tardis

(Temporal Analysis of Relative Distances)
Software manual

TARDIS (Temporal Analysis of Relative Distances) is a software created to get unbiased and accurate single-particle tracking (spt) information directly from localization data. It is especially powerful in difficult spt conditions, such as heavily blinking particles, high density of localizations, or with a lot of additional noise. The full overview of the scientific implementation of the algorithms underlying TARDIS can be found in the manuscript:

"Temporal analysis of relative distances (TARDIS) is a robust, parameter-free alternative to single-particle tracking" by Martens et al: https://rdcu.be/dv1sr or doi.org/10.1038/s41592-023-02149-7

When using this software, please cite us as "Martens et al., Temporal analysis of relative distances (TARDIS) is a robust, parameter-free alternative to single-particle tracking, Nature Methods (2024)".

For more information about the TARDIS software, also read *Getting started with TARDIS*.

Version history
V1.21.1 (2024-01-18): anaDDA-related bugfixes
V1.21 (2024-01-15): Updates to article links
V1.20 (2023-11-22): Visual updates, bugfixes
V1.10 (2023-09-20): Added custom PDF fitting and spatialTARDIS
V1.03 (2023-08-01): Improved prior parameter obtaining
V1.01 (2023-04-03): Initial version

Contents

Required toolboxes
- Signal Processing Toolbox
- Curve Fitting Toolbox
- Optimization Toolbox

User interface overview – basic

The TARDIS UI can be opened by using the shortcut created after installation, or from within MATLAB by running **TARDIS_app**.

After opening the UI for the first time, a screen similar to this will appear (without results). To start an analysis, briefly follow the following four steps.
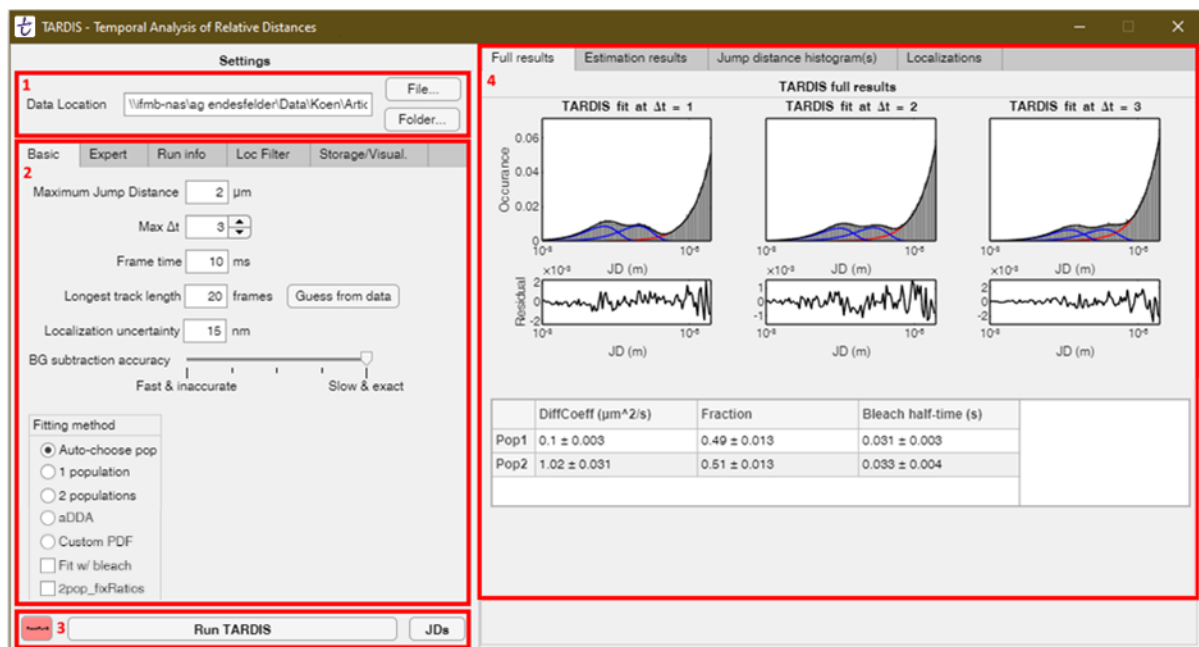
Overview of the TARDIS UI. **(1)** Area to select one or more input datasets. **(2)** Area to choose settings, and look at the run info during analysis. **(3)** Buttons to run TARDIS. **(4)** Results tab.

1. **Choose location for the data.**
   - The data should be either a **.csv containing ThunderSTORM** output (with at least column headings 'frame', 'x [nm]' and 'y [nm]'), a **.csv containing DECODE** output (with at least columns 'x', 'y', 'frame_ix') or a **.mat file containing a 'pos' array** with at least 3 columns (1st column frame, 2nd column x position, 3rd column y position; positions should be in meter units).
   - You can find a file with the 'File…' button, or alternatively do batch-analysis on a whole folder via the 'Folder…' button.
     - After choosing a .csv file, the Loc Filter tab will open, allowing you to filter the localizations if you so want.
2. **Choose the correct settings**
   - There are five tabs at the top. Only the **'Basic'** settings are normally required to be changed between experiments, as these have to do with experiment-dependant

settings. The **'Expert'** tab is normally not required. The **'Storage/Visual.'** tab provides some options for visualisation and storage.

- o Note that you can hover over the different settings with the mouse to get a brief overview of the parameter.

- Settings will be stored between uses of TARDIS.
- Here is a brief description of the settings that you should likely change, or at least look at in the 'Basic' tab:
  - o **Maximum Jump Distance**: the maximum jump distance that TARDIS analyses. This should be larger (at least approximately 50% higher than) the maximum expected 'true' jump distance in the experiment.
  - o **Max Δt**: The maximum frame-delay over which TARDIS should run.
  - o **Frame time**: The frame time in milli-seconds.
  - o **Longest track length**: What you expect is the longest track in your dataset. Precisely, TARDIS will start creating the inter-emitter (i.e. 'background') dataset from frame-delays bigger than this value. It's better to over-estimate than under-estimate this. You could use the 'Guess from data' button to estimate this from the data via a Wilcoxon statistical test (see the manuscript for more detail). Note that this guessing might take up to 1-5 minutes.
  - o **Localization uncertainty**: The (expected) localization uncertainty of your dataset in nm
  - o **BG subtraction accuracy:** How many Δt-bins (starting at the 'longest track length') will be used. More bins will give smoother results, but also longer analysis times. This scale is logarithmic, and scales from 1 Δt to 50 Δt.
  - o **Fitting method**: This determines which analytical formula will be fitted. TARDIS currently supports single- and double-populations fit, as well as analytical diffusion distribution analysis (aDDA).
    - o Alternatively, you could opt for the Auto-choose pop option, which tests both 1 and 2 populations, and gives the result with the best adjusted R^2 value (more details in manuscript).
    - o If aDDA is chosen, a new sub-menu for aDDA will be opened for aDDA-specific settings, such as number of species and possible cellular confinement
    - o If CustomPDF is chosen (expert setting), a new sub-menu appears. In this submenu, you can choose a custom PDF, and give start/lower/upper fitting bounds for this custom PDF. Please open and look through CustomPDF_Example_1pop.m for detailed info on how to use this. **Note**: CustomPDF does not work with Windows-installation, only with running TARDIS(-app) via MATLAB.
    - o The 'Fit w/ bleach' option enables TARDIS to assume (and fit) a uniform bleaching of all populations.
    - o The 2pop_fixRatios checkmark can be marked if the bleach rate is constant for both populations (assuming 2 populations are fitted).

3. **Run TARDIS**

- TARDIS analysis starts by pressing the 'Run TARDIS' button.
  - o Alternatively, the 'JDs' button only subtracts the inter-emitter (i.e. 'background') data from the full data, to give a good estimation on the intra-emitter (i.e. 'true') jump distance distribution.
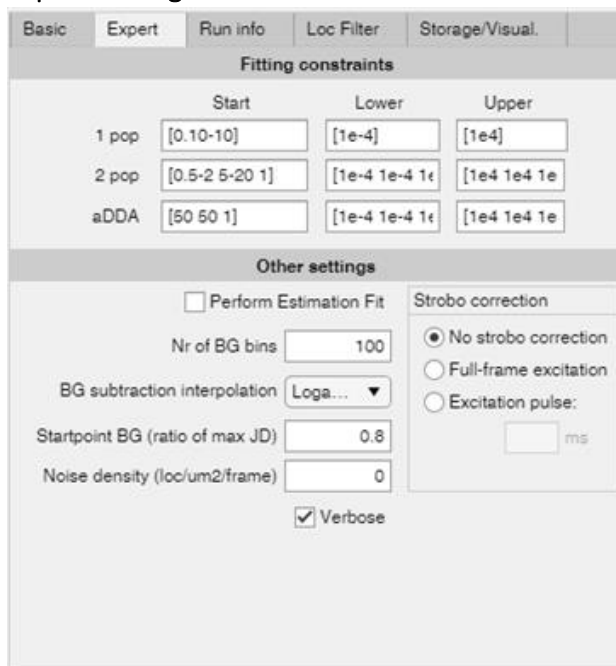
- o The red button can be pressed if MATLAB has crashed, and the UI needs to be reset. Please note that pressing this button does not interrupt the analysis.
- As soon as analysis starts, the results-section (on the right) will be cleared, and will show the localizations found in the loaded file. This can be a quick check that the correct file is chosen.
  - o While TARDIS is running, the **'Run info'** tab will show info of the fitting procedure.
- After TARDIS has been completed, the output will be presented in the results-section (number 4 in the image above).
- Additionally, TARDIS result data will be stored in a 'TARDIS_Results' subfolder at the location of the data, based on the settings in the **'Storage/Visual.'** tab
  - o It is recommended that you at least store the JSON fit data (containing the numerical final fit details), and/or the MAT fit data (similar to the JSON fit data).
  - o Additionally, it is recommended to store the intermediate fit and final fit images, which are always stored both as .png and .svg.
  - o If wanted, you could also store a very big .mat file containing all data ('store all data'; such as all jump distances). This .mat file will be substantially larger than the original localization data size, and can go up to GB sizes.

4. **Review results**
   - The result-section is subdivided in four parts.
     - o The **'Full results'** tab will provide the final TARDIS results, and gives both a visual fit (top), and a quantitative result of the fitting parameters (bottom).
     - o The **'Estimation results'** tab will provide some details on a first, estimative fit. This estimation is only used to obtain starting parameters for the final fit to prevent fitting local maxima.
     - o The **'Jump distance histogram(s)'** provides an overview of the jump distances found by TARDIS after a first estimation of the inter-emitter (i.e. 'background') population.
     - o The **'Localizations'** tab provides an overview of the localizations loaded in the data, and can act as a quick check that the data is loaded correctly.

User interface overview – continued

Expert settings

The settings that can be changed in the expert tab rarely have an influence on TARDIS uses, but nonetheless can be changed by the user.

Fitting constraints: used to determine start values and lower/upper bounds. Note that a dash (i.e. 0.1-20) can be used to have a number randomly chosen between those values. All values need to be between square brackets as shown above. Variables are ordered as follows:

- 1 population: [Diffusion coefficient]
- 2 populations: [Diffusion coefficient 1; Diffusion coefficient 2; ratio between populations]
- aDDA: [k_{on}; k_{off}; Diffusion coefficient]

Other settings: other settings used in TARDIS

- **Estimation fit:** an estimation fit can be done before doing a full fit. This does not translate to better fitting results, but could result in a better start point for the final fit, and/or decrease analysis time.
- **Nr of BG bins:** The amount of bins used for determining the background. Does not have a strong effect on results
- **BG subtraction interpolation**: what interpolation is used between background bins. Does not have a strong effect on results
- **Startpoint BG (ratio of max JD)**: at which point TARDIS assumes only background information, **only** for the estimation fit. A value of 0.8 would mean that TARDIS assumes that everything higher than 80% of the max distance (basic settings) isn't diffusional data.
- **Verbose:** whether or not textual output is given in the UI.
- **Strobo correction:** Correction for stroboscopic illumination according to Berglund 2010.

Run info

Provides basic verbose information on the status of TARDIS during the analysis.

Loc Filter

Allows for basic filtering of localizations before analysed by TARDIS. Names in the editable fields should be exact matches for the data headers. Note that for ThunderSTORM, this requires the syntax as used above.

Storage/Visusalisation

This tab should be used to determine what data should be saved after analysis. The data will be stored in a sub-folder in the data folder, called 'TARDIS_Results'. Please note that data will be overwritten if TARDIS is run multiple times on the same dataset.

- **Store JSON fit data**: Store the results of the analysis as a JSON file
- **Store MAT fit data:** Store the results of the fits as a MAT file. This should be used if you would want to repeat the visualisation of the results at a later time
- **Store all data**: Storing all data used by TARDIS, also during its analysis. This is not recommended since it will create datasets much larger than the original dataset's size.
- **Store JD histogram image**: Store the image showing the jumping distances without applying any fitting as a PNG and SVG image
- **Store intermediate fit image**: Store the estimation fit (see expert tab) as a PNG and SVG image
- **Store final fit image:** Store the final fit as a PNG and SVG image
- **Store swift/MHT parameters:** Store expected swift/MHT starting parameters from the TARDIS analysis (see the manuscript for more details)

Additionally, you can choose whether the visualisation is shown on a linear or logarithmic axis (does not influence the fit).

Finally, you can re-visualise previous TARDIS runs if the MAT fit data is saved by using the 'Visualise old results' button.

Command-line implementation

TARDIS can be incorporated in MATLAB scripts with the following syntax. Alternatively, you can follow the syntax in **`Call_TARDIS_example.m`**.

In general, the following workflow should be adhered to:

1. Load a dataset

An array (e.g. 'pos') with columns 'frame' (integer values), 'x', 'y' coordinates (both in meter units) should be loaded.

2. Define TARDIS settings

TARDIS settings are easily defined by running **`GenerateTARDISsettings()`**. In that script, all settings are initialised and shortly explained, but are also fully explained in Appendix 1.

3. Run TARDIS

TARDIS can then simply be run by calling **`TARDIS(pos, TARDISsettings).`** Alternatively, you could run **`[parameters, parametersCI] = TARDIS(pos, TARDISsettings)`** to store the fit parameters and confidence interval in the set output. For a full exploration of the output parameters, read the documentation on the TARDIS function.

4. Re-visualise TARDIS output

Depending on the TARDIS settings, running TARDIS as in the previous step will result in visual output (when using the default settings, visualisation is performed). Alternatively, you could run **`Visualisation_FF([],[],Visualisation_FF_outputCell)`** to repeat the visualisation, where **`Visualisation_FF_outputCell`** is one of the outputs of TARDIS.

Appendix 1: Full description of TARDIS input, settings and default values

Required input: poslist and settings structure

Poslist should be a N-by-4 matrix with columns indicating [frame]-[x position]-[y position]-[z position]. The frame index should start at 1. X and Y positions should be in the meter-unit. The z-position is not used in the current TARDIS implementation, so this can be defaulted to a list of zeros.

Settings structure

The settings structure should contain the following fields. If they are not found, TARDIS defaults these to the default value specified below (specifically, determined in lines 32-80 of URDA_HO_FF_function.m).

The settings are ordered by importance:

- **Crucial** indicates you should take a look at it and likely set them specifically for your experimental conditions, because they are often directly involved in getting the correct fitting etc out.
- **High** can have a big effect on performance and results, but are normally fine in typical sptPALM conditions.
- **Medium** are settings that only have an effect in specific conditions.
- **Low** indicate settings that have very little effect throughout TARDIS testing and/or involve debugging information.
- **Verbose** indicates verbose (i.e. textual/visual output or debugging) related settings.

Setting name	Description	Default value	Importance
`frame_time`	Frame time of the spt experiment in seconds	0.01	Crucial
`loc_unc`	Localization uncertainty (estimation) of the spt experiment. Unit is meters.	15e-9	Crucial
`populations`	Number of populations to be fitted. This is overwritten by AutoChoosePop if that is True. Accepted values are 1, 2 (for 1 or 2 populations), and 0 (for aDDA fitting).	1	Crucial
`AutoChoosePop`	Boolean whether or not an adjusted R^2 is calculated for both 1 and 2 population fits, after which the best fit is chosen and further used. Note that this calculates both fits, and thus increases analysis time.	False	Crucial
`frame_dist_BG`	An array of frame-delay(s) that are used to find the inter-emitter (i.e. 'background') distribution. These should be frame-delays where it is expected that no true track is present. If for example you expect that the longest track you have in your dataset will be 20 frames long, this array should start at minimum at 21 frame-delays. While a single frame-delay is good enough for TARDIS, adding more frame-delays increases the	[20:40]	Crucial

	available data for inter-emitter curve determination. Adding more data logically also decreases analysis rate. The minimum value of this can also be determined statistically via the script BGLengthDetermination_Wilcoxon.m		
dt_arr	Array describing which frame-delays should be used for fitting. It is a [1-by-N]-array with increasing values, starting from 1.	[1 2 3]	Crucial
performsecondfit	Boolean whether or not the maximum-likelihood estimation fit should be done. If this is set to false, only the estimation fit is performed (assuming **performestimationfit** is set to true). If both are set to false, no fitting is performed.	True	Crucial
performestimationfit	Boolean whether or not an estimation fit should be done. For an estimation fit, the found inter-emitter distribution (i.e. 'background') is simply subtracted from the original jump-distance histogram, and the resulting histogram is fitted. The results from this estimation fit are then used in the second (maximum-likelihood estimation) fit.	False	High
maxdist	The maximum distance TARDIS will analyse. It is important that **maxdist** will be longer than the longest jump distance in the spt dataset at the highest analysed frame-delay. If this value is too low, the fitted jump-distance curve can be confined to too low values. While there is no real downside to having this value too large, it is not recommended to increase this further than ~4x maximum jump distance in the spt data.	3e-6	High
fitWithBleach	Boolean whether or not the decrease in localizations over dt-bins is fitted with bleaching-information, or left free for fitting. If this is set to True, a starting bleach value randomly chosen between 0.01 a 0.1 frame is used as starting point, with lower and upper bounds of 0.0001 and 5, respectively. These can be changed in the changeFitParams_populations_Bleach.m script.	False	High
fixRatios2pop	Boolean whether or not the 2 populations have the same bleaching rate. Only applicable when both **fitWithBleach == True** and **populations == 2** (or **AutoChoosePop == True**)	False	High
stroboFrameTime	Duration of stroboscopic excitation. Should be lower than (or equal to) frame_time. Alternatively, use '0' for no	0	High

		stroboscopic frame-time correction, or '-1' for full-frame stroboscopic duration correction.		
`startpointBG`		The distance (in m) where TARDIS expects only inter-emitter information. This is used for the first 'estimation' step, but in the final fitting, it is allowed to have intra-emitter (i.e. 'true positive') jump distance data in this regime.	1e-6	Medium
`start_1pop`		Starting positions of fitting a 1-population fit. Array containing [Diffusion coefficient in um2/s]. Alternatively, random start position(s) between two extremes can be indicated by [X-Y], indicating that a random start position between values X and Y is chosen. This value should be a string.	'[0.1-20]'	Medium
`lb_1pop`		Lower bound of all values specified in start_1pop. This value should be a string.	'[1e-5]'	Medium
`ub_1pop`		Upper bound of all values specified in start_1pop. This value should be a string.	'[50]'	Medium
`start_2pop`		Starting positions of fitting a 2-population fit. Array containing [Diffusion coefficient in um2/s – Diffusion coefficient in um2/s – Ratio between the populations]. For the ratio, a value of 1 would mean 1:1 ratio between the populations, a value of 0.1 would mean a 1:10 ratio, etc. Alternatively, random start position(s) between two extremes can be indicated by [X-Y], indicating that a random start position between values X and Y is chosen. This value should be a string.	'[0.1-20 0.1-20 0.1-10]'	Medium
`lb_2pop`		Lower bound of all values specified in start_2pop. This value should be a string.	'[1e-5 1e-5 1e-5]'	Medium
`ub_2pop`		Upper bound of all values specified in start_2pop. This value should be a string.	'[50 50 100]'	Medium
`start_3pop`		Not (yet) implemented	-	
`lb_3pop`		Not (yet) implemented	-	
`ub_3pop`		Not (yet) implemented	-	
`start_aDDA`		Start positions of an analytical Diffusion Distribution Analysis (aDDA) fit. Array containing [k_{on} - k_{off} - D_{free} (in μm^2/s)]. Alternatively, random start position(s) between two extremes can be indicated by [X-Y], indicating that a random start position between values X and Y is chosen. This value should be a string.	'[10-80 10-80 0.1-10]'	Medium
`lb_aDDA`		Lower bound of all values specified in start_aDDA. This value should be a string.	'[1e-5 1e-5 1e-5]'	Medium
`ub_aDDA`		Upper bound of all values specified in start_aDDA. This value should be a string.	'[1e4 1e4 1e2]'	Medium

`stroboFrameTime`	Timing of stroboscopic illumination (i.e. illumination for a shorter time than the frame time). This is used to correct diffusion coefficients according to Berglund, 2010. Set to 0 for no stroboscopic illumination (i.e. ignoring this factor), -1 for full-frame illumination, or a specific time in seconds.	0	Medium
`AlternativeLookupPosList`	A lookup list, alternative to the required poslist. If AlternativeLookupPosList is different than poslist, TARDIS will link between poslist and AlternativeLookupPosList – e.g. if using a sub-FoV, TARDIS in this manner can lookup all distances starting in a sub-FoV, but linked to the whole FoV. If left at 0, it's ignored, and poslist is used as lookup-poslist.	0	Medium, Added in v1.10
`customPDF`	Information about a customPDF, if a customPDF should be used. For example, use 'CustomPDF_Example_1pop' to run that CustomPDF (Found in Main-Functions-Core-Fitting Functions-CustomPDFs). Also see customPDF_start,_upper,_lower. Leave 0 for no custom PDF.	0	Medium, added in V1.10
`customPDF_start`	Start fit parameters for customPDF. Size depends on customPDF.	[]	Medium, added in V1.10
`customPDF_upper`	Upper bound fit parameters for customPDF. Size depends on customPDF.	[]	Medium, added in V1.10
`customPDF_lower`	Lower bound fit parameters for customPDF. Size depends on customPDF.	[]	Medium, added in V1.10
`freefit_locunc`	Experimental! Tries to fit localization uncertainty rather than uses the one provided. Only applicable with D < 0.3 μm^2/s or so. Only works when using log BG subtract. Boolean value	False	Experimental, Medium
`bgbinningnr`	Number of bins used in the inter-emitter (i.e. 'background') distribution. Since the inter-emitter distribution cannot be described analytically, a list of [x,y] coordinates is used to describe this distribution. This setting indicates how many [x,y] points are used. There is very little effect on using > 100 points, but also very little downside	100	Low
`linorlogBGsubtract`	A string that determines whether the inter-emitter (i.e. 'background') distribution will be sub-divided in logarithmically or linearly spaced bins. Should be either 'lin' or 'log'. Also look at `bgbinningnr`.	'log'	Low

`minlogpoint`	Starting point of logarithmic bins if `linorlogBGsubtract` is set to 'log'. No real reason to change this to a different value, unless you expect very very low JDs. Units in m.	10^-8.5	Low
`norm_bins`	TARDIS needs to perform some normalization somewhen. This is the number of bins used for this normalization. Normally no reason to change this, it seems to work with everything above ~1000.	50000	Low
`createJDsonly`	Only create a JD list of the intra-emitter (i.e. 'True') jump-distances, rather than trying to fit a distribution. Overwrites `performestimationfit` and `performsecondfit`.	False	Low
`verbose`	Whether or not the software outputs basic verbose information to the MATLAB command line or TARDIS UI.	True	Low
`verboseReal`	Whether or not the software outputs basic verbose information about the fitting to the MATLAB command line or TARDIS UI.	False	verbose
`StoreSWIFTparameters`	Whether or not TARDIS outputs a JSON with first-guesses for SWIFT analysis (Endesfelder, work in progress)	False	verbose
`debug`	Generating debugging information. Might be useful to pinpoint problems. Not used for general analysis.	False	verbose
`vis`	Visibility of the subtraction of inter-emitter distribution in separate figure plots.	True	verbose
`fitvisHO`	Visualisation of estimation fit in new (separate) figure	False	verbose
`visualisationMLEIntFit`	Boolean whether or not a separate image of the final fit should be created.	False	verbose
`callfromUI`	Either 'False' when run from command line or scripts, or the full information on the UI if run from the UI.	False	verbose