DyHDER: Dynamic Habitat Disturbance and Ecological Resilience

This document is intended to orient users to the different input and output files and model parameters necessary to run DyHDER ("Die Harder"; Dynamic Habitat Disturbance and Ecological Resilience), a MATLAB model developed by Dr. Brendan P. Murphy (Utah State University; last updated as of January 2020).

The DyHDER scripts (written in MATLAB R2018a and available for download at: https://github.com/bpmurphy) are capable of, but not limited to, running the model and reproducing the results (with some additional work by the end-user) described in the following publication. For more information on DyHDER, or if you use DyHDER in your work, please cite the following open-source, peer-reviewed publication:

Murphy, B. P., T. E. Walsworth, P. Belmont, M. M. Conner, and P. Budy. 2020. Dynamic Habitat Disturbance and Ecological Resilience (DyHDER): modeling population responses to habitat condition. *Ecosphere* 11:1–26. DOI: 10.1002/ecs2.3023.

Table of Contents

A. Brief Description of DyHDER

DyHDER (pronounced "Die Harder") is a spatially explicit, stage-structured matrix population model. The model can be used in a PVA framework, can be applied to metapopulations of any size or species, and was specifically designed with a structure capable of evaluating mechanistic metapopulation responses to habitat condition (e.g., disturbance). This is accomplished through the use of species-specific habitat suitability relations. DyHDER also includes a metapopulation dispersal model that drives probabilistic emigration and immigration rates between subpopulation sites as a function of changing habitat conditions and site densities. DyHDER is intended to provide a framework for evaluating ecological population-level responses of probabilistic or theoretical habitat disturbance scenarios (e.g., flood, drought, wildfire, habitat restoration, fragmentation) of varying location, magnitude, and duration.

DyHDER is coded in MATLAB R2018a and all scripts are open-access and open-source. To download the DyHDER model, please visit: https://github.com/bpmurphy

The model is run using the 'DyHDER_Master.m' script file. All model parameters (described in detail in *Setting Up A Model Run*) can be adjusted in the opening sections of this annotated master script. A subfolder of Excel spreadsheets that define the model inputs (see included subfolder of examples: 'Inputs_LoganRiverExample') must be located within the same directory folder as the DyHDER_Master.m script. A subfolder of DyHDER function scripts (default name: 'DyHDER_functions') must also be located within the same directory folder as the DyHDER_Master.m script.

INPUTS: The model requires a minimum of 7 Excel spreadsheets (.xlsx) to inform the demographics, habitat conditions, and dispersal within DyHDER (described in more detail below), but can optionally read-in additional spreadsheets that describe additional habitat suitability relations to be included, as well as stocking/harvesting scenarios. These files must all be saved within the same subfolder that is located within a parent folder containing both the master code ('DyHDER_master.m') and an unzipped subfolder containing all the required DyHDER function (available for download as: 'murphy_DyHDER_scripts.zip').

OUTPUTS: The model default plots a number of figures within MATLAB, including but not limited to metapopulation abundance time-series of each stochastic simulation, histograms of all simulation abundances in pre and post-disturbance years of the user's choice, population density time-series of both the metapopulation and subpopulations. If optionally selected by the user, the model will create a subfolder with the prefix 'Output_' and a suffix that includes the date and time stamp at the time of model completion. This folder will save MATLAB variable files (.mat) of the input data structures and tables that define the parameters and inputs to the model, as well as the key population output variables (described in more detail below). Finally, a table with headers will be saved to 'Results.xls' that contains the mean, 5th and 95th percentile of population densities from the last *n* number of years (where *n* is prescribed by user) across all simulations for both the metapopulation and each subpopulation. For reference: the mean results from this table were used to create Figure 8 in the results of Murphy et al., (2020).

B. Descriptions of Excel Spreadsheet Inputs

The following describes how to set up the Excel spreadsheets necessary to read data into DyHDER. Example spreadsheets are included with the scripts that include the data inputs used in Murphy et al. (2020). To run the model, the desired input files must all be located in a single subfolder of the directory folder with a name that is the same as that prescribed by the input variable 'InputsFolder'.

1. Demographic Data

Filename: 'Demographics_*.xlsx' (*any suffix can be added)

Brief Description:

Contains all required subpopulation demographic data to build Lefkovitch matrices.

Contents:

Must have 6 sheets (do not alter sheet names):

1. 'Survival' – Apparent Survival Rates – The first column are the ϕ_i variables names representing survival rates for each of the *i* life-stages in the model, ordered from youngest to oldest and starting on the second row. The first row are the subpopulation site IDs (no spaces), starting on the second column. Matrix can be any size (i.e., number of life-stages and number of sites). Fill in estimated apparent survival rates for each stage at each site.
2. 'Transition' – Stage Transition Rates – The first column are the $\gamma_{i,i+1}$ variable names representing transition rates between the *i* life-stages in the model starting on the second row and ordered from youngest to oldest (number of variables = # of stages – 1). The first row are the subpopulation site IDs (no spaces), starting on the second column. Matrix should be one row smaller than 'Survival'. Fill in transition rates for each stage at each site.
3. 'Reproduction' – Reproduction Rates – The first column are the F_i variables names representing reproduction rates for each of the *i* life-stages in the model, ordered from youngest to oldest and starting on the second row. The first row are the subpopulation site IDs (no spaces), starting on the second column. Matrix should be the same size as 'Survival'. Fill in reproduction rates for each stage at each site.
4. 'Variance' – Temporal Variance – The first column, starting on the second row, are the $\sigma_{\phi i}$ variables names representing temporal variance (standard deviation) of survival rates for each of the *i* life-stages in the model ordered from youngest to oldest. This is immediately followed by the σ_{Fi} variables names representing temporal variance (standard deviation) of reproduction rates for each of the *i* life-stages in the model ordered from youngest to oldest. The first row are the subpopulation site IDs (no spaces), starting on the second column. Matrix should have twice as many data rows (i.e., not including headers) as 'Survival'. Fill in temporal variance for each respective rate at each site.
5. 'CarryingCapacity' – Site Carrying Capacities – The first column is just the *K* variable name on the second row. The first row are the subpopulation site IDs (no spaces), starting on the second column. Fill in the carrying capacities in the second row below each subpopulation site ID.
6. 'Offspring' – Density-Dependent Offspring Survival Estimates – The first column are the variables: $\phi_{01}(\frac{N}{K} = 1)$ and $\phi_{01}(\frac{N}{K} \approx 0)$, respectively, representing the two offspring survival probabilities to stage-1 as a function of local population density. The first row are the subpopulation site IDs (no spaces), starting on the second column. Fill in each offspring survival rate in their respective row below each subpopulation site ID.

2. Habitat Time-Series Data

Filename: Habitat_*.xlsx (*any suffix can be added – recommend using the habitat metric name, e.g., temperature)

Brief Description:

Contains multiple sheets (one for each site location) describing complete time-series for all habitat metrics to be modeled in a DyHDER simulation.

Contents:

A variable number of sheets depending on number of subpopulation sites in the model (each sheet must be named with respective site ID names – no spaces, consistent with headers in Demographics*.xlsx).

The first row of each sheet must contain consistent headers across all sheets.

The first column of the header row should be 'Years', and all subsequent columns (any number) contain the names of each of the habitat metric used in the model (no spaces). Starting in the second row, the first column ('Year') should iterate from 1:Inputs.dt:Inputs.tmax consistent with the input parameters of the DyHDER_master.m script.

All subsequent columns should then be filled with time-series data describing metrics of habitat condition (e.g., temperature, dissolved oxygen, gravel size) respective of the timesteps detailed in the first column. The habitat metrics can have any units, as long as they are consistent with those in the respective habitat suitability relations. The header names for these columns must also be consistent with habitat identifiers in other input data files, as they are used to link this data to the habitat suitability relations and to assign metrics during habitat-based adjustment (ex: 'Inputs.surv_adj'). For example, if temperature was a metric used in the model to adjust transition rates, then the following inputs should all contain the same ID identifier (ID bolded here for emphasis):

Header Name in Habitat_*.xlsx = **Temperature**

In DyHDER_master.m, input parameter 'Inputs.adv_adj' should then = {'**Temperature**'}

Habitat Suitability Filename = 'Suitability_**Temperature**.xlsx'

Header Row in Suitability_*.xlsx = **Temperature**

3. Habitat Suitability Relations

<u>Filename(s)</u>: Suitability_*.xlsx (*any suffix can be added but must be identical to the habitat metric name used in the header row Habitat_*.xlsx – there should be one file for each habitat metric)

<u>Brief Description</u>:

Each file contains data describing the relationship between habitat metric values and respective suitability values (0 – 1) specific to the modeled species and spanning the range of habitat metric values to be explored within the application of the model.

<u>Contents:</u>

Each file should contain just one sheet.

The second row of the first column must contain the habitat metric name (no spaces – identical to the header in 'Habitat_*.xlsx' file. All subsequent rows in the first column should contain the values of the independent variable (habitat metric values) for the specific habitat suitability relation.

The second column of the second row should contain the word "ALL" and then subsequent cells in this row should contain the stage-names of incrementally increasing stage (these should be the same names as the LifeStages cell array from DyHDER_master.m parameters).

There are three options for completing the Suitability files. The selected approach is critical as it directly affects how habitat adjustment is applied to different life-stages within the DyHDER model:

1) Habitat-Adjustment Applied Uniformly to All Life-stages
    - If there is only one habitat suitability relation that the user intends to apply uniformly to all modeled life-stages, then only the values of the dependent variable (suitability values) need to filled into the rows below the 'ALL' column header to correspond with the independent variable values in the first row.
2) Variable Habitat Suitability by Life-stage
    - If the user intends to apply different suitability relations for different life-stages, then the 'ALL' column should be left completely unfilled and every life-stage column should be filled with their respective suitability values.
3) Habitat-Adjustment to Some but not All Life-stages
    - If the user intends for habitat-based adjustments to apply only to some life-stages and not others, the suitability values in columns for the life-stages that are not affected should be filled in completely with a value of 1.

4. Site-to-Site Distances

Filename: Distance_*.xlsx (*any suffix can be added)

Brief Description:

This file contains a square distance matrix, with cell elements describing the path distance (d) between every subpopulation site within the model.

Contents:

This file should contain just one sheet following the format of the template. Columns in the matrix represent migration destinations, while rows represent origination. As a distance matrix, the main diagonal (i.e., distance between a site an itself) should be zeros and the upper and lower triangular parts of the matrix should be reflections of each other. The second row of the spreadsheet should list subpopulation site IDs (no spaces), starting in the third column. There may be any number of sites included, but the number of and ID names must match those in the Demographics_*.xlsx file. The second column of the spreadsheet should list subpopulation site IDs as well, starting in the third row. The vertical order of sites ID names should match those in row 2. Add or remove rows and columns as needed for your study, but the upper leftmost element of the matrix should always be located in the third row, third column of the spreadsheet. All internal elements of the matrix should be values of distance, in kilometers, representing the path travel distance between every origin (row) and destination (column) in the study.

5. Physical Site-to-Site Connectivity

Filename: Connectivity_*.xlsx (*any suffix can be used)

Brief Description:

This file contains a square matrix, formatted similar to the distance matrix, with cell elements describing the relative physical connectivity (c = 0 – 1) between every site within the model.

Contents:

This file should contain just one sheet following the format of the template. Columns in the matrix represent migration destinations, while rows represent origination – allowing for users to include directionally variable connectivity through the differentiation of upper and lower triangular parts of the matrix. The second row of the spreadsheet should list subpopulation site IDs (no spaces), starting in the third column. There may be any number of sites included, but the number of and ID names must match those in the Demographics_*.xlsx file. The second column of the spreadsheet should list subpopulation site IDs as well, starting in the third row. The vertical order of sites ID names should match those in row 2. Add or remove rows and columns as needed for your study, but the upper leftmost element of the matrix should always be located in the third row, third column of the spreadsheet. All internal elements of the matrix should be values between 0 and 1 and represent the relative degree of physical passage from origin (row) to destination (column), with 1 equaling unobstructed passage.

6. Dispersal Rates by Life-Stage

<u>Filename</u>: Dispersal_*.xlsx (*any suffix can be added)

<u>Brief Description</u>:

This file contains data describing the probability of dispersal for each life-stage (0 – 1) assuming conditions of optimal habitat suitability ($\alpha = 1$) and a subpopulation density at carrying capacity.

<u>Contents:</u>

This file should contain just one sheet following the format of the template. The second row of this spreadsheet should list the life-stage names in order from youngest to oldest, starting in the first column. You may add as many life-stages as required for your species and study. The third row of this spreadsheet should include the corresponding dispersal probability for each life-stage (as listed in row 2), assuming the conditions of optimal habitat suitability ($\alpha = 1$) and a subpopulation density at carrying capacity.

7. Life-Stage Dispersal Distance Scalars

Filename: ScalarDistance_*.xlsx (*any suffix can be added)

Brief Description:

This file contains data describing the distance scalar of migration (δ) for each life-stage based on the exponential migration function described in Murphy et al. (2020).

Contents:

This file should contain just one sheet following the format of the template. The second row of this spreadsheet should list the life-stage names in order from youngest to oldest, starting in the first column. You may add as many life-stages as required for your species and study. The third row of this spreadsheet should include the corresponding distance scalars for each life-stage (as listed in row 2), with units of length that must be consistent with those used in the distance matrix.

8. Stocking & Harvesting (*Optional*)

Filename: Stocking_*.xlsx (*any suffix can be added)

Brief Description:

This is an optional file, based on user study design. This file contains the data describing the location, year, life-stage, and abundance of individuals stocked or harvested within the simulation.

Contents:

This file should contain just one sheet following the format of the template. The first row of this spreadsheet should be a header row with 4 columns listing (in this order): 'Stocking Site', 'Model Year', 'Life-Stage', and 'Abundance'. Each subsequent row filled in below must include attributes for all columns. Stocking/harvesting in the same year and site but of different life-stages must be on separate lines with respective abundances listed for each. Stocking is applied through entering positive abundance values, and harvesting is applied through entering negative abundance values. There is no limit to the number of stocking/harvesting rows that can be input, but the 'Model Year' listed must match to a value in the array [1:Inputs.dt:Inputs.tmax] from the input parameters of the DyHDER_master.m script. Site IDs must exactly match those input to Demographics_*.xlsx file, and life-stage IDs must exactly match those input to model variable 'LifeStages' in DyHDER_master.m script.

The DyHDER model checks and compares prescribed harvest abundances against the present abundances for the specific life-stage in the target subpopulation site. If harvest abundances exceed present abundance, then abundance of that life-stage will be set equal to 0. However, at present DyHDER will provide no indication to the user if this scenario occurs.

In the order of operations within the DyHDER model, stocking/harvesting occurs after both population projection and metapopulation dispersal have been computed.

To include stocking/harvesting scenarios in model runs, Inputs.Stock must be set to = 1 in master script.

C. Setting Up A Model Run

The model is run using the 'DyHDER_Master.m' script. All model parameters (described below) can be adjusted in the opening sections of this annotated master script. A subfolder of Excel spreadsheets that define the model inputs (see included subfolder of examples: 'Inputs_LoganRiverExample') must be located within the same directory folder as the DyHDER_Master.m script. A subfolder of DyHDER function scripts (default name: 'DyHDER_functions') must also be located within the same directory folder as the DyHDER_Master.m script.

The following MATLAB variables must be modified in the opening sections of the master script: 'DyHDER_Master.m'. Users should go through and set these parameters before running the model. Some of these variables represent basic modeling parameters (e.g., number of simulations, length of simulations), some represent options for the population modeling (e.g., turn dispersal on or off), and others are responsible for prescribing key ecological factors for the system (e.g., type of density dependence model, habitat metric assignment). Users should be careful to check the parameters they set before each model run, as they will be saved based on those used in the previous run. In the descriptions below the appropriate data type is defined in brackets [ ] for each variable.

1. Model Parameters

InputsFolder: [string] : name of directory subfolder containing input spreadsheets

DyHDERFunctions: [string] : name of directory subfolder containing all subscripts

LifeStages: [cell array of comma-delimited strings] : list of life-stage names ordered from youngest to oldest (no spaces allowed)

Inputs.SimNum: [integer] : number of model simulations (if stochastic)

Inputs.tmax: [integer] : maximum model run time for each simulation, in years – value must be a multiple of Inputs.dt

Inputs.dt: [double] : size of each model timestep, in years – possible that bugs and instabilities could arise if using values other than 1

Inputs.PreDisturbance: [integer] : value between 1 and Inputs.tmax used to set the model year to evaluate the distribution of metapopulation abundance across all simulations – intended for characterizing abundance before a disturbance is introduced in habitat time-series data

Inputs.PostDisturbance: [integer] : value between Inputs.PreDisturbance and Inputs.tmax used to set the model year to evaluate the distribution of metapopulation abundance across all simulations – intended for characterizing abundance after a disturbance is introduced

Inputs.CompileYears: [integer] : the number of model timesteps, in years, desired to statistically characterize population abundances at the end of each simulation, e.g., if Inputs.tmax = 100 and Inputs.CompileYears = 25, then mean, 5th and 95th percentiles of population abundance are calculated from years 76:100 across all simulations

Inputs.Nx: [double] : fraction of the carrying capacity (for both sub- and metapopulation levels) that defines the quasi-extinction threshold for the model

Inputs.QEx: [integer] : absolute number of individuals (for both sub- and metapopulation levels) that defines the quasi-extinction threshold for the model

2. Model Options [On = 1, Off = 0]

Inputs.terminate: option to terminate a model simulation if the metapopulation abundance falls below the prescribed quasi-extinction threshold, which is determined by whichever of the two inputs (Inputs.Nx or Inputs.QEx) represents the larger number of individuals

Inputs.Halt: option to suspend population projection in any subpopulation if its abundance falls below the prescribed quasi-extinction threshold – if dispersal is on and the abundance increases above the threshold from immigration, projection will continue – included to avoid Allee effects

Inputs.EnvStochasticity: option to include stochasticity in model simulations

Inputs.Dispersal: option to allow dispersal between metapopulation sites

Inputs.Stock: option to include stocking and harvesting scenarios (if turned on, a spreadsheet of Stocking_*.xlsx must be included in the subfolder of input files)

Inputs.SaveData: option to save record of model inputs and output data to a new subfolder

3. Parameters for Population Dynamics & Habitat Assignment

Inputs.DD_model: [integer] : assign desired density dependence model, where there are currently two options that are prescribed through numerical assignment:

1 = Ricker Model
2 = Beverton-Holt Model

Inputs.fuzzytype: [string] : assign the desired fuzzy aggregation method for calculating habitat suitability index (α) and for aggregating variables if there are multiple habitat metrics influencing single demographic parameters. There are three options prescribed by respective strings:

- Product of suitability metrics = 'product'
- Minimum of suitability metrics = 'minimum'
- Geometric mean of suitability metrics = 'geomean'

For the following inputs, if no habitat metric is meant to influence a particular demographic rate, then user must still include cell array brackets but leave the array empty = { }. Additionally, the name and case of the input string used in these inputs must exactly match those from the spreadsheets.

Inputs.surv_adj: [cell array of comma-delimited strings] : habitat metrics influencing survival rates

Inputs.adv_adj: [cell array of comma-delimited strings] : habitat metrics influencing transition rates

Inputs.fec_adj: [cell array of comma-delimited strings] : habitat metrics influencing reproduction rates

Inputs.hsi_params: [cell array of comma-delimited strings] : habitat metrics aggregated to calculate the habitat suitability index at each site (α)

D. Model Data Outputs

If input option 'Inputs.SaveData' is set = 1 in the master script, then records of model input parameters and outputs will be saved to an output folder that is created within the directory and has an automatically-generated name that includes the date and time at the completion of the model run.

Inputs are saved to 'Inputs.mat' and include a suite of structures that make a record of:

1. Filenames: input filenames,
2. Inputs: user-selected model options and parameters,
3. Demographics: initial demographics for each subpopulation,
4. DispersalMetrics: dispersal parameters and matrices,
5. HabitatMetrics: time-series of habitat conditions for each site,
6. SuitabilityRelations: defined habitat suitability relations,
7. SuitabilityValues: site and life-stage specific time-series of suitability values created for each habitat metric,
8. HIS: site and life-stage specific time-series of the aggregate habitat suitability

Outputs are saved to 'Outputs.mat' and include a suite of structures of:

1. Population: time-series of the total metapopulation abundances through each simulation,
2. Record: the complete record for each site of time-series data of the adjusted Lefkovitch matrices, stage-based abundances, stage-based proportions, population density, growth rates calculated as interannual changes in abundance (.GrowthRate), as well as the dominant eigenvalue of the adjusted site matrix (.Lambda) for each simulation,
3. prob_ext: the probability of extinction calculated across all simulations