Developing a Coupling Model System of Global Rainfall Data and Open-Source Model for Water Forecast in Poorly Gauged Basins

Duong Du BUI,∗, Duc Minh TRAN, Huong Thi VU, and Nuong Thi BUI

1National Center for Water Resources Planning and Investigation, Hanoi, Vietnam
2Vietnam Academy for Water Resources, Hanoi, Vietnam
3Hanoi University of Natural Resources and Environment, Vietnam

∗Corresponding Author: Duong Du BUI (Email: duongdubui@gmail.com)
(Received: 19-Nov-2018; accepted: 25-Mar-2019; published: 31-Mar-2019)
DOI: http://dx.doi.org/10.25073/jaec.201931.215

Abstract. Water security is under severe pressures from human interventions and climate change in all over the world and improved water forecast is essential for water management. HYPE is a semi-distributed hydrographic model, running on Windows or Linux operating systems. The code of the model is written by the Fortran programming language and open source as Lesser GNU Public License. HYPE has been becoming a widely used tool in the forecasting of transboundary flows. However, the application of HYPE encounters many difficulties in processing input data and serving the construction, calibration, and validation of the model. This article introduces the development of the V-HYPE tool that helps to couple global rainfall data and HYPE model for operational use. V-HYPE allows developing a user-friendly interface and setting parameters of the HYPE model as well as evaluating errors and transforming and visually displaying the results of the model. Besides, the V-HYPE has the ability to show related maps (i.e. sub-basins, river network, lake, and dams, etc), set up input data, automatically download global rainfall data, and visually display results on WebGIS. V-HYPE also can generate bulletins supporting for operational water resources warning and forecasting works in Vietnam. The utilities of this tool are demonstrated in the case study of Serepok river basin.

Keywords

1. Introduction

Vietnam is naturally "a water deficient nation", given that the current average surface water per capita generated within the territory is 3,480 m³/year - below the international standards (4,000 m³/year) of International Water Resources Association (IWRA) (see [10]). Additionally, approximately two-thirds of Vietnam's water resources inflow from outside the country. Across different regions and times of the year, water resources are not evenly distributed (see [12]). Characterized by water-related development pressures and cross-boundary problems, increasing attention has been recently paid to Vietnam, particularly where there are limited ground monitoring stations (Kawasaki et al. [4], Tran et al. [7, 8]; Vu et al. [9]).

In recent years, the development of remote sensing and geographic information systems GIS has brought new strength in the collection of input data for hydrological models in the flow forecast. Radar and satellite systems have been
actually changing the receiving methods of information in the simulations and predictions of flow based on rainfall. Meteorological, hydrological and land use data collected from the analysis of satellite images has helped to monitor and predict the climate and water features before a long period of time and in a large area. Thus, global data having good quality and being accessible, and easy-to-access are an important part of the global hydrological model, becoming an effective solution to improve the efficiency of simulation and prediction of surface water resources in the river basins, especially where lacks monitoring data and transboundary river basins.

Recent advances in input data help improve the quality of flow simulation based on rainfall by permitting the use of distributed and semi-distributed hydrological models, especially models of the grid or hydrographic response unit with inputs such as information of geography, remote sensing, estimated rain from satellite or predicted rainfall. According to this trend, the HYPE model which is an open source model with semi-distributed parameters is developed by the Swedish Meteorology and Hydrology Institute (SMHI). With its strength as an open source model that allows users updating or changing code and being able to integrate with global data, the HYPE model allows forecasting on the large area (international, transboundary river basins) based on the ability of flexible description of spatial resolution and time steps. For example, Donnelly et al. [2] used new methods for evaluating a number of aspects of E-HYPE multi-basin model performance across Europe. The results show that this model catches the spatial variability of flow and is therefore suitable for predictions in ungauged basins; Pechlivanidis et al. [6] used a recent HYPE hydrological model set-up across 6000 subbasins for the Indian subcontinent, named India-HYPE v1.0 to identify challenges and recommend ways to overcome them. Therefore, it is recognized that the integration of the open source model, specifically the HYPE model with global rainfall data is essential to improve the quality of surface water forecast in river basins. However, because the HYPE model is an open source model written on the basis of programming language, the interface is not user-friendly.

In addition, the extraction of global data as inputs to the HYPE model faces many difficulties, requiring the user to have knowledge of satellite image, Remote Sensing, or Java Scrips. Therefore, it is imperative to have a tool that helps integrate global rainfall data and HYPE model easily and quickly and make HYPE model more friendly to user.

This study presents a user-friendly hydrological modeling tool that helps develop HYPE model, which is named V-HYPE tool. V-HYPE has been developed within the Visual Basic.Net framework accessible through an intuitive graphical user interface and designed to enable users to navigate the model building processes confidently. The automated processes include most of the steps needed to develop the HYPE model: exporting TRMM rainfall data, climate, and hydrological and geographical input set-up and output generation in the types of maps and exporting warning bulletins and displaying on the Webgis.

2. METHOD AND MATERIAL

2.1. Required softwares

a. HYPE model

HYPE (Hydrological Predictions for the Environment) model is a semi-distributed hydrological model, running on Windows or Linux operating systems. The code of the model is written in the Fortran programming language and an open source in the format of the Lesser GNU Public License. HYPE simulates water and nutrient concentrations in the landscape at the catchment scale. Its spatial division is related to catchments and sub-catchments, land use or land cover, soil type and elevation. Within a catchment, the model will simulate different compartments; soil including shallow groundwater, rivers, and lakes (Fig. 1). It is a dynamical model forced with time series of precipitation and air temperature, typically on a daily time step. Forcing in the form of nutrient loads is not dynamical (see Lindström et al. [3]).
All input and output files are in the form of text files located in one or two folders. The input may be grouped into time series input data, which is forcing data and observation time series, and geographical input data. The geographical data is coupled to the subbasin division. Outputs are simulated time series of water and nutrients and performance criteria. In addition, information to run the simulation is given as text files. The error checking is very limited, so be careful to follow the file format as described in the File Reference.

b. Other softwares

In this research, V-HYPE is developed on the basis of popular open source tools as the system programming language: VB.net, PHP, R, JavaScript; Map format: Shapefile: *.shp; GIS technology: MapServer;

Based on the languages of Visual Basic.net, PHP, R, JavaScript, V-HYPE has capacities of storage of information and extraction of global data into text files which are input data to develop HYPE model and predict water resources. After that, the results of water resource forecast are displayed and exported into maps and water resources forecasting and warning reports being compatible with web browsers by MapServer and mmapper technology. WebGIS technology allows integrating and exchanging of geographic information on the Internet and allows users to simply export, manipulate and exploit GIS data from the web browser without the support of GIS software.

2.2. Used data

To setup HYPE modeling, a huge of diversity data was collected, which are Hydro-meteorological data (Precipitation, air temperature, and river discharge), Geographical data (Digital Elevation model, sub-basin area, soil type, land use, hydrographical network), Lake information (Depths, regulation rules, rating curve, water level). A summary of input data and source as shown in Tab. 1.

3. V-HYPE TOOL AND PILOT APPLICATION IN THE SEREPOK RIVER BASIN

In this section, the study systematically presents the detailed functionalities of V-HYPE and results of pilot application in the whole Serepok river basin (18,500 km$^2$), including parts of Gia Lai, Dak Lak, Dak Nong and Lam Dong with water discharging to Cambodia, with a total number of 86 sub-basins. Sub-basin areas ranged between <30 km$^2$ and >400 km$^2$ with an average of 218 km$^2$. Hydraulic network in
Table 1. Summary of input data of HYPE model for Serepok river.

<table>
<thead>
<tr>
<th>No</th>
<th>Variables</th>
<th>Detail/resolution</th>
<th>Data source</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Topography and routing</td>
<td>15 arc-second, 30m, Average = 218 km², median = 170 km²</td>
<td>Hydroshe [5], and Hydro 1K (USGS) [17], Hydrographical network from National database</td>
</tr>
<tr>
<td>2</td>
<td>Land cover</td>
<td>250 m</td>
<td>ESA Climate Change Initiative - Land Cover project (ESA [15]) <a href="https://www.esa-landcover-cci.org/">https://www.esa-landcover-cci.org/</a></td>
</tr>
<tr>
<td>4</td>
<td>Lake and reservoir information</td>
<td>Depths, regulation rules, rating curve of 4 main reservoirs</td>
<td>GLOBAL LAKES AND WETLANDS DATABASE, Lehner et al. [5], National database, Government document no. 1201/QD-TTg [13]</td>
</tr>
<tr>
<td>5</td>
<td>Temperature</td>
<td>0.5 degree (9 stations), daily</td>
<td>[14], HydroGFD (Berg et al. [1])</td>
</tr>
<tr>
<td>6</td>
<td>Precipitation</td>
<td>13 stations, daily, (1980-2015) &amp; 0.5 degree (9 stations), daily (1980-2015)</td>
<td>National Centre for Hydro-Meteorological HydroGFD (Berg et al. [1])</td>
</tr>
<tr>
<td>7</td>
<td>River discharge</td>
<td>4 stations, daily, (1980-2015)</td>
<td>National Centre for Hydro-Meteorological Forecasting (NCHMF)</td>
</tr>
<tr>
<td>8</td>
<td>Evaporation</td>
<td>1000 km, 8 days</td>
<td>MOD16A2, MODIS</td>
</tr>
<tr>
<td>10</td>
<td>Crop and water uses</td>
<td>Crop calendar, water use</td>
<td>National Center for Water Planning and Investigation (NAWAPI) [11]</td>
</tr>
</tbody>
</table>

The basin is dense and there are 5 lakes as results of hydropower production, storing water. The information of daily river discharge (Q) was obtained from four on-site stream stations which are marked (red triangles) in Figure 3. The collection data time is from 2000 - 2009. Daily rainfall data were recorded by 13 terrestrial rain gauges (Blue circles-Fig. 3), and the daily maximum, minimum and averaged temperature of 9 Global Forecast System (GFD) stations (blue square-Fig. 3) stations are used for the inputs with the same period of 10 years (Detail in Tab. 1). Here, the study describes several general features of the tool prior to the step-by-step guide to the V-HYPE functionalities. V-HYPE is designed to meet requirements of user interface and setting up parameters for HYPE model such as capacity of working on a single PC system, storage of information, extraction of information under text file as input files for HYPE model; ensuring user-friendly interface to easily train for user groups at different levels and not requiring high-level computer skills. At the same time, the integration system has the low cost of deploying and developing the system and compatibility with various devices and configurations while it still ensures the quality of information and is consistent with current conditions and the
development in the future. In addition, the system helps to export the results of flow forecast into the geographic information format (GIS), process and allow easily exploiting and searching of information through popular WEB browsers; however, it must ensure the security, confidentiality of information and has clear rules about data security level.

3.1. Automatic download and adjustment of TRMM data

Rainfall data used in the study is daily data of TRMM 3B42 (Tropical Rainfall Measuring Mission) from the source of "https://giovanni.gsfc.nasa.gov/giovanni/" [16]. These data are firstly downloaded in the ".nc" format. By the R language, V-Hype will automatically change the format of data files from ".nc" to ".tiff", then exporting the data to the location of the points (like rain gauges) and converting these files into the format of input files for Hype model as Pobs.txt. At the same time, V_HYPE will automatically compare the correlation level between the satellite data after processing and the observation data at the stations through the comparative charts of day and month data (Fig 4).

3.2. Management of input files

Management of meteorological and hydrological data: File Pobs, Tobs, Qobs contain data of rainfall and temperature, discharge, respectively. They are compulsory inputs for the model, defines the time, daily value at each location (like rain gauges). If it needs to add data to the model such as PET, these files can be created by the "info" toolbar (Fig. 5). Link from SUBID in HYPE to the square in the input file generated from the WHIST tool or GIS software. The data files of Pobs and Tobs must have a SUBID as the column header or a separate PobsID / TobsID. Management of geography and land use data: The geographical input data (area, average height ...) for each subbasin is given in "GeoData", while information about land use data (characteristics of land and land use class) is given in "GeoClass". The same classes exist for all subbasins in a model set-up, but they can be zero in area for some subbasins. One subbasin’s information is given on a row in "GeoData". Necessary information is an identification number (subid), subbasin area (area), and class area fractions (slc_nn). Other information that is often included in GeoData is the coupling, i.e. subid of downstream subbasin (maindown), main river length (rivlen), and outlet lake average depth (lake_depth) (Fig. 6).

For each class land use type, soil type, stream drainage depth, number of soil layers and soil-
layer thickness is necessary. This information is given in GeoClass with one class per row. For lake classes, which are special from land classes, this also needs to be given in GeoClass. GeoClass is column order dependent, so the information needs to be put in the right column. Another column contains information that can be used but is not necessary. GeoClass.txt defines information on soil properties and land use classes and describes special classes (lakes and glaciers) with information on the quantity and depth of soil layers.

**Management of lake data:** LakeData.txt is used to design the performance curve and regulation chart of the lake. LakeData contains specific information on water depth of lake, a relationship of water level and discharge, flow regulation. In the HYPE model, the lakes must be linked to the SUBID and the Geodata data is linked to the lake data via the "Lakedataid" column (Fig. 7).

3.3. **Supporting user interface and HYPE model operation**

Requirements for these functionalities must be user-friendly, easy to manipulate such as editing, adding new data, adjusting and adding input parameters or input data. The main functions are as follows:

- **Parameters update:** Model parameters determine the function of the model. These are given in the menu "par" including general parameters, land use, soil characteristics (Fig. 8). In particular, some parameters can be adjusted; some parameters are designed for specific areas, in addition, there are some parameters which are featured for the whole region. The model parameters are described in more detail in the File Reference and Model Description. For a soil runoff model compartment: In a cold climate snowmelt is necessary (cmlt). It is a land use dependent parameter and need one value per land use defined in GeoClass. Evapotranspiration is governed by a land use dependent parameter (cevp). The default values of all model parameters are zero, so if these are not set, no snowmelt or evapotranspiration will occur. For a model with more than one soil layer, the max-
minimum amount of percolation (mperc1, mperc2) needs to be set for percolation to occur. The available storage of water in the soil needs to be set. The model parameters for "field capacity" (wcf) and effective porosity (wcep) depend on soil type. Runoff from the soil is determined by runoff coefficients. The runoff coefficient of the topsoil layer is necessary (rrcs1), and depend on soil type.

For surface water model compartment: Flow in rivers has a peak velocity (rivvel). This parameter determines the delay in the system. If you have lakes you can give a general lake depth (gldepi) and rating curve coefficients (gratp, gratk). Typical intervals for above mentioned model parameters: cmlt=2-5, cevp=0.1-0.3, mperc1, mperc2=5-100, wcf=0.05-0.5, wcep=0.05-0.5, rivvel=0.5-2, gldepi=5-10, gratp=1-100, gratk=1-2.

- Visually displaying and evaluating the simulation results: The calibration and validation of HYPE model were implemented in two periods: 2000-2005 and 2006-2009, respectively. Based on the comparison between simulated and observation value, the HYPE parameters which are the best fit with this Srepok area can be found. In this process, we used the 13 ground rain gauges for calibration the discharge at four stations of Giang Son, Cau 14, Ban Don and Duc Xuyen for the 2000-2005 periods.

There are several options for output of V-HYPE simulation result. The example given here is the time series for simulated outflow (cout) from each subbasin in the model set-up as columns (defined by subid). There are options for printing time series of several variables for one subbasin in one file and printing period mean values of a variable for example. When observations are given as input to the model set-up, e.g menu "Qobs" with observed discharge, HYPE may calculate performance criteria. The variables given in the menu "info" for criteria calculation are compared. Performance criteria will be calculated for every subbasin based on the values of RMSE (Root-Mean-Square Error), NSE (Nash Sutcliffe Error coefficient), R² (Coefficient of determination) and KGE (Kling-Gupta Efficiency).

V-HYPE software also allows displaying results in the forms of diagrams and maps which are visual and convenient for evaluation and analysis of model results (Fig. 11). Maps allow zooming in and zooming out of attributed infor-
mation, moving, searching, measuring distance, etc. The maps here are the results extracted from the calibration, validation, and prediction of the water resource by HYPE model. From the model accuracy was not homogeneous entire the watershed, the overall errors demonstrated the model performance was acceptable.

Fig. 9: HYPE model accuracy evaluation using error coefficients: (a) RMSE, (b) NSE, (c) $R^2$, and (d) KGE.

Fig. 10: Daily and monthly HYPE discharge compared to measured data in the model validation for the 2006-2009 periods at Giang Son, Cau 14, Ban Don and Duc Xuyen stations.

Fig. 9 and Fig. 10, Duc Xuyen station was the weakest point of the model where all the coefficients showed low accuracy such as NSE, $R^2$ and KGE under 0.5. In contrast, Ban Don and Giang Son presented the most reliable parts of the model with the highest values of NSE, $R^2$ and KGE. Although the accuracy assessment showed

3.4. Developing the water resource forecasting and warning bulletins and exporting the results and presenting on WebGIS

- Developing the water resources forecasting and warning bulletins: The forecast results of HYPE model will automatically be extracted into the water resources warning and forecasting bulletins for the river basin. The interface of the pilot water resources warning bulletin is shown in Fig. 12:

- Displaying and sharing the model result on WebGIS: Maps and the warning and forecasting bulletins of water resources for the river basin will be presented on the Internet (Fig. 13). Users can easily explore, search, reference information through popular WEB browsers. This facilitates connection and exchange of information between government agencies and access to information of people and businesses.
4. CONCLUSION

Hydrological modeling for a poorly gauged transboundary river basin with impacts of human development is always a huge challenge because of limited observed information. Lack of reliable information might prevent the appropriate water management plans, particularly for the poorly gauged river basin. The HYPE model is being widely used in the world; however, the model is developed on the open-source platform and run on DOS interface making it difficult to apply this model in practice. Therefore, to facilitate the application of the HYPE model in practice, the authors have studied and developed the V-HYPE tool. Integration system of global rainfall data and the HYPE open source model (V-HYPE) is a useful tool which can automatically extract and evaluate global rainfall data by the R language. It is a very significant contribution because it is the first time V-HYPE tool has developed in Vietnam to help timely forecast water resources, overcome the lack of observed rainfall data at present.

Besides, the tool automatically develops the water resources forecasting and warning bulletins, based on HYPE's simulation results, which has practical significance to shorten the work of the internal work and provide news to serve the information need of the society in timely manner.

In addition to researching on the theoretical foundations, model structures, data structures, and programming, the authors have also successfully applied V-HYPE integration software to the Srepok basin and are used in forecasting water resources at the Center for Forecasting Water Resources under the National Water Resources Planning and Investigation Center.

Acknowledgements

This work was carried out as a part of the project "Study to integrate open-source model with global rainfall data for surface water resources forecast in transboundary river basins. A case study in the Serepok river basin" funded by the Vietnam Ministry of Natural Resources and Environment (MONRE) under grant number TNMT.2016.02.01 [11]. We would like to express our gratitude to government officials and experts for their helpful advice.

References


[13] Prime Minister, Decision 1201/QD-TTg on the inter-reservoir operation regulation in the Serepok river basin. (Report)


About Authors

Duong Du BUI currently is a Director of Department of Water Resources Monitoring at the National Center for Water Resources Planning and Investigation (NAWAPI) at the Ministry of Natural Resources and Environment (MONRE) of Vietnam. The key responsibility
of his department is to maintain the operation of national network with 1000 water monitoring stations and water forecasting throughout the country, towards providing timely water information including climate impacts on water for socio-economic activities in Vietnam. He is a Hydrologist with over 15 years’ experience in both national and international projects.

**Minh Duc TRAN** graduated from the University of Mining and Geosciences with major of information technology. His career so far, he has participated on many government projects in Vietnam. Especially those project relating to apply and develop information technology in the issues of the natural resources and environment sector. He is an expert in programming and developing of mapping information systems. He has experienced several leadership positions; currently he is deputy of department of information technology at NAWAPI.

**Thi Huong VU** graduated from Hanoi University of Natural Resources and Environment with major of Hydrology. Her career so far, she has participated on many international and government projects related to hydrology and water resources in Vietnam. Currently she is working as a young researcher and consultant on hydrological and hydraulic models in the filed of water resources at Vietnam Academy for Water Resources (VAWR).

**Nuong Thi BUI** After receiving her Doctoral Degree of Civil and Environmental Engineering from Tokyo Metropolitan University, Japan, she continued her work as a lecturer at the Environment Faculty of Hanoi University of Natural Resources and Environment, Vietnam. Her research interests are applied math and computer programming in environmental and natural resources valuation and management. She has published more than ten articles and book chapters in both domestic and international peer-reviewed journals.

"This is an Open Access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium provided the original work is properly cited (CC BY 4.0)."