Questions & Answers Session 3

Please type your questions in the Question Box. We will try our best to get to all your questions. If we don't, feel free to email Amita Mehta (amita.v.mehta@nasa.gov) or Kel Markert (kel.markert@nasa.gov) your question(s).

Homework is available at https://arset.gsfc.nasa.gov/water/webinars/VIC18

Question 1: Does VIC model works on hourly time scale data?
Answer 1: Yes - the VIC model can simulate any time step <1 day. The minimum time step is daily data. Any kind of data after that, the VIC model will be able to run, as long as you specify what time step that is in the global parameter file.

Question 2: How can we analyse for changes in terms of time series?
Answer 2: There's a lot of research that has been conducted looking at time series analyses. It really depends what kind of information you're trying to get from your analysis. If you're looking at, say, long-term changes, just looking at % change from a climatology would suffice. Or if you want to look at trends, you can use your time series to calculate the change per year (or time step). It depends on your analysis, but I would suggest getting an idea of what you want to present as results and then searching literature for what other people have done.
Amita: if you want to quickly view time evolution, you can use Panoply tool Brock mentioned. That's easy to install on your computer for analysis. (Instructions for installation: https://www.giss.nasa.gov/tools/panoply/

Question 3: Are the python scripts used in preprocessing public or custom made?
Answer 3: Both! Kel created the scripts as part of graduate research to make model setup easier, at least from his perspective. So in that case they're custom made, but publicly available. The following link can be used to access the scripts: https://github.com/KMarkert/servir-vic-training/ If you have trouble, let Kel know so he can make it better and fix any issues that come up.

Question 4: What initial parameters are you guessing about?
Answer 4: Quite a bit - the soil parameters we guess about are the infiltration curve, the fraction of max soil moisture where non-linear base soil occurs, the fraction of maximum base flow where non-linear baseflow occurs, and we also estimate or

"guess" about some parameterizations for land cover, such as model resistance - important for evapotranspiration - calculations and also some other parameters that go into that like minimum soil radiation before ET occurs. There's all these that kind of get guessed, but there is published research with these parameterizations and that's what we're using. Specific, local data would be best to insert into the model. Here are the links where you can find the initial guess for soils and vegetations parameters

Question 5: It takes days to run the weather time-series?
Answer 5: I'm assuming this is talking about creation of meteorological forcing - in some cases, yes. What the script is doing is looping through the entire model grid that you've set up and then it loops through the entire time series. If you have a very large basin and a large time series, the script can take quite a while. The setup time for the model can be somewhat time consuming. Once you have a model setup, at least for meteorological forcing side, it becomes easy to adjust parameters, but the meteorological forcing is somewhat static info - at least if you're doing historical analysis. If you're running in a NRT sense and you want to append new data to it, the script runs pretty quickly. Since the Mekong Basin is a large basin, and we're running it for 30 years, it just takes quite a while.
Amita: It also depends on spatial resolution you choose to run at.

Question 6: could we get the video file for model input and processing steps
Answer 6: The recording of today's webinar will be available on the ARSET website. The video file showed in the recording is quite large, but it doesn't have any audio, so the best option might be looking at the recording.

Question 7: Does panoply software come with VIC? or is it independent?
Answer 7: Panoply is an independent software and was created by NASA. Info is available at https://www.giss.nasa.gov/tools/panoply/

Question 8: is it necessary install additional python libraries for running VIC model?
Answer 8: For running the VIC model - the actual model, no. But for setting up the model, like what we showed in today's session, yes - it requires a couple of packages such as GDAL, NetCDF 4, and a couple other ones. If you go to the website with all the scripts, there's documentation.
https://github.com/KMarkert/servir-vic-training/blob/master/documents/software_installation_instructions_final.pdf

Question 9: Where can we find the scripts for VIC?

Answer 9: All of the scripts and a tutorial to walk through the setup of the VIC model (underneath the documents and tutorials folders) which can be found at the following link: https://github.com/KMarkert/servir-vic-training . When using the scripts, please be sure to acknowledge Kel Markert/SERVIR/NASA Capacity Building Program

All these are free and open. All scripts are available at the github link - there's quite a bit of info. There're scripts, documentation, and test data and so - what we did is packaged this into a virtual machine and you can download that. There's a link on the github page to download the virtual machine. You can go through an example case that's in eastern africa to explore the model using the tutorial in the package.

Question 10: Is snow parameter file necessary? My study don't have any snow region. So how to deal such situation in VIC model.

Answer 10: The snow parameter file is not necessary. If you do not use snow parameter file, the VIC model will assume constant elevation across the entire grid cell. That means you'll be simulating the hydrologic fluxes at that one elevation across the grid cell. The snow parameter file lets you simulate different elevation changes within the model. This takes into account temp as well, which is important for energy balance. If your region doesn't have snow or much terrain, then it is not necessary to use the snow parameter file but is still suggested to use.

Question 11: Where I can get the calibration scripts? Can you help me for calibration setup?

Answer 11: The calibration scripts are on the github page (above) under the scripts folder. There's two scripts there - one is "calibrate_vic.py" and the other is calibrate_vic_sceua.py". The calibrate VIC uses a Monte Carlo simulation to randomly selects parameters and provides you with the best parameter set.

The SCEUA uses an optimization algorithm where it randomly picks parameter sets and then it takes that initial guess and fine-tunes parameters. The way the scripts is setup is using the workflows Kel presented where we create the soil parameter file, turning our flux data into NetCDF and running the routing model, and checking those results. The calibration scripts are somewhat custom, but if you do have questions, let Kel know and he'll try and help out as best he can. But the calibration setup is somewhat customized for your own application and where your files are.

Amita: suggestion about contacting Kel or any of us - please do try and run the script according to the instructions you go today. If you have a specific question - or if you're stuck at a specific point - it's good to ask those questions. It's more difficult to walk everyone through the calibration process. Watch the recording, go through the

process, and when you get stuck at a certain point - that's a more helpful question for us.

Question 12: How can we take weeks to run parameterization and be done by mid March (Homework assignment)?

Answer 12: In the homework, you aren't actually running the model, so it shouldn't take weeks. The questions are based on the webinars you watched - the info you received based on that. They're all multiple choice. You don't need to download data or run the model to complete the homework.

Question 13: I am calibrating the VIC model by energy balance. In my case the peak flow is too large and I tried a lot by manually calibration but peak does not down to match with observed flow. Any suggestions?

Answer 13: Yes! Peak flow is really dominated by surface runoff. In the model the surface runoff is most sensitive to B infiltration parameter.A higher B infiltration parameter yields higher surface runoff. Also, check the baseflow parameters because if the soil becomes saturated quickly the model will yield high baseflow as well.

Amita: would elevation band setup affect runoff?

Kel: I would say yes for certain cases. The reason being that you're - depending on your region, if you have a lot of terrain in the region, you can have snow accumulation at the peaks of mountains. By having accumulated snow, that's adjusting your timing of runoff. If you think about it, snow is storage and it has to melt before it's considered runoff. If you have a lot of terrain ,there's snow accumulation and you aren't capturing it, that would probably give you more runoff than you're expecting.

Amita: Even if there is no snow - if you, rather than using mean terrain, if you use actual terrain, would that affect streamflow? It would, right? But maybe not total runoff? But streamflow will change depending on how you represent your terrain even if there's no snow

Kel: Correct - elevation bands will also adjust the evapotranspiration simulations which will either increase or decrease runoff.

Amita: Like peak flow - elevation, how you represent it, will have some impact in addition to infiltration

Question 14: Can we find all the commands of the vic model including parameter creation and running?

Answer 14: Yes, there is a tutorial on the github page with explanations at the following link:

https://github.com/KMarkert/servir-vic-training/blob/master/documents/tutorials/VICsetup_commands.pdf

Question 15: Where do I get parameters required in vegetation library file for different vegetation classes in modis data?

Answer 15: I would look at literature. What I did was skim through a bunch of published research and created a parameter set for IGBP landcover classifications. That's also available on the github page here. I would suggest using that and looking at research and use their parameterizations.

Question 16: Is possible run sub-daily using the scripts share?

Answer 16: Yes - the vegetation, snow, and soil parameter files are all static or are not dynamic in time. The meteorological forcing is the only one that would be dynamic in time and for the parameter setup you'd have to customize the script to read in your data at sub-daily timesteps and write the data at sub-daily timesteps.

Question 17: I have run VIC model, but correlation between observed and predicted data is only 0.3m. how we can get good correlation between observed and predicted data?

Answer 17: Depending on the application and watershed you are modeling, 0.3 m can be pretty good. In the mekong basin if we get an error of 30 cm, we'd call that really good. But - it depends on your region. And I would try performing some calibration processes on that to make sure you're trying to capture - you're getting the best parameter set. It also depends on your input dataset. The model is pretty sensitive to errors in precipitation. So if you have precip data either doesn't capture peak precip or high intensity storms or doesn't get the right amount of precipitation that occurs, then that can cause error in the results.

Question 18: Vegetation and soil data using in the example are available at global area?

Answer 18: Yes they are - in session 2, we went through and talked about how we can - where to get this data - the soil data and MODIS data. And what all the data is. The file types, what kind of classes they are, and here was showing how we can use that data as inputs into the model.

Question 19: Can I calibrate the model on window cygwin? By your given calibration scripts?

Answer 19: Don't know. I would need to try that out since I wrote the scripts for a Linux environment and not a cygwin environment. So you'd need to make sure we have all the necessary packages installed before you can do the calibration process.

Question 20: Do we have to use Panoply to read the files produced and transformed?
Answer 20: Not necessarily - it's just easy. It reads netcdf and hdf files and provides maps as well as navigation through the time plots - it can go back and forth. You can make time series. It's just convenient to use, but if you want to use - if you have any other analysis or graphic packages that work with NetCDF, you can use that as well.

Question 21: my question is related to precipitation input. If I use a local gridded precipitation product, do I have to resample this product to the same resolution of the other inputs or it is not necessary?
Answer 21: Yes, you do have to resample it to the inputs or grid cells that you're running the model at. The model assumes that your precipitation is the average areal precip for that grid cell. What the model does on the backend is it takes that avg areal precip and uses the area of the grid cell and other info to calc your volume of water later on. That's how we know how much water is coming into the region. If you do resample as long as you have an accurate mean, the model should capture all those dynamics.

Question 22: Can the calibration based only on discharge data ensure the realistic individual hydrological components simulation since it is important for hydrological processes characterization?
Answer 22: That's a really good question. So - this is somewhat of a research topic - I believe that the calibration process on discharge only solves for water balance. So, when you calibrate your model on discharge, you're only getting your components of surface runoff and baseflow correct. It's assumed the other balance components equal out to being correct. Whether or not it's actually modeling those processes (snow water equivalent or evapotranspiration) - the model doesn't care about that. As long as it's able to model the runoff and baseflow correctly. There are ways to do a multi-objective calibration process. That way you're calibrating your model on discharge, for eg or ET or soil moisture. That way you're using multiple pieces of info to calibrate the model.

Question 23: Do you have multi-objective script including for example signature measure (flow duration curve), soft data (surface runoff, baseflow) beside of discharge (e.g. performance statistic such as Nash and percent bias) for model calibration?

Answer 23: Right now - no. I am working on a calibration process that's taking into account soil moisture, ET, and discharge so that we're capturing all those components, but it's still in development. If you're interested, follow up with Kel in about a month or two.

Question 24: What is minimum area of a basin that should be applied VIC?
Answer 24: I would suggest that you have at least 3 or 4 grid cells. That way you're capturing the spatial components in the basin ...
Amita: Highest resolution you can run VIC is 3 km - you need at least 4 grid points that are at least 3 km - the smallest watershed is ~27-36 sq km is the *minimum* - the smallest watershed you can work with some reasonable accuracy.

Question 25: Before calibration of the model, is sensitivity analysis necessary so we can limit the parameters to calibrate?
Answer 25: Not necessary - but if you do a sensitivity analysis, that will definitely help with the calibration process. You can calibrate on the parameters that are only going to really affect your results for that basin. There are methods to do sensitivity analysis - I haven't done one myself, but it'd be really good to do.

Question 26: Has the VIC model applied in watersheds in Arctic region?
Answer 26: Univ of Washington has developed a regional model of arctic region for VIC. The link to look at will be http://www.oc.nps.edu/NAME/RASM.htm - if you search for VIC for Arctic you'll be taken to a presentation univ of washington has put together

Question 27: can we transform the output files into a different format, to be read by other programs?
Answer 27: Yes - I like using NetCDF datasets - it doesn't mean you have to use them. If you have a specific format you like to use, then please use that. Make sure you understand how flux files are formatted. Flux files are a time series at each grid cell. If you want to use any other kind of dataset like a GeoTIFF or HDF5, just make sure you understand how flux files are formatted so you can properly format your other dataset.

Question 28: What interpolation method do you suggest re-grid all the data into a common grid?
Answer 28: It depends on the dataset. What Kel uses is for discrete data - like soil classes or the land cover classes - use nearest neighbor. It doesn't make sense to interpolate or get half a class value or something, or for elevation - since the VIC model needs mean elevation for the grid cell, I aggregate each grid cell to the mean value. In

the case of other datasets, if you're trying to interpolate or snap high resolution elevation data needed for the snow file, I use bilinear interpretation, and that typically does the trick. Just use a method that makes the most sense.

Question 29: Is it possible to have the setup of the model for a specific basin in West Africa? As the model has been used for the flood and drought monitoring project of princeton. In case it is possible to get it, who can I contact?

Answer 29: Yeah - you can set up the model for, theoretically, any basin. I would - Justin Scheffield is the PI for the flood and drought project for Africa. I would suggest contacting him if you would like to learn more about the model setup he has. If you can access that model. hydrology.princeton.edu/

You can also look at the data from FLDAS: https://ldas.gsfc.nasa.gov/fldas/ You can download the parameter files and the meteorological forcing dataset to set up your model from the website.

Question 30: How long data is necessary to perform automatic calibration using the SCE-UA algorithm in regional small watersheds with lakes and reservoirs to get optimal parameters?

Answer 30: The SCE-UA calibration program actually converges fairly quickly compared to other calibration processes. I would say it's about 5000 to 10,000 iterations. It really depends on how many parameters you're trying to calibrate and your region.

Question 31: Is VIC available as R script?

Answer 31: Distinguish here - the VIC model itself is written in C programming language. Right now, I don't think there's any plans to make it into R scripts. They do have Python bondings for the model, but those are in development. For the VIC model, don't see it being written in R, but for the setup process - what Kel demo'ed earlier, we can make those available in R, but it would take quite a while to translate between Python and R.

Question 32: At what point would you consider your surface water runoff calibration "good enough" for the Mekong basin example. I have used root mean square and other statistical methods comparing real to simulated data, but would be really interested to hear what is generally used for your use of the VIC model.

Answer 32: What I (Kel) use is an article as a standard guideline. Moriasi et al., 2007 - is a really good article that creates a standard for evaluating models. It gives guidelines

on which thresholds to use that you can use to be considered an acceptable, good, or excellent model setup. Those are the guideline I follow.

Question 33: If you don't need to work in an area that does not have snow, do you have to make a snow parameter file or can you skip this step?
Answer 33: The snow parameter file is not necessary if you do not have snow in your region but it is still suggested to use.
It is an optional parameter. But that also assumes that the elevation is uniform in your grid. Even if there is no snow, specifying elevation would affect temperature, runoff,and ET.

Question 34: Is the input file free format?
Answer 34: Yes, the input files are in ASCII (text) format but have to be written for each grid point in time series format.

Question 35: Are the parameter/arguments for each of the python scripts, presented somewhere e.g. in a user's manual?
Answer 35: Yes - they are. The parameters needed are all defined - please see Week-2 presentations for that information and web-links. On the GITHUB page there is a tutorial for this. https://github.com/KMarkert/servir-vic-training

Question 36: How efficient is an automatic calibration than manual calibration?
Answer 36: Manual is not recommended. This model has so many parameters and by adjust the parameters manually it may not be intuitive how to get the best parameters. The automatic calibration is more efficient at finding the best parameter set and much quicker.

Question 37: Can you use the VIC model in watersheds with regulated flows (upstream dams with more or less regular releases)
Answer 37: No, the VIC model does not include parameterization of water management - or man-made sources/sinks. The version we are using here responds to natural weather variability.

Question 38: Could VIC model be used for multiple watersheds in one go?
Answer 38: Yes the model can cover more than one watershed. As long as your grids cover all of them.

Question 39: Which statistics are used for calibration and validation of streamflow? Apart from NSE.
Answer 39: for more information on this, I refer you to a paper by [Moriasi et al. 2007](#)

Question 40:Which are emerging fields (Projects) where VIC model is useful? Apart from drought and flood?
Answer 40: Overall water resources management and also for near-surface energy balance. In general assessing basin-scale water availability.

Question 41: Can vic be used for carbon cycle?
Answer 41: Yes, it has a simple carbon cycle model - it is not the primary goal of the model though and needs detailed input information. The carbon cycle output is net primary productivity.