

## MSI-tool

**- a tool for calculating Multi Species Indices and trends therein by Monte Carlo simulation -**

Statistics Netherlands  
July 2016

*The tool is an R-script that combines the yearly indices of species into Multi Species Indices (MSI). In addition, a linear and smooth trend through the MSI are calculated and the trends are classified in terms like 'strong increase', 'stable' etc. A detailed description of the algorithm can be found in Soldaat et al. (subm.). To calculate an MSI, three steps are to be taken: (A) create the data file, (B) specify the settings of the analysis and (C) run the tool.*

### A. Create the input file

The input file is a text file with at least the columns 'species', 'year', 'index' and 'se'.  
For instance:

species	year	index	se
1340	2005	100	0
1340	2006	110	10
1340	2007	120	15
1340	2008	125	20
1340	2009	120	10
1340	2010	130	25
3040	2005	100	0
3040	2006	95	10
3040	2007	90	5
3040	2008	100	15
3040	2009	85	10
3040	2010	NA	NA
...	...	...	...
...	...	...	...

- Missing values may be left out OR should be inserted with values for species code and for year and 'NA' in the columns 'index' and 'se'.
- The base year (with index 100) may differ between species and has standard error zero. If the indices are expressed with index 1 in the base year, all indices and standard errors should be multiplied by 100.
- Species codes may also be names, but without a space (*Passerdomesticus* instead of *Passer domesticus*).

### B. Specify settings of the analysis

Under "Data import and specifications" in the first lines of the script the following parameters are set:

1. setwd: Specify the path for input and output files. Dependent on your ITC-environment a forward (/) or backward (\) slash is needed to separate folders.
2. rdata: Specify the name of the input file.
3. jobname: Specify a name that will be used in the output files.

4. nsim: Specify the number of Monte Carlo simulations. Usually 1000 simulations will give reliable results, but with 10.000 simulations the accuracy and reproducibility will increase (at the expense of longer calculation time).
5. SEbaseyear: Specify the year for which the standard error of the MSI should be zero. Usually the first year of the time series will be chosen, but any year is allowed.
6. plotbaseyear: Specify the year for which the index in the plot should be set to 100. This is not necessarily the same year as SEbaseyear.
7. index\_smooth: Specify whether the MSI (INDEX) or the smoothed trend value (SMOOTH) in plotbaseyear should be set to 100. Specifying the smoothed trend value will make the plot less sensitive to the annual variation in the MSI.
8. lastyears: Specify the number of years at the end of the time series for which you wish a separate trend estimate.
9. maxCV: Specify the maximum coefficient of variation (CV) of the indices of a species that is allowed. Species with higher CVs will not be included in the MSI. This is only a rough method to exclude species with probably unreliable indices. The CV is calculated only for indices >10, as small indices may have very high CVs without having strong influence on the confidence interval of the trends in the MSI. It is recommended to start with a high value for maxCV, e.g. 10 (which will usually include all species) and then inspect the plot with CVs that is produced in R in the command line:  

```
plot(uspecies, mnCV, type="p", pch=19, col="black", main=jobname, xlab="species code", ylab="meanCV")
```

This can be done by running the script up to this line. Decide whether species with high CV should be excluded or not. Comparing the effect of excluding a species on the MSI and its CIs by running the script with and without a species may help decide.
10. changepoint: Specify a year for which trends before and after that year should be calculated and compared.
11. truncfac: Specify the maximal year-to-year ratio in indices of a species that is allowed. This option is built in to level off undesired effects of highly variable yearly indices. Such indices may indicate unreliable estimates or indices of eruptive species. In the Living Planet Index a truncfac of 10 is applied.
12. TRUNC: Specify the minimal index value of the species that is allowed. Lower indices will be truncated to TRUNC. Low indices (e.g. <1) indicate marginal population abundances, but may have undesired strong effect on the MSI. For instance, a species increasing from 1 to 2 individuals has the same (and probably undesired) effect as a species changing from 100 to 200 individuals.

### **C. Run the MSI-tool:**

CTRL+A (select all)

CTRL+ENTER (run the selected script)

## OUTPUT

After running the tool, which dependent on the number of species and the number of simulations may take several minutes, the MSI-tool produces the following output files:

### 1. jobname\_RESULTS.csv

A csv-file with the main results: the MSI with standard error and confidence interval, the smoothed trend with its confidence interval and an indication of the difference between each year and the last year. In the example the SEbaseyear is set to 1990 (standard error 0), but the MSI in 1990 is not 100, because all results are rescaled by setting the smoothed trend value in 1990 to 100 (plotbaseyear = 1990 / index\_smooth = SMOOTH).

year	MSI	sd_MSI	lower_CL_MSI	upper_CL_MSI	Trend	lower_CL_trend	upper_CL_trend	trend_class
1990	105.88	0	105.88	105.88	100	97.24	103.25	moderate_decline
1991	93.15	2.86	87.71	98.93	98.99	97.24	101.21	moderate_decline
1992	97.46	3	91.76	103.52	97.99	96.27	100.2	moderate_decline
1993	98.02	2.96	92.38	104	96.96	95.31	99.2	moderate_decline
1994	95.57	2.71	90.4	101.03	95.92	94.36	98.22	moderate_decline
1995	91.45	2.64	86.43	96.77	94.86	92.5	97.24	moderate_decline
1996	89.26	2.63	84.27	94.56	93.67	91.58	95.31	moderate_decline
1997	96.53	2.69	91.41	101.94	92.47	90.66	94.36	moderate_decline
1998	94.82	2.62	89.83	100.1	91.25	89.76	93.43	moderate_decline
1999	89.61	2.65	84.57	94.95	89.78	87.99	91.58	moderate_decline
2000	89.68	2.53	84.85	94.79	87.98	86.24	89.76	moderate_decline
2001	84.68	2.39	80.11	89.49	85.96	84.54	87.99	moderate_decline
2002	81.46	2.34	77	86.17	83.78	82.04	85.38	moderate_decline
2003	79.2	2.22	74.95	83.69	81.12	79.61	82.86	moderate_decline
2004	80.36	2.35	75.89	85.09	78.17	76.49	80.41	moderate_decline
2005	80.98	2.52	76.19	86.07	75.69	73.49	77.26	moderate_decline
2006	71.61	2.56	66.76	76.82	73.82	72.04	75.73	moderate_decline
2007	70.11	2.49	65.4	75.15	72.14	69.91	74.23	moderate_decline
2008	69.57	2.62	64.63	74.89	70.55	68.52	72.76	moderate_decline
2009	66.89	2.71	61.78	72.41	69.05	67.17	70.61	moderate_decline
2010	65.07	2.96	59.51	71.14	67.65	65.84	69.21	moderate_decline
2011	75.3	2.34	70.86	80.03	66.39	64.53	67.84	moderate_decline
2012	66.74	2.03	62.88	70.84	65.27	63.89	67.17	moderate_decline
2013	59.11	1.82	55.66	62.79	64.27	62	66.5	stable
2014	64.57	1.96	60.84	68.53	63.38	60.77	65.84	NA

### 2. jobname\_SIMTRENDS.csv

A csv-file containing a long-term (V1) and short-term (V2) additive trend value for each simulation. These values can be used to test for differences with another MSI (not included in the script).

V1	V2
-0,0217	-0,0208
-0,0219	-0,0174
-0,02	-0,0233
-0,0205	-0,0283
-0,0204	-0,0177
-0,0213	-0,0217
-0,021	-0,022
....	....
....	....
....	....

### 3. jobname\_TRENDS.csv

A csv-file containing all linear trend estimates. In addition, the percentage change (based on the smoothed trend value) is given and tested for the whole time series and for the last X years (the -36.6% change is the difference between trend values 63.38 in 2014 and 100 in 1990 in jobname\_RESULTS).

	value	significance
overall trend	0,9793	moderate decline
SE overall trend	8,00E-04	
trend last years	0,9795	moderate decline
SE trend last years	0,0035	
change point	2003	n.s.
trend before change point	0,9841	moderate decline
SE trend before change point	0,0017	
trend after change point	0,978	moderate decline
SE trend after change point	0,0025	
% change	-36,599	p<0.01
SE % change	1,647	
% change last years	-12,308	p<0.01
SE % change last years	1,861	

#### 4. jobname\_GRAPH.jpg

A jpg-file showing the MSIs (dots) and the smoothed trend. In the example below the plotbaseyear is set to 1990 and the index\_smooth parameter is set to SMOOTH (the smoothed trend value in 1990 is rescaled to 100).

