# Human Influence in SDMs: Literature Review (Part I)

## Veronica F. Frans (verofrans@gmail.com)

## July 20, 2023

## Contents

1	Sun	nmary	3	
<b>2</b>	R setup			
	2.1	Libraries	3	
	2.2	Directories	3	
3	ISI	Web of Science search	4	
	3.1	Upload Web of Science records	4	
4	Abs	stract screening	5	
	4.1	Data cleanup	5	
	4.2	Abstract screening: round 1	7	
	4.3	Abstract screening: round 2	8	
	4.4	Abstract screening: round 3	9	
	4.5	Abstract screening: round 4	11	
	4.6	Abstract screening: round 5	12	
	4.7	Abstract screening: round 6	13	
	4.8	Abstract screening: round 7	14	
	4.9	Abstract screening: round 8	18	
	4.10	Abstract screening: round 9	19	
	4.11	Abstract screening: round 10	21	
	4.12	Abstract screening: round 11	22	
	4.13	Abstract screening: round 12	27	
	4.14	Abstract screening: round 13	28	
	4.15	Abstract screening: round 14	29	
	4.16	Abstract screening: round 15	31	
	4.17	Abstract screening: round 16	32	

	4.18 Abstract screening: round 17	33		
	4.19 Abstract screening: round 18	34		
	4.20 Abstract screening: round 19	35		
	4.21 Abstract screening: round 20	37		
	4.22 Abstract screening: round 21	38		
	4.23 Abstract screening: round 22	39		
	4.24 Abstract screening: round 23	40		
	4.25 Abstract screening: round 24	41		
	4.26 Abstract screening: round 25	42		
	4.27 Abstract screening: round 26	43		
	4.28 Abstract screening: round 27	44		
	4.29 Abstract screening: round 28	45		
5	Abstract Screening Summary	47		
3	Setting up for full article review			
7	Save	49		

## 1 Summary

This is the first R script of the literature review and synthesis for the article entitled, "Gaps and opportunities in modeling human influence on species distributions in the Anthropocene," by Veronica F. Frans and Jianguo Liu.

We examine the Web of Science search results and screen abstracts to prepare for full article screening, review, and the qualitative synthesis.

Here, in Part I of the qualitative synthesis, the following is accomplished:

- (1) Upload of Web of Science search results and cleaning of records
- (2) Abstract screening
- (3) Export of articles assigned for full article screening

Afterwards, the full article screening is filled out in an Excel spreadsheet and the data for the qualitative synthesis is extracted from all accepted articles. The data from the review are then processed in R Scripts Parts II to V.

## 2 R setup

We are using R version 4.3.0 (R Core Team 2023).

#### 2.1 Libraries

Load libraries

```
# load libraries
 library("bibliometrix") # for biblio-analytics
 library("revtools")
                         # for reviewing abstracts
 library("dplyr")
                          # for table manipulations
 library("scales")
                          # for scales and formatting
 library("kableExtra")
                         # for table viewing in Rmarkdown
 library("tidyr")
                          # for table manipulations
 library("plyr")
                          # for table manipulations
 library("tidyverse")
                          # for graphics/table management
 library("ggplot2")
                          # for graphics
 library("RColorBrewer") # for graphics
```

#### 2.2 Directories

Create and name directories for file retrieval and storage. The primary directory is the folder where the R project file hum\_sdm\_litrv\_r.Rproj is stored.

```
# create image folder and its directory
dir.create(paste0("images"),recursive = TRUE)
image.dir <- paste0("images\\")

# create data folder and its directory
dir.create(paste0("data"),recursive = TRUE)</pre>
```

```
data.dir <- paste0("data\\")

# bibtex file directories
bib.dir <- paste0(data.dir, "saved_recs_FINAL\\")</pre>
```

#### 3 ISI Web of Science search

The articles to review were selected under the following search terms in *ISI Web of Science* (https://www.webofscience.com/):

For all SDM papers up to the end of year 2021.

```
___TS=(("species distribution model" OR "environmental niche model" OR "species niche model" OR "bioclimatic niche model" OR "habitat suitability model" OR "ecological niche model" OR "habitat model") AND DT=(Article) AND PY=(1900-2021)*___
```

The document type was "articles" only, in order to avoid review articles and book chapters (any remaining reviews or book chapters are removed manually).

This yielded 12,854 articles.

This search was done on September 14, 2022.

The ISI Web of Science results (all data except citations) were saved in bibtex format (.bib) for analysis here in R.

#### 3.1 Upload Web of Science records

Upload records and convert to readable files

```
# all SDM-human papers
# load all records
bibs <- pasteO(bib.dir,list.files(bib.dir))

# conversions
bibs.df <- convert2df(c(bibs),dbsource = "isi",format = "bibtex")</pre>
```

Save as CSV for use in the PRISMA summary.

```
# save for PRISMA framework
write.csv(bibs.df, paste0(data.dir,"bibtex_dataframe_RAW.csv"),row.names=FALSE)
```

Get counts.

```
# get new number of articles
paste('number of articles from Web of Science:',nrow(bibs.df))
## [1] "number of articles from Web of Science: 12854"
```

```
# get number of duplicates
paste('number of duplicates:',nrow(bibs.df[duplicated(bibs.df$TI),]))
```

```
## [1] "number of duplicates: 8"
```

Remove duplicates

```
# get duplicates
dups <- (bibs.df[duplicated(bibs.df$TI),])

# remove duplicates
bibs2.df <- bibs.df[!duplicated(bibs.df$TI),]

# get new number of articles
paste('Final number of articles without duplicates:',nrow(bibs2.df))</pre>
```

## [1] "Final number of articles without duplicates: 12846"

Save as CSV for use in the PRISMA summary.

## 4 Abstract screening

The 12,846 articles' abstracts were screened to accept research articles that modeled species' distributions AND acknowledged human influence on species' distributions. Such an acknowledgement could be direct (e.g., mentioning humans, people, anthropogenic drivers, etc.) or indirect (e.g., mentioning land cover or land use change, etc.).

The abstracts were screened using the revtools package, which provides a GUI for screening and loads the results to a data frame.

#### 4.1 Data cleanup

Duplicates need to be removed, and also we need to remove any 2022 articles.

```
# reload articles to revtools bibliography format for screening
abs.df <- read_bibliography(bibs)
nrow(abs.df)</pre>
```

```
## [1] 12854
```

Remove articles published in 2022 (early access date of 2021 is why they were included in the original Web of Science search results).

```
# remove 2022 articles
abs.df <- abs.df[!abs.df$year==2022,]

# get count
nrow(abs.df)</pre>
```

#### ## [1] 12691

Remove duplicates.

```
# remove duplicates from this dataset
dups.df <- find_duplicates(abs.df, match_variable = "title")
abs2.df <- extract_unique_references(abs.df, dups.df)

# get count
nrow(abs2.df)</pre>
```

#### ## [1] 12683

There are actually 12,683 articles to be screened once duplicates and year 2022 articles are removed.

Next, sort the data.

```
# Sort the data by year, followed by the title (most realistic of an online search)
abs2.df$year <- as.integer(abs2.df$year)
abs2.df <- abs2.df[order(-abs2.df$year, abs2.df$title),]

# Add a column to give it a unique ID
abs2.df$UID <- seq.int(nrow(abs2.df))</pre>
```

Here is the list of all column names:

```
# get column names
options(width=85) #ensure width
colnames(abs2.df)
```

```
##
  [1] "label"
                                      "type"
  [3] "author"
##
                                      "title"
  [5] "journal"
                                      "year"
## [7] "pages"
                                      "abstract"
## [9] "volume"
                                      "month"
## [11] "publisher"
                                      "number"
## [13] "address"
                                      "language"
## [15] "affiliation"
                                      "doi"
## [17] "issn"
                                      "research_areas"
## [19] "keywords"
                                      "web_of_science_categories"
## [21] "keywords_plus"
                                      "author_email"
## [23] "number_of_cited_references"
                                      "orcid_numbers"
## [25] "researcherid_numbers"
                                      "eissn"
## [27] "times_cited"
                                      "funding_acknowledgement"
## [29] "usage_count_last_180_days"
                                      "funding_text"
## [31] "usage_count_since_2013"
                                      "journal_iso"
## [33] "doc_delivery_number"
                                      "web_of_science_index"
## [35] "unique_id"
                                      "article_number"
## [37] "earlyaccessdate"
                                      "note"
## [39] "organization"
                                      "editor"
## [41] "booktitle"
                                      "book_group_author"
## [43] "isbn"
                                      "series"
## [45] "origin"
                                      "book_author"
```

```
## [47] "oa" "da"
## [49] "registered" "filename"
## [51] "n_duplicates" "UID"
```

Save as CSV for use in the PRISMA summary.

#### 4.2 Abstract screening: round 1

Next, screen the articles. A GUI will pop up for manual screening.

```
# screen abstracts
scr_abs <- screen_abstracts(abs2.df)

# updated screening (9-14-2022)
scr_abs <- screen_abstracts(scr_abs)
scr_abs_1 <- scr_abs #duplicate for editing</pre>
```

While screening 300 abstracts, a list of key words were written down, to help filter the abstracts. We then use these key words to accept abstracts, and then the remaining papers are once again filtered manually.

Note that this process only helps to broaden what is included, but no articles are being excluded by this text-mining procedure. Any articles that are not selected by this process still have to be handled manually and rejected manually. All accepted articles, whether manually or mined, are also ultimately read in the full article screening process. Thus, all 12,683 abstracts and/or full articles were read at some point in the review process.

```
# get number of abstracts for each
count(scr_abs, "screened_abstracts")
```

122 out of 300 articles were selected in this round.

```
'land cover change','habitat conversion', 'land conversion',
    'habitat degradation','oil spill*',
    #human-wildlife conflict
    'collision','human-wildlife conflict','human interaction*',
    'road kill','roadkill',
    #artificial structures
    'artificial surface*', 'hydropower', 'artificial light',
    'artificial night light'
    #policy interventions
)
```

Get new summary.

```
# get number of abstracts for each
count(scr_abs_1, "screened_abstracts")
```

With these key terms learned during screening, 2,133 new articles were selected, including 11 that were originally excluded.

#### 4.3 Abstract screening: round 2

Next, screen abstracts again manually, and see what other terms come up through the next 300 papers (2723 screened total).

```
# updated screening (9-14-2022)
scr_abs_2 <- screen_abstracts(scr_abs_1)
scr_abs_3 <- scr_abs_2 #duplicate for editing</pre>
```

```
# get number of abstracts for each
count(scr_abs_2, "screened_abstracts")
```

```
## screened_abstracts freq
## 1 excluded 419
## 2 selected 2304
## 3 <NA> 9960
```

Next, make a list of terms that were key words from the "selected" articles in this round.

```
abs terms <- c(#terms related to...
               #resources
               #human influence in general
               'anthropogenic change', 'human use', 'anthropogenic activit*',
               'shipping', 'anthropogenic global change', 'human activit*',
               'human population', 'anthropogenic stressor*', 'human footprint*',
               'anthropogenic driver*', 'human pressure*', 'synanthrop*',
               'anthropic intervention', 'anthropogenic cause*', 'anthropogenic habitat',
               'human-dominated',
               #human degradation or pollution
               'exploitation', 'exploited',
               #human-wildlife conflict
               'domestic animal*',
               #artificial structures
               'highway*', '^mines',
               #policy interventions
               'reintroduc*', 'relocat*', 'translocat*'
```

Fill in rows as 'selected' if they have these terms (this doesn't affect the papers already included manually, but it may affect papers manually excluded; i.e., sensitivity is increased).

Get new summary.

#### 4.4 Abstract screening: round 3

Next, screen another 300 abstracts manually again manually, and see what other terms come up for filtering more of the articles with text mining.

```
# updated screening (9-15-2022)
scr_abs_4 <- screen_abstracts(scr_abs_3)
scr_abs_5 <- scr_abs_4 #duplicate for editing</pre>
```

Next, make a list of terms that were key words from the "selected" articles in this round.

```
abs_terms <- c(#terms related to...</pre>
               #resources
               'fishing', 'human water use', 'cattle', 'whaling',
                'market demand*', 'anthropic pressure*', 'harvest*',
               'aquarium', 'grazing', 'anthropogenic pressure*',
               'global demand',
                '^logging','reservoir*',
               #human influence in general
               'land use', 'globalization', 'globalisation',
               'land cover', 'landcover', #indirectly related to human influence
               'distance from village*', 'anthropogenic variable*',
               #human degradation or pollution
               'reclamation', 'human development', #'habitat loss',
               'anthropogenic litter', 'land use change*',
               #human-wildlife conflict
               'entanglement', 'wind energy development', 'human persecution',
               'habitat encroachment',
               #artificial structures
               'developed landscape*'
               #policy interventions
               #'restoration'
```

Fill in rows as 'selected' if they have these terms (this doesn't affect the papers already included manually, but it may affect papers manually excluded; i.e., sensitivity is increased).

## 1

## 2

## 3

#### 4.5 Abstract screening: round 4

excluded 903

selected 3982

<NA> 7798

Next, screen another 300 abstracts manually again, and see what other terms come up for filtering more of the articles with text mining.

```
# updated screening (9-14-2022)
scr_abs_6 <- screen_abstracts(scr_abs_5)
scr_abs_7 <- scr_abs_6 #duplicate for editing

# get number of abstracts for each
count(scr_abs_6, "screened_abstracts")</pre>
## screened abstracts freq
```

Next, make a list of terms that were key words from the "selected" articles in this round.

```
abs_terms <- c(#terms related to...

#resources

#'food security','commercially important',

#'socio-economic importan*',

#human influence in general

'human related activit*', 'human-related activit*',

'human-mediated','farmland','anthropogenic and',

'anthropogenic ecological change*','human settlement*',

'village','villages','anthropogenic expansion*',

'human infrastructure*','land-use change*',

#human degradation or pollution

'abandoned field*','abandoned area*','developed area*'

#human-wildlife conflict

#artificial structures

#policy interventions

)
```

Fill in rows as 'selected' if they have these terms (this doesn't affect the papers already included manually, but it may affect papers manually excluded; i.e., sensitivity is increased).

#### 4.6 Abstract screening: round 5

Next, make a list of terms that were key words from the "selected" articles in this round.

```
abs_terms <- c(#terms related to...
               #resources
               'rice field*', 'public land',
               #human influence in general
               'public use*',#'import*'
               'human influence index', 'closeness to ports',
               'domestic pig census', 'human-related change*',
               #human degradation or pollution
               'land-use data', 'land use data', 'land cover data', 'landcover data',
               'human-altered landscape*', 'built environment*', 'built-up area',
               'habitat destruction', 'land-use variable*', 'human-caused habitat loss',
               'human-induced global environmental change',
               #human-wildlife conflict
               'human-carnivore conflict*',
               #artificial structures
               '^city','^cities',
               #policy interventions
               'effectiveness of protected areas', 'effectiveness of PA',
               'eradicat*','forest management practice*','ability of PAs to conserve',
               'park boundar*', 'nonreserved area*'#, 'management stress',
```

Fill in rows as 'selected' if they have these terms (this doesn't affect the papers already included manually, but it may affect papers manually excluded; i.e., sensitivity is increased).

Get new summary.

#### 4.7 Abstract screening: round 6

```
abs_terms <- c(#terms related to...

#resources

'crop area*',

#human influence in general

'human-induced variables','land-use realism',

'Socio-Economic Connectedness ','human behavior',

#human degradation or pollution

'disturbed environment*','ski slope*','clear-cut logging',

'human modification','abandoned land*','abandonment','^mining activit*',

#human-wildlife conflict

'Human-sun bear conflict','human propert*','conflict area*',

'conflict hotspot*',

#artificial structures

'shelterbelt*', 'human-made habitats','farmstead'

#policy interventions
)
```

Get new summary.

#### 4.8 Abstract screening: round 7

```
abs_terms <- c(#terms related to...

#resources

'abstraction','demand for','commercial logging',

#'food security','commercially important',

#'socio-economic importan*',

#'social value*',

#human influence in general

'human and natural variable*','human and natural predictor*',

'landscape-Use','land-use pattern','anthropogenic factor*',

'land-cover','natural *land*','National Land Cover Data',

'*NLCD*','forest availability',

#human degradation or pollution

'oil development','gas development','*restored habitat',
```

```
'human-caused',
#human-wildlife conflict
'distance to anthropogenic','density of vehicle roads',
'road density','density of roads','avoid* road*','road* avoid*',
'avoid* of road*','avoid* of trail*','avoid* trail*',
'attract* to road*', 'attract* to trail*',
#artificial structures

#policy interventions
'undeveloped','human intervention*'
#,'management stress',
)
```

Get new summary.

<NA> 6585

## 3

We realized that the grepl() default is ignore.case=FALSE. This means that the search terms were CASE SENSITIVE. Prior to proceeding to another round of abstract screening, all terms from round 1 thru 7 will be rerun to the current dataset scr\_abs\_13 to remove case sensitivity.

```
abs_terms <- c(#terms related to...

### round 1

#resources
    'agricultur*', 'aquacultur*', 'hunting', 'viticultur*', 'hunted',
    'plantation*', 'livestock', 'game bird*', 'palm oil', 'cropland',
    #human influence in general
    'anthropogenically-transformed', 'urban*', 'anthropogenic impact*',
    'human impact*',
    #human degradation or pollution
    'pollutant*', 'insecticide*', 'pesticide*',
    'land abandonment', 'deforest*',
    'landuse change', 'land use change', 'landcover change',</pre>
```

```
'land cover change', 'habitat conversion', 'land conversion',
'habitat degradation','oil spill*',
#human-wildlife conflict
'collision', 'human-wildlife conflict', 'human interaction*',
'road kill', 'roadkill',
#artificial structures
'artificial surface*', 'hydropower', 'artificial light',
'artificial night light',
### round 2
#human influence in general
'anthropogenic change', 'human use', 'anthropogenic activit*',
'shipping', 'anthropogenic global change', 'human activit*',
'human population', 'anthropogenic stressor*', 'human footprint*',
'anthropogenic driver*', 'human pressure*', 'synanthrop*',
'anthropic intervention', 'anthropogenic cause*', 'anthropogenic habitat',
'human-dominated',
#human degradation or pollution
'exploitation', 'exploited',
#human-wildlife conflict
'domestic animal*',
#artificial structures
'highway*','^mines',
#policy interventions
'reintroduc*', 'relocat*', 'translocat*',
### round 3
#resources
'fishing', 'human water use', 'cattle', 'whaling',
'market demand*', 'anthropic pressure*', 'harvest*',
'aquarium', 'grazing', 'anthropogenic pressure*',
'global demand',
'^logging','reservoir*',
#human influence in general
'land use', 'globalization', 'globalisation',
'land cover', 'landcover', #indirectly related to human influence
'distance from village*', 'anthropogenic variable*',
#human degradation or pollution
'reclamation', 'human development', #'habitat loss',
'anthropogenic litter', 'land use change*',
#human-wildlife conflict
'entanglement', 'wind energy development', 'human persecution',
'habitat encroachment',
#artificial structures
'developed landscape*',
### round 4
#human influence in general
'human related activit*', 'human-related activit*',
'human-mediated', 'farmland', 'anthropogenic and',
```

```
'anthropogenic ecological change*', 'human settlement*',
'village', 'villages', 'anthropogenic expansion*',
'human infrastructure*', 'land-use change*',
#human degradation or pollution
'abandoned field*', 'abandoned area*', 'developed area*',
### round 5
#resources
'rice field*', 'public land',
#human influence in general
'public use*',
'human influence index', 'closeness to ports',
'domestic pig census', 'human-related change*',
#human degradation or pollution
'land-use data', 'land use data', 'land cover data', 'landcover data',
'human-altered landscape*', 'built environment*', 'built-up area',
'habitat destruction', 'land-use variable*', 'human-caused habitat loss',
'human-induced global environmental change',
#human-wildlife conflict
'human-carnivore conflict*',
#artificial structures
'^city','^cities',
#policy interventions
'effectiveness of protected areas', 'effectiveness of PA',
'eradicat*', 'forest management practice*', 'ability of PAs to conserve',
'park boundar*', 'nonreserved area*',
### round 6
#resources
'crop area*',
#human influence in general
'human-induced variables', 'land-use realism',
'Socio-Economic Connectedness ', 'human behavior',
#human degradation or pollution
'disturbed environment*', 'ski slope*', 'clear-cut logging',
'human modification', 'abandoned land*', 'abandonment', '^mining activit*',
#human-wildlife conflict
'Human-sun bear conflict', 'human propert*', 'conflict area*',
'conflict hotspot*',
#artificial structures
'shelterbelt*', 'human-made habitats', 'farmstead',
### round 7
#resources
'abstraction', 'demand for', 'commercial logging',
#human influence in general
'human and natural variable*', 'human and natural predictor*',
'landscape-Use', 'land-use pattern', 'anthropogenic factor*',
'land-cover', 'natural *land*', 'National Land Cover Data',
'*NLCD*','forest availability',
```

```
#human degradation or pollution
'oil development','gas development','*restored habitat',
'human-caused',
#human-wildlife conflict
'distance to anthropogenic','density of vehicle roads',
'road density','density of roads','avoid* road*','road* avoid*',
'avoid* of road*','avoid* of trail*','avoid* trail*',
'attract* to road*', 'attract* to trail*',
#policy interventions
'undeveloped','human intervention*'
)
```

Get new summary.

#### 4.9 Abstract screening: round 8

```
abs_terms <- c(#terms related to...
               #resources
               'timber production',
               #'food security', 'commercially important',
               #'socio-economic importan*',
               'social value*',
               #human influence in general
               'ballast','^cities','orchard*','pasture*','human-induced',
               'companies that import', 'imported by compan*', 'company import*',
               'companies import', 'human occup*', 'land-use', 'human-pressure*',
               #human degradation or pollution
               'landscape chang*', 'chang* landscape*',
               #human-wildlife conflict
               'poach*', 'access* roads', 'access to road*',
               'road* access*', 'industrial cent*', 'power development',
               'wind energy',
               #artificial structures
               'anthropogenic structure*'#,
               #policy interventions
               #'reconstruction'
```

Get new summary.

#### 4.10 Abstract screening: round 9

```
# updated screening (9-19-2022)
scr_abs_17 <- screen_abstracts(scr_abs_16)
scr_abs_18 <- scr_abs_17 #duplicate for editing</pre>
```

Next, make a list of terms that were key words from the "selected" articles in this round.

```
abs_terms <- c(#terms related to...
               #resources
               #'food security','commercially important',
               #'socio-economic importan*',
               #human influence in general
               'anthropogenic* introduc*', 'forest edge*', 'changes that human*',
               'changes that people', 'human influence*',
               'anthropogenic environmental change*', 'human incidence*',
               'international travel', 'freight transport*',
               #'travel',
               #human degradation or pollution
               'pitlake', '^pit lake', 'anthropogenic stress', 'human-made',
               #human-wildlife conflict
               'roadside', 'tourism development'
               #artificial structures
               #policy interventions
               #'reconstruction'
```

Fill in rows as 'selected' if they have these terms (this doesn't affect the papers already included manually, but it may affect papers manually excluded; i.e., sensitivity is increased).

Get new summary.

excluded 2250

selected 4599

<NA> 5834

## 1

## 2

## 3

```
# get number of abstracts for each
count(scr_abs_18, "screened_abstracts")
## screened_abstracts freq
```

#### 4.11 Abstract screening: round 10

Next, make a list of terms that were key words from the "selected" articles in this round.

```
abs_terms <- c(#terms related to...
               #resources
               #'farming',
               #human influence in general
               'anthropo* parameter*', 'anthropo* feature*',
               #human degradation or pollution
               'pollution', 'pollutant*',
               '^escape*','air traffic',
               #human-wildlife conflict
               'distance* from road*', 'distance* from *trail*', 'distance* from highway*',
               'distance* from settlement*', 'distance* from village*', 'fisher* bycatch',
               #'bycatch',
               'ship strike*', 'anthro* sound*',
               #artificial structures
               #policy interventions
               #'interven*',
               'anthropo* *nourishment','renourishment event',
               'anthropic restoration', 'uninterrupt* area*', 'forest thinning',
               'thin* forest*', 'habitat management', 'manag* habitat*'
```

Fill in rows as 'selected' if they have these terms (this doesn't affect the papers already included manually, but it may affect papers manually excluded; i.e., sensitivity is increased).

#### 4.12 Abstract screening: round 11

Screen another 300 articles and add new terms.

Next, make a list of terms that were key words from the "selected" articles in this round.

Fill in rows as 'selected' if they have these terms (this doesn't affect the papers already included manually, but it may affect papers manually excluded; i.e., sensitivity is increased).

```
ignore.case = TRUE,
    #search column

scr_abs_22$abstract),

#fill row in designated column
'selected', scr_abs_22$screened_abstracts)
```

Get new summary.

```
# get number of abstracts for each
count(scr_abs_22, "screened_abstracts")
```

After screening the abstracts, We found that the wildcard symbol (\*) being used in the grepl() command was not registering. We are thus re-applying all the terms again, so that the wildcards can be applied.

```
abs_terms <- c(#terms related to...
               ### round 1
               #resources
               'agricultur*', 'aquacultur*', 'hunting', 'viticultur*', 'hunted',
               'plantation*', 'livestock', 'game bird*', 'palm oil', 'cropland',
               #human influence in general
               'anthropogenically-transformed', 'urban*', 'anthropogenic impact*',
               'human impact*',
               #human degradation or pollution
               'pollutant*', 'insecticide*', 'pesticide*',
               'land abandonment', 'deforest*',
               'landuse change', 'land use change', 'landcover change',
               'land cover change', 'habitat conversion', 'land conversion',
               'habitat degradation','oil spill*',
               #human-wildlife conflict
               'collision', 'human-wildlife conflict', 'human interaction*',
               'road kill', 'roadkill',
               #artificial structures
               'artificial surface*', 'hydropower', 'artificial light',
               'artificial night light',
               ### round 2
               #human influence in general
               'anthropogenic change', 'human use', 'anthropogenic activit*',
               'shipping', 'anthropogenic global change', 'human activit*',
               'human population', 'anthropogenic stressor*', 'human footprint*',
               'anthropogenic driver*', 'human pressure*', 'synanthrop*',
                'anthropic intervention', 'anthropogenic cause*', 'anthropogenic habitat',
               'human-dominated',
               #human degradation or pollution
```

```
'exploitation', 'exploited',
#human-wildlife conflict
'domestic animal*',
#artificial structures
'highway*','^mines',
#policy interventions
'reintroduc*', 'relocat*', 'translocat*',
### round 3
#resources
'fishing', 'human water use', 'cattle', 'whaling',
'market demand*', 'anthropic pressure*', 'harvest*',
'aquarium', 'grazing', 'anthropogenic pressure*',
'global demand',
'^logging','reservoir*',
#human influence in general
'land use', 'globalization', 'globalisation',
'land cover', 'landcover', #indirectly related to human influence
'distance from village*', 'anthropogenic variable*',
#human degradation or pollution
'reclamation', 'human development', #'habitat loss',
'anthropogenic litter', 'land use change*',
#human-wildlife conflict
'entanglement', 'wind energy development', 'human persecution',
'habitat encroachment',
#artificial structures
'developed landscape*',
### round 4
#human influence in general
'human related activit*', 'human-related activit*',
'human-mediated','farmland','anthropogenic and',
'anthropogenic ecological change*', 'human settlement*',
'village', 'villages', 'anthropogenic expansion*',
'human infrastructure*', 'land-use change*',
#human degradation or pollution
'abandoned field*', 'abandoned area*', 'developed area*',
### round 5
#resources
'rice field*', 'public land',
#human influence in general
'public use*',
'human influence index', 'closeness to ports',
'domestic pig census', 'human-related change*',
#human degradation or pollution
'land-use data', 'land use data', 'land cover data', 'landcover data',
'human-altered landscape*', 'built environment*', 'built-up area',
'habitat destruction', 'land-use variable*', 'human-caused habitat loss',
'human-induced global environmental change',
```

```
#human-wildlife conflict
'human-carnivore conflict*',
#artificial structures
'^city','^cities',
#policy interventions
'effectiveness of protected areas', 'effectiveness of PA',
'eradicat*', 'forest management practice*', 'ability of PAs to conserve',
'park boundar*', 'nonreserved area*',
### round 6
#resources
'crop area*',
#human influence in general
'human-induced variables', 'land-use realism',
'Socio-Economic Connectedness ', 'human behavior',
#human degradation or pollution
'disturbed environment*', 'ski slope*', 'clear-cut logging',
'human modification', 'abandoned land*', 'abandonment', '^mining activit*',
#human-wildlife conflict
'Human-sun bear conflict', 'human propert*', 'conflict area*',
'conflict hotspot*',
#artificial structures
'shelterbelt*', 'human-made habitats', 'farmstead',
### round 7
#resources
'abstraction', 'demand for', 'commercial logging',
#human influence in general
'human and natural variable*', 'human and natural predictor*',
'landscape-Use', 'land-use pattern', 'anthropogenic factor*',
'land-cover', 'natural *land*', 'National Land Cover Data',
'*NLCD*', 'forest availability',
#human degradation or pollution
'oil development', 'gas development', '*restored habitat',
'human-caused',
#human-wildlife conflict
'distance to anthropogenic', 'density of vehicle roads',
'road density', 'density of roads', 'avoid* road*', 'road* avoid*',
'avoid* of road*', 'avoid* of trail*', 'avoid* trail*',
'attract* to road*', 'attract* to trail*',
#policy interventions
'undeveloped', 'human intervention*',
### round 8
#resources
'timber production',
'social value*'.
#human influence in general
'ballast','^cities','orchard*','pasture*','human-induced',
'companies that import', 'imported by compan*', 'company import*',
```

```
'companies import', 'human occup*', 'land-use', 'human-pressure*',
#human degradation or pollution
'landscape chang*', 'chang* landscape*',
#human-wildlife conflict
'poach*', 'access* roads', 'access to road*',
'road* access*', 'industrial cent*', 'power development',
'wind energy',
#artificial structures
'anthropogenic structure*',
### round 9
#human influence in general
'anthropogenic* introduc*', 'forest edge*', 'changes that human*',
'changes that people', 'human influence*',
'anthropogenic environmental change*', 'human incidence*',
'international travel', 'freight transport*',
#'travel',
#human degradation or pollution
'pitlake', '^pit lake', 'anthropogenic stress', 'human-made',
#human-wildlife conflict
'roadside', 'tourism development',
### round 10
#human influence in general
'anthropo* parameter*', 'anthropo* feature*',
#human degradation or pollution
'pollution', 'pollutant*',
'^escape*','air traffic',
#human-wildlife conflict
'distance* from road*', 'distance* from *trail*', 'distance* from highway*',
'distance* from settlement*', 'distance* from village*', 'fisher* bycatch',
#'bycatch',
'ship strike*', 'anthro* sound*',
#artificial structures
#policy interventions
#'interven*'.
'anthropo* *nourishment','renourishment event',
'anthropic restoration', 'uninterrupt* area*', 'forest thinning',
'thin* forest*', 'habitat management', 'manag* habitat*',
### round 11
#resources
'shade tree*','coffee farm*',
#human influence in general
'anthrop* dominated', 'anthrop*-dominated',
'human dominated', 'human-dominated',
'dominated by human*', 'dominated by people',
'anthrop* indices', 'anthrop* index',
#human degradation or pollution
#human-wildlife conflict
```

```
'barrier effect','main road*',
#artificial structures
#policy interventions
'vector control','manage* factor*',
'protect* area*'
)
```

Fill in rows as 'selected' if they have these terms.

```
# get number of abstracts for each
count(scr_abs_23, "screened_abstracts")
```

```
## screened_abstracts freq
## 1 excluded 2802
## 2 selected 4741
## 3 <NA> 5140
```

#### 4.13 Abstract screening: round 12

Screen another 300 articles and add new terms.

Get new summary.

#### 4.14 Abstract screening: round 13

Screen another 300 articles and add new terms.

```
# updated screening (9-23-2022)
scr_abs_26 <- screen_abstracts(scr_abs_25)
scr_abs_27 <- scr_abs_26 #duplicate for editing

# get number of abstracts for each
count(scr_abs_26, "screened_abstracts")</pre>
```

Next, make a list of terms that were key words from the "selected" articles in this round.

```
abs_terms <- c(#terms related to...
               #resources
               'fishery catch',
               #human influence in general
               'accessible to human.*', 'anthropogenic predictor.*',
               'anthropogenic parameter.*', 'anthropogenic extinction',
               'anthropogenic feature.*', 'demand.* for timber',
               'preindustrial period', 'industrialization', 'industrialisation',
               #human degradation or pollution
               'anthropogenic degradation', 'edge.* of fields', 'field edge.*',
               #human-wildlife conflict
               'conflict with human.*', 'human safety',
               'human disturbed', 'human-disturbed', 'disturbed by human.*',
               'anthropo.* disturb.*',
               #artificial structures
               'Ski-resort.*','ski resort.*'
               #policy interventions
```

Fill in rows as 'selected' if they have these terms (this doesn't affect the papers already included manually, but it may affect papers manually excluded; i.e., sensitivity is increased).

Get new summary.

```
# get number of abstracts for each
count(scr_abs_27, "screened_abstracts")
```

#### 4.15 Abstract screening: round 14

Screen another 300 articles and add new terms.

Next, make a list of terms that were key words from the "selected" articles in this round.

```
abs_terms <- c(#terms related to...
               #resources
               'planting area.*', 'intentional.* introduc.*',
               'trade pattern.*',
               #human influence in general
               'socioeconomic factor.*',
               #human degradation or pollution
               'industrial source.*', 'municipal source.*',
               'anthropogenic fire.*', 'human induced threat.*', 'human-induced threat.*',
               'threat.* by human.*', 'anthrop.* threat.*',
               #human-wildlife conflict
               #artificial structures
               'non-ecological', 'nonecological',
               #policy interventions
               'introduction site.*',
               'occurs in protected areas'
```

Fill in rows as 'selected' if they have these terms (this doesn't affect the papers already included manually, but it may affect papers manually excluded; i.e., sensitivity is increased).

```
# get number of abstracts for each
count(scr_abs_29, "screened_abstracts")
```

```
## screened_abstracts freq
## 1 excluded 3626
## 2 selected 4880
## 3 <NA> 4177
```

#### 4.16 Abstract screening: round 15

Screen another 300 articles and add new terms.

Next, make a list of terms that were key words from the "selected" articles in this round.

```
abs_terms <- c(#terms related to...
               #resources
               'silvicultur*',
               #human influence in general
               'socio-economic covariate.*', 'socioeconomic covariate.*',
               'accidentally introduced', 'accidentally-introduced',
               'unintentionally transported', 'transport hub',
               'inadvertent.* introduc.*',
               #human degradation or pollution
               'post-mining landscape',
               #human-wildlife conflict
               'intentional feeding', 'unintentional feeding',
               'transport pressure',
               #artificial structures
               'dam construction','^on road.*',
               #policy interventions
               'prescribed fire', 'fire suppression',
               'euthan*'
```

Fill in rows as 'selected' if they have these terms (this doesn't affect the papers already included manually, but it may affect papers manually excluded; i.e., sensitivity is increased).

#### 4.17 Abstract screening: round 16

Screen another 300 articles and add new terms.

Next, make a list of terms that were key words from the "selected" articles in this round.

```
abs_terms <- c(#terms related to...
               #resources
               '^hunted area',
               #human influence in general
               'human presence', 'humans created niches', 'human economic activity',
               #human degradation or pollution
               'due to energy development',
               #human-wildlife conflict
               'fishery has the potential to compete with', 'collector.* interest',
               'distance to road',
               #artificial structures
               'settlements', 'night-time light.*', 'nighttime light', 'artificial light',
               'ecological impact of hydroelectric power plants',
               #policy interventions
               'effectiveness of conservation areas',
               'expanding protected areas',
               'measures to protect has led to',
               'protection measures',
               'conservation and economic development'
```

Fill in rows as 'selected' if they have these terms (this doesn't affect the papers already included manually, but it may affect papers manually excluded; i.e., sensitivity is increased).

Get new summary.

#### 4.18 Abstract screening: round 17

Screen another 300 articles and add new terms.

```
abs_terms <- c(#terms related to...

#resources

'threat from fisheries','fisher.* threat',

#human influence in general

'human and spatial variable.*','human and environment.* variable.*',

'human and spatial predictor','human and environent.* predictor',

'human-related variable','human related variable','human variable',

'human-related predictor','human related predictor','human predictor',

'human induced pressure','human-induced pressure','exacerbated by human',

'human degradation or pollution

#human-wildlife conflict
```

```
#'stray cat','stray dog','domestic cat','domestic dog',
#' human and dog densities'
#artificial structures
#policy interventions
'during ecological restoration.*',
'wolf (Canis lupus) restoration'
#'species restoration'
#'assisted migration'
)
```

Get new summary.

## 3

```
# get number of abstracts for each
count(scr_abs_35, "screened_abstracts")
```

```
## screened_abstracts freq
## 1 excluded 4470
## 2 selected 4963
## 3 <NA> 3250
```

#### 4.19 Abstract screening: round 18

Screen another 300 articles and add new terms.

<NA> 2950

Get new summary.

#### 4.20 Abstract screening: round 19

Screen another 300 articles and add new terms.

```
# updated screening (9-26-2022)
scr_abs_38 <- screen_abstracts(scr_abs_37)
scr_abs_39 <- scr_abs_38 #duplicate for editing

# get number of abstracts for each
count(scr_abs_38, "screened_abstracts")</pre>
```

In this round, many of the papers are very specific. Not really common phrases being used here.

Next, make a list of terms that were key words from the "selected" articles in this round.

```
abs_terms <- c(#terms related to...
               #resources
               'olive cultivation', 'prevalence in vineyards',
               #human influence in general
               'human introduc.*', 'biocontrol agent',
               'anthropogenic landscape.*',
               'anthropogenic vector.*', 'anthropic', 'role played by human.*',
               #human degradation or pollution
               'amount of cleared land', 'marine litter', 'microplastic',
               'mesoplastics', 'macrolitter', 'anthropogenic modification',
               #human-wildlife conflict
               'coincid.* with human', 'dive tourism', 'conflict.* with human.*',
               #artificial structures
               #'hydropeaking',
               'major harbor.*',
               'airport.*',
               'entry through ports',
               #policy interventions
               'manipulated organism*',
               'political division.*',
               'Water Framework Directive', 'sensitive to forest management'
```

Fill in rows as 'selected' if they have these terms (this doesn't affect the papers already included manually, but it may affect papers manually excluded; i.e., sensitivity is increased).

```
# get number of abstracts for each
count(scr_abs_39, "screened_abstracts")
```

```
## screened_abstracts freq
## 1 excluded 5024
## 2 selected 5022
## 3 <NA> 2637
```

# 4.21 Abstract screening: round 20

Screen another 300 articles and add new terms.

```
# updated screening (9-27-2022)
scr_abs_40 <- screen_abstracts(scr_abs_39)
scr_abs_41 <- scr_abs_40 #duplicate for editing</pre>
# set number of shatmasts for each
```

```
# get number of abstracts for each
count(scr_abs_40, "screened_abstracts")
```

In this round, many of the papers are very specific. Not really common phrases being used here.

Next, make a list of terms that were key words from the "selected" articles in this round.

Fill in rows as 'selected' if they have these terms (this doesn't affect the papers already included manually, but it may affect papers manually excluded; i.e., sensitivity is increased).

Get new summary.

#### 4.22 Abstract screening: round 21

Screen another 300 articles and add new terms.

Next, make a list of terms that were key words from the "selected" articles in this round.

Fill in rows as 'selected' if they have these terms (this doesn't affect the papers already included manually, but it may affect papers manually excluded; i.e., sensitivity is increased).

```
ignore.case = TRUE,
    #search column
    scr_abs_44$abstract),
#fill row in designated column
'selected', scr_abs_44$screened_abstracts)
```

Get new summary.

# 4.23 Abstract screening: round 22

Screen another 300 articles and add new terms.

Next, make a list of terms that were key words from the "selected" articles in this round.

Fill in rows as 'selected' if they have these terms (this doesn't affect the papers already included manually, but it may affect papers manually excluded; i.e., sensitivity is increased).

Get new summary.

#### 4.24 Abstract screening: round 23

Screen another 300 articles and add new terms.

Next, make a list of terms that were key words from the "selected" articles in this round.

```
abs_terms <- c(#terms related to...

#resources

'land resource use',

#human influence in general

'human access','socioecological framework',

'human-environment relationship','human means',

'arrive by human','human assisted','human-assisted',

#human degradation or pollution

'development project',

#human-wildlife conflict

'pet store',

#artificial structures
```

```
'powerline',
#policy interventions
'restoration success','restoration design','after restoration'
)
```

Fill in rows as 'selected' if they have these terms (this doesn't affect the papers already included manually, but it may affect papers manually excluded; i.e., sensitivity is increased).

Get new summary.

## 3

#### 4.25 Abstract screening: round 24

Screen another 300 articles and add new terms.

<NA> 1118

Next, make a list of terms that were key words from the "selected" articles in this round.

```
#human degradation or pollution
'^mining activit*','trawling impact',
#human-wildlife conflict
'cost of conflict resolution','human-cougar conflict',
'resource utilization with human','co-occurrence of human',
#artificial structures
#policy interventions
#'flow management',
'restorations strateg','restoration strateg'
)
```

Fill in rows as 'selected' if they have these terms (this doesn't affect the papers already included manually, but it may affect papers manually excluded; i.e., sensitivity is increased).

Get new summary.

### 4.26 Abstract screening: round 25

Screen another 300 articles and add new terms.

Next, make a list of terms that were key words from the "selected" articles in this round.

Fill in rows as 'selected' if they have these terms (this doesn't affect the papers already included manually, but it may affect papers manually excluded; i.e., sensitivity is increased).

Get new summary.

#### 4.27 Abstract screening: round 26

Screen another 300 articles and add new terms.

Next, make a list of terms that were key words from the "selected" articles in this round.

```
abs terms <- c(#terms related to...
               #resources
               'geographic expansion of cultivated populations',
               #human influence in general
               'anthropogenic land', 'socioeconomic variable',
               'human and historic variables', 'impact of humans',
               'social factor',
               #human degradation or pollution
               'development and population', 'land clearing',
               #human-wildlife conflict
               'anthropogenic mortality',
               #artificial structures
               'density of road', 'human-modified',
               #policy interventions
               'historical land management', 'MPA zoning',
               'with irrigation', 'without irrigation', 'distance to protected area'
```

Fill in rows as 'selected' if they have these terms (this doesn't affect the papers already included manually, but it may affect papers manually excluded; i.e., sensitivity is increased).

Get new summary.

### 4.28 Abstract screening: round 27

Screen another 300 articles and add new terms.

```
# updated screening (10-30-2022)
scr_abs_55 <- screen_abstracts(scr_abs_54)
scr_abs_56 <- scr_abs_55 #duplicate for editing</pre>
```

Next, make a list of terms that were key words from the "selected" articles in this round.

Fill in rows as 'selected' if they have these terms (this doesn't affect the papers already included manually, but it may affect papers manually excluded; i.e., sensitivity is increased).

Get new summary.

```
# get number of abstracts for each
count(scr_abs_56, "screened_abstracts")
```

# 4.29 Abstract screening: round 28

Screen last 300 articles and add new terms.

Next, make a list of terms that were key words from the "selected" articles in this round.

Fill in rows as 'selected' if they have these terms (this doesn't affect the papers already included manually, but it may affect papers manually excluded; i.e., sensitivity is increased).

Get new summary.

# 5 Abstract Screening Summary

Make copies and separate table

```
# copy final abstract screening result
screened_final <- scr_abs_58

# separate by yes vs no
screened_yes <- screened_final[screened_final$screened_abstracts=='selected',]
screened_no <- screened_final[screened_final$screened_abstracts=='excluded',]</pre>
```

Get summaries for PRISMA framework

```
# summary for PRISMA
  paste('Total articles in WoS search up to 2021:', nrow(abs.df))
  paste('Total duplicates removed:', nrow(abs.df)-nrow(abs2.df))
  paste('Total articles screened:', nrow(screened_final))
  paste('Total articles selected:', nrow(screened_yes))
  paste('Total articles excluded:', nrow(screened_no))
  paste('Percent articles selected:',
        round((nrow(screened yes)/nrow(screened final))*100,2),'%')
  paste('Percent articles excluded:',
        round((nrow(screened no)/nrow(screened final))*100,2),'%')
## [1] "Total articles in WoS search up to 2021: 12691"
## [1] "Total duplicates removed: 8"
## [1] "Total articles screened: 12683"
## [1] "Total articles selected: 5177"
## [1] "Total articles excluded: 7506"
## [1] "Percent articles selected: 40.82 %"
## [1] "Percent articles excluded: 59.18 %"
Next, save as CSV.
# save as CSV
  write.csv(screened_final, paste0(data.dir, "screened_final.csv"), row.names=FALSE)
  write.csv(screened yes, pasteO(data.dir, "screened yes.csv"), row.names=FALSE)
  write.csv(screened no, paste0(data.dir, "screened no.csv"), row.names=FALSE)
```

# 6 Setting up for full article review

Next, we will go through each of the selected 5177 articles from abstract screening (files from screened\_yes.csv) to determine whether human predictors are being used in the SDMs. Accepted full articles for the synthesis will require that human predictors are at least within the SDM training stage. This is a critical point, as some papers may list human predictors in their model methods, but will use human predictors post-hoc for e.g., masking predictions of distributions without human influence (such as masking out urban areas). This full article review stage will also include the review of any supplementary materials, as methods are sometimes described there, instead.

For this, a new CSV will be made, with only the selected articles and the following data fields:

uid

- year
- title
- author
- journal
- doi

```
# Create a new table with only UID, year, title, author, journal, and DOI.
  full_screen <- subset(screened_yes,</pre>
                         select=c('uid','year','title','author','journal','doi'))
# add a column for use when screening full articles
 full_screen$relevant <- ''</pre>
# Re-order the columns to move UID to the front
 full_screen <- full_screen[,c('uid','year','title','author','journal','doi','relevant')]</pre>
# Sort the data by uid (originally uid is most recent year, followed by the title)
 full_screen <- full_screen[order(full_screen$uid),]</pre>
# rewrite DOI to an HTML link
  #(note that some of these will be NA, so new HTML will be manually added if relevant)
 full_screen$doi <- paste0('http://doi.org/',full_screen$doi)</pre>
# Save as a csv
  write.csv(full_screen, paste0(data.dir, "full_screening.csv"), row.names=FALSE)
# show example
 options(width=85) # ensure width
 head(full_screen)
         uid year
           2 2021
## 4435
## 11471
           4 2021
## 7980
          7 2021
## 1715
          9 2021
## 5645
        10 2021
## 11175 13 2021
##
## 4435
                                     'The Best of Two Worlds'-Combining Classifier Fusion and Ecological
## 11471
                                  A Bayesian framework for assessing extinction risk based on ordinal ca
                        A Bioclimate-Based Maximum Entropy Model for Comperiella calauanica Barrion, Ali
## 7980
## 1715 A Case Study in Connecting Fisheries Management Challenges With Models and Analysis to Support
                                            A comparative study on habitat models for adult bigeye tuna
## 5645
## 11175
                A comprehensive approach to assessing the future persistence of the endangered rainfore
##
## 4435
## 11471
## 7980
## 1715 Tommasi, Desiree and deReynier, Yvonne and Townsend, Howard and Harvey, Chris J. and Satterthw
## 5645
## 11175
```

```
##
                              journal
                                                                               doi
                       REMOTE SENSING
                                                 http://doi.org/10.3390/rs13163287
## 4435
## 11471
              BIOLOGICAL CONSERVATION http://doi.org/10.1016/j.biocon.2020.108866
## 7980
                              INSECTS
                                           http://doi.org/10.3390/insects12010026
                                         http://doi.org/10.3389/fmars.2021.624161
## 1715
         FRONTIERS IN MARINE SCIENCE
## 5645
               FISHERIES OCEANOGRAPHY
                                                 http://doi.org/10.1111/fog.12539
                                                    http://doi.org/10.1071/BT20160
## 11175 AUSTRALIAN JOURNAL OF BOTANY
         relevant
##
## 4435
## 11471
## 7980
## 1715
## 5645
## 11175
```

# 7 Save

```
# save progress
save.image("SDMs_human_lit_review_I.RData")
```

#### THIS IS THE END OF THE SCRIPT.

See "Human Influence in SDMs: Literature Review (Part II)" for next steps.