**Modeling Notebook notes for TRACE – BLT Model**

This is not a TRACE document, but a TRACE notebook. General guidelines for TRACE are found on [Grimm et al 2014](http://dx.doi.org/10.1016/j.ecolmodel.2014.01.018) (structure on Table 1).

Interface gráfica do usuário, Diagrama

Descrição gerada automaticamente com confiança média

|  |  |
| --- | --- |
| **TRACE element/MN entry tag** | **MN keyword** |
| 1. Problem formulation | Model purpose; Research questions |
| 2. Model description | Model development; Design decisions |
| 3. Data evaluation | Parameterization; Patterns |
| 4. Conceptual model evaluation | Conceptual design decisions |
| 5. Implementation verification | Debugging  Software verification/Testing  Usability tools design |
| 6. Model output verification | Output verification/Goodness- of-fit Calibration; Tests on environmental drivers |
| 7. Model analysis and application | Sensitivity analysis; Uncertainty analysis Robustness analysis; Simulation experiment |
| 8. Model output corroboration | Output corroboration/Validation |

Implementation 1 refers to Step length + turning angles

implementation 2 refers to Resource visitation rules

implementation 3 refers to Territoriality

**Examples of topics to be included each day and a brief description:**

/purpose

To validate a simulation model of BLT dispersing seeds

/data

Check .xlsx

/verification

/calibration

/calibration/direct parametrization

Obtaining parameter values directly from the literature or experts

/calibration/inverse parametrization

Obtaining parameter values inversely by calibrating the model to observa- tions

/sensitivity analysis

/sensitivity analysis/submodels/energy

/sensitivity analysis/local

Varying one parameter at a time

/sensitivity analysis/global

Varying several or all parameters over their whole ranges

/alternatives

\_\_\_\_\_\_\_

\_\_\_\_\_\_\_

**2. Model description**

**February 04th 2022**

Discussion with Ronald about implementing turning angles and step lengths. We realized we need to check the empirical data a little bit further in order to implement a rule based on sl and ta in the model. I decided to check in [a published ABM](https://doi.org/10.1111/oik.07431) how it was done. Turning angles was implemented with von Mises distribution base on rvm-function from {CircStats} package, however I didn’t have a code for step length. Other two papers did similar implementations (Raghunathan et al 2020 and Gazagne et al 2020) with an HMM model but used also a gamma distribution to step lengths. With this I’ve checked the “Check if the impl. 1 needs model parameters” task. The answer is: Yes, it does. Gamma distribution for step lengths and von Mises distribution for turning angles. **For DPL and Path Twisting: don't know yet**.

After meeting with Ronald at 13:30 we decided to first have the nlrx workflow (to check results after running experiments) ready before making an implementation. Most importantly, we discussed more about step length/turning angles relation, I showed him the results from the ATBC and runned emptyspace analysis, explaining again the idea of taking time windows of resources and clumping measurements to make the home range emerge together with a DPL and Path Twisting based on a random point process specifying the tree aggregation. For last, I explained to him why I didn’t believe tamarins were going after resources and cited the mechanistic home range/central place forager hypothesis.

**February 09th 2022**

*Keywords: Model development; Design decisions*

From Feb 07th to now I’ve been troubleshooting Mayara’s code to match with the nlrx workflow. Two problems have risen:

1. Model output with nlrx to set workflow with R: even though it works with the Wolf-Sheep model, when I try with the BLT model, the px and py positions of the simulations don’t come out with **unnest\_simoutput()** function. Alternatively, [the nlrx vignette (see Step 3) explaining how to take manual output](https://docs.ropensci.org/nlrx/articles/manual-output.html) and relate to the nlrx object requires a reporter in NetLogo code, but when I insert it as a global, NetLogo gives me an error saying I should define it.
2. to-sleeping-trees procedure does not work when using the code for empirical sleeping trees. Thus I’ve created a Chooser **“sleeping-trees-scenario”** of whether simulated sleeping and resting trees are used or not (but only the code for simulated sleeping/resting trees work -> check **to-sleeping-trees** and **search-sleeping** procedure and sleeping-trees-here object is unused)

I also did the following:

* settled a **“tree-scenario”** Chooser for selecting which .shp with feeding trees was being used as input;
* settled an /tempdir directory in model interface.

**5. Implementation verification [?]**

/extensions/gis

Got the code of nlrx running for Mayara’s model. Problem: It is generating one file for each day, thus it dumps everything in github

Files and code: [path](file:///D:\Data\Documentos\github\BLT_IBM-Model\Model_development)

Script: 00\_start-nlrx.R

Left to fix/decide:

* If I’m keeping model output as external files

/extensions/gis

**Feb 2nd 2022**

Debugged shapefiles of the three areas (Guareí, Suzano and Taquara) to initialize with a chooser and and if condition.

Files and code: [path](file:///D:\Data\Documentos\github\BLT_IBM-Model\Model_development\gis-extension)

Included a procedure to check if turtles are inside the shapefile: ‘check-agent-in-fragment’

Tried looking how to scale the different locations. Didn’t find anything.

Left to fix/decide:

* Better .shp files to NetLogo (for Suzano and Taquara, too big)

**6. Model output verification/calibration**

**Day 1 – January 14th**

Six steps for making a good calibration (Railsback & Grimm 2012 Chap 20):

1. Identify a few good parameters

Important, uncertain and independent parameters. To know convincingly which parameters are important, we need sensitivity analysis. It is prudent to reconsider model calibration after this sensitivity analysis. A parameter that is highly uncertain but has little effect on results should not be used for calibration.

Decision for the BLT model: activity budget (e.g. time spent travelling?)

1. Choose Categorical vs. Best-Fit Calibration

Categorial gives a range of values (e.g. mean number of agents between 120-150), while Best-Fit Calibration gives single values (e.g. mean number of agents = 127).

Decision for the BLT model: Categorical seems the best option

1. Decide Whether and How to Use Time-Series Calibration

“If our model's purpose includes representing how results change over time (e.g., how long does it take the system to recover from some perturbation? How is the system affected by changes in its environment over time?), then it usually does make sense to use time-series calibration. But some ABMs (e.g., the woodhoopoe model of section 19.4.3) are intended to explain long-term average conditions, so they intentionally do not contain all the processes that cause the real system to change over time and use no input data to represent how the agents’ environment changes over time. In such cases, time-series calibration may not be useful or necessary.”

Decision for the BLT model: As I’ll only be focusing on the mean and sd numbers at the end of the runs, I believe I don’t need Time-Series Calibration

1. Identify Calibration Criteria

* Calibrate the model against all interested patterns
* Of apples and oranges: the observations need to match the same time and space scale
* There are kinds and kinds of variation: Calibrating variability measures (CV, sd) should be done carefully
* Often our data is inaccurate and we need to know if it is inaccurate by 10, 20 or 50%. Using uncertain data is OK and unavoidable: when the calibration patterns are more uncertain, we don't worry so much about matching them exactly, and we recognize that calibrated parameter values are less certain. But we need to have at least some idea how accurate or certain the observations are, so we know how much information they really contain
* We must specify how we will compare the observed patterns to model results. Of particular concern is how to calibrate several different kinds of model results at once: if we want to calibrate a model to reproduce the number, size, and wealth of agents, how do we decide between a set of parameters that reproduces number and wealth well but not size, and a parameter set that reproduces size and wealth well but not the number of agents?
* At the end of this step of defining calibration criteria, we should have a specific algorithm for quantifying how well a set of model results reproduces the selected observations

Decision for the BLT Model:

* Calibrate against SDD, DPL, Home range and Activity Budget
* (In order of importance) prioritize SDD, DPL, Home range and Activity Budget
* Algorithm: the mean (min and max too?) SDD, the Home range and Activity Budget evaluated after 30 running days and the DPL of each of these days compared to data of Guareí (PEMD and Suzano to be independent observations serving as validation?)

1. Design and Conduct Simulation Experiments

This calibration experiment executes the model many times, using combinations of values over the feasible range of all parameters. The results of this experiment will tell us what ranges of parameter values produce results that meet the calibration criteria. Steps:

* Select values for the non-calibration parameters and input data (if any) that represent the conditions (the same time period, environment, etc.) under which calibration patterns were observed.
* Define *parameter space* (It is usually good to include values that bound the range of feasibility) (see Figure 20.3)
* If the model is stochastic, analyze means seems right, but plot model results against parameter values [? p. 357]

Decision for the BLT model:

* Select all non-calibration parameters an input data from Guareí (using values that bound the range of feasibility, e.g. energy-from-fruit 0)

1. Analyze Calibration Experiment Results

* If your ABM does meet all the calibration criteria for one or more combinations of parameter values, then you can in fact complete this step and move on to the kinds of analysis we address in part IV
* What should you do if you cannot meet all the calibration criteria at once?

1. Screen for errors (code, submodels, etc)
2. If there isn’t any error, your model is too simple or too simple in the wrong ways, to reproduce the observed patterns you chose as calibration criteria. You could consider going back to the theory development stage and seeing if you can reproduce the observed patterns you chose as calibration criteria. You could consider going back to the theory development stage and seeing if you can improve the traits for agent behavior, and you could consider adding processes or behaviors that were left out the first time. But keep in mind that there are costs to adding complexity to your model, especially if it is not very clear what change needs to be made.
3. [Alternatively] It is very common for good modelers to keep their model simple instead of adding stuff to it until the model can reproduce all the calibration criteria. Keep in mind the overfitting issue: it can be risky to try too hard to make the model reproduce a limited set of observations. It may make sense to revise your calibration criteria so they are not as restrictive. If you choose not to revise the model to make it fit more of the calibration criteria, simply document your calibration results and the extent to which the model does not meet some criteria under your “best” parameter values, and your decision not to revise the model. Then, when you use the model to solve problems, keep in mind which results are less certain as indicated by the calibration experiment.

Decision for the BLT Model: I can’t make any decision right now, but the comment 3) seems very important. I’ll probably proceed if I find a set of values that meet the first 2 variables (SDD, DPL), and then Home Range size (not core home range as it is mostly not stable throughout time)

Left to fix/decide: