

GEO-REFERENCED DATA TO INFORM EARTH OBSERVATIONS MODELING FOR AGRICULTURE: A DISCUSSION AMONG COLLECTORS, USERS, AND AGGREGATORS

QUESTION AND ANSWER TRANSCRIPT

MAY 29, 2019

PRESENTERS

Philomin Juliana, International Maize and Wheat Improvement Center

Jawoo Koo, International Food Policy Research Institute

Robert Hijmans, Geospatial and Farming Systems Consortium

Estefania Puricelli, GEOGLAM and NASA Harvest

Narendra Das, NASA Jet Propulsion Lab

Raghavan Srinivasan, Texas A&M University

MODERATOR

Julie MacCartee, USAID Bureau for Food Security

Kiersten Johnson, USAID Bureau for Food Security

Paul Tanger, USAID Bureau for Food Security

Paul Tanger: Well thanks so much. This is Paul Tanger again. I'm reconnected and apologies for dropping off a bit there, but thanks Julie and others for helping to resolve those issues. And I'm glad you all stuck around for this. I wanted to ask a question to Narendra or Robert and you both mentioned this in your presentations and it's also a bit related to [inaudible 00:00:41] question about the data required for crop modeling. And so essentially we often hear that there are specific parameters needed for data, for proper modeling and as modeling approaches get refined that may change or ... I think essentially the question is how best can we parametrize and communicate better ways to collect data or design a specific program to collect data that would fulfill exactly the needs for the modeling approaches?

Narendra: Hi, this is Narendra. So far our modeling efforts specially per crop modeling and the type of end user we are serving, their basic needs are basically how well we can predict or forecast the crop yield and how well we can tell them what is the situation on ground right now. And they're really interested especially about the crop specific parameter, the yield, the flowering date maybe or like the what is the health of the crop, like whether the what crop is experiencing water stress or not, those kind of a scenarios. They want to know also that they can do better resource allocation and water resource planning.

Narendra: So are what constraints be face basically especially the crop modeling. Our hydrologic modeling is up to the mark and we depend a lot on the remote sensing as in information we assimilate them and which is working great. But when it comes to crop modeling, it's a very human oriented. Like lots of interventions are there. It's not like a nature take its course. So when human intervention is there, we need lots of data basically to make an optimal on model run.

Narendra: The data which we really look for, is the planting date. So planting date, we have up with an approach basically using our now Sentinel or upcoming NISA mission, we would be able to detect what are the planting dates in the [region 00:00:03:08]. So that is one data we look for as a critical data for input because it depends the ultimate yield depend upon when we planted and where we planted. So right now the local agencies are trying to help, but this is one of the big Lacuna.

Narendra: And then the next data we look for is the fertilizer application rate. What a farmer is applying, what kind of fertilizer is applying and what date is applying. And the third thing is irrigation rate. We have some historical data literature and all that, but it changes year to year. So we need some kind of data like what is the irrigation being applied in the region, like countywide average or we can look into that and get the better estimate. And then the fourth thing is basically the seed cultivar.

Narendra: It's not like very organized especially in these developing countries and underdeveloped countries where people don't buy seeds especially from an agency or private company, the Monsanto and all that. So it's a cultivars are a lot like the seed genetic characteristics. So that is where also we need like what is the basic genetic characteristics?

Narendra: So we are trying to calibrate our model. If we have like 20 or 30 years of field data, we are trying to do that. But it is a very big effort when you have to go into each and every county district and do that and there are so many of them. So that is where we need help also like what kind of seeds or

genetic coefficients of the seeds we require? If we get some good [inaudible 00:05:07] we will be able to calibrate them. So that is where we are struggling a little bit.

Narendra: *But these are the four type of data. We are actually looking for a very optimal crop modeling basically for now cast and the forecast. So our agencies with whom we are working, they are trying their best. But I think extra support or will make this project very successful. So far we have been very successful with hydrologic modeling, crop modeling as an upcoming process there. So that's my answer.*

Paul: *Thank you very much. That's very interesting that it's more of a matter of just getting the local agronomic data rather than anything specific about the other, the remote sense data. Robert, I don't know if you had [crosstalk 00:05:58].*

Narendra: *For remote sensing data, we are pretty capable actually in JPL is custodian of many data. And we are very capable to get access and use that data as assimilate that into our system especially for hydrologic model as well as for crop model. So we extensively planning to use Sentinel map upcoming NISA mission and we use all kinds of remotely sent precipitation data which are by selected. So we only lack basically the ground economic data as you mentioned. Yeah.*

Paul: *Great. Thanks. Maybe Robert, if you had a quick comment on this and then I'd like to turn it to Julie to ask a comment from the question from the chat box. So Robert, I don't know if you had any comments from your modeling side on this question of parameters of data that could help improve modeling. And it's not-*

Julie: *We just promoted Robert back to presenter. He got kicked down. So we'll see if Robert is ready to chime in if you're able to reconnect Robert.*

Robert: *It's hard to generalize, there's different needs for different purposes but then again there is of course a lot of general, if he would have one question is indeed do you want aggregated data at a county or district level, whatever the name may be. Or do you want a specific to reference plot data.*

Robert: *County data is very useful by all means. The problem is often that you don't really know how good it is, how is it derived, how ... and so forth. So to me, so from the remote sensing site, if you want to estimate what's crop is grown where and what the yields are, so that the two main things, then ideally you would have field level, high quality observations. And that's a real difficult problem.*

Robert: *It's easy to say, right. Obviously yeah we've like thousands fields in every district or so. Well, we really don't have the technology even to do that efficiently. Like people do crop cops that's very expensive and even those are very small subsets off bluffs. So I think a lot of innovation there has to happen to even collect the data, you really liked to have to evaluate to it, to train and test your remote sensing models of crop performance. But if anything I would say, what's crop is grown where and what is it productivity? After that there's other things you need like okay, management's information planting and fertilizer what have you. Thank you.*

Paul: Great. Thanks so much. Julie do you have a question from the participants?

Julie: We've got a few come in and we don't have a ton of time left and so we are still collecting the questions and we'll do our best to get them answered in a post event blog post and send those out with the recording of this webinar and other resources. So if your question didn't get answered today, we'll still do our best to get answers for you. Let's see, we had a question from Ross hunt who asked in terms of catching rate of change, what data set has the highest refresh rate? I suspect that [Lancet 00:09:37] which can give report every 14 days or faster, but is that actually correct? Is there something with the best refresh rate?

Narendra: MODIS could be better faster than Lancet depending on the solution.

Robert: And Sentinel is also pretty ... It's an alternative, it's similar.

Julie: Great. Excellent. And while I'm on perhaps I'll ask another kind of quick question from Tomaso Cozzarelli, which is, are their vegetation index alternative to NDVI such as incorporating red edge, something that has been tested and can perform better?

Speaker 3: Yes. If I can quickly ask that question. Yeah, we do try other vegetation and this is for different objectives. For like optimum irrigation, we have like we do NDVI at an estimate of biomass for an optimum irrigated environments. But then we have [stress 00:10:40] and vitamins and we do have [endosis 00:10:42] like photochemical reflectance and that's another endosis that we test. But sometimes endosis are quite highly correlated and so they could be just correlated predictors and if we combine a lot of different monitors together, we don't get very high increases in prediction accuracy.

Speaker 3: So we have tried that and depends on the context, depends to be objective with you're trying to project and what environment are you trying to predict and we haven't seen ... It could be a standard deviation of 0.1 but then we haven't seen any great improvements using different other endosis. Thank you.

Narendra: One of the upcoming technology or the studies people are doing, especially for vegetation, which you can use is basically Sentinel cloth called Signature, which is a VH Signature, which is the highly correlated with a vegetation structure as well as vegetation, water content. And it could be correlated well with NDVI. The another advantage of using Sentinel-1A or 1B crop for Signature is that they are ... You can have that kind of data even in cloud conditions, there have microwave [feed band 00:00:12:01] until unless it's a very, very heavy rainfall. So it's mostly available with a refresh rate of around the world at 12 days and in over Europe, it's six days.

Narendra: This is another thing which you can look at basically if you are looking at some of the vegetation aspects. And it's mostly, it's accessible in any condition, clouds and all that doesn't matter. Especially when we talk of crop, cereal crops, rice and all that. Most of the countries at that time of the year have cloud cover. So I hopefully this data, which is data can help in those conditions.

Julie: Excellent. Thank you all. Paul, do you want to do another one of your predetermined questions and then we can come back to an audience question?

Julie: I'm not hearing Paul jumping in. So perhaps I'll ask one more question from the audience in the meantime, let's see. So a question Jojey Varientos, in order to build the necessary datasets from the ground besides technology, how do you create educational models for farmers such? As in how to maintain and keep their equipment running?

Speaker 7: But even also I think just in terms of providing standards for collecting the data in the first place.

Julie: Oh yeah. Can someone weigh in on that?

Narendra: Well, I think they've got some technology becoming available less the same as telling the cloud sourcing data services my people can push data from the mobile applications directly to a server and receive information at the same time that our technology in the US now for the farmers are using to send and receive data to the crop consultants for example. So that they can take a picture and then that can be translated and transmitted to a central system that could be reviewed and provide the feedback in terms of what's happening on the field without needing to visit the field and so on. So there are some technologies becoming available, is still very premature in terms of standardization or in terms of making that available widely. But there are a lot of researchers working on this right now.

Paul: Great. Thanks. I'm back on now. So a question and there's been a couple of comments related to this. I think this is maybe a question for the Estefania and Jawoo as thinking about platforms for managing and collecting some of this information, how do you think about managing curation, time and resources? Like who's responsible in determining what's the most useful data to prioritize sharing and and curation of ... ?

Estefania: [inaudible 00:15:29] I can start. So in our case and our platform, I mean we are all the time in contact with our partners. We try to discuss with them, which is the better way to gather the data, to show it. And we try to put data that there is already online somewhere else or that they are okay with the copyright and all that. So everything is very communicated. It's not that we decide when unilaterally what data is going to be app in the have or we change anything of the data. This is something very collaboratory and our partners are the ones who in the end the decision if they want their data to be uploaded or not. I hope that that's answers the question from my side.

Paul: Great. Thanks and I don't know Jawoo, had any comments, if he is on. I'm not sure I see him. I think someone needs to enable him to speak. Maybe in the meantime, let's see if we have another question from the chat box. So there's a question from Pascal, is it true that SWAT has been incorporated under free watch software or it's still works independently? I think this is a question for [inaudible 00:17:21].

Speaker 9: Yeah, no, there is a ... SWAT is not included in the FREEWAT program. It is still an independent, but SWAT has been linked with MODFLOW which is in FREEWAT. So there is a link between SWAT and MODFLOW the surface and ground water model. So that may be what you're referring

to, but SWAT itself is not in FREEWAT. But it's just a product of SWAT is called Epic, which is in FREEWATER. So which is really a crop growth model, not a hydrological watershed model actually, but only a crop growth model is in FREEWAT. I hope that answers the question.

Paul: Great, thanks. And I think that Jawoo might have it speaker enabled now or his microphone enabled now.

Speaker 3: Yeah. [inaudible 00:18:10].

Jawoo: Excellent question for [inaudible 00:18:12] if we it is to make this data, we are open and curative and making out there. This is very a key challenging questions, we all trying to get better answers there. From our program, the data [inaudible 00:18:30], we actually work creative as of this issue. And we think this is institutional responsibility. It should be part of something we do as a research organization on another he make like publications and jumping nothing out there with it to make data and all the information we have been collecting during the process out there and make it more useful.

Jawoo: [inaudible 00:18:54] is supported financially and also a technical resources such that we are doing now and I think we will get better and better sense or how we can institutionalize internally. But yeah, it's a really challenging issue. So all this PII and responsible data mention data are also adding even more like what I have and can the institutional part and the [inaudible 00:19:21] manage their part on the researchers in a minute if we ever managing data and information. So yeah, I don't have a clear and good answer yet, but yeah, we are trying to find the best way possible and maybe give you something as a community we can all work with together. And your second question for or what kind of data is most useful. And this also a lot depends. And this is one of the reason we ran the survey, I quickly flash it in one of my slides.

Jawoo: The one you could better understand what kinds of data are being used and how they are being used and other types of opportunities there. And yeah, it should really, I think this really wide and various out there. And I think different are getting more and more interesting use cases are being developed by linking different types of data. We've got to see not just kind of one trend or data from one source. People really learn a lot by linking different types of data. We keep seeing socioeconomic data being linked to nutrition data linked to agriculture date also linked to climate and other type of biophysical data.

Jawoo: So I think to be able to support that moving ones or big data and better organized data of moving forward, it's nearly difficult to pre determine this is the important data until we see that whole picture collectively combating together. I saw one of the comments from Brian kind of commenting because of that ontology or notation and better kind of semantic organization under important and I totally agree on. So we have to do just a better job in maintaining our better data. However boring if found like, I mean we have to keep in mind that they're really use cases. There'll be insight beyond what we already saw. Or what the data was initially developed for and prepared for. So I think we have to be prepared for that.

Paul: [crosstalk 00:21:27]. Thanks a lot.

Kiersten: Okay, so thanks everybody. Just to wrap up, I just want to say that this has been a fantastic discussion and thanks everybody for your patience with our technical difficulties. But thank you even more for a really rich, enlightening exchange. At USAID we're in a fortunate position where we have the opportunity to talk with different groups and individuals all working on different aspects around a similar domain. And so we have the opportunity to get different perspectives on things like availability of and access to the data that are needed to feed the earth observation based models that we all feel hold quite a bit of promise in terms of helping us better describe and understand the function of agricultural systems around the world.

Kiersten: So on the one hand we would hear from our colleagues and partners who need those ground reference data to inform their earth observation based models. And as you heard today, we would always hear there's clearly an urgent need for more high quality and preferably public data on things like crop yield, but also on things like geo-reference agricultural plot perimeters, which kind of in the aggregate constitute a really important type of data infrastructure for national ministries of agriculture, crop classification information, directly observed soil characteristics data. There's a lot of demand for ground reference data on these subjects so that they can also be set into earth observation based models.

Kiersten: And then on the other end we also hear a lot from our other colleagues and partners like summit and others that they've got lots of data that are available. But as was pointed out today, even even making those data available isn't always enough. I mean it's certainly the first and a very large step in the right direction, but kind of moving towards making the data more easily findable, accessible, interoperable and reusable is kind of what we want to help the field aim for. And so then at the same time we're also aware of efforts like those, that GEOGLAM AMA and the CGI are a big data platform that are ramping up.

Kiersten: So really what we wanted to accomplish today was to bring representatives of those kind of three groups together so we could learn about our respective work and share it with everybody who's joined us today, to hear from our colleagues, both everybody who's joined the Webinar, kind of about what their interests in the topic are and to share insights about the kinds of data this community needs to advance the science in our respective domains. And kind of how can we help each other move forward in our respective work. And I think in order to do that, it really does require kind of a community of practitioners working in this area who also make the time and space to come together as a community and exchange information, which as we know, isn't always easy to do.

Kiersten: But I hope that this has been a good initial step in that right direction and we hope to be able to continue the conversation. So there's a poll question up there that kind of asks for your insights or opinions on how you might like to continue engaging in this community in the future. One of the things that we'll also do is when Evelyn sends out the link to the recorded Webinar, we're also going to send out an opportunity for you to opt in to a community like a list serve so that we can ... That that will help us kind of facilitator, continue the conversation and keep this community of interested researchers involved in together around this topic.

Julie:

Great. Okay. Thank you so much Kirsten. Thank you all for joining. Thanks to the [inaudible 00:25:14] team for hosting, the Agrilinks team for hosting this webinar, and we'll be in touch for future Agrilinks events. Okay. See you soon. Thank you.