Monday, June 2, 2014

Monday, June 2, 2014 9:50 AM

looking at generalizing population allocation

Skibbe - @ the parcel level - but over the entire treasure valley

Dave Stuve - viz guy O.S. hired,
looking at the Env. arch currently
using D3.js

AFRI grant

John suggested:
Classes Online - students use envision, develop

BSU June 2015 - modeling course

SVN Repository
- O.S. holds the core Env. tightly.
- nightly's in the trunk.

With no policies defined, default actor group okay...

* Use cases for the web service,
* What to add to it?
  - dynamically selected policies.
* want to show the UI,
  P/INVOKE layer?
  ⇓
  to simple Envsim interface

Monday, June 02, 2014

Monday, June 02, 2014 2:29 PM

High level ||-ism

Lots of runs of a scenario (for monte carlo)
Run lots of different scenarios

Mid-level ||-ism
- Need to understand sub-comp dependencies
  - Too hard to do this?
- Envision can run different plugins in || as long as you can define the dependencies between them
  - They don't use this much though.

Low-level
- Parallel algorithms

Use cases:

- Publish/component manager
  - Publish your components to the server
    - Server has repos for data, projects and the model components
- Thin client/web service/web site to specify components, data, projects to use
  - Could share projects and results
  - Web site used to generate
  - Envision as a Service

Four types of data output
- Maps (shape file output)
- CSV files (time series output)
  - Which polies used, frequencies used, etc.
  - Plugins can specifiy output variables, which are also put in a csv file
  - DataObj's can be output as CSV files--each DataObj in its own separate CSV file.
- Flow binary files (hydrology data)
- DeltaArray
  - All DELTA's are tracked. 10-20GB data. 100,000,000 changes.

  - 1st dim is IDU indexes
  - 2nd dim is attributes
  - 3rd dim is time

  - Slice these 3 dims however you want to get info about temporal, spatial location.

  - Use case: look at temporal surroundings of fire event
    - Saved delta's used to give context.

Computer science

Look globally/locally
- If looking outside node domain then probably hit the key-value store
- Otherwise, hit local cache.

Send deltaarrays enmass

Getdata/setdata

EnvEngine todo:
- Few execution management functions to interface
- Remove mfc dependencies.

Cuahsi model

SpatialAllocator

Actor based decision making (bottom up at IDU level)
But, more realistically, decisions made globally
- In order to meet a target, along a set of guidelines

How can resources be allocated across the landscape w/in constraints to meet specified targets. Resources are semi-optimally allocated, respectful of the preference weights, and in a way that meets the target specifications.

Allocation
- To stuff on larget than idu scale

Constraint
- Specify a constraint query on the alocation

Preference
- "soft constraint"
- Has weight attribute that is applied if the query is true.
  - The weights are relative, they aren't percentages or anything.

Algorithm:
- Walk down list of pointers (to what?)
  - Figure out top score in list
  - Expand if needed
  - Decrement the resources available
  - Allocate resources to the IDU
  - Go again, for the next highest vlaue in the list.

Spatial pattern of water usage in 50 years?

For tomorrow:
Hydro in the morning

Tuesday, June 03, 2014

Tuesday, June 03, 2014 10:49 AM

Cropsyst
Recsim model -- army core of engineers
Envest

Envision has "recsim light"
For using recsim models in
Must export csv files from recsim and use them
Goes through list of constraints, applying them at control points to determine the amount of

1st look at water pool height
if above guide curve, then release
other wise can't release water
.....

recsim might be a good use case for distributed Envision
Read about recsim for a little understanding on what might be going on in WW2100 plugin.

HBV

Tags:
<controlPonts>
points in the system being targeted
Reservoirs are the controlpoints

cp_Min_flow_at_Salem_23791083.csv

Hardware

R820
128GB RAM
4 8 core HT
disk is regular stuff, 7200RPM
$18K
req. 16-24 GB for a big simulation
2 or 3 of these

ibm tool purify for instrumentation

Visualization

currently ASP.NET MFC idea
at first, would be an interface to display envision output.
later, a more comprehensive interface to control of an "EnvisionEngine" backend core.

engagement with citizens
vis needs to be notional, but not unrealistic
"this is what a nieghborhood with these attributes would look like"

Leaflet mapbased rendering tool

GeoJSON

FOSS4G -- Portland

Dave is working on the BigWood data set.
stuved@onid.oregonstate.edu

Urban Development in Envision

variation per project, depending on geo, etc.

Adding population to coverage
Changing the coverage through development, city change/development
CA approaches not worked well before.

Target

use target to "rain" people on the landscape

Developer

modeling urban expansion --
unlike Target, takes into account the ability of the landscape to support population
one job is to mapping pop dens. into structures
upon certain threshold reached, an expansion event is triggered, modeling urban expansion
expansion modifies the undrlying spatial data,
rules for where this happens is defined by spatial queries
realistic areas, aka Areas of Impact.

want to make zones demand driven
expansion in t=t+1 is a function of:
based on targets capacity, place people on the landscape proportional to the difference between the population curve

see bsh's "TV_Population.xml" <-- the Target curve is defined in there completely.

John's team often writes custom code for Developer.

NCLD classifications

Envision can also track financial resources for policy application. So, if a policy can not be supported financially, then it would not be applied.
For example, maintinance costs
This is subtracted from a budget
This is defined in a policy definition.

Target and Developer are interactive with each other.

An actor tag:
used as a template to stamp out one actor per idu
weights are adjusted via value standard deviation
internally though of as an actor group

Policies

Possible Journals of interest:
Ecological Applications
Ecological Modeling
Environmental Modeling and Software
Journal of Artificial Societies and Social Simulation

More domain specific:
Water Resource Research
Journal of Fire Sciences

for jim, IEEE journal might be easier.

Wednesday, June 04, 2014

Wednesday, June 04, 2014 11:56 AM

Josh has meeting with IT on 6/17 <-- I need to get in on this!

Follow up from yesterday's journal discussion (Possible Journals of interest)

Josh also IEEE visualization journal

Visualization

Envision can only do thematic maping, but that might have limitations relating to stakeholders.
communityviz
cityengine

Kelly Reservoir Question

Thursday, June 05, 2014

Thursday, June 05, 2014 2:43 PM

Job scheduler for Envision.

WaterMaster (flow plugin) 25% of execution time for flow :)
Needed for flow:
- Points of diversion are GPS coordinates, which may/maynot line up with the stream!
- Priority dates, dates when you can pump, max pump rate w/in a period

- Waterrights for watermaster plugin
- Maximum and minimum flow requirements -- through "straws" that act as the pumps.

Urban areas associated with certain waterrights.

Envision executes existing modflow model, applies at timesteps.
- Exchange processes in modflow need to be worked out
- Non-channel exchange processes (pumping into/out of the ground at a point source)

Integration with cropsyst, epic
- Talk to Claudia

HBVContainer

Crop choice models
- Does an IDU have a water right?
- For a given IDU,
  - Calc prob, normalize probs
  - Check water rights,
  - Sample from prob distribution, and pick a crop.

    Ie, create a PMF, and randomly draw

Process based representations

- Spaitialallocator for statusquo scenario

- Cropchoice model for failure, select different crops based on crop failure

- Anything available in the IDU coverage can be a factor in decisions.

Friday, June 06, 2014

Friday, June 06, 2014 11:41 AM

Where does the crop coefs come from?

Agrimet crop cooefs

Graphical and tabular format

crop_coefficients.txt

bw_landcover_coeeficients.csv file

Big wood study area

Planting methods column

Several planting methods -- are these methods represented as policies in Flow?

Paper:

Authors: Rick Allen and Robbinson? Evapotransporation in Idaho (report)

1. **T30 method**

   If average temp in last 30 days meets a threshold, then plant that crop, but not earlier than a certain date.

   Primarily, this method is used.

   , this is **method 1**

2. **Cumulative $GDD_{January}$ (**Growing degree days**)**

   Threshold on accumulated heat units

   For GDDEquation column, number is corn equation, otherwise use equation 1.

   Method 2

CGDD P to T

Cumulative growing degree days, planting to termination

Killing frost temp [C]

Threshold, kills crop

Earliest planing date column

0 based -- day of year

-1 is what? Crops that aren't crops

Term Date Method

2 -- killing frost

3 -- which ever comes first (heat units or killing frost)

1 -- CGDD since planted

4 -- the July 15th method

Min growing season [days]

Don't kill the crop if a kill threshold occurs w/in # of days. (will have at least that many days)

0 -- no min growing days

Columns beyond excel colum AN are not used except for cross checking calculations.

How to handle municiple

Using -1 for plant date method,

Note to self: what is type of m_pCropTable ?

Documentation of architecture -- John's team will get his team to put some together
- Likely multiple groups plugging into the parallelization idea
- Docs for July 15

Envision output --> XML docs from Daves backed --> D3 frontend

Miles Van der weeten
- UG CS student @ o state
- Working to possibly

Usecases for policy modification
- Sensitivity analysis

Lejo's Notes from Google Drive Doc

Monday, June 09, 2014 6:43 PM

Tentative Schedule:
Monday:
Khila: David Hulse at University of Oregon (Wednesday morning? - Department of Landscape Architecture)
Josh:
Notes:
OSU: Vizualization hire: Dave Stuve - tasked with thinking about visualization in a broader sense, as well as particular to Envision.
D3… JavaScript library for visualization. Classification and typology of data.
High-level architecture discussion.
John…
Monday -
10-12pm
Lab meeting 3pm in conference room
Tuesday
Fairly open
Wednesday -
On coast
Thursday - Friday
Here mostly
Kellie...
Wednesday - pretty open
Otherwise… 50/50
Monday AM Notes:

- Version control discussion… all source code in one branch of the repository, datasets in another. Basically primarily geared toward internal development.
- Core Envision - source held onto fairly tightly.
- Plugin Envision - source much more loosely.
- Philosophy… much less of an issue when a plugin is broken than with core Envision.
- We're the first use case for external interactors with the code.
- Fairly simple version control… single branch. Code checkin made almost daily. Revisions only published maybe weekly. Build Envision about once a week. Every morning you do an upgrade, every evening doing a commit.
- Flow worst offender because it's recent and fairly developed frequently. Last couple weeks Flow has been fairly stable. When they check in code, they are also checking in input datasets.
- Flow documentation always lags development.
- Two use cases: (1) coder level, (2) general purpose, general programmer type.
    - Two very different models. Which one are we? Which one is Josh - we don't know the answer to this yet but should think about it?
- Data management…. can we collectively make progress on data publishing. Issues of provenance.

Monday PM Notes:

- Architecture discussion…
- 3 levels of parallelism
- High-level: run different scenarios, parameter sets across different devices
- Mid-level: run individual modules differently across different environments
- Low-level: Traditional distributing matrices, etc. across a cluster
- Three types of data:
- Time series output
- Flow binary files
- Delta arrays (10-20GB of data)... tracking changes in the landscapes at every time step.
    - Data cube…. IDU location vs. attribute vs. time.
    - Can visualize slices of data through cubes… i.e., what happens to a particular location throughout time…. this is probably a good place for application of visualization for Josh.
- Windows dependencies being removed from Envision (GUI) was the tricky part. Weeks not months left of work…

Tuesday AM notes:
Prompts for John:

- Make sure that email to David Hulse gets sent… tomorrow at 1:30pm
- Names of other social science end users of Envision for Khila to hook up with…

- Control points: Points along that a reservoirs operation will respond to (i.e., minimum flows in some reach downstream)....
- Flow has "version of HEC-ResSim" http://www.hec.usace.army.mil/software/hec-ressim/ they call it ResSimLite
- Can we get HEC-ResSim model from the Corps? It would be really useful to have their model to compare out ResSimLite simulations to and get their buy in.
- Right now reservoirs assume that we're simulating the inflows to Lucky Peak or that we'll be providing input to Lucky Peak. We may want to think about actually simulating the upper basin too… one reason for this is that if we wanted to get into ecosystem services related to recreation (e.g., serviceshed type of stuff… fishing, rafting, boating on lucky peak), then you want to simulate the whole reservoir system… Anderson Ranch, Arrowrock, Lucky Peak, etc. It's also a really interesting water resources engineering problem where three large reservoirs are managed for a variety of services (flood control, water supply, recreation).
- Mores Creek as an important watershed where changes to fire and precipitation regimes are particularly risky for Lucky Peak Reservoir because Lucky Peak is the only reservoir to buffer flood waves from Mores Creek.

Hardware:

- Buying up Dell R820… 128 GB RAM… 4 x 8 core processors, hyperthreaded gives 64 cores. $18K big simulations (16-24 GB of RAM needed)... allows 3-4 to run in parallel.
- Remote Desktop login… Windows Server doesn't have a limit on number of RD connections (or if there is one, they haven't reached it).
- Scheduling? Informal agreement to use resource collaboratively. Not controlled at one point… undergrad CS student to create a queuing and scheduling app? Can this student work with Jim P. to advance the queuing and scheduling app that he is writing?

Minimum specs:

- 8 GB RAM, 100-200 GB free disk space, 4 cores, decent processor.
- Monte Carlo runs… turn off keep Delta Arrays between runs… saves a lot of RAM. One known bug… related to Windows Resources. When you update the

screen a million times…. will run out of Windows Resources. Turn off screen updates for long runs or multiple runs! Flow is especially problematic.
- Purify… IBM. Good investment. Catches memory leaks, bad pointers, broken pointers, etc. But it's pricey ($2000 for subscription).
- Much more compute bound than disk-space bound.
- Platform independence… string handling is an issue 85% of the issue. Jim P is really interested in helping to port to Linux.

Tuesday PM:
To-dos:
- Visualization
- Development

Visualization discussion

- ASP.NET controller… develop interface to

Free open-source
https://2014.foss4g.org/

Development (plugin) discussion:

- Use 'target' to "rain people down on the landscape"
- Next step is how to allocate people within the landscape… how the landscape will change as you add people to it. Using a relatively new plugin called developer.
- Urban expansion… changes to underlying spatial information.
- Changes modifying some landscape qualities references in target.
- Changes to zoning "should be" demand driven… specify within a particular part of the landscape (e.g., a city)... attach a set of rules dictating how that city expands. When you reach certain capacity thresholds… that triggers an "expansion event."
- In Oregon… cities have to maintain 20 years of land capacity in urban expansion boundary. So when population density reaches 80% of buildout, an expansion event is created. Rules for where urban expansion happens are triggered by queries… e.g., required that any expansion has to be contiguous with the expansion area. Defined through queries where that expansion can occur (e.g., not in areas of steep slope, not in hazard areas, not in protected areas). We should find and map "Areas of Impact" as defined by city urban planning into IDU coverage

- Talk to planners and COMPASS about urban growth representation and how to change zoning.
- Can attach budgets to policies… can assign pots of money to implement particular policies (e.g., land acquisition, conservation projects, etc.). It keeps track of previous

Decision algorithm…

1. See if there are any IDUs where decisions will be made
2. Run pre-timing autonomous processes
3. Actors do their thing:
    a. See what IDUs correspond to a particular actor
    b. See if it's time to make a decision
    c. See what policies are available to make decisions on
    d. Which of those policies are applied (at least zero, at most one)
    e. Evaluate policies against weight
        i. How does this policy score against all of the decision elements
        ii. Calculate a agent personalized score by considering the weights of each weights
        iii. Generate a random number and select a policy based on the normalized personalized agent weights.
4. Post timing autonomous processes
5. Evaluative models run

Publication Ideas

Journals where integrative modeling approaches appear commonly:

1. Environmental Modelling and Software
2. Ecological Modeling > integrative articles still stand out
3. Ecological Applications > integrative articles still stand out
4. Journal of Artificial Societies and Social Simulation
5. Specific domain-specific journals:
6. Water Resources Research
7. Journal of Fire Science
8. IEEE journals (visualization - maybe, data management, parallel stuff)

Questions for when we get back…
Coordinate with Josh before meeting on the 17th (me, Nancy, Shawn, Jim P.) with Research Computing people.

Thursday PM Notes:
Platforms

Friday AM notes:

- WaterMaster… place of use and point of use data. Note that for the Treasure Valley the number of unique PlaceOfUse records is less than the number of water rights in the shapefile.
- WRExists is a little bit tricky because it's encoded as a 32 bit integer. But James has some IDL scripts to do this.
- Crop water demand calculations… in definition table need to specify how growing season is determined (i.e., growing degree day threshold exceedance, etc)
- Where can Boise State make a contribution? Friday from 10a-12p to discuss.
- BigData issues… using MapReduce approaches to find patterns in DeltaArrays??? Some ongoing work on this with some folks at Purdue?

Work plan for the next 6-9 months:
Where do we want to be then?
Think about prioritizing the following in terms stakeholder engagement:

- **Irrigation:** Need to talk James Salzman… global methods in flow is called watermaster. Implementation of prior appropriation. Required is a database related to points of withdrawal to points of use. Attributing a reach in the channel network to the point of use. Polygons corresponding to points of use. Use the polygons to modify IDU coverage. Type of water right… ag, irrigation, urban, etc. Need priority date. Dates during year when you're allowed to pump and how much you're allowed to pump. As flow is calculating discharges, watermaster looks at demands, priority dates, and flows; makes decisions about whether demands can be satisfied based on discharge. Can define minimum flows. Works pretty well. Other ways we can provide straws.
- **Urban water use:** Per capita water consumption rate as a function of income, population density in urban area, total of water consumption. Regression

model.

- **Groundwater:** coupling between MODFLOW and Envision still "in progress." There are still things that need to be done. Seems doable… but need a grad student pretty well-versed in MODFLOW and surface water modeling. Would need to know FORTRAN. Now using deeper layers of HBV as the "groundwater" reservoir. Use a linear or reduced-order reservoir approach? Solve the groundwater equations on a non-rectangular grid. This needs discussion once we get back. There are some stubs for groundwater model in Flow. Can we use the groundwater model used in tRIBS
    - Canal losses?
    - Need to discuss and make a decision: (1) quick and dirty groundwater linear reservoirs connected with straws while working on a tRIBS-like implementation of groundwater, or (2) try to kludge together a MODFLOW model
- **Canal routing… nothing has been done on this?** This is a longer term issue. Look into coverages of canals. Additionally, the
- **Crop water demand:** CropSyst… phenology, growth and development. C++. Writing a wrapper wouldn't be that bad??? Contact Claudio???
- **Crop choice models:** No standard way to do this. A couple projects which crop choice models were needed… Willamette. Economists put together a choice model (linear regression) that depends on does IDU have water right? Is it ag use? Soil capability, mean elevation. Generates a probability for each crop… gets evaluated for each Ag IDU and crop. 8 crops in Willamette work. Basically generate the Probability Mass Function for each IDU and crop, randomly sample to assign crops. Orchards ignored. Look in EconModels.cpp. Used data from NASS data. Second approach for AgCanada…. big crop group does crop forecasts. Looking at global prices, etc. They generated forecasts for 44 years to be used as targets in SpatialAllocator. Then for non-status quos you apply changes to distributions as articulated by stakeholders. This approach is what's being done for Big Wood. Extension person for Big Wood. Started generating a lot of climate-related metrics… e.g., how much rain in May, etc. Potentially use regression trees.
    - A simple approach… use NASS data to calculate the fraction of land assigned to each crop and have spatial allocator use as targets, have someone like IDWR tell us what crops they "worry about" (e.g., water hungry crops, etc.).

Other items:

Developer:

Khila can contact John for help.

Visualization:

- A lot of progress expected.
- Committed to D3 at this point.
- Putting together infrastructure to specify what you want a set of plots to look like, 3 months.
- Deliverable for Dave Stuve is a tool to "publish" Envision simulations set of inputs, set of outputs, set of scripts that do stuff including visualization. Where can Josh contribute? Not sure…. let's see where the conversation goes and identify areas where we can collaborate.
- Dave is doing a lot and will get a lot done in the next 10 months or so. Can potentially workshare with Josh where it's appropriate and beneficial for all, but probably need to be a little careful here so as not to duplicate efforts.

Architecture:

- Trying to get documentation together on how to look at issue to articulate what OSU is working on and what they're not.
- Another group at Illinois that they wrote a proposal with.
- Fairly detailed documentation to be collated into a single document. Over the next month… middle of July for distribution. Will include us in articulating where they're going and where we can contribute.
- Miles Vander Wheaton? Undergrad CS student working on potentially develop an app to do CPU monitoring, help with job scheduling and load balancing.

Contacts:
Developer - John
Flow - Kellie
WaterMaster - James Salzman
ET - Cynthia