratio between land grid and land routing grid) from both “hydro.namelist” and “dhydro_update_Fulldom.csh”.
2) Edit the “list.txt” to input all observation data. If the “list.txt” file is missing, the mask will be generated for the whole domain. See the format of “list.txt” in the following section.
3) Run the command “./dhydro_update_Fulldom.csh”, it will generate the template netcdf file and update “DOMAIN/Fulldom_finemesh.nc”. This command only needs to run one time.
4) At the end of the script “./dhydro_update_Fulldom.csh” , the command “river_geo.exe” has been commented out. The user needs to manually run this command to create the river geometry file. This is an optional feature, when considering the river status for overbank flow or land to river discharge.

3.7 Run the model/tune the parameter

1) After running the command “./dhydro_update_Fulldom.csh”, it is ready to tune the parameters based on the observation data. “station_list.txt” was created with the model id there (the third column), this will be used in the observation data (latitude longitude model_id obs date). See the following about the format.
2) The “namelist.hrldas” will be used to run “wrf_hydro.exe”. When selected the parameters, “SUBRTSWCRT=0”, “OVRTSWCRT=0”, “CHANRTSWCRT=0”, the “wrf_hydro.exe” will fall back to NoahMP land model only. The output file “*.forcing.nc” will be used to drive “dhydro.exe” directly. If the user wants to use “*.LDASOUT_DOMAIN1”, the user needs to

rename them as "*.forcing.nc". To run "wrf_hydro.exe", the suggested "REFKDT_DATA" from the file "GENPARM.TBL" is around 0.8 (recommended to be 0.6 to 1.5). After the parameters have been tuned, the user can use "dhydro.exe" or "wrf_hydro.exe" to do simulation or forecast, the results are consistent.

3) "namelist.dhydro" is used for "dhydro.exe" and "NoahMP_dhydro.exe". When "KHOUR" is negative, it will use the "end_date", otherwise it will not use "end_date" for the forecast length. When "dabiaLoop" is negative, it runs the model and outputs the pair of the model and observation under the folder "pair_out". When "dabiaLoop" is 0, no observation pairs are output. When ""dabiaLoop" is positive, it is the maximum iteration for the parameter adjustment, which runs the model repeatedly to update the parameters so that the model fits the observation.
4) To get the time series for a station, use the command such as "grep model_id *.iteration_time.pair.txt". For example, the file "pair_out/201701041445.00001.06.pair.txt", "06" is the output at the iteration "6". "00001" is the output from process 1.
5) Stations for the parameter adjustment or pair output are from "station_list.txt". If the user does not want the information for the specific stations, the user can delete the related stations from "station_list.txt".

4. Detailed procedures/functions

4.1 Create the geogrid file, wrfinput and Fulldom file

To run the WRF-Hydro or hydro-extension, the user needs to create the geogrid (from WPS), wrfinput (from WRF, or the tool we provided here), Fulldom files (from ARC-GIS) and forcing data. WRF-Hydro release document provides the instruction to manually create those files step by step.

To create a geogrid file, refer to WRF user guide.

1) Compile the code to generate the "geogrid.exe".
2) Download the land surface type data "WPS_GEOG"
3) Under the WPS_exe directory, edit the "namelist.wps" to set the correct domain: domain size, central latitude and longitude information.
4) Run ./geogrid.exe command, it will generate the "geo_em.d01.nc" file upon successful finish.

To create a Fulldom file from ARC-GIS, refer to WRF-Hydro user guide.

To create a wrfinput, the user can run the command: "./cwrfInput.exe geo_em.d01.nc wrfinput_d01" instead of running real.exe from WRF.

4.2 Create forcing data

4.2.1 Create the forcing data for wrf_hydro.exe or NoahMP_dhydro.exe

WRF-HYDRO and hydro-extension require the forcing data interpolated into the geogrid domain (land grid). WRF-Hydro document has a description and template about how to use the "ncl esmf regrid" function for data interpolation from one domain to another. To manually create the forcing data to run WRF-Hydro , please refer to the WRF-Hydro document.

4.2.2 Create the forcing data for dhydro.exe

The forcing data for hydro-extension can be: surface runoff, accumulated surface runoff, precipitation rate, or the hybrid mode (some watersheds use surface runoff, and some watersheds use precipitation rate).

1) Manually create the forcing data to run hydro-extension.
   The forcing data file name format: "yyyymmddhhMM.forcing.nc". It is on the land grid.
   Where yyyy is year, mm is month, dd is day, MM is minutes.
   The variable names are : SFCRNOFF, UGDRNOFF or RAINRATE (case sensitive).
   SFCRNOFF is accumulated surface runoff from the land model (mm).
   UGDRNOFF is accumulated subsurface runoff from land model (mm).
   RAINRATE is the precipitation rate (mm/s).

Those variables can be hourly, 30 minutes accumulated variables, or accumulated from model start time. The namelist option will tell different forcing types.

2) The forcing data used for WRF-Hydro or community released NoahMP can be used directly by renaming as "yyyymmddhhMM.forcing.nc", if use the precipitation as forcing variable,
3) Users can run "./wrf_hydro.exe " as offline NoahMP, and it will automatically create the surface runoff variables with the file names as "yyyymmddhhMM.forcing.nc", which can be used to drive dhydro.exe. To run "wrf_hydro.exe" as offline NoahMP, user needs to set the following variables to be zero from "hydro.namelist": SUBRTSWCRT, OVRTSWCRT, CHANRTSWCRT, GWBASESWCRT
4) The output from other land model/atmosphere models needs to be regrid to the correct domain. The variable name (either runoff or precipitation name, or both) and file name will follow what is addressed above.
5) Above namelist file and parameters can be used to run "NoahMP_dhydro.exe". The "factor.nc" has different meanings for "NoahMP_dhydro.exe" and "dhydro.exe". It needs to be tuned separately. The rest of the parameters and files are the same. In "NoahMP_dhydro.exe", the "factor.nc" is used to tune the parameter in the NoahMP land surface model. In "dhydro.exe", the "factor.nc" is used as a weight for the precipitation or surface runoff.

4.3. Prepare the streamflow observation data.

File name format: yyyymmddhhMM.obs.txt. For example: 201901170845.obs.txt
There are two formats:

1) latitude longitude stationID obs_value date

When “addsOnSwitch=1”, it will use this format.
For example: 38.64184848 121.3827302 9 0.02973264 201901170845
The model will look for the date/timing information based on the file format. All stations at the same time will be saved to the same file. The station ID (9 in this case) will be the id inside the model. See section 4.11 about how to get this station ID from the “find.exe” command.
The observation file directory is defined from “hydro.namelist”: obsDir

2) latitude longitude stationID obs_value date order1 order2

When “addsOnSwitch=2”, it will use this format.
Order1 and order2 are the minimum and maximum of the estimated river order. Those two columns will be used when stationID is negative. If Order1 is negative, the model will use the obs_value to estimate the order range. If order1 is positive and order2 is negative, the order2 will be 10 from the model.

4.4. Run the model

“Hydro-extension” provides the capability to run the model with the distributed parameters, which are tuned based on observations. A set of parameters need to be created for tuning purposes.

There are three running interfaces: wrf_hydro.exe, hydro.exe and NoahMP_dhydro.exe.

1) wrf_hydro.exe : Using NoahMP to drive the hydro_extension. The forcing data used for WRF-Hydro can be used directly. When “OVRTSWCRT, SUBRTSWCRT, CHANRTSWCRT and GWBASESWCRT” from “hydro.namelist” are set to be zero, the wrf_hydro.exe will fall back to land surface model only. It will generate the forcing data to drive “dhydro.exe”.
2) dhydro.exe: The new interface of hydro-extension, used for either running the water model or tuning the parameters.
3) NoahMP_dhydro.exe: combine the above two together.

4.5 Create KML file for google earth

The program “hydro2kml.exe” is a tool to dump out the channel node information into a kml file so that users can use google earth to find the node information for specific locations, such as lakes or reservoirs. The latitude and longitude information for lakes/reservoirs or gauges may not be correct/matched with GIS. This tool can help.

```
./hydro2kml.exe yyyymmddhh00.CHRTOUT_DOMAIN out.kml  order seg_info.nc station_list.txt
```

order: is the stream network order. The nodes will not be printed out if its order is smaller than this value.

yyyymmddhh00.CHRTOUT_DOMAIN: Created by running model. Any of the files should be fine.

out.kml: The output file, to be displayed with google earth.

seg_info.nc : created from the DA procedure.

station_list.txt: created from DA procedure. The example of the format is:

38.64184848 -121.3827302 101 ……..

4.6 Nesting

Hydro-extension provided two nesting options for the nested domain. In real time application, users can have the coarse resolution running on a large domain, providing the inflow boundary condition for a fine mesh grid. When data assimilation is turned on, the boundary input for large rivers will be more accurate. Right now, we only consider the boundary river inflow. The terrain routing input on coarse resolution can be ignored.

a. Create the coarse domain and do the forecast first.
We need this to create the boundary data as fine mesh domain boundary flow input.

b. Create the mapping file for fine mesh grid
Modify the namelist.nesting file and run the command: ./mapping.exe
The output file is "mapping.nc"
As long as the domains (coarse and fine mesh) do not change, this only needs to run one time.

c. Create the time series boundary input data for nested domains based on the coarse domain model output.

```
./nesting.exe yyyymmddhh00.CHRTOUT_DOMAIN0 yyyymmddhh_bdy.nc
```

Here yyyymmddhh00.CHRTOUT_DOMAIN0 is the output from the coarse grid. yyyymmddhh_bdy.nc is the created boundary condition. Users need to use this naming convention so that the model can read in the data.

d. Run the model
Edit the namelist file "hydro.namelist" for "HYDRO_nesting" sector.
bdyDt is the time interval for boundary input.
inDir is the directory saving *_bdy.nc files.
Example of running the model: mpirun -np 4 ./wrf_hydro.exe

Another choice is to write the boundary input data as observation data format (yyyymmddhhMM.obs.txt), and use the data assimilation method (addsOnSwitch=1 or 2, prefers 2) to add the boundary input into the model.

4.7 Create the reservoir

Refer to the section of Lake/Reservoir data assimilation
This part is the most complicated part. Users can use this to 1) create the watershed, 2) find the segment among reservoirs, 3) create the new reservoir or lakes, 4) create the bucket for wrf-hydro. The user can do all or some of them.

Description of the basic procedures:

i. Create the Fulldom file from arc-gis first.
ii. Run the model to create the files *.CHRTOUT_DOMAIN. One time level file is enough.
iii. Run the model with the option of “outStatic=1” from hydro.namelist. It will create the file “staticData.txt”.
iv. If the user wants to create the new reservoir, now it is ready to update the newly created Fulldom file.

Create the selected watershed

Use the script “dhydro_update_Fulldom.csh” from 4.4.2 b section to update the Fulldom file.

Create the reservoir

```
./mk_reservoir.exe reservoir_node height cellArea(m^2) Fulldom.nc
```

Here, height is the reservoir bank relative to the height of the dam.
dst is the node id, output from running “find.exe”, fort.66. The reservoir starts from this node, and calculates upstream.
Reservoirs need to be added one by one. More specifically, the user first runs find.exe for one reservoir to find the node number (dst) in the Fulldom.nc, and then uses mk_reservoir.exe to create one.

Reservoir management

This is controlled by the available discharge data from the reservoir. The control is in the “WATERMANAGEMENT” sector. The reservoir obs data is defined under “inDir” directory. Here is the name convention of the reservoir obs data: yyyymmddhhMM.obs.txt, for example: 201809030600.lakeobs.txt
The format of this file is:
K, H, Discharge
Here K is the lake label: 1, 2, ….
H is the lake/reservoir water depth
Discharge is the lake/reservoir discharge.

4.8 City Flooding

In city and residential areas, the soil infiltration and surface discharge are different from rural areas. Hydro-extension has bug fixes to accurately describe the overland and river flooding over residential and city areas.

4.9 Coastal area/Storm surge inundation feature

The ocean boundary has been added into the model. The code has the hook to simulate the impact of storm surge or ocean wave related land inundation. Here it is the tool to process the observed or forecasted wave height time series data.

4.9.1 Create the wave height forcing data

1) Create the wave height forcing data from single or multiple ocean observations. Under run_template directory, run the command:

```
./processWaveH.exe DOMAIN/geo_finemesh.nc DOMAIN/Fulldom_finemesh.nc obs.txt yyyymmddhh00.waveheight.nc 1
```

Here 1 is the aggregation factor between the land grid and terrain routing grid.
yyyymmddhh00.waveheight.nc is the created wave height data at the routing grid.
Users need to use this naming convention so that the model can read in the data.
For example, 201810011200.waveheight.nc is for Oct 1, 2018, 12Z, 00 minutes.
The format of the wave height file (obs.txt) is as:
latitude(degree) longitude(degree) wave_height(meters)

2) Users can create a series of forcing data yyyymmddhh00.waveheight.nc from the ocean model output.

4.9.2 Run the model

Edit the hydro.namelist file, for the "HYDRO_wave" sector.
inDir is the directory saved the files of *.waveheight.nc.
forcDt is the time interval (seconds) for wave height input.
For example, if forcDT=3600, the user needs to create the yyyymmddhh00.waveheight.nc every hour.
Example of running model command: `mpirun -np 4 ./NoahMP_dhydro.exe`

4.10 Data Assimilation/Calibration

The data assimilation is required for the real-time application. "Hydro-extension" combines three parts/functions for the data assimilation. Each part can run separately. The design of the model framework considers the nowcasting feature, in which all the available observations will be applied to provide the best model results.

1) The fall back running is driven with the observed precipitation. The streamflow observation and distributed parameters are not used with "DABE=0" from "hydro.namelist". This is also called the spin-up process.
2) The fall back running is driven with the observed precipitation, including the parameter adjustment for upstream river grids. The streamflow observation is used to parameter

adjustment for upstream watersheds. With “DABE=1” from “hydro.namelist”. “diaLoop > 0” from “namelist.dhydro”. This part builds a connection between the surface runoff and the related streamflow observations and can be applied to land surface model data assimilation or land surface model calibration. The parameter adjustment only happens to the stations defined from “obsAdjFile”.

3) The fall back running is driven with the observed precipitation, including the data assimilation for downstream river grids. The streamflow observation is used to modify the downstream river status. With “DABE=1” and “addsOnSwitch=1 or 2” from “hydro.namelist”, diaLoop<=0 from “namelist.dhydro”. When “addsOnSwitch > 0”, the observed streamflow discharge will be applied to affect the downstream river flow.

4.11 Create/tune the distributed parameters

Hydro-extension is based on the distributed parameters defined on the land routing grid. The observed stations are mapped to these high resolution grids. For this purpose, we create the “seg_info.nc” file (The template is created from running the model with outStatic=1.) and update the original Fulldom file (created from ARC-GIS). The watershed for each station is defined as a segment and the related parameters are adjusted separately.

a. Create the distributed parameters

1) General procedure

This procedure will overwrite the Fulldom file. It is better for users to save the original Fulldom file first. Create a file named as “list.txt”, with the following format: (station lat, lon, id)

```
34.2322397 -89.9141667 1
34.3322397 -88.9141667 2
 …………...
```

Here 1 is the station id, from 1 to n according to the sequence. The output "station_list.txt" will be used to create the mapping file. The user can add all of the observation stations in this file. If there is no observation, the user needs to find a point which is inside the watershed.

In the file "station_list.txt", the first two columns are the lat-lon of the river node inside the model, consistent with that from the Fulldom file. The third column is the global Id of the river node, the index of the point data output from “*.CHRTOUT_DOMAIN1”. The fourth column is the order of the river node defined in the model. The seventh and eighth columns are the lat-lon information of the gauge location defined.

After creating “list.txt”, use the scripts from run_template/dhydro_update_Fulldom.csh. Before running the command “dhydro_update_Fulldom.csh”, remember to review and change the aggregation factor and land routing domain resolution. If the user needs to add additional observation data later, the user needs the following steps. Otherwise, it is ready to run the model.

2) More features for “find.exe”.

```
./find.exe orderIn line_number_of_staticData.txt
```

Here: orderIn is the threshold value for the output of a CSV file.
List.txt:
28.178665, 101.629141 ----> station list, can be obs. The first one is the station of the most downstream one.
The “staticData.txt” is created when run (wrf_hydro.exe, or dhydro.exe, or NoahMP_dhydro) when set outStatic=1 from “hydro.namelist”. When “outStatic=1”, it will create the template data (seg_info.nc, factor.nc) for parameter adjustment and the river node logic information (staticData.txt).

After running the find.exe command, it will create the CSV files for the selected watershed/segment or the whole domain. You can use these CSV files to create the bucket as descriptive in “dhydro_update_Fulldom.csh”. When you run the hydrology model, you can either use this bucket information or bias correction to take care of the subsurface runoff.

The full format of the list.txt is:
Lat, Lon,sId, order1, order2
Lat and lon are the latitude and longitude of the stations.
sId is the id of the station, such as the station id (15 integer) from the USGS station. If there is no such value, the user can give the integer index from 1 to N. order1 is the estimated minimum river node order for this station. Order2 is the estimated maximum rivier node order for this station.
When orderIn > 0, read the first two columns, and look for the river nodes, whose order is >= orderIn.
When -10< orderIn <0, use the order1 to find the river node. Only list the station, whose river node order >= abs(orderIn).
When orderIn < -10, use the order1 and order2 to find the river node, and only list the station whose river node order >= abs(orderIn) -10.

b. ***<u>Add the additional observation data</u>***

The following procedures are for later modification based on the modified Fulldom file from above step. It provides the convenience if the new station needs to be added later. Note: if no “seg_info.nc” exists, users can create the template file by running wrf_hydro.exe/dhydro.exe with “outStaic=1” from “hydro.namelist”.

1) Append new station information to the “list.txt” file.
   If the station has already been registered in “seg_info.nc”, it will be automatically jumped when running the commands: find.exe, updateSeg.exe, createSegMask.exe.
2) running command: find.exe order_thresh_hold maximum_node, aggFactor, geo_f
   order_thresh_hold: can be 0.
   maximum_node: Total node from Fulldom file (or the total lines of staticData.txt: wc -l staticData.txt).

aggFactor: the aggregation factor between the land grid (geogrid file) and landRouting grid (Fulldom file).

geo_f : geogrid file

3) update information to "seg_info.nc" file

```
./updateSeg.exe DOMAIN/geo_finemesh.nc  DOMAIN/Fulldom_finemesh.nc $dxrt -99
```

Here dxrt is the horizontal resolution of terrain routing (meters)

This command will add new stations from "station_list.txt" to the list inside "seg_info.nc". If the station existed in the list, it would do nothing. This command only needs to run one time, unless having new observation stations, which need to be added later.

The default input file is "station_list.txt".

The format of the "station_list.txt" is:

Latitude longitude station_id -9999 ii jj

"ii" is the index of (west-east) direction and jj is the index of y (north-south) direction. When those two numbers are provided, the utility program will use them to find the location of the station instead of the latitude and longitude information. The staion_id is an integer from the output. User needs to update it later, since the observation files need to have the correct value.

4) Create the segment mask

```
./createSegMask.exe $HOME/hydro_extension/lsu_exe/DOMAIN/Fulldom_finemesh.nc $aggFact
```

Here aggFact is the aggregation factor between land model grid and terrain routing grid.

Note: Above steps set the parameters for the selected domain and only need to run one time.

5) Assign "station_list.txt" to "obsAdjFile" from "hydro.namelist".

The parameter adjustment only happens to the stations defined from "obsAdjFile". If the user does not want to modify the parameters for specific stations, drop them from the "obsAdjFile".

c. Tune the distributed parameters.

The parameter adjustment is integrated in the hydro-extension model framework so that it can run in parallel. In order to get the best model performance, hydro-extension created this feature so that the model will automatically adjust the parameters. For example, change the "dialoop=6" from "namelist.dhydro". Change "DABE=1" "from hydro.namelist".

Run the model, for example: `mpirun -np 2 ./dhydro.exe`

Check the output from "pair_out/". For example, the last output is "06". `"cat *.06.pair.txt`
The format is : "date, model, obs, station id, station id, id" . Here "06" is the interaction at 6 times. If the model does not match very well with the observations, repeat the command "mpirun -np 2 ./dhydro.exe".

When "dialoop >1", it will iterate until 1) the model output and observations are consistent; or 2) the maximum iteration (dialoop) reached. The parameters in the file "seg_info.nc" and "factor.nc" will be updated based on the automatic adjustment procedure.

When dialoop = 0, it will do a forecast without parameter adjustment.

When dialoop < 0, it will do a model simulation without parameter adjustment and output the pairs between the model and observation for all observed stations.

In this release, "NoahMP_dhydro.exe" has the same function to tune the parameter. "dhydro.exe" will be used, if we do not want to use the NoahMP land model.

4.12 Create the averaged parameters for the rest of the domains.

The parameters are tuned based on limited gauge stations. For the watershed without the gauges, it can be updated with the averaged parameters.

```
./averageParm.exe station_list.txt
```

The default modified files: seg_info.nc and factor.nc.

It uses the stations, liststed from station_list.txt to calculate the averaged parameters, and assign the parameters to the rest of the river grids.

5. Lake/Reservoir data assimilation and parameter adjustment

There are three parts of possible lake/reservoir observation data. The lake/reservoir input, lake/reservoir water height and lake/reservoir discharge. The lake/reservoir input can be used to tune the upstream parameters. The lake/reservoir water height and discharge can be used for both hindercast and future forecast. The discharge can be either observation or a future discharge plan. Furthermore, if no lake discharge observation is available, the related parameters can be adjusted based on the downstream observations.

5.1 Create the lake/reservoir

This part is the most complicated part. Users can use this to 1) create the watershed, 2) find the segment among reservoirs, 3) create the new reservoir or lakes, 4) create the bucket for wrf-hydro. The user can do all or some of them.

Description of the basic procedures:

v. Create the Fulldom file from arc-gis first.

vi. Run the model to create the files *.CHRTOUT_DOMAIN. One time level file is enough.

vii. Run the model with the option of “outStatic=1” from hydro.namelist. It will create the file “staticData.txt”.

viii. If the user wants to create the new reservoir, now it is ready to update the newly created Fulldom file.

Create the selected watershed

Use the script “dhydro_update_Fulldom.csh” from 4.4.2 b section to update the Fulldom file. This procedure adds the factors (in the fine grid resolution) to Fulldom files so that they can be used in data assimilation or parameter adjustment.

Create the reservoir

```
./mk_reservoir.exe dst height cellArea(m^2) Fulldom.nc
```

Here, height is the reservoir bank relative height.
dst is the node id, output from running “find.exe”, fort.66. The reservoir starts from this node, and calculates upstream.
Reservoirs need to be added one by one. More specifically, the user runs find.exe for one reservoir to find the node number (dst) in the Fulldom.nc first, and then uses mk_reservoir.exe to create one.

Reservoir management

This is controlled by the available discharge data from the reservoir. The control is in the “WATERMANAGEMENT” sector. The reservoir obs data is defined under “inDir” directory. Here is the name convention of the reservoir obs data: yyyymmddhhMM.obs.txt, for example:
201809030600.lakeobs.txt
The format of this file is:
K, H, Discharge
Here K is the lake label: 1, 2, ….
H is the lake/reservoir water depth
Discharge is the lake/reservoir discharge.

./mk_reservoir.exe can be repeated for all available reservoirs/lakes. When no more lakes/reservoirs are created, users need to follow section 4.4.2 to create new parameter templates, since the Fulldom file has been changed.

5.2 Use lake/reservoir input data

The lake_list.txt provides the model id and lake id. Users will use the model id for the lake input observation, and write it as the format of regular streamflow observations.

Latitude longitude model_id (-1)*lake_id i_location j_location latitude longitude lake_id

5.3 Use lake/reservoir discharge

This is controlled by the available discharge data from the reservoir. The control “WATERMANAGEMENT” sector. The reservoir obs data is defined under “inDir” directory. Here is the name convention of the reservoir obs data: yyyymmddhhMM.obs.txt, for example:

201809030600.lakeobs.txt
The format of this file is:
K, H, Discharge
Here K is the lake label: 1, 2, ….
H is the lake/reservoir water depth (in meter)
Discharge is the lake/reservoir discharge (m^3/s).

5.4 Use the downstream observation

The lake discharge parameters can be adjusted based on the downstream observations.
obsLakeWeig=0…. 1. 0: no impact to lake discharge.
For example, obsLakeWeig=0.8. The 80% difference at the downstream observation station is caused by the lake discharge. The related lake discharge parameters will be adjusted based on this estimation.

5.5 Transfer the lake netcdf file to ascii format

The ARC-GIS tool from WRF-Hydro created the netcdf format for the lake parameters. We prefer to use the TBL file so that users can manually modify the related parameters.

```
./lakenc2txt.exe LAKEPARAMETER.nc
```

The output will be a TBL file.

6. River geometry data

With the stream geometry, the water model can provide additional features: the interaction between the land and river, flooding risk (flooding area mask during the forecast period), and the feedback of the river (soil moisture) to the atmosphere/land model.
To create the stream geometry file over the model domain, the user needs to set up the model domain first, and then use the command "./river_geo.exe" to create the river geometry, matched with the Fulldom river network. It will take some time to run, since it will look through the big database. The original shape data can be downloaded from http://gaia.geosci.unc.edu/rivers/. The script "shape2nc.ncl" converts the shape data into netcdf format. The "shape2nc.ncl" is under the hydro/utils/drv/ncl/ directory.

Example to run the command: ./river_geo.exe narivs.shp.nc DOMAIN/Fulldom_finemesh.nc DOMAIN/geo_finemesh.nc out.nc 100 seg_info.nc
"River_geo.exe" will use the stream geometry information from narivs.shp.nc and match it to the model domain. "Narivs.shp.nc" is the shape data for North America, created from "shape2nc.ncl". If the domain is between the two "narivs.shp.nc", such as North America and Central America, run the river_geo.exe twice with two different "narivs.shp.nc". Here, if the out.nc existed, it will pick up the old running information. Here seg_info.nc is optional and will check all rivers to fill the missing values. You should have this option at the last run if you need to run "river_geo.exe" more than one time.

7. Map projection

The map projections used here are: LC(1), ME(2), CE(3), ST(4).
The number "1 to 4" is defined as MAP_PROJ from the WPS created geo_grid file.

8. Bias correction

There are some extra water resources, such as the input of spring and underground water. Those water resources are not included from the forcing data or land model. To mitigate the related impacts, hydro-extension provided a function as "bias correction". Assume those water resources do not change too frequently, a bias can be applied to the related watershed. For research, we prefer users to update manually and the related tool has been provided.

```
bias2Model.exe textFile  biasNc  segFile
```

Only update the bias of the watershed from the selected stations in the "textFile".
The format of "textFile": obs_model_Id, bias_value, Opt(0:new, 1:combine with old bias)
"obs_model_Id" is the obs ID used in streamflow observation.
"biasNc": a netcdf file, "dabe.nc" created with seg_info.nc together when using "outStatic =1" from hydro.namelist file. Here we will update the related bias value.
segFile: seg_info.nc, which has already been updated with the observation information.

9. Baseflow

There are three options to set the baseflow with hydro-extension.

1) Use the bucket model as described from WRF-Hydro.
2) Use the weighted subsurface runoff. Reading the surface runoff as forcing data. The weight is set from "namelist.dhydro".
3) Use the bias correction function, described as above.

10. Selected namelist variables

1) **sys_cpl:**
   1: offline run with either Noah or NoahMP. Noah is a sequential code, NoahMP is parallelized.
   2: Coupled with the WRF model.

2) **dhydroCplOpt** : from hydro-extension, about the feedback of the surface head to land model.

0: no surface head feedback to the land grid.
1: all of the surface head will feed back to the land grid for re-infiltration.
2: partial surface head feedback to the land grid for re-infiltration.

3) **cplTime :** from hydro-extension, time interval for coupling feedback in seconds. Should be consistent during the offline run and coupled run (with WRF-Hydro or WRF) when the land routing is on.

4) CHRTOUT_DOMAIN
Always choose 1 for the output of river flow information.

5) rt_option : overland routing option.
Suggest to choose 1. It has a bug fix.

6) **GWBASESWCRT** : from WRF-Hydro, switch to activate baseflow bucket model
0=none, 1=exp. bucket, 2=pass-through.
There are three other options in hydro-extension instead of using the bucket model: data assimilation, system bias correction, and weighted subsurface runoff.

7) **outStatic** : from hydro-extension.
0: none.
1: output the template of the distributed parameters and then stop.

8) **Overbank** : from hydro-extension.

9) **domain_msk_opt** : from hydro-extension. When it is on (=1), the land routing will be off outside the watershed.

10) **DABE** : from hydro-extension. Turn on (=1) to use the distributed parameters and data assimilation.

11) **factorIn** : from hydro-extension. Weights for overland runoff or rainfall. Sugge choose 1, and use the model to automatically adjust the parameters.

12) **baseflowFactor** : from hydro-extension. The weighted subsurface can be dumped into the river network directly.

13) **minSlope** : from hydro-extension. The minimum difference of the terrain gradient between two river nodes. Default is 0.002. Suggest to use 0.002.

14) **maxSlope** : from hydro-extension. The maximum terrain gradient between two nodes. Suggest to use 3.

15) **minDTCT0** : from hydro-extension. The minimum time step for channel routing, when instability happens. When using the grid-routing option, the timestep will be automatically reduced when instability happens. It will take a longer time for the model to run.

16) corOption: 1: use correlation for phase adjustment.
Suggested for realtime application. For a research paper with a single wave, suggest to use corOption = 0 and it will have a better phase alignment.

17) watBalanceVal: when watBalanceOpt >0, watBalanceVal will be a threshold. For a station, the averaged input of the whole segment (watershed) is bigger or equal to watBalanceVal (unit: m), it will do the parameter adjustment. Otherwise, not. When the precipitation and gauge observed streamflow are not matched, or in other words, when the precipitation is wrong and could not generate the surface runoff, the parameter adjustment could have problems. This feature will prevent this situation from happening.

11. New Added Important Variables

accmaxsfchead : The maximum land surface ponded water for each grid cell during the forecast time.
maxsfchead : The maximum land surface ponded water for each output time interval.
acc_streamflowMax: the maximum streamflow at each river grid during the forecast time.
streamflowMax: the maximum streamflow at each river grid for each output time interval.

12. Fully coupled with WRF

The river flow information from the coupled atmosphere-water model and offline water model is consistent. The impacts of the feedback from the water model to the atmosphere model are clearly defined. First, the subsurface routing feedback is a long term process. The impact is not clear. Second, we assume that there is a bias in the generation of the surface runoff from the atmosphere. The feedback from the surface routing is mainly to correct this bias and the related distributed parameters are derived based on the streamflow gauge observations. As far as the impact of the water redistribution over the land surface is not clear and needs further study. Third, the sub-grid river information does not exist in the atmosphere model. In the new water model, we introduced the river geometry information, the subgrid river feedback will affect the soil moisture in the atmosphere model. This can be set by “river2land=1” from “hydro.namelist”. The inundation caused soil moisture change is also included.

1) Configure the WRF model first, with WRF_HYDRO=1 set from the environment.
2) Untar the hydro-extension binary package and use “hydro1/” to replace the directory: WRFV3/hydro/. Under the hydro/ directory, run the “./install_dhydro.sh”.
3) Compile the wrf model: ./compile em_real
   It will create the wrf.exe after the compiling finishes successfully.
4) Copy the files from offline run of wrf_hydro.exe or hydro.exe to the same directory for wrf to run. Do not overwrite the files, which WRF already has.
5) Run the “wrf.exe”, which includes the fully coupled function with the hydrology part. It will use the namelist.input for WRF and hydro.namelist for the hydrology part.
6) The “factor.nc” is from the parameter adjustment using “dhydro.exe”. The “NoahMP_dhydro.exe” generated factor.nc is used by the offline NoahMP land model only.

13. Set up hybrid forcing

No need to use “NoahMP_dhydro.exe”. The “factor.nc” has already considered the situation that hybrid forcing is needed.

1) Create the template file.
   Run the wrf_hydro.exe or dhydro.exe without “factor.nc”, when set DABE=1. It will create the “factor.nc” file with the default value to be 1.
2) Run the command “geo2urbanMask.exe”. It will use the urban type from the geo_grid file and create the value of “0” for the variable “forcingMask”.
   When choosing the forcing type 13, the value of forcingMask 1: using runoff; 0: using rain rate.
3) Run the command “updateUrbFactor.exe” to update the value of the factor for urban value. For example, if the factor is 0.3, it means the 30% of the runoff value or precipitation value will be applied in land routing and channel routing process, and the rest (70%) of the value will be removed or applied to soil moisture directly (when coupled with land model or atmosphere model).

14. Derive averaged parameters

In a real-time system, the user may want to keep the parameter to be adjusted all the time and also avoid the possible sharp change of the adjusted parameters. Here we provide a function for this purpose.
New_av = (old_av * N +new_val)/(N+1)
The parameter can be updated iteratively with the periodic running of the parameter adjustment.
For example: ./update_statis.exe factor.nc bias.nc seg_info.nc 1 0 1
It will update the information for factor.nc and seg_info.nc. “1” means update, “0” means not update. It will check the current value and the averaged one. If different, it will use the above formula to calculate the new average and the current one will be set as the new average also.