Technical report
The LPR Belgium Project

Lead author: Olga Szczodry

Input and review: Hilde Eggermont, Marc Herremans, Jean-Yves Paquet, Sofie Ruysschaert, Hendrik Segers

Version 9 September 2020

Glossary

- BD: Biodiversity
- BE: Belgium
- NGO: non-governmental organization
- LPR: Living Planet report
- LPI: Living Planet index
- NL: The Netherlands
- NP: Natuurpunt
- PCM: Project Cycle Management
- RBINS: Royal Belgian Institute for Natural Sciences
- MoU: Memorandum of Understanding
- SOM: Site-Occupancy Modelling
- TRIM: Trends and Indices for monitoring data
- FRL: Flemish Red List
- EBCC: European Birds Census Council
- STI: Species Temperature Index

STEP 6: DATA MAPPING

The first step into developing an LPR was to answer the question: *What data do we have in Belgium?* For this purpose, the different BD datasets existing in Belgium were mapped. The resulting data mapping exercise was presented for the first time during the kick-off workshop (07/09/2018): it presented per taxonomic group: the existing data sets, the region concerned, type of data, time period and analysis of the limitations.

6.1 BIRDS

							Flaws						
project	Flanders	Wallonia	Brussels	type	start	years	Small sample size	Representative ?	Area not constant	No protocol	Treatment	Method changed	Short time span
Common breeding bird monitoring	ABV	x	STOC	abundance	1990/2007	11 / 29			x				x
Rare bird monitoring	BBV	x	x	abundance	?	?	x	x	x	x			
migration counts	x	x	x	meas. effort	1980	39		x	x	x		x	
ringing data ?	x	x	x	meas. effort	c.1960	c.60		x	x	x	x	x	
winter counts	PTT			abundance	1989	30	x	x	x				
waterbirds	IWC	IWC	IWC	trends	1992	27			x	x			
single species: Peregrine	x	x	x	breed. pairs	1997	20				x	x		x
Black grouse		x		breed. pairs	c.1970	c.50			x	x	x		
Eagle owl	x	x		breed. pairs	c.1980	c.40			x	x			x
Barn owl	x	x	x	breed. pairs				x	x	x	x		
Little owl	x	x	x	breed. pairs				x	x	x	x		x
House sparrow	x			abundance				x	x				x
Corn bunting	x			singing males		c.10			x	x			x
House martin	x	x	x	nests	c.1990	c.30		x	x				x
swallows	x			nests				x	x				x
Red List	x	x	x	trend	c.1990	c.30						x	
EU Birds Directive report	x	x	x	trend	c1973	45							x
repeated atlasses	x	x		occur./abund.	c.1990	c.30		x				x	
waarnemingen.be/observation.be	x	x	x	occupancy	2008	10		x	x	x			x

Figure 4 Result of the data mapping exercise for Birds, as presented in September 2018

6.2 MAMMALS

							Flaws						
project	Flanders	Wallonia	Brussels	type	start	years	Small sample size	Representative ?	Area not constant	No protocol	Treatment	Method changed	Short time span
single species: Hamster	x	x		estim. abund.					x	x	x	x	x
Badger		x		dens					x	x			
Dormouse	x	x		abundance					x		x		x
Edible dormouse	x	x		abundance							x		x
martens	x	x		occurrence				x	x	x			x
bats: winter counts	x	x	x	abundance				x	x				
summer counts	x	x	x	relat. abund.			x	x	x				x
Red List	x			trend									
repeated atlasses	x		x	occurrence	1987	30		x				x	
owl pellet analyses	x	x		relat. abund.	2008	10		x	x				x
waarnemingen.be/observation.be	x	x	x	occupancy	2008	10		x	x	x			x

Figure 5: Result of the data mapping exercise for Mammals, as presented in September 2018

6.3 REPTILES AND AMPHIBIANS

							Flaws						
project	Flanders	Wallonia	Brussels	type	start	years	Small sample size	Representative ?	Area not constant	No protocol	Treatment	Method changed	Short time span
trap monitoring	x			abundance	ca. 1970	50	x	x	x				
"paddenoverzet"	x			abundance	1988	30		x	x	x	x		
rare species		x		populations									
Red List	x			trend									
repeated atlasses	x			occurrence	1965	ca. 55		x	x	x		x	
waarnemingen.be/observation.be	x	x	x	occupancy	2008	10		x	x	x			x

Figure 6: Result of the data mapping exercise for Reptiles and Amphibians, as presented in September 2018 (yellow indicates important gaps)

6.4 FISH

							Flaws						
project	Flanders	Wallonia	Brussels	type	start	years	Small sample size	Representative ?	Area not constant	No protocol	Treatment	Method changed	Short time span
monitoring network (INBO)	x			abundance									
Red List	x			trend									

Figure 7: Result of the data mapping exercise for Reptiles and Fish, as presented in September 2018 (yellow indicates important gaps)

6.5 BUTTERFLIES AND MOTHS

							Flaws						
project	Flanders	Wallonia	Brussels	type	start	years	Small sample size	Representative ?	Area not constant	No protocol	Treatment	Method changed	Short time span
diurnal butterflies:													
montly garden counts	x			abundance	2007	11	x	x	x	x	x		x
Red List	x			trend									
repeated atlasses	x			occurrence	1965	ca. 55		x	x	x		x	
moths: monitoring network	x			abundance	2008	10	x	x	x				x
waarnemingen.be/observation.be	x	x	x	occupancy	2008	10		x	x	x			x

Figure 8: Result of the data mapping exercise for Butterflies and Moths, as presented in September 2018 (yellow indicates important gaps)

6.6 INVERTEBRATES

							Flaws						
project	Flanders	Wallonia	Brussels	type	start	years	Small sample size	Representative ?	Area not constant	No protocol	Treatment	Method changed	Short time span
DRAGONFLIES													
Red List	x			trends									
repeated atlasses	x			occurrence	ca 1970	ca. 50			x	x			
BEES													
BELBEES	x	x	x	occurrence					x	x			
KBIN (Alain Pauly)	x	x	x	occurrence				x	?				
private (M. Jacobs & K. Janssen)	x	x	x	occurrence		15		x	x	x		x	x
GRASHOPPERS													
Red List	x			trends									
BEETLES													
Red List ladybugs	x			trends									
Red List Saproxyle scarab beetles	x			trends									
Red List ground beetles	x			trends									
Red List water beetles	x			trends									
SPIDERS													
Red List				occurrence									
ANTS													
Red List				occurrence									
FLIES													
Red Lists Empididae+Dolichopodidae				occurrence									
waarnemingen.be/observation.be	x	x	x	occupancy	2008	10		x	x	x			x

Figure 9: Result of the data mapping exercise for remaining invertebrates, as presented in September 2018 (yellow indicates important gaps)

STEP 7: ANALYSIS AND SELECTION OF DATA

This step involved an assessment of which data could be used for the LPR. Scope definition happened in two phases, keeping in mind the available time and resources to perform the exercise, and overall feasibility.

7.1 SETTING CRITERIA FOR SELECTING TAXA: DEFINING GUIDING RULES

The first phase of setting criteria took place during the first scientific committee meeting. During that meeting, guiding rules for defining the criteria were concluded (taking into account that it is the first exercise of its kind in BE and that regionalization of BD data complicates the undertaking):

- Due to limited time and means: focus on "low-hanging fruit", then evaluate the data sets that require more effort and decide in which cases it is economic to attempt to include them in the analysis ("middle-hanging fruit") or hardly ("high-hanging fruit").
- Aim is to have the most comprehensive indicator of BD possible in BE; thus, we want to go beyond the vertebrates (as done in the global LPI) and include every taxon for which the dataset is sufficiently comprehensive and of high-quality.
- Include species that occur in only one of the regions (advantage of having regional LPI).
- Include taxonomic groups that are not representative of the whole country (in line with the ZSL procedure).
- Exclude invasive/exotic species from the LPI (see below).
- Include reintroduced populations for the LPI to reflect the effect of management strategies.
- Focus on native species only: so wintering counts are excluded since they comprise data of migrating (foreign) species. Trends of wintering migratory species are difficult to interpret as they are erratic (many species choose different, geographically separated, overwintering sites) or may depend too strongly on factors beyond the reach of management in BE.
- Exclude groups with too few data to allow building trends: ants, spiders, beetles and remaining invertebrates
- Exclude 'extra'[2] groups with relatively large number of species that would unduly affect the LPIs: mosses, mushrooms
- Focus on groups were good quality, comprehensive (over time and area) data are known to be available: dragonflies, birds, butterflies, reptiles and amphibians
- Keep exploring possibilities for: plants, marine species, bumblebees, moths, bats, mammals

7.2 SELECTING DATASETS

The second step aimed at choosing which datasets would be selected for the LPI. This step was done in parallel and after the methodology for data analysis was defined (Step 6).

In order to proceed, the following questions needed to be answered:

i) *Are data meeting the criteria* (required by the analysis models – see step 6)?
ii) *How much preparatory work do they require before they can be analyzed?* (e.g. transcription from old book, cleaning doubles, standardizing location info.)
iii) *Are they accessible?*

[2] Extra groups are groups that were not included in previous LPI, hence for which no clear methodology exists for integrating them.

The datasets were categorized into low/medium/high-hanging fruit to help prioritization. The low-hanging fruit were given priority 1 (high) and medium 2 (middle). High-hanging fruit were not considered here because of time constraints.

Low/medium/high-hanging fruits determined based on:

- Data Type (abundance data are favored)
- Amount of cleaning/preparing data needed (e.g. some owl pellets are still in a freezer)
- Ease of access (e.g. data owners willing to share, responsive)

7.3 OVERVIEW OF FINAL DECISIONS PER SPECIES GROUP

Species Group	Rationale	Final Conclusion
Marine Species	Several issues associated with the marine ecosystems: - Different biomes (air, pelagic, benthic, sand) with different types of data - Difficult to distinguish between federal and Flemish territory - SOM not applicable because the closure assumption (*i.e.* the assumption that the occupancy status of a species is constant during the season) is easily violated in the marine environment (species very mobile). - Not enough abundance data available. - Marine mammals: impossible to use by-catch data because there is no measure of effort.	Exclude the marine ecosystems from the current LPI, but consider it for a next edition; still include in the LPR as case-study
Freshwater species	- Necessary to split between exotic and native species (as for all the data) - Fish: data exist for Flanders, but data not accessible. Data exist for Brussels, but incomplete. - Macro-Invertebrates: not included in the first edition	Not included in the current LPI

Mammals	- Occupancy data from waarnemingen.be for day-active mammal species.
	- Use 2005 in Flanders and 2008 in Wallonia as starting year which corresponds to the first data recorded on waarnemingen.be/observations.org.
	- Exclusion of hunting data because: no measure of effort and populations are artificially managed: exclusion of *red deer, wild boar*
	  - Exclusion of species with strong detection bias in the data that cannot be solved (lack of a standardized detection protocol for signs of presences):
	  - Nocturnal species : *fox, hamster*
	  - Diurnal forest-dependent species: *red squirrel*
	- Did not manage to access hamster and badger count data.
		Inclusion of mammals only for which there are full lists available for day-active species from waarnemingen.be (e.g. resulting from targeted search for the species) since it is not possible to correct for the searching effort (thus for detection probability in opportunistic data)

Bats	- Abundance data exist: winter and summer count datasets - Include both summer and winter counts? Or present separately? - Very large debate among experts with regard to the reliability of data (e.g. technologies have drastically changed over time).	Exclude bats from the LPI as there is lacking expert trust in a.o. the relevance of the trends (first conclusion was to include them in mammal LPI, then do a separate case study, then exclusion)
Birds	- Good quality abundance data on breeding birds - Difficult to build SOM models (few full lists and migratory birds). - Exclusion of birds in winter or migratory season, because this does not concern (certain) local breeding populations	Include only breeding birds and bird species, which can be modelled with TRIM
Amphibians & Reptiles	Suggestion to use a finer scale for the SOM data, but not feasible with the current models to be used (follow the protocol of the NL)	Include in the LPI
Butterflies	Data available for time period	Include in the LPI
Moths	There is standardized data available for moths, but only for Flanders.	Include in the LPI
Dragonflies	- Enough data for many species - Use adult stage data only	Include in the LPI
B(umbelb)ees	- No data suited to perform either TRIM or SOM analysis for bees in Belgium (Belbees dataset are counts, but not repeated over time and no lists available because not opportunistic data) - Alternative: Perform List Length Analysis to generate trends for bees to include separately using the BELBEES dataset, but the data received was not suitable to perform any type of trend analysis.	Include a case-study based on published literature to illustrate the state of pollinators in Belgium
Beetles	- Isolated datasets that could be used (stag beetles), but does not meet the criteria for this first edition	Exclude from LPI
Ants, Spiders,	Not enough data	Exclude from LPI
Plants	- Very large number of species thus should not be included in the LPI with animal species (would skew the trend) and/or some species should be selected. - No abundance data available (no TRIM possible). - Probably not enough replicated occupancy data (no SOM possible) - It is important to include plants in the report at some level.	Include a case-study based on published literature to illustrate the status of plant diversity in Belgium: Use some trends generated by INBO (W. Van Landuyt & H. Van Calster) as illustration for Flanders

Mushrooms & Mosses	Large number of species that should be balanced to prevent skewing of the LPI	Exclude from this version of the LPR
Exotic species	Not native, so out of scope Suggestion to calculate some trends to be included in the LPR	Exclude from LPI, because not native species, but expand on the topic in a case-study

STEP 8: DEFINING THE METHODOLOGY FOR DATA ANALYSIS

The project initially intended to follow the methodology of the Zoological Society of London, as used for the developed for the LPR International. However, this method relies on abundance data only and in Belgium, abundance data are only available for Birds, Moths (for Flanders only) and Bats (and perhaps a few more datasets that are limited to single species, or in geographical range and time, such as on hamsters or badgers).

Having a national LPI covering only data from Birds, bats, and Flemish moths would not fulfill the objective of the LPI development, which was ***to develop the most representative indicator of Belgian biodiversity using the LPI.***

The methodology opted for in the Netherlands, which used abundance data (through TRIM modelling) as well as opportunistic data (through Site-Occupancy Modelling) - which are available in Belgium for several well-surveyed taxonomic groups– thus seemed the best suited.

Hence, the Belgian methodology (depicted in figure 11) uses abundance data wherever possible, and occupancy data for groups for which abundance data were not available (the majority). For certain groups such as bees, where even occupancy models could not be applied, List Length Analysis was considered as an alternative. However, combining three methods in a LPI would introduce a totally novel approach, which would need to be validated. It was decided to follow the already-validated methods used by the Netherlands.

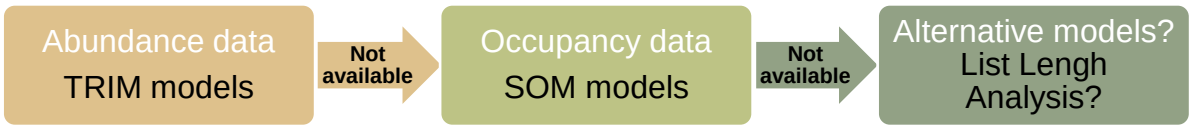

Figure 10: Graphical outline of the methodology considered for the development of Belgian LPIs

8.1 TRENDS AND INDICES FOR MONITORING DATA (TRIM)

Trim is a freeware stand-alone program that enables to analyze ecological data with many missing values, developed by the 'Centraal Bureau voor de Statistiek' (CBS) in the Netherlands[3]. It is currently available as an R package: RTRIM (see e.g. *https://cran.r-project.org/web/packages/rtrim/index.html*) more info on the CBS website).

TRIM generates population trends using abundance data. It is a well-validated (peer-reviewed publications), widely used (Dutch LPI, EU reporting, Bird census, etc.), and freely accessible (Package in R) method.

[3] Pannekoek and Van Strien (1991)

The requirements data that need to be fulfilled for applying TRIM are:

- Data need to be at species level (e.g. data at genus level, as in many insects, are disqualified)
- Data available for at least 2 years over the whole study period
- Data collected at the same location over the whole study period (to ensure the same population is studied over time)
- Data collected with the same method over the whole study period (to ensure comparable results over time)

The type of data fulfilling these criteria are thus standardized data: population counts, population estimates, densities, indices, proxies, measures per unit effort.

Typically, data not fulfilling the above-mentioned criteria are recruitment data, catch or hunting data without a measure of effort, opportunistic sighting data, and data from atlases[4].

8.2 SITE OCCUPANCY MODELLING (SOM)

Site-occupancy modelling uses Bayesian statistics (in JAGS) to generate trends from presence data. For the studied species, the number of occupied sites is estimated for every year. The model infers absence/non-detection from the length of the List reported by the observer. It also explicitly takes into account the search effort, which is known to have increased tremendously in the recent years, and detectability bias (higher detectability in fragmented landscapes and public parks than in woodlands).

This methodology offers thus the possibility to use unstandardized data, as opportunistic sightings, and thereby citizen-science (thus platforms like waarneemingen.be/ obervations.be) to generate species trends, which match well with trends obtained from abundance data (Van Strien *et al.* 2013). These trends are thus used as proxies for population trends for species lacking abundance data.

The requirements data that need to be fulfilled for applying SOM are:

- Data must be collected in > 50 sites (1x1km grid cell)
- The data must be collected with replicated visits of the same site within the year/season (the same 1*1 km grid cell is visited multiple times over the year/season)
- Data must be available for at least two years in the study period
- Full species list must be reported

In order to quickly assess whether a dataset qualifies for SOM, the following questions have been asked to data providers:

1) Assess for which species groups replicated visits occur (meaning that a grid cell or site is visited several times in one and the same year).
   → *If there are no replicated visits anyway >> no SOM possible*
2) If there are replicated visits, in how many sites this occurs per year?
   → *If less than 50 sites per year >> no SOM feasible*
3) If there are replicated visits in >= 50 sites per year, how many sites were visited in more than one year? If few (< 50) >> SOM makes not much sense

[4] Atlases, which constitutes an abundant source of standardized data in Belgium, are usually not replicated in time, or are based on different methodologies in different locations.

8.3 STARTING YEAR OF THE LPI

In the Dutch report, 1990 was used as starting year, because it is from that year onwards that most datasets started being collected. In Wallonia, 1990 is also the start of the Breeding Bird Census. For Belgium, 2008 is the year since when many data became available.

However, as the aim is to have the best picture of the evolution of the status of biodiversity over time, it is preferred to have the earliest starting date possible (here 1990).

Moreover, it is not recommendable to drop some available data (that are already scarce) by choosing a 'late' starting point (like 2008).

Therefore, the scientific committee agreed to use 1990 as starting date, thereby being consistent with the NL methodology. For time series starting after 1990, their missing values are imputed from the other trends.

STEP 9: DATA PREPARATION

9.1 DATA SOURCES/PROVIDERS

Institute	Region	Taxonomic group	Type of Data	Contact name
INBO	FL	Butterflies	Abundance	Dirk Maes
INBO	FL	Breeding birds	Abundance	Glenn Vermeersch
Natuurpunt	FL	Moths	Abundance	Kristijn Swinnen
INBO	FL	Dragonflies	Occupancy	Geert de Knijf
Natuurpunt		Amphibians and reptiles (HYLA)	Occupancy	Jeroen Speybroeck
Natuurpunt	FL	Mammals, butterflies, dragonflies, grasshoppers, etc. from waarnemingen.be	Occupancy	
Natagora	WAL	Breeding birds	Abundance	Anne Weiserbs
SPW/DEMNA	WAL	- Mammals (ca. 50 species) - Reptiles & Amphibians (ca. 25 species) - Butterflies (ca. 90 species) - Dragonflies (ca. 60 species) - Grasshoppers (ca 30 species)	Fusion Abundance/Occupancy data	Yvan Barbier

9.2 DATA PREPARERS

Institute	Taxonomic group	Type of Data	Data preparer
INBO	Butterflies, Moths Grasshoppers	Occupancy Occupancy	Hans Van Calster
Natagora	Amphibians and reptiles Birds	Occupnacy Abundance	Anne Weiserbs
Natuurpunt	Mammals Dragonflies	Occupancy Occupancy	Kristijn Swinnen

Data curators, knowing the limits and characteristics of their data, are the most appropriate persons to prepare the data for analysis, while downloading huge sets of data from an automatic system like GBIF is suboptimal. Here, some of the data preparers are also data curators for several datasets.

9.3 METHODOLOGY AND OUTPUT

The "data-preparation" step consisted of grouping all datasets selected for the species group of interest and performing the following steps:

1. Collection the needed datasets from data providers
2. Harmonization of species names
3. Harmonization of spatial reference formats
4. Harmonization of field names/headers
5. Addition of data source name
6. Filtering for duplicates
7. Determination of closure period for each species (= period in the year during which the site is assumed to be either occupied or unoccupied, but does not become permanently abandoned or colonized – for insects, reptiles and amphibians only).
8. Exclude observations that fall outside the closure period for a species

Detailed steps as to how to perform data preparation can be found in the **Annex 3-STEPS**.

Datasets were provided by the institutes holding the data. These data were filtered for errors and observations not qualifying for the analysis; and have standardized date formats.

The data preparation results in three output files in the format required for running the models:

1) **Observation_file**: Presenting all records of one and the same specific species group in one file for the entire country (i.e. records from Flanders, Wallonia and Brussels together).
2) **Species_Name file** (as .xls): For each species, presenting the species ID and its corresponding scientific and vernacular names in NL/FR as well as the closure period.
3) **Extra file (as .xls),** when appropriate: for Butterflies, Dragonflies, Amphibians coding the different stages of their lifecycle.

Specifications as to how to perform data preparation can be found in the **Annex 4-PREPA**. The codes to get the final output of the data preparation are available in the GitHub environment.

9.4 DEVELOPMENT OF BELGIAN GRID SYSTEM

Belgian data are referenced in different geodetic systems from UTM or Lambert2008 to coordinates in WGS84. One common grid system, dividing Belgium in squares of 1x1km needed to be established for the LPR[5]. As a standard grid of 1x1 was not available at the National Institute of Geography (IGN), a novel grid for Belgium has been developed. For this purpose, the ETRS89 1x1km grid system for Belgium developed by the European Environment agency has been used as reference. Each grid has received a unique site-ID code and all occurrences were assigned to this site ID following the procedure below:

- Geographic information of old data was converted into the centroid of the square in the UTM grid
- The UTM 1x1 centroid coordinates were converted into the ETRS89 grid system and assigned a site ID.

There is a possible issue with a slight shift when converting from one grid to the other. It is best to have a bias (thus the same error for all data) than an imprecision (thus different treatment for different data) that may end up assigning adjacent observation to different squares (see Figure 12). Therefore, it has been decided to follow the same procedure for all data. This includes data in which accurate coordinates are available and could have been assigned directly to the ETRS89 1x1 grid.

[5] Many old occupancy data (from before 2008) are only localized in an UTM square. This UTM grid is not suitable for SOM data preparation, because all UTM square in Belgium are not of the same area. So it needed to be changed to a regular 1x1 km grid. To avoid possible difference in treatment, it was decided to first convert all new data into the UTM grid.

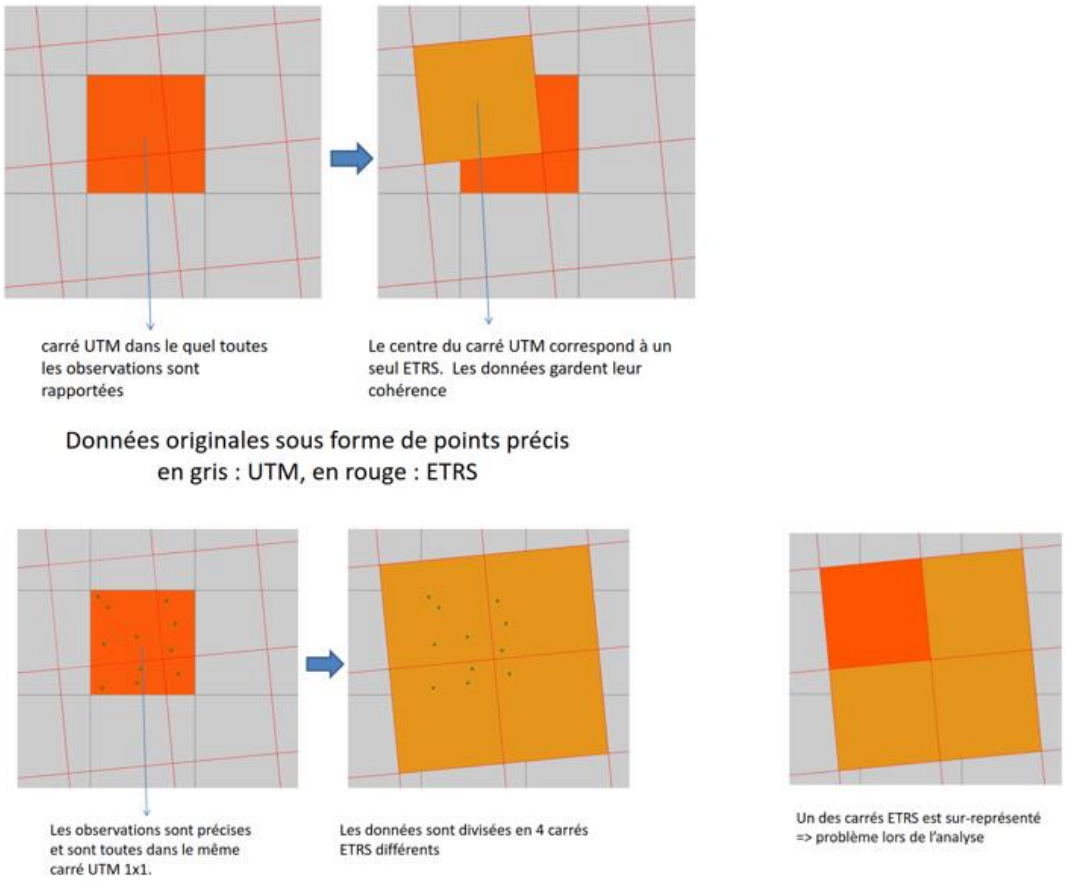

Figure 11 : Schematic representation of the arising issue when converting precise location points from one grid system to another. This issue arising explain why the choice has been made to convert the precise location points into the centroid of the UTM square before transposing it in the ETRS grid.

The final document of the Belgian gird system with its corresponding site ID, can be found in **Annex 5-GRID**.

STEP 10: RUNNING THE MODELS

10.1 TRIM – FOR ABUNDANCE DATA

TRIM Analysis has been performed by Antoine Derouaux for Birds and Kristijn Swinnen for Moths from Flanders.

The procedure is described in ANNEX TRIM_MSI-Tools.

10.2 SOM – FOR OCCUPANCY DATA

SOM Analysis has been performed by Arco Van Strien for Mammals, Amphibians, Reptiles, Grasshoppers, Dragonflies, and Butterflies and Moths.

Non-standardized ("opportunistic") data of observations of animal species were analysed using site-occupancy models (SOMs) per species to estimate trends in occupancy (= number of occupied grid cells of 1 km x 1km).

The principle of SOM is that any systematic change in recorder effort per site leads to changes in detection probability, but not in occupancy (Van Strien *et al.* 2013. J. Applied Ecology).
A SOM requires data with replicated visits per year and site and uses presence records and absence (or non-detection) records. Non-detections for study species were generated from the observation of all other species of the same species group.

The basic model of the SOMs applied included two covariates:

1) Julian date of the visit, to take into account that detection varies between visits in the season.
2) The number of species observed per visit, to take into account that detection varies with recorder effort; the number of species recorded is considered a proxy for this effort. The number of species is implemented as a categorical variable (single records, 2-3 species, more than 3 species).
3) location (site) is included as a random effect
4) Year is included as a fixed effect

For many species the number of grid cells is too high to run the SOM analysis in practice. Therefore, a two-step approach has been applied:

1) SOMs were run for separate provinces or – depending on the amount of data – combination of provinces and the results are aggregated per region
2) results were aggregated to estimate annual occupancy in Flanders (without Brussels), Wallonia and Belgium (including Brussels).

The first year of the dataset used is 1990, or - in case of few data - the first year in which observations of the species group came from more than 200 grid cells.

STEP 11: VALIDATING SPECIFIC TRENDS

The output generated by the model was species trends. Figures 13 depicts the process from output to LPI calculation.

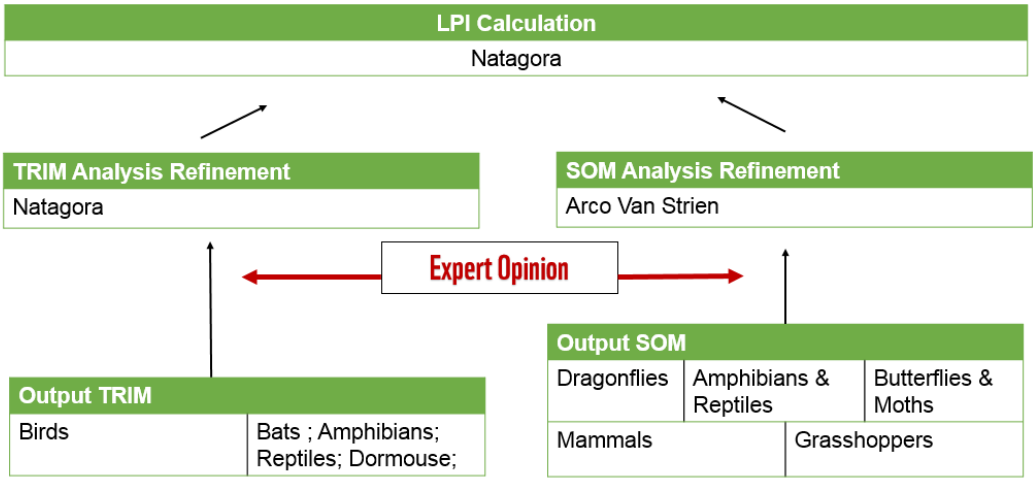

Figure 12: Schematic representation of the validation step in the process of LPI calculation

The output for SOM is available on request as a WinRAR folder.

Explanatory terms for understanding the SOM output:

- **No. of grid cells**: number of grid cells of 1x1 km where the species have been found at least once during the entire study period;
- **Multiplicative trend**: trend expressed as a factor (1: stable; 0.99: 1% annual decline; 1.01:
- 1% annual increase);
- **Standard trend** = magnitude of change in occupancy per year;
- **Record number 1** indicates indices, **2** standard error of indices, **3** occupancy (varies between 0 and 1), **4** se of occupancy, **9** time totals = number of occupied grid cells, **10** se of time totals, **11** detection probability for single records, **13** detection for longer lists.

Belgian Experts for each taxonomic group were invited to join a data workshop aimed at validating the trends and refining analysis for certain species. This step was foreseen in order to build the LPI on a qualitative set of species trends that are trustworthy and robust.

Trends generated with SOM (see SOM Output files) were sent out to the experts a few days before the workshop along with a set of questions. For all groups, experts were asked:

a. *Does the trend estimate of the species make sense to you? Are they as expected or not?*
b. *Are there any species whose trend should be considered as not reliable or not very informative and which better should not be used for the LPR? Think of e.g. migratory butterflies, vagrant species and species which are hard to identify by many observers.*
c. *Are there some species whose trend could be recalculated (in another manner) to produce a more reliable trend?*
d. *Are there any species that require a specific field method and for which the observation of other species is hardly or not informative to generate non-detections?*
e. *Are there any species missing in the list while you expect there must be sufficient data?*
f. *Do some trends seem to be flawed by a problem relating to data?*
g. *Do you have other remarks/ feed-back to share?*

During the workshop, trends were discussed in groups for each taxon to be included in the LPI. The project manager (Olga Szczodry) and the SOM analyst (Arco van Strien) joined each session for at least 10 minutes in turns to answer methodological questions. When experts could not be present at the workshop, they were given the opportunity to send feedback by email to be discussed during the sessions. In each meeting, the attribution of a Habitat to each species was also examined. A first attribution, coming from the Flemish Red Lists, was presented and discussed to see whether it was valid for the whole country. Further details on habitat assignment are detailed further.

List of Species Experts present in each taxonomic group meeting:

Session	Experts	Function	Affiliation
Amphibians/reptiles	E. Graitson	Feedback by email	Natagora
	T. Kinet	Expert Wallonia	Natagora
	S. van de Poel	Expert Flanders	Natuurpunt
Grasshoppers	JM. Couvreur	Expert Wallonia	DEMNA
	P. Goffart	Expert Wallonia	DEMNA
	H. Van Calster	Expert Flanders	INBO
	A. Heughebaert	Expert Brussels	RBINS
	H. Segers	Chair	RBINS
Dragonflies	H. Segers	Chair	RBINS
	RM. Lafontaine	Expert Brussels	
	K. Swinnen	Data provider	Natuurpunt
	G. de Rnijf	Expert Flanders	INBO
	N. Mayon	Expert Wallonia	GT-Gomphus
Mammals	K. Swinnen	Expert Flanders	Natuurpunt
	C. Rousseau	Expert Wallonia	WWF
	M. Herremans	Expert Flanders	Natuurpunt
	V. Schockert	Feedback by email	ULiege
Birds	JY. Paquet	Chair	Natagora
	A. Derouaux	Expert Wallonia	Natagora
	M. Herremans	Expert Flanders	Natuurpunt
	K. Swinnen	Data provider	Natuurpunt
	RM. Lafontaine	Expert Brussels	
Butterflies/moths	H. Van Calster	Expert Flanders	INBO
	P. Goffart	Expert Wallonia	DEMNA
	A. Heughebaert	Expert Brussels	RBINS
	M. Herremans	Expert Flanders	Natuurpunt

Main issues discussed:

1. Strong variability in the first years of the trend due to lack of data. However, despite the high stochasticity, the overall trend coincided with expert judgment (what they see on the field) and since the aggregation of multiple trends tends to smoothen the trendlines, they were kept as such. Using another starting year would have been possible for the species group as a whole, but not for separate regions. However, it would be sub-optimal to voluntarily discard some available data and choosing a later start year would result in higher number of imputations, which is not desirable for the robustness of the resulting aggregated trend.
2. The impossibility to use another starting year for a whole taxonomic group, because it means the trends could not be aggregated anymore or would bias the final trend. Final decision: keep the trend with variations if the trend is generally coinciding with expert observations) because the stochasticity will be smoothed when the trends are aggregated. Only in few butterfly species the starting date will be different.
3. SOM output appears to be more optimistic than expert's appreciation in some species: although these trends may be a bit on the positive side for declining species, it appears rather limited. There is no way to resolve this issue
4. In dragonflies, although there is a general increase in this taxonomic group as a whole, experts underlined that this is not the case for stenotopic species.
5. An alternative occupancy modelling method for rare species would be perhaps better suited (Outhwaite *et al.* 2018). But is has not yet been widely used so we stick to validated methodology (no room for methodological research is the LPI project).

6. Species trends for newts needed to be recalculated to improve the methodology of generating zeros

The overall outcome of the workshop was that the majority of trends were considered reliable and sound by the experts.

Some species were removed, others were re-calculated following expert advice.

Further details can be found in the minutes of the workshop in **Annex 6-WORK**.

STEP 12: REFINING SPECIFIC TRENDS

The table below presents, per taxonomic group, the number of species for which an action (deleting or refining the trend) was necessary.

Species group	Trends to be recalculated	Trends to be removed
Amphibians	1	3
Reptiles	-	1
Butterflies	13	5
Birds	23	20
Dragonflies	1	5
Grasshoppers	-	2
Mammals	-	2
Moths	-	10

To summarize, besides deleting the designated species from the subset, the following actions have been undertaken:

- For several species, another starting year has been used (for all regions – impossible to have different starting dates between different regions because no combination tool, such as MSI-tool for TRIM, exists for SOM) in order to erase the high variability in the first years in certain species trends
- For butterfly species, data have been prepared again to allow taking into account the phenological peak with the highest number of occurrences, instead of the systematically selecting the first peak.
- Add the TRIM output for moths
- Solve a bug in RTRIL Shell
- Adapting indices for some species for which a different starting year is needed, because they were not present in Belgium during the first years of the time series used. *e.g.* some dragonfly and grasshopper species.
- Trends for newts have been recalculated generating zeros from observations of other newt species, but results were not much different, so kept original methods (see **Annex 7-NEWT**)
- Population estimates of *Bombina variegata* from mark-recapture data have been converted into annual indices.

The full list of refined trends is available in **Annex 8-REFI**.

STEP 13: BUILDING LPIS

Figure 14 below presents the relationship between the different LPIs developed. It can be seen that the National LPI is not only composed of the aggregated LPIs for Flanders and Wallonia, but also

contains some Brussels data. Similarly, the Habitat LPIs do not come from a strict disaggregation of the National LPI, as eurytopic species are not considered. Finally, the LPI for southern and northern species is based on a subset of species included in the LPI Belgium.

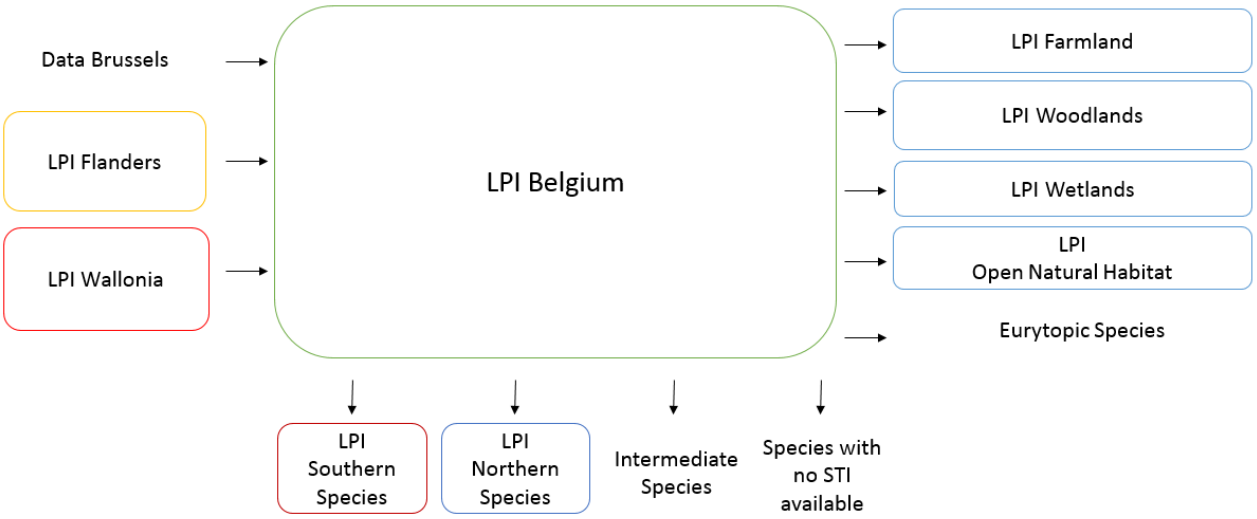

Figure 13 Schematic representation of the articulation of the different LPIs towards one another. It appears clear that the National LPI is not merely the aggregation of the regional LPIs. It is also clear that the Habitat LPIs are not a simple disaggregation of the National LPIs

The LPIs were built using the MSI-tool in R package. For a detailed methodology please consult **Annex 9-RTRIM_MSITOOLS** and codes on GitHub.

13.1 NATIONAL AND REGIONAL LPIS

13.1.1 National LPI

Producing a Living Planet Index for Belgium was the core of the LPR Project. Indeed, in every Living Planet Report, a general LPI is at the heart of the publication (global LPI at international level, and national LPI at national level) as it is a powerful tool to communicate.

13.1.2 Regional LPIs

With biodiversity topics being largely a regional competence in Belgium, the scientific committee (at the first meeting) considered it relevant to produce regional LPIs for policy purposes, an exercise that was facilitated because data are collected independently in the regions, using different methodologies.

Building regional LPIs also allows representing species data that are only available for one region.

E.g. Data for moths are only available in Flanders. The group is included in the Flemish LPI, but not in the Belgian LPI. The group is too big (contains too many species) to be included in the national LPI without knowing the representativity of the trends for the whole country (i.e. whether Flemish trends can be extrapolated to Wallonia)

For the region of Brussels, it has not been possible to calculate a separate LPI for the following reasons:

- No TRIM was possible due to very few abundance data are available for the region, for most of the groups except for a few bats, and birds.
- No SOM was possible because, due to the limited size of the region, not many datasets would have qualified to reach >50 sites visited at least twice within the season and at least for two years. It has been envisaged to reduce de size of the site, but it was considered unacceptable to have inconsistent resolution within the studied territory (Belgium).

However, data for Brussels were taken into account in the Belgian LPI (calculated with data of Flemish Brabant in SOM).

13.2 LPI PER HABITAT

13.2.1 Assessment of different methodological options

To build the LPI by habitat, several methodologies have been considered:

- **Using habitats assigned with occurrences** as was done in the Dutch LPR (preferred option). However, in Belgium it was impossible to use habitat types from the occurrences or the habitat in which species occurred most because (1) there are insufficient records in the Belgian dataset and (2) the habitat is not associated with the observations and would have had to be extracted using maps at 1km² resolution. This was unachievable within the context of the present LPI project.
- **Assign a site to a habitat: t**his option was deemed undesirable by the scientific committee, because 1x1 km resolution is not sufficiently detailed to correctly account for habitat diversity (e.g. a site with extensive grassland crossed by a river could not have been identified as natural open habitat and wetland).
- **Linking species to habitats:** the scientific committee concluded that this was the only feasible option for this edition of the LPR. Three options were considered to achieve that:
    - ***Using group specific information*** from international resources; e.g. for birds, the habitat suitability classification from the European Birds Census Council (EBCC) could be used. However, these are not sufficiently detailed (only three classes: species characteristic of farmland, forest and other species) and not available for all taxonomic groups.
    - ***Using the IUCN Red List habitat suitability list***: as done for the global LPI. All suitable habitats for all species have been extracted, however:
        - Approximately 90% of Belgian species do not have data available on the RL.
        - There are multiple habitats for each species, so it is difficult to evaluate to which habitat the species is specific to.
        - The level of detail for each habitat is high, creating complexity. However, the first level of detail at the global scale is not sufficiently detailed for the Belgian scale.
    - ***Using the Flemish Red Lists***: after the attempt of using IUCN red list, it was concluded that the Flemish red lists are a much better starting point for assigning Belgian species to Belgian habitats. This classification was further refined through expert opinion at the data workshop for validating species trends, to ensure correct extrapolation of the species-to-habitat association to the whole country.

13.2.2 Process of assignment of species to habitat

The assignment of species to habitat was done using a combination of the habitat classification obtained from the validated Flemish Red List (*https://bdj.pensoft.net/article/34089/ Tabs > Table 7*') and expert opinion.

For all species included in the LPI calculations, the habitat classification from the validated Flemish Red List (FRL) was extracted. FRL used quantitative criteria and a representative sample of occurrences across all ecological regions in Flanders for Red List assessment. The habitat given in the database is based on broad classifications used in the land use map of Flanders (Gobin *et al.* 2009) and in Corine Land Cover (Version 18.5.1) or Natura 2000 habitats (Council Directive 92/43/EEC). For species occurring in two different habitat types, both are given in the dataset. We classified habitats into different types to be used for the LPR according to the following cross-walking table:

Habitat/biotope (Flemish List)	Habitat (LPR)
Agriculture	Agriculture
Grasslands	Agriculture/Nature Open land
Heathlands	Nature open land
Shrubs	Woodland
Woodlands	Woodland
Marshes	Wetland
Running Waters	Wetland
Salt Marshes	Wetland
Standing Waters	Wetland
Urban	Urban
Marine	NA
Dunes	NA
Eurytopic	No specific habitat

Because this FRL is based on the preferred habitat in Flanders (which may not be the same as in Wallonia) and because FRL does not cover all species (only species that occur in Flanders), experts from Wallonia, Flanders and Brussels were asked to review the list and propose alternatives if deemed necessary. This was done during the workshop of 06/06/19 and following exchanges by e-mail.

As the aim was to have further insights on the trends in particular habitats, the assigning methodology was rather conservative, and aimed at assigning only specialist species, i.e. species specific to the habitat. Generalist species were assigned to an 'eurytopic' group, and no LPI was calculated for this group. Similarly, when a species occurs in more than one habitat, it was assigned to the 'eurytopic' group.

Some butterflies, birds and grasshoppers are classified in the Flemish Red List as 'grassland' species. This is problematic as grasslands can be either extensively managed, in which case they should be classified in Nature Open Land, or intensively managed, in which case they should be classified in Agriculture. For these species, the decision was taken to create an approach per taxonomic group:

- For birds: Use the EBCC's list of Farmland Species. The 'grassland' species classified as farmland fall into the Agriculture Habitats whereas the others in Nature Open Land
- For Butterflies: Use the EU Butterfly indicator for Grassland species. Species that are 'specialists' are then classified into Nature Open Land while species that are 'widespread' are classified into Eurytopic. Species which are not included in the EU report were further classified by the butterfly expert group set up for the data workshop. Table 1 presents the classification of grassland butterfly species in the EU Indicator and the LPR classification.
- For Grasshoppers: expert opinion was sought.

Table 1 Butterfly species classified as 'grassland' species in the Validated Flemish Red List, with their corresponding category in the EU Grassland Indicator and their final assignation in an LPR Habitat.

Species name	**EU grassland Indicator**	**LPR Habitat classification**
Aphantopus hyperantus	Not included	Eurytopic
Aricia agestis	Not included	Open Nature Land
Coenonympha pamphilus	Widespread	Agriculture
Cupido minimus	Specialist	Nature Open Land
Cyaniris semiargus	Specialist	Nature Open Land
Erynnis tages	Specialist	Nature Open Land
Lycaena phlaeas	Widespread	Agriculture
Lycaena tityrus	Not included	Open Nature Land
Melanargia galathea	Not included	Open Nature Land
Melitaea cinxia	Not included	Open Nature Land
Spialia sertorius	Specialist	Nature Open Land
Thymelicus acteon	Specialist	Nature Open Land
Polyommatus/Lysandra coridon	Specialist	Nature Open Land
Euphydryas aurinia	Not included	Nature Open Land

The final assignation of each species to a habitat can be found in **Annex 10-RES** giving an overview of the results.

The habitat LPIs are thus based on trends of species specific to the given habitat. The inherent problem with this methodology is that it is impossible to discriminate whether the trends observed are an effect of the habitat or the taxonomic group that is prevailing in the set of species included in the list of species specific to the given habitat.

13.3 LPI FOR SOUTHERN & NORTHERN SPECIES

In order to explore the possible effect of climate change on Belgian biodiversity, we looked in more details into the trends of species having a 'southern' or 'northern' distribution. This was performed for species groups for which an STI (Species Temperature Index) was available dragonflies (Termaat *et al.*, 2019) or could be calculated using the same methodology (birds, butterflies), namely bird, dragonfly and butterfly. The Species Temperature Index is the average of the yearly mean temperature across the species range in Europe.

Following the method used in Termaat *et al.* (2019), we calculated a 25 and 75 percentile STI per taxonomic group (Butterflies, Dragonflies and Birds) then classified species as "cold" (i.e. Northern) when their STI < 25 percentile of group STI and as "hot" (i.e. Southern) when their STI > 75 percentile of the group STI. The rest was classified as 'medium' with no LPI calculated. Scripts are available in Github.

STEP 14: REFINING LPIS

After the LPI has been calculated, the results were sent to the scientific committee for review in the form of a PDF document. This PDF comprises all results linked to the trends for each species. The first results were sent on the 17/10/2019 along with a set of questions. The questions aimed at informing the choices regarding the methodology, discussing and refining the results and receiving input on interpretation (see example below and complete set of questions in **Annex 11-QUEST**).

Example of questions sent to the scientific committee:

About the methods

- Should species with trends classified as "uncertain" be included (i.e. trends not classified as increasing/ decreasing/stable)?
- Do you have any objections regarding the methodology followed for assigning species to habitats?
- *Do you have some objection against combining abundance (Trim) and occupancy (SOM) output into one LPI for Belgium as done in the Dutch LPR?*

About the results

- Do the results make sense? If not, what is not OK?
- Are the results of the species representative for the group? If not is the current level of representativeness a problem regarding the objectives of the project?
- Are the data of each group representative for what is really going on, and if not, is the current level of representativeness problematic?

About the messages

- How to communicate around the increasing LPI for both Belgium and Flanders/Wallonia?
- Do you believe it is possible to compensate the potentially positive message of a slightly increasing trend with a good accompanying text and communication package stating that this outcome does not imply biodiversity is doing well?
- Do you agree with the following message developed in view of the results?

The points raised by the scientific committee and their feedback was then discussed within the Board and corrective actions were taken where deemed necessary. Consequently, several rounds of results' presentation and feedback took place (updated results versions were released on 07/11/19, 15/11/19, 25/11/19 and final results on 29/11/19)

The main points of discussion during the validation & refinement process were:

- **The classification of Grasslands** in LPR Habitat: Following the scientific committee's advice, the method was adapted (see point above)
- **The classification of Hot/Cold Species**: following advice by the scientific committee, the method has been adapted (see point above)
- **Combination of SOM & TRIM:** There were reservations expressed as to combining TRIM output and SOM output into one LPI. Occupancy being only a proxy for abundance, it is not optimal to combine both, especially as some scientists expressed doubts on the extent to which the sharp increase in detectability over the past years is taken into account in SOM, guaranteeing thereby that the latter increase is not influencing the results. Keeping this limitation in mind, the Board decided to take a pragmatic approach and 'not to throw the baby with the bathing water' which means to develop the LPI as a combination of TRIM and SOM output for the report, but with carefully chosen and formulated key messages and a disclaimer stating clearly all limitations (i.e. technical annex).
- **Inclusion of 'uncertain' trends:** The trends have been classified as moderately or strongly decreasing/increasing or stable or uncertain depending on the value of the slope and the width of the Confidence Interval (CI). If the CI is wide (comprises 1 and + or -5%) the trend in deemed uncertain whereas when the is CI is narrow but comprises 1, it is deemed 'stable'. A significant increase/decrease happens when the CI does not comprise 1, and it is qualified as 'strong' when the slope is more than + or -5%/year (see figure 15 for a graphical

representation). The category “uncertain” is thus an appreciation of the trend; it does not pertain to quality or “trustworthiness” of the data. Several species showing fluctuations are qualified “uncertain”, because fluctuations hamper the mathematical distinction of a trend. However, this does not mean the fluctuations are not real. Moreover, the trend could be "uncertain", because the standard error of the slope is high; the LPI integrates this uncertainty around the annual index (the standard error for each annual index is used in the multi-species trend estimation). For these reasons it has been judged valid to retain “uncertain” trends in the LPI.

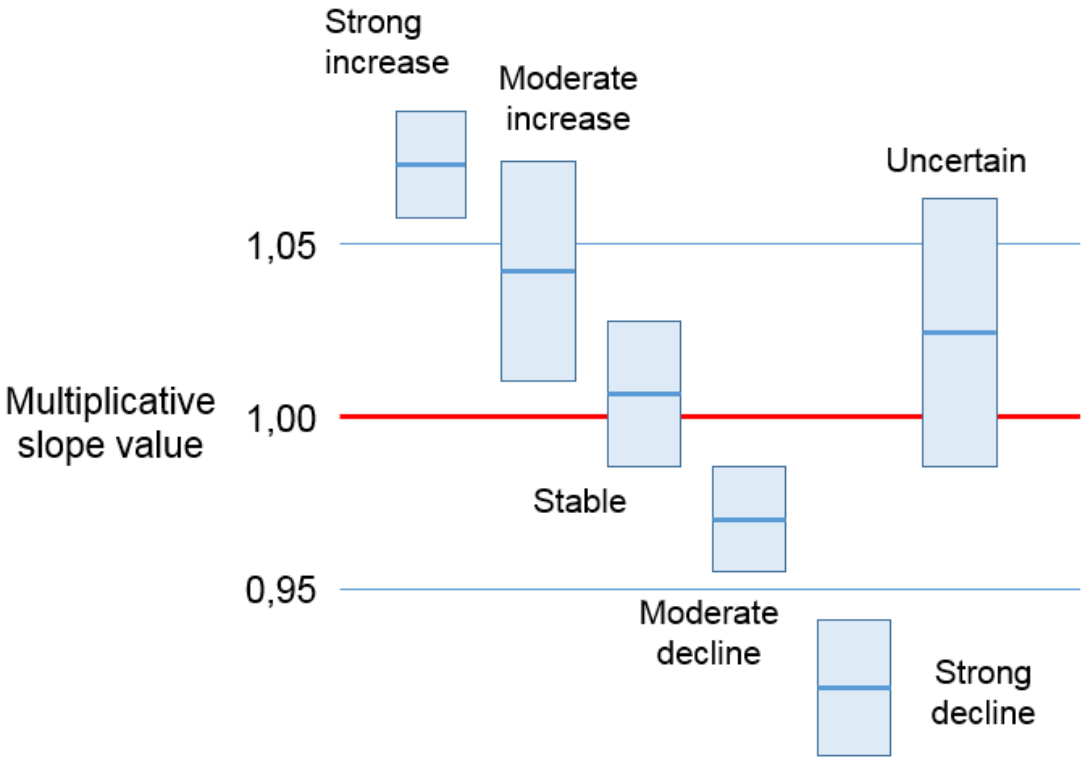

Figure 14 Visual representation of the trend categorization depending on the slope and the width of the confidence interval.

- **Inclusion of bats:** Whereas bats were at first positively considered for inclusion the LPI, discussion arose regarding the interpretation of data. Bat experts considered it difficult to judge whether current increasing trends are an artefact (e.g. decreased in number of available sites so increase in colony numbers, big effect of observer’s expertise and used technique) or a true increase. Data are relative for wintering populations, which means the date pertain to a mix of indigenous and foreign (migrating) species. Moreover, wintering sites are protected and managed (potential artificial component). For these reasons it was finally decided not to include them in the LPI.
- **Inclusion of Urban LPI:** The urban LPI was based on very few species, which are ‘using buildings’. This does not mean that these species are city specialists and live only in the cities (they could also live in countryside buildings). Consequently, it was decided to discard an urban LPI.
- **Inclusion of moths:** There was intense discussion on whether to include the Flemish moth data set. The final decision was to exclude moth data from the national LPI, because the data may not be representative for the whole country. For Flanders, it has been decided to calculate an LPI both with and without moths, since they may have a non-negligible weight in the LPI due to its high number of species.

See PDF with final results in **Annex 12-LPI**.

STEP 15: INTERPRETING LPIS

As seen in Annex 11-QUEST, questions were asked to the scientific committee on the interpretation of the results and writing of text describing the LPIs in the report. All points received as feed-back from the board were included in either the text of Chapter 2, or the Technical Annex in the LPR.

For the sake of interpretation, robustness tests (running the model without a specific taxonomic group and comparing the result) were run to inform which components drive the overall trend. Similarly, a correlation between the slope and the number of occupied sites was performed to inform whether 'rare' species had generally decreasing trends (see point on 'rarity' in the Limitation section).

The Technical Annex of the LPR gives guidelines on how to read the results and presents the limitations of the chosen methodology. One should read this Annex to have full information.

15.1 LIMITATIONS

Limitations of the chosen methodology are explained in the Technical Annex of the report. Here, only a brief description is given for each limitation.

Data Collection

- Bias in species monitored and prospected sites. This is exacerbated in citizen science where certain taxonomic groups and observing locations are preferred.
- Bias to data collection: variability in detection method (sampling tools, inter-observer variability, intra-observer variability), sampling choice.
- Representativity issue: are the species included representative for their taxonomic group group/location?

Site-Occupancy Modelling

- Detectability issue with non-standardized data: these are exclusively presence data so inferring absence is a difficult (truly absent, non-detected or non-reported?)
- The use of a proxy: list lengths as a measure of search effort
- Assumption that species have the same detectability over the season (closure assumption), but this is easily violated (e.g. by highly mobile species)
- Assumption that species have the same detectability between sites: this is not the case when species are not equally spread or if there is focused research
- Many experts believe that the noticed increases can result from increased sampling effort, a factor which could not be fully taken into account in the SOM
- SOM models are more sensitive to increases than declines, rendering the resulting LPI probably slightly optimistic. However, there are no means to correct for this asymmetry

LPI building

- Conclusions are more reliable when all trends are going in the same way, which is not the case in the Belgian LPI.
- There is an unbalance between taxonomic groups in the LPI suggesting that some groups may drive the LPI trend.
- Habitat assigning: the classification method into habitat is relying on species so it is not possible to exclude confounding factors (e.g. In the LPI for Southern and Northern species, it

is difficult to discriminate what factor from the taxonomic group, the habitat, or climate change, is mostly explaining the trend).

- An issue arises when 'rare' species are decreasing because, by definition, rare species may be underrepresented in the set of data. This is mostly problematic for some multivoltine butterflies.
- The resulting LPI could thus be biased towards the positive end (too optimistic). However, in the current dataset, no correlation was found between the slope and the number of occupied sites (proxy for rarity): (r=-0,068, p-value=0,336), there is no indication that the assumption of an "over-optimistic LPI" is correct.

15.2 RECOMMENDATIONS FOR FUTURE WORK

For next editions of the LPR it would be interesting to explore:

- GIS Assignation method for habitats: using Corine Land cover to extract habitat type and assign the observations to each habitat using their location
- Model everything through occupancy: some information will be lost, but it would render everything more comparable, solving thereby the problem of combining SOM and TRIM into one index
- Explore (and import to Belgium) the method of Outhwaite *et al.* (2018) for rare species, as these models appear to be taking more rareness into account. This would allow the integration of more species (i.e. the rare species) and therefore counter the potential bias of underrepresenting rare species. Note that A. Van Strien ran comparisons in the past, which appeared to refute an effect of rarity.
- Explore other occupancy models that take extra factors into account, such as, increase in detectability within the observer (learning curve).