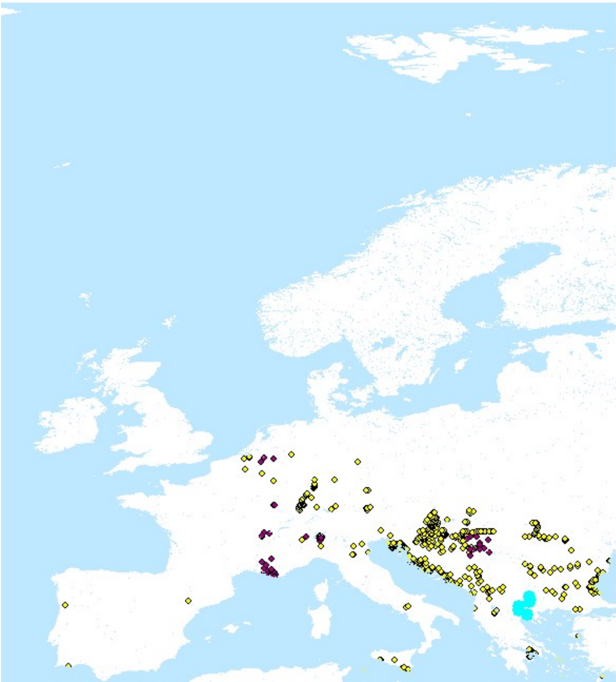
(WHAT) DATA FOR MAPS

Funded by the Horizon 2020 Framework Programme of the European Union

BASICS OF TURNING DATA INTO MAPS



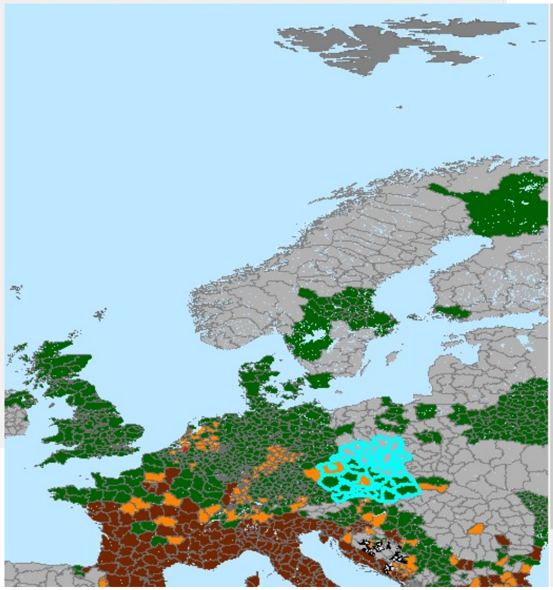
Table

Albopictus

	FID	Shape *	rowid_	source	albopa	statusterm	latitude	longitude	lat10k	lon10k	lat1
▸	977	Point	978	VN19DAT	1	Established	40.59221	23.05472	40.6	23.1	40.
	978	Point	979	VN19DAT	1	Established	40.59358	23.05281	40.6	23.1	40.
	991	Point	992	VN19DAT	1	Established	40.83193	22.44997	40.8	22.4	40.
	1008	Point	1009	VN19DAT	1	Established	40.93351	22.46952	40.9	22.5	40.
	1116	Point	1117	VN19DAT	1	Introduced	41.16538	22.49795	41.2	22.5	41.
	1247	Point	1248	VN19DAT	1	Established	41.396	23.217	41.4	23.2	4
	1248	Point	1249	VN19DAT	1	Established	41.396	23.217	41.4	23.2	4
	1249	Point	1250	VN19DAT	1	Established	41.39848	23.21952	41.4	23.2	4
	1250	Point	1251	VN19DAT	1	Established	41.39848	23.21952	41.4	23.2	4
	1251	Point	1252	VN19DAT	1	Established	41.39848	23.21952	41.4	23.2	4
	1252	Point	1253	VN19DAT	1	Established	41.39848	23.21952	41.4	23.2	4
	1278	Point	1279	VN19DAT	1	Established	41.56719	23.28093	41.6	23.3	41.
	1279	Point	1280	VN19DAT	1	Established	41.56719	23.28093	41.6	23.3	41.
	1310	Point	1311	VN19DAT	0	Absent	41.78276	23.15564	41.8	23.2	41.
	1311	Point	1312	VN19DAT	0	Absent	41.78276	23.15564	41.8	23.2	41.

Simple maps are just spreadsheets with pictures attached to each row:

Points and countries



14.691014	15.94094	AT12	Austria	Absent	15/07/2016		<Null>	
15.733306	16.953785	AT12	Austria	Absent	15/07/2016		<Null>	
15.712512	16.97662	AT12	Austria	Absent	15/07/2016		<Null>	
16.181901	16.576992	AT13	Austria	Absent	15/07/2016	Present	28/11/2019	
13.396288	15.532582	CZ02	Czech Republic	Introduced	26/04/2018	Presavia	02/12/2019	
13.535966	15.602685	CZ03	Czech Republic	Absent	01/10/2010		<Null>	
14.343548	15.632219	CZ05	Czech Republic	No Data	01/10/2010		<Null>	
15.104145	16.583434	CZ05	Czech Republic	Absent	01/10/2010		<Null>	
15.362709	16.865996	CZ05	Czech Republic	Absent	01/10/2010		<Null>	
14.889696	16.416133	CZ06	Czech Republic	Absent	01/10/2010		<Null>	
15.54245	17.64693	CZ06	Czech Republic	Introduced	01/01/2015	Presavia	02/12/2019	
16.711997	17.916857	CZ07	Czech Republic	Absent	01/10/2010		<Null>	
17.11208	18.415687	CZ07	Czech Republic	Absent	01/10/2010		<Null>	
17.147395	18.859236	CZ08	Czech Republic	No Data	01/10/2010		<Null>	
18.074521	19.312516	PL116	Poland	No Data	01/10/2010		<Null>	
18.612386	19.832163	PL224	Poland	No Data	01/10/2010		<Null>	
18.565684	19.480765	PL225	Poland	No Data	01/10/2010		<Null>	
18.035061	18.775039	PL227	Poland	No Data	01/10/2010		<Null>	
18.478264	19.134466	PL228	Poland	No Data	01/10/2010		<Null>	
18.357744	18.86234	PL229	Poland	No Data	01/10/2010		<Null>	
18.618048	19.253811	PL22C	Poland	No Data	01/10/2010		<Null>	
17.128033	18.481648	PL416	Poland	No Data	01/10/2010		<Null>	
16.80734	17.175163	PL514	Poland	No Data	01/10/2010		<Null>	
14.817373	16.480191	PL515	Poland	No Data	01/10/2010		<Null>	
16.067617	17.164896	PL517	Poland	No Data	01/10/2010		<Null>	
16.331656	17.79527	PL518	Poland	No Data	01/10/2010		<Null>	
16.907925	18.068069	PL523	Poland	No Data	01/10/2010		<Null>	
17.466549	18.695487	PL524	Poland	No Data	01/10/2010		<Null>	
16.833229	17.529213	SK010	Slovakia	Absent	01/10/2010		<Null>	
16.933688	17.98447	SK021	Slovakia	Absent	01/10/2010	Presavia	02/12/2019	
17.353044	18.826351	SK022	Slovakia	Absent	01/10/2010		<Null>	
17.705437	19.072107	SK023	Slovakia	Absent	01/10/2010		<Null>	
18.322436	20.06043	SK031	Slovakia	Absent	01/10/2010		<Null>	
18.478677	20.4699	SK032	Slovakia	Absent	01/10/2010		<Null>	

Maps are usually 1 (colour) or at most 2 parameters only (e.g. size and colour)

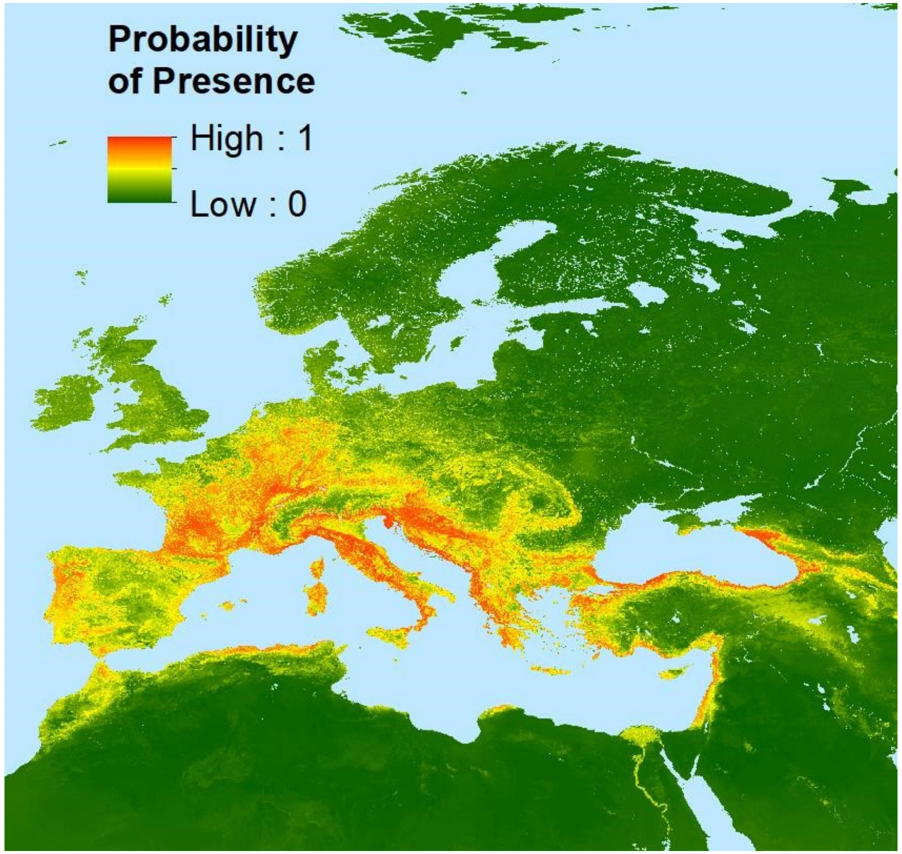
So vector data needs to be well structured to be mappable

Funded by the Horizon 2020 Framework Programme of the European Union

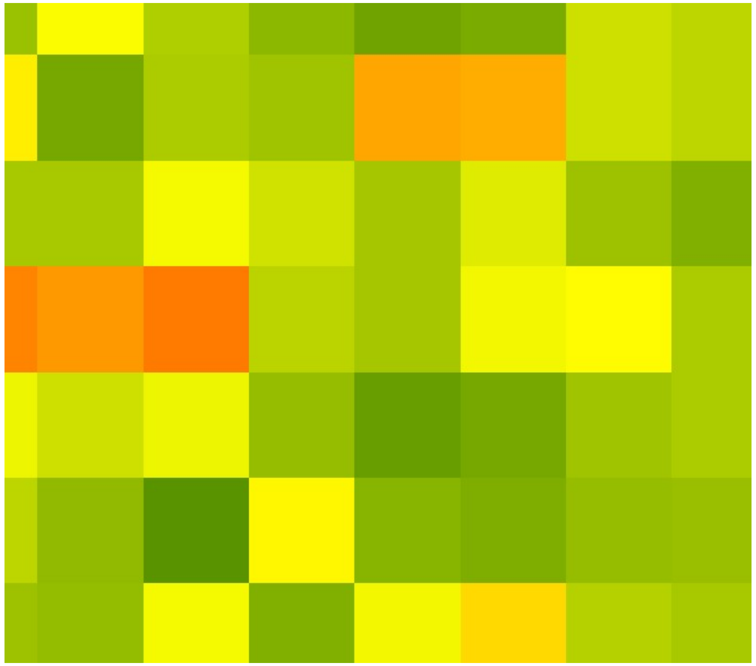
AIM Aedes Invasive Mosquitoes

cost EUROPEAN COOPERATION IN SCIENCE & TECHNOLOGY

BASICS OF TURNING DATA INTO MAPS



'Raster' maps are in principal similar except a single value per pixel. And no readily accessible table



Particular colour assigned to particular values

BASICS OF TURNING DATA INTO MAPS

Geo referencing is a requirement

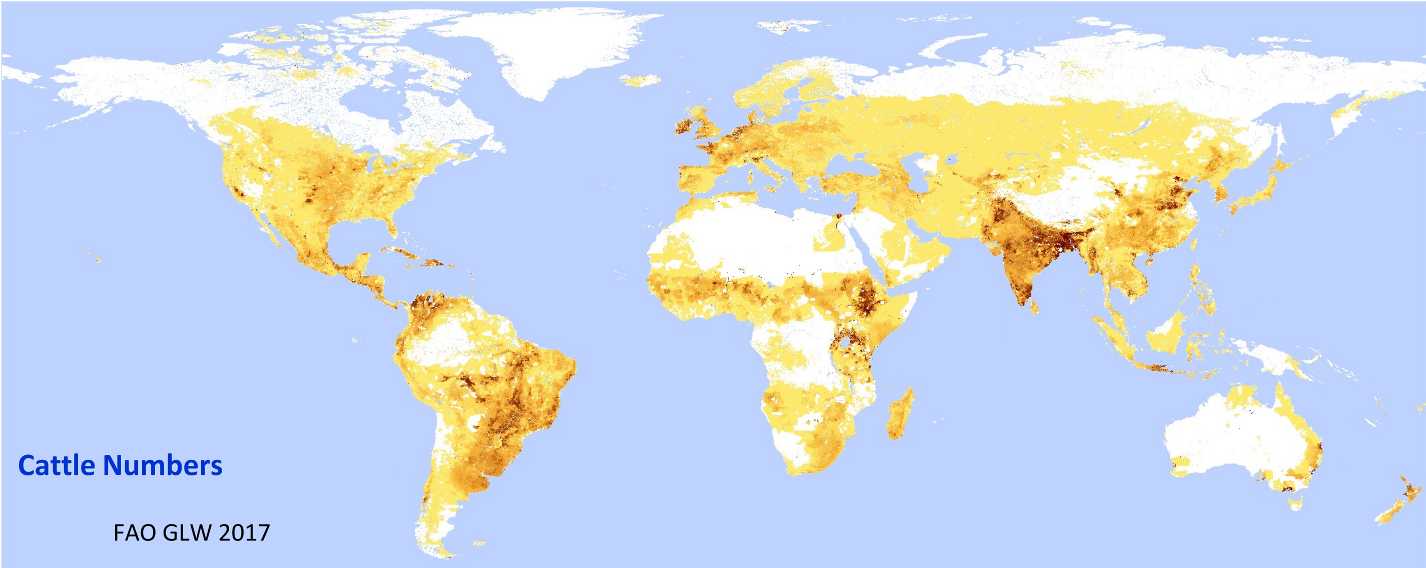
Maps need 'standard' data from everywhere because they are effectively comparing values for different places

Standardised sample number

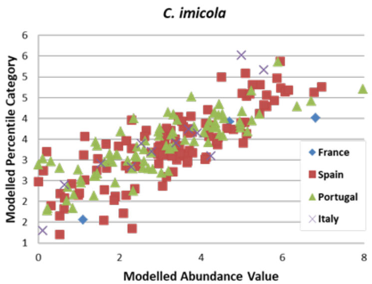
Zero is a (necessary) value

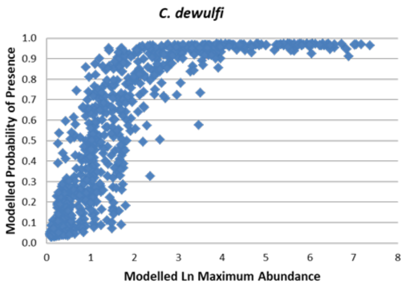
WHAT VECTOR MEASURES ARE NEEDED

Presence absence or abundance?

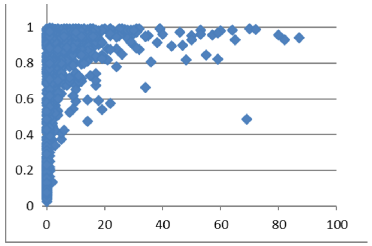


European Midges



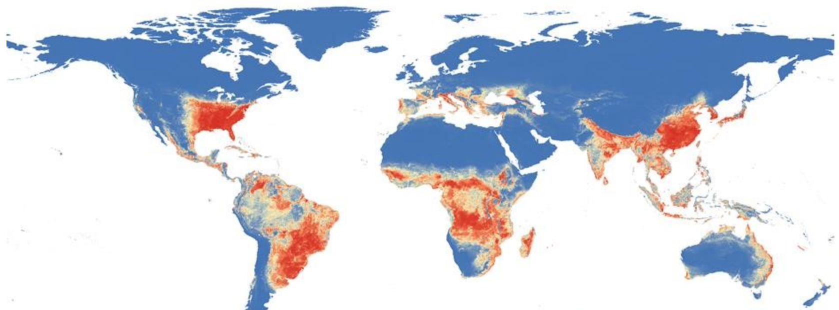


Tick



Wint et al, ECDC Technical Opinion, Dec 2018

Tiger Mosquito probability



Kraemer et al, Nature 2015

Aedes Invasive Mosquitoes

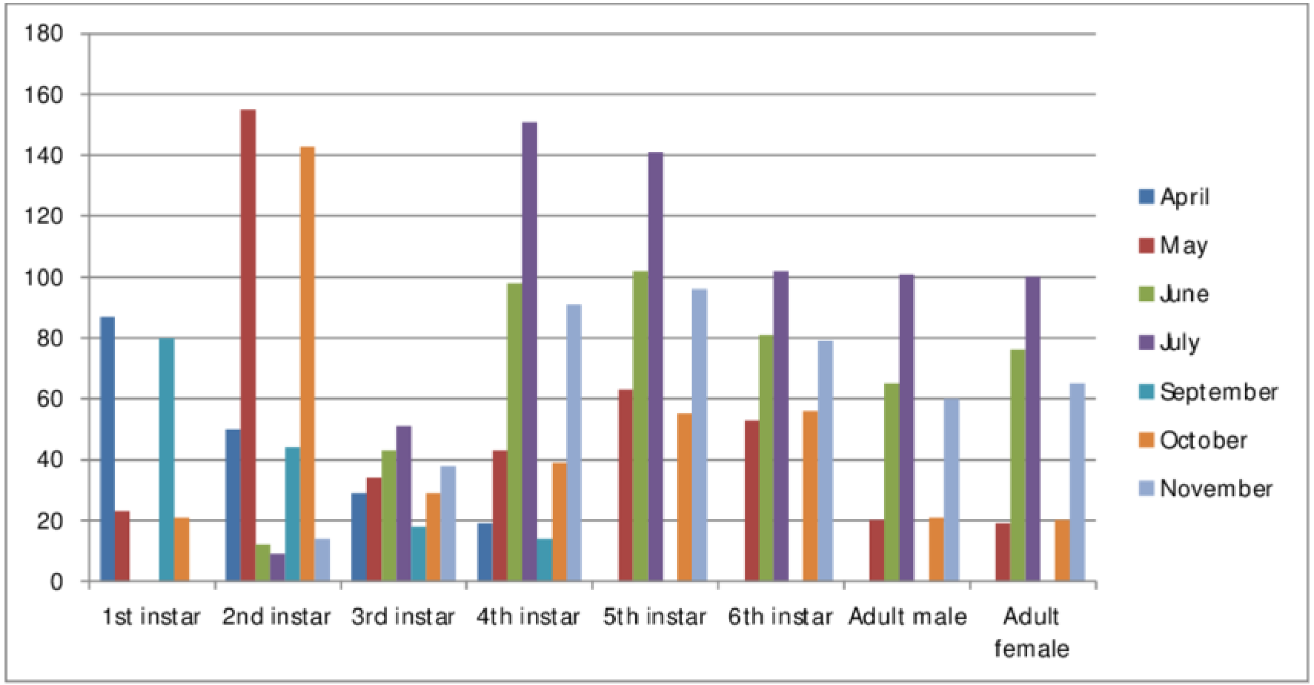
EUROPEAN COOPERATION IN SCIENCE & TECHNOLOGY

WHAT VECTOR DATA ARE NEEDED

What targets: eggs, larvae, adults,

What traps used - what is sample unit?

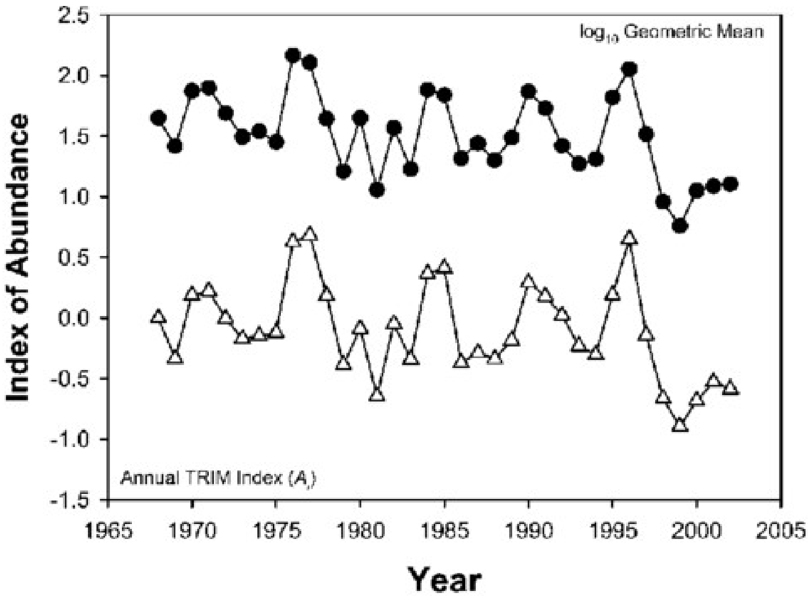
When should the samples be taken?



AIM Aedes Invasive Mosquitoes
cost EUROPEAN COOPERATION IN SCIENCE & TECHNOLOGY

WHAT VECTOR DATA ARE NEEDED
ABUNDANCE

when: annual, monthly, weekly, once



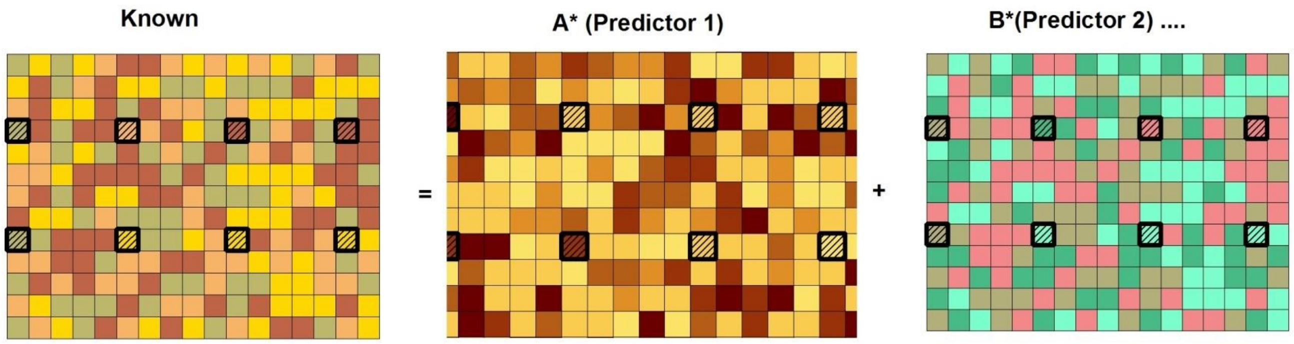
Once – Presence, early warning, not absence

Weekly or monthly seasonal activity annual comparisons

Maximum value or dates starts and end

Smoothed a second possibility

WHERE IS DATA NEEDED FROM



1) Convert all data maps to images with same pixel size (resolution). Then extract values for each data type at fixed sample points (hatched squares). NB one of these must be the 'known' values.

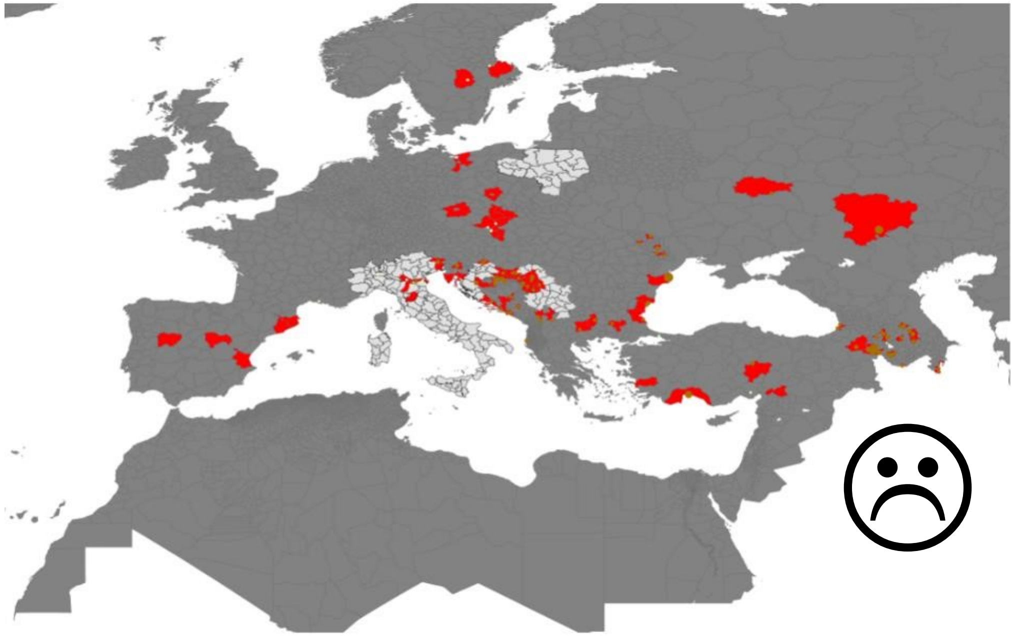
2) Calculate a 'regression equation' of the form:
Known=Constant + A*(Predictor1) + B*(Predictor2) ...
NB There can be several predictor variables in the equations.

3) Providing the equation is statistically significant (i.e. reliable), apply the right hand side of the equation to all the pixels in the images, not just the ones sampled.

4) Repeat the process for each of a series of analysis zones (e.g. ecozones)

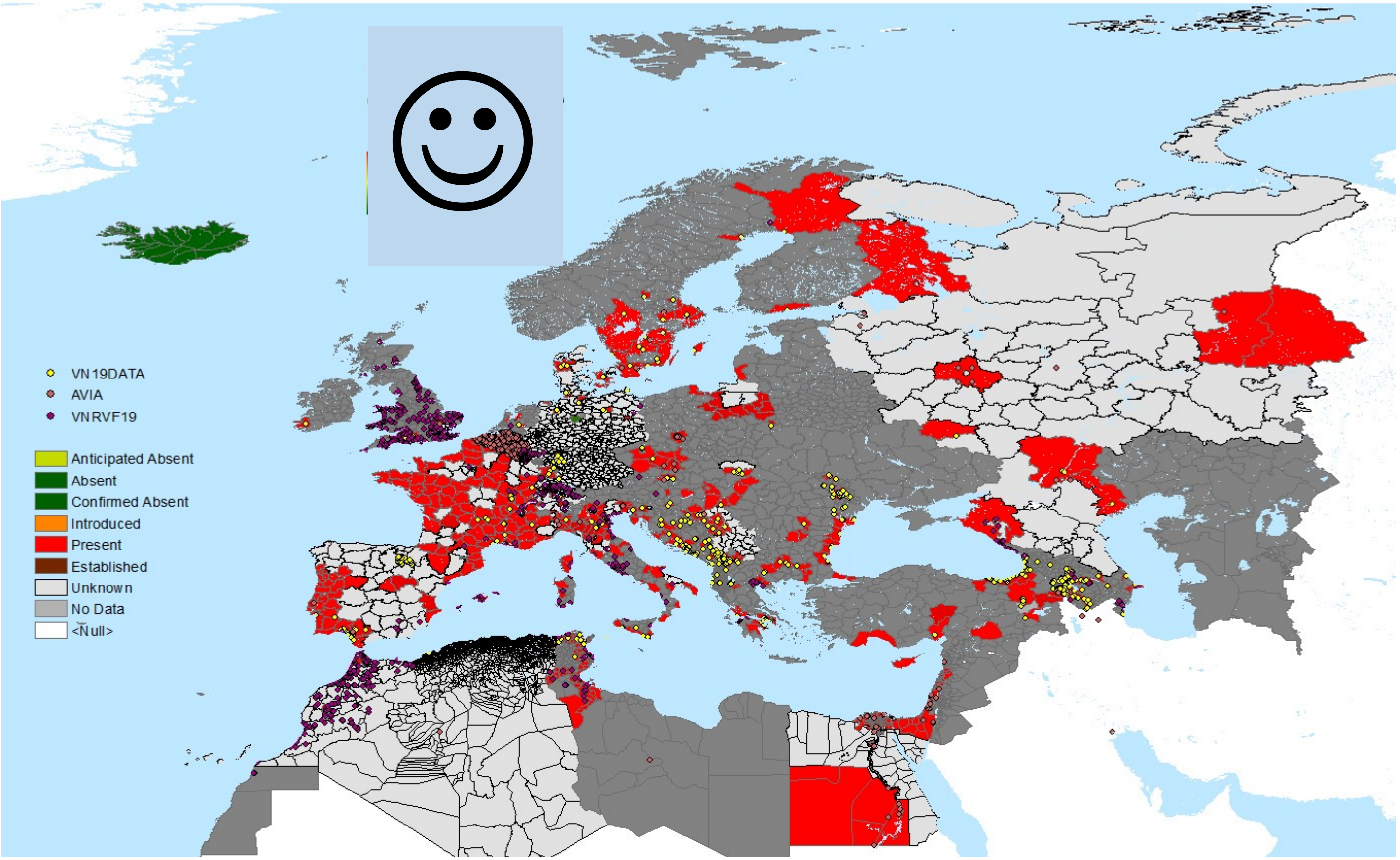
Most models rely on some sort of exptrapolation:
This means that its important to decide not only what data but which locations are sampled

WHAT DATA ARE NEEDED
CLUSTERING, SCALE



Points sparse, many gaps

WHAT DATA ARE NEEDED
CLUSTERING, SCALE



Pipiens

BUT

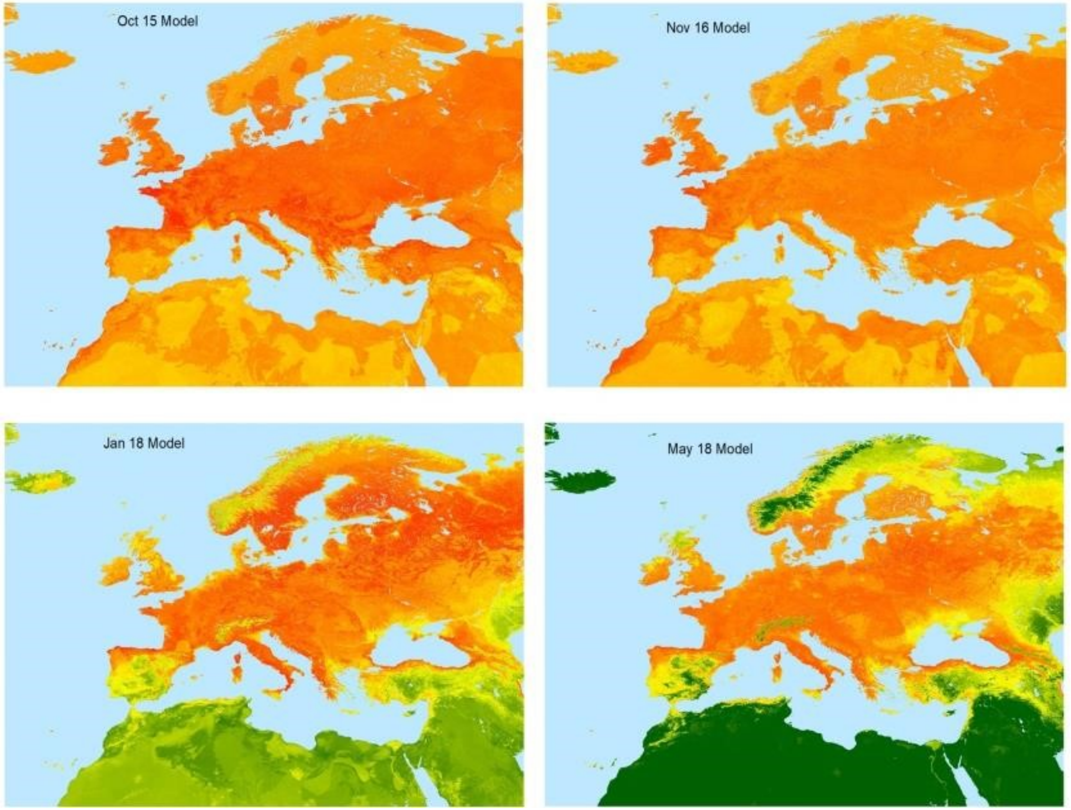
VERY clustered in some places

If these are (reliable) abundance data then good

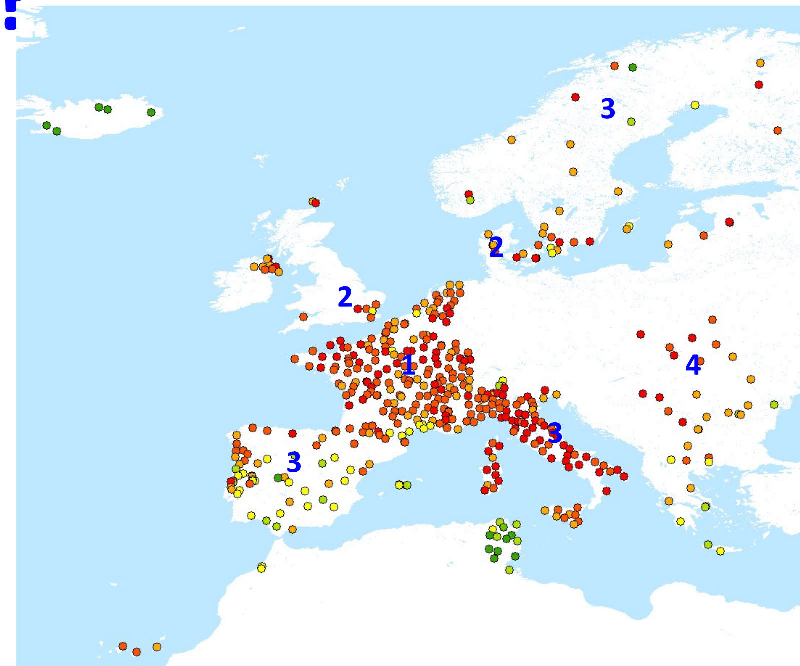
What about absence

EXTRAPOLATION?

Evolution of changing midge models



obsoletus

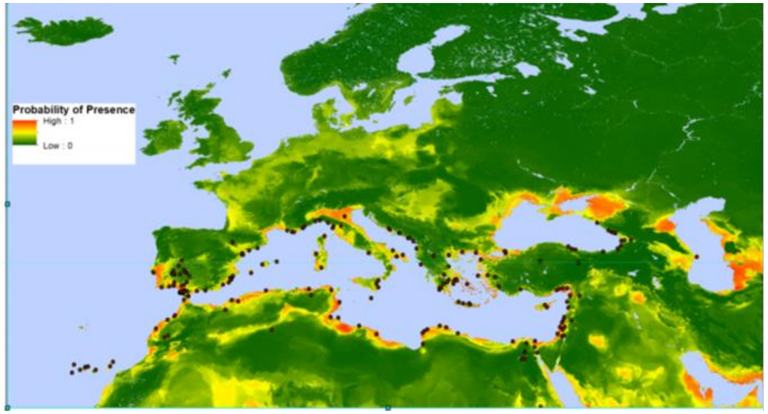


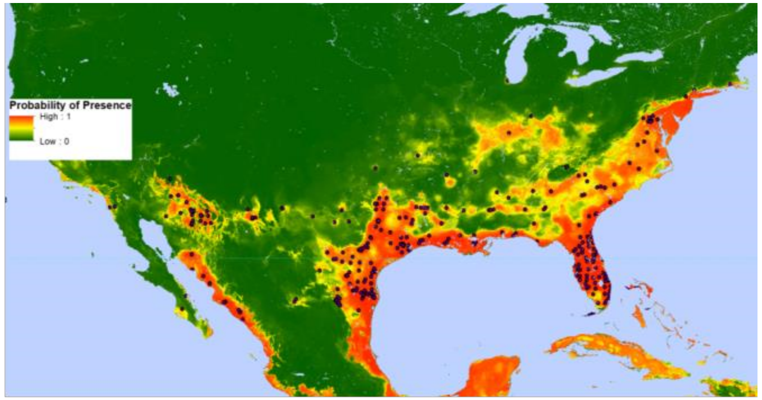
VectorNet

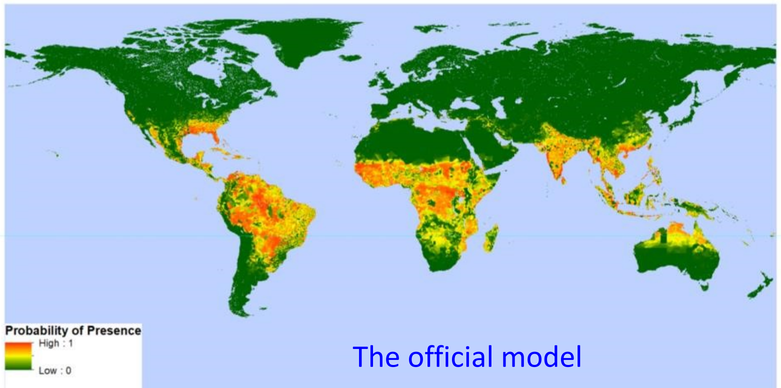
AIM
Aedes Invasive Mosquitoes
cost
EUROPEAN COOPERATION IN SCIENCE & TECHNOLOGY

HAVING SAID THAT!

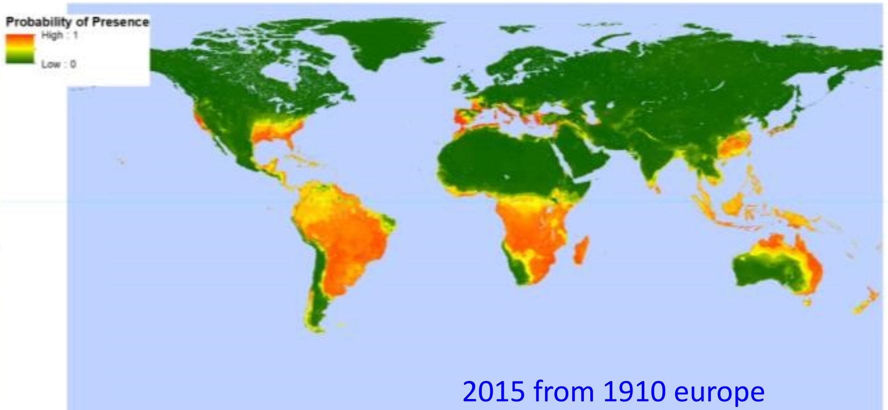
Europe from US DATA, with 1910 EU on top







The official model



2015 from 1910 europe

Aedes Invasive Mosquitoes
EUROPEAN COOPERATION IN SCIENCE & TECHNOLOGY

RECORD STRUCTURE

This means that for modelling, the data you collect needs to have a minimum number of components:

Record number

Coordinates, preferably decimal latitude and longitude. If other reference systems (e.g. UTM) then clear statements to say so

(Location name, admin unit, country)

Date of collection (DD/MM/YYY)

Trap type, number of traps

Species

Number caught, of each stage (including zero).

Contact/reference

ARCHIVE RECORDS

If they are to be compatible with VectorNet, many of these data values should have specific codes – trap types/collection method, and a large number of other metadata are required:

FOR current PA ECDC Vector Maps

Mandatory:

- Species
- Mode
- Status
- Collection method
- Start date
- End date
- Expert
- Coordinates or Admin Units
- Identification method
- Life Stage

Optional:

- Publication
- Life Stage
- Sex
- Abundance
- Sample effort
- Study name (not on web form)
- Study identification (not on web form)
- Coordinate precision
- StudyContext
- TaxanomicRange
- TrapIdentifier and type (depending on the collection method)

Format should be something that can be converted to Excel: dbf, csv, tab delimited, or of course xls.

CASE STUDIES

How do these data look -

Cleaning/interpreting a dataset.

Preparing the clean dataset for modelling

Data used VectorNet:

Data sources

survey

published literature

unpublished records

A REAL DATASET

1	Country	admin	lat	long	prec	status	meth	start	end	numb	stage	sex	ident	trno	comments
2	Slovenia		46.680361	########	Exact locati	Present	Larval sam	12/6/2013	12/6/2013	1	larvae		morphological		CONTAINER POS n= NA TOT n= NA
3	Slovenia		46.157176	########	Exact locati	Present	Larval sam	26/6/2013	26/6/2013	1	Adults	male	morphological		CONTAINER POS n= NA TOT n= NA
4	Slovenia		46.157176	########	Exact locati	Present	Larval sam	26/6/2013	26/6/2013	2	Adults	female	morphological		CONTAINER POS n= NA TOT n= NA
5	Slovenia		46.157176	########	Exact locati	Present	Larval sam	26/6/2013	26/6/2013	16	larvae		morphological		CONTAINER POS n= NA TOT n= NA
6	Slovenia		45.681509	########	Exact locati	Present	Larval sam	28/9/2015	28/9/2015	6	Adults	female	morpholog	80	CONTAINER POS n= 10 TOT n= 80
7	Austria	Niederösterreich-Süd			Region	Present	Ovitrap	1/7/2017	30/10/2017	21	Eggs		Malditof		
8	Austria	West- und Südsteiermark			Region	Present	Ovitrap	1/8/2017	30/10/2017	114	Eggs		Malditof		
9	Austria	Südburgenland			Region	Present	Ovitrap	1/8/2017	30/10/2017	12	Eggs		Malditof		
10	Austria	Nordburgenland			Region	Present	Ovitrap	1/8/2017	30/10/2017	16	Eggs		Malditof		
11	Austria	Linz-Wels	48*17.001'N	14*16.663'	Exact locati	Present	Other	10/7/1905	10/7/1905	7	Adults	female	PCR		light trap and Gravid trap; 1 trap-night
12	Austria	Graz	47*04.995'N	15*27.865'	Exact locati	Present	Other	10/7/1905	10/7/1905	17	Adults	female	PCR		light trap and Gravid trap; 2 trap-nights
13	Germany	Bitburg-Prum	6.164760	49.987799	Exact locati	Present	Larval sam	2/8/2018	2/8/2018	8	Larvae		morpholog	1	
14	Germany	Rhein-Hunsru	7.511174	49.981248	Exact locati	Present	Larval sam	13/8/2018	13/8/2018	2	Larvae		morpholog	1	
15	Germany	Bernkastel-W	7.132197	49.818163	Exact locati	Present	Larval sam	13/8/2018	13/8/2018	31	Larvae		morpholog	1	
16	Germany	Bernkastel-W	6.743988	49.910280	Exact locati	Present	Larval sam	13/8/2018	13/8/2018	6	Larvae		morpholog	1	
17	Luxembourg	Luxembourg	49.962348	6.168536	Exact locati	Present	Human bai	5/7/2018	5/7/2018	3	Adults	Female	morpholog	1	
18	Luxembourg	Luxembourg	49.962024	6.167952	Exact locati	Present	Larval sam	1/8/2018	1/8/2018	27	Larvae		morphological		
19	Luxembourg	Luxembourg	49.962066	6.167973	Exact locati	Present	Larval sam	1/8/2018	1/8/2018	1	Larvae		morpholog	1	
20	Luxembourg	Luxembourg	49.962257	6.168106	Exact locati	Present	Larval sam	1/8/2018	1/8/2018	12	Larvae		morpholog	1	
21	Luxembourg	Luxembourg	49.961676	6.167431	Exact locati	Present	Larval sam	1/8/2018	1/8/2018	36	Larvae		morpholog	1	
22	Luxembourg	Luxembourg	49.962733	6.168749	Exact locati	Present	Larval sam	1/8/2018	1/8/2018	10	Larvae		morpholog	1	
23	Luxembourg	Luxembourg	49.957706	6.189971	Exact locati	Present	Larval sam	1/8/2018	1/8/2018	22	Larvae		morpholog	1	
24	Luxembourg	Luxembourg	49.926524	6.217778	Exact locati	Present	Larval sam	1/8/2018	1/8/2018	2	Larvae		morpholog	1	
25	Luxembourg	Luxembourg	49.982247	6.123373	Exact locati	Present	Larval sam	2/8/2018	2/8/2018	1	Larvae		morpholog	1	
26	Luxembourg	Luxembourg	49.952061	6.017269	Exact locati	Present	Larval sam	14/8/2018	14/8/2018	10	Larvae		morpholog	1	
27	Luxembourg	Luxembourg	49.882011	6.258663	Exact locati	Present	Larval sam	15/8/2018	15/8/2018	2	Larvae		morpholog	1	
28	Luxembourg	Luxembourg	49.815946	6.417555	Exact locati	Present	Larval sam	15/8/2018	15/8/2018	7	Larvae		morpholog	1	
29	Luxembourg	Luxembourg	49.790678	6.302217	Exact locati	Present	Larval sam	15/8/2018	15/8/2018	4	Larvae		morpholog	1	
30	Luxembourg	Luxembourg	49.873726	6.156638	Exact locati	Present	Larval sam	15/8/2018	15/8/2018	18	Larvae		morpholog	1	
31	Luxembourg	Luxembourg	49.707217	6.427252	Exact locati	Present	Larval sam	15/8/2018	15/8/2018	1	Larvae		morpholog	1	
32	Luxembourg	Luxembourg	49.849343	6.093564	Exact locati	Present	Larval sam	16/8/2018	16/8/2018	2	Larvae		morpholog	1	
33	Luxembourg	Luxembourg	49.706832	6.370539	Exact locati	Present	Larval sam	4/10/2018	4/10/2018	0	Larvae		morpholog	1	
34	Luxembourg	Luxembourg	49.679666	6.251942	Exact locati	Present	Larval sam	5/10/2018	5/10/2018	12	Larvae		morpholog	1	
35	France	Meurthe-et-N	48.461592	6.847366	Exact locati	Present	Larval sam	08/16/2018	08/16/2018	8	Larvae		morpholog	1	
36	France	Moselle	49.448238	6.359291	Exact locati	Present	Larval sam	4/10/2018	4/10/2018	4	Larvae		morpholog	1	

AIM Aedes Invasive Mosquitoes

cost EUROPEAN COOPERATION IN SCIENCE & TECHNOLOGY

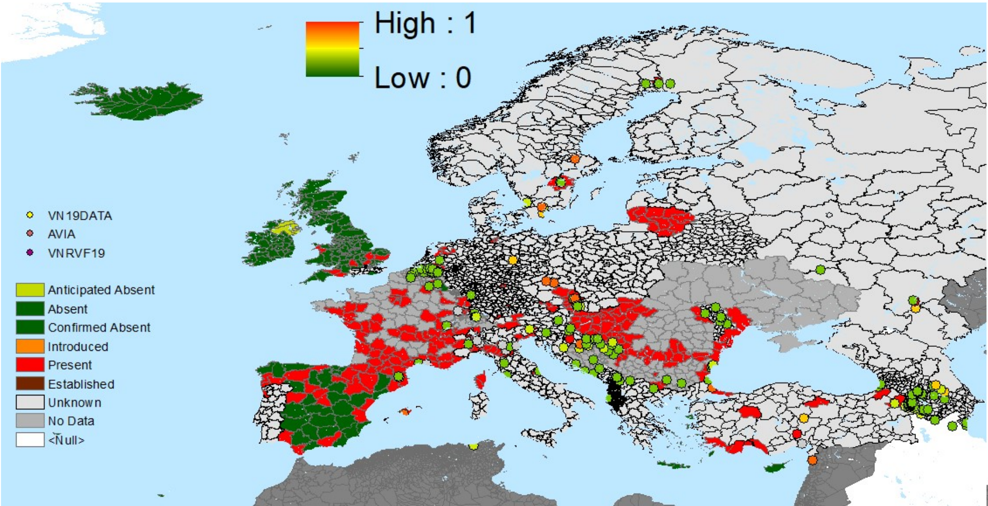
Funded by the Horizon 2020 Framework Programme of the European Union

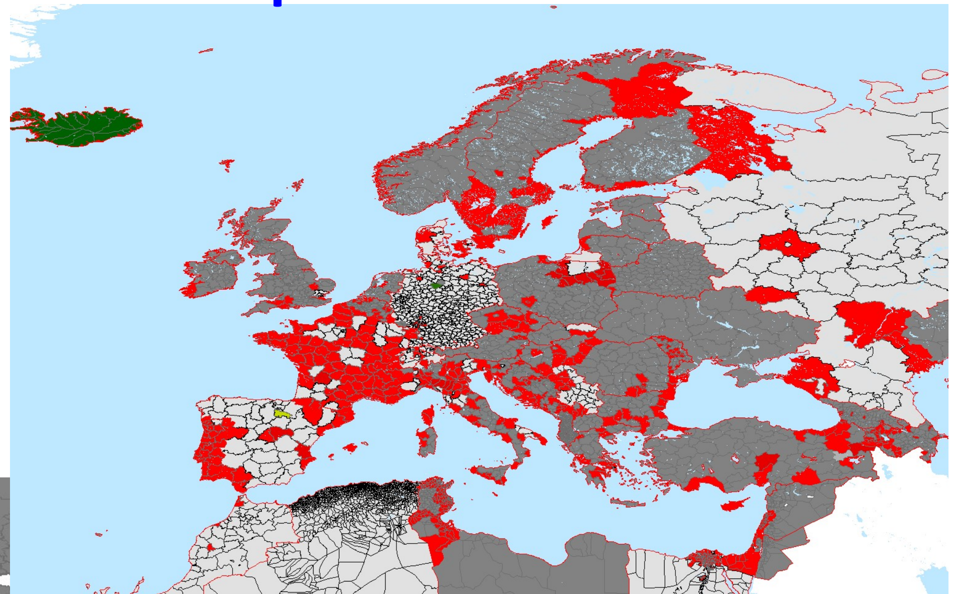
A REAL DATASET: SOME ISSUES

	Country	admin	lat	long	prec	status	meth	start	end	numb	stage	sex	ident	trno	comments
2	Slovenia		46.680361	#########	Exact locati	Present	Larval sam	12/6/2013	12/6/2013	1	larvae		morphological		CONTAINER POS n= NA TOT n= NA
3	Slovenia		46.157176	#########	Exact locati	Present	Larval sam	26/6/2013	26/6/2013	1	Adults	male 8	morphological	9	CONTAINER POS n= NA TOT n= NA
4	1 Slovenia		46.157176	#########	Exact locati	Present	Larval sam	26/6/2013	26/6/2013	2	Adults	female	morphological		CONTAINER POS n= NA TOT n= NA
5	Slovenia		46.157176	#########	Exact locati	Present	Larval sam	26/6/2013	26/6/2013	16	larvae		morphological		CONTAINER POS n= NA TOT n= NA
6	Slovenia		45.681509	#########	Exact locati	Present	Larval sam	28/9/2015	28/9/2015	6	Adults	female	morpholog	80	CONTAINER POS n= 10 TOT n= 80
7	Austria	Niederösterreich-Süd 2			Region	Present	Ovitrap	1/7/2017	30/10/2017	21	Eggs		Malditof		
8	Austria	West- und Südsteiermark			Region	Present	Ovitrap	1/8/2017	30/10/2017	114	Eggs 7		Malditof		
9	Austria	Südburgenland			Region	Present	Ovitrap	1/8/2017	30/10/2017	12	Eggs		Malditof		
10	Austria	Nordburgenland 3			Region	Present	Ovitrap	1/8/2017	30/10/2017	16	Eggs		Malditof		
11	Austria	Linz-Wels	48°17.001'N	14°16.663'	Exact locati	Present	Other	10/7/1905	10/7/1905	7	Adults	female	PCR		ligh 10 ap and Gravid trap; 1 trap-night
12	Austria	Graz	47°04.995'N	15°27.865'	Exact locati	Present	Other	10/7/1905	10/7/1905	17	Adults	female	PCR		light trap and Gravid trap; 2 trap-nights
13	Germany	Bitburg-Prum	6.164760	49.987799	Exact locati	Present	Larval sam	2/8/2018	2/8/2018	8	Larvae		morpholog	1	
14	Germany	Rhein-Hunsru	7.511174	4 49.981248	Exact locati	Present	Larval sam	13/8/2018	13/8/2018	2	Larvae		morpholog	1	
15	Germany	Bernkastel-W	7.132197	49.818163	Exact locati	Present	Larval sam	13/8/2018	13/8/2018	31	Larvae		morpholog	1	
16	Germany	Bernkastel-W	6.743988	49.910280	Exact locati	Present	Larval sam	13/8/2018	13/8/2018	6	Larvae		morpholog	1	
17	Luxembourg	Luxembourg	49.962348	6.168536	Exact locati	Present	Human bai	5/7/2018	5/7/2018	3	Adults	Female	morpholog	1	
18	Luxembourg	Luxembourg	49.962024	6.16 49.987799	locati	Present	Larval sam	1/8/2018	1/8/2018	27	Larvae		morphological		
19	Luxembourg	Luxembourg	49.962066	6.16 49.981248	locati	Present	Larval sam	1/8/2018	1/8/2018	1	Larvae		morpholog	1	
20	Luxembourg	Luxembourg	49.962257	6.16 49.818163	locati	Present	Larval sam	1/8/2018	1/8/2018	12	Larvae		morpholog	1	
21	Luxembourg	Luxembourg	49.961676	6.16 49.910280	locati	Present	Larval sam	1/8/2018	1/8/2018	36	Larvae		morpholog	1	
22	Luxembourg	Luxembourg	49.962733	6.168749	Exact locati	Present	Larval sam	1/8/2018	1/8/2018	10	Larvae		morpholog	1	
23	Luxembourg	Luxembourg	49.957706	6.189971	Exact locati	Present	Larval sam	1/8/2018	1/8/2018	22	Larvae		morpholog	1	
24	Luxembourg	Luxembourg	49.926524	6.217778	Exact locati	Present	Larval sam	1/8/2018	1/8/2018	2	Larvae		morpholog	1	
25	Luxembourg	Luxembourg	49.982247	6.123373	Exact locati	Present	Larval sam	2/8/2018	2/8/2018	1	Larvae		morpholog	1	
26	Luxembourg	Luxembourg	49.952061	6.017269	Exact locati	Present	Larval sam	14/8/2018	14/8/2018	10	Larvae		morpholog	1	
27	Luxembourg	Luxembourg	49.882011	6.258663	Exact locati	Present	Larval sam	15/8/2018	15/8/2018	2	Larvae		morpholog	1	
28	Luxembourg	Luxembourg	49.815946	6.417555	Exact locati	Present	Larval sam	15/8/2018	15/8/2018	7	Larvae		morpholog	1	
29	Luxembourg	Luxembourg	49.790678	6.302217	Exact locati	Present	Larval sam	15/8/2018	15/8/2018	4	Larvae		morpholog	1	
30	Luxembourg	Luxembourg	49.873726	6.156638	Exact locati	Present	Larval sam	15/8/2018	15/8/2018	18	Larvae		morpholog	1	
31	Luxembourg	Luxembourg	49.707217	6.427252	Exact locati	Present	Larval sam	15/8/2018	15/8/2018	1	Larvae		morpholog	1	
32	Luxembourg	Luxembourg	49.849343	6.093564	Exact locati	Present	Larval sam	16/8/2018	16/8/2018	2	Larvae		morpholog	1	
33	Luxembourg	Luxembourg	49.706832	6.370539	Exact locati	Present	Larval sam	4/10/2018	4/10/2018	5 0	Larvae		morpholog	1	
34	Luxembourg	Luxembourg	49.679666	6.251942	Exact locati	Present	Larval sam	5/10/2018	5/10/2018 6	12	Larvae		morpholog	1	
35	France	Meurthe-et-N	48.461592	6.847366	Exact locati	Present	Larval sam	08/16/2018	08/16/2018	8	Larvae		morpholog	1	
36	France	Moselle	49.448238	6.359291	Exact locati	Present	Larval sam	4/10/2018	4/10/2018	4	Larvae		morpholog	1	

REAL DATASETS

Real datasets: what can you see that might cause a problem





AIM Aedes Invasive Mosquitoes
cost EUROPEAN COOPERATION IN SCIENCE & TECHNOLOGY

TOPIC 2: DOING THE MODELLING

Intro to spatial modelling

What else is needed to produce models

Doing the modelling

How (not) to make maps useful

Funded by the Horizon 2020 Framework Programme of the European Union

AIM Aedes Invasive Mosquitoes

EUROPEAN COOPERATION IN SCIENCE & TECHNOLOGY

WHAT ARE THE ISSUES

Producing risk maps is often to a deadline for planning requirements set by people who just want answers and for whom risk is a black box component, NOT a research project

Issues are : **Data acquisition, networks and sharing**

Objectives: what is good enough

Data not just disease/vector risk

Using appropriate methods

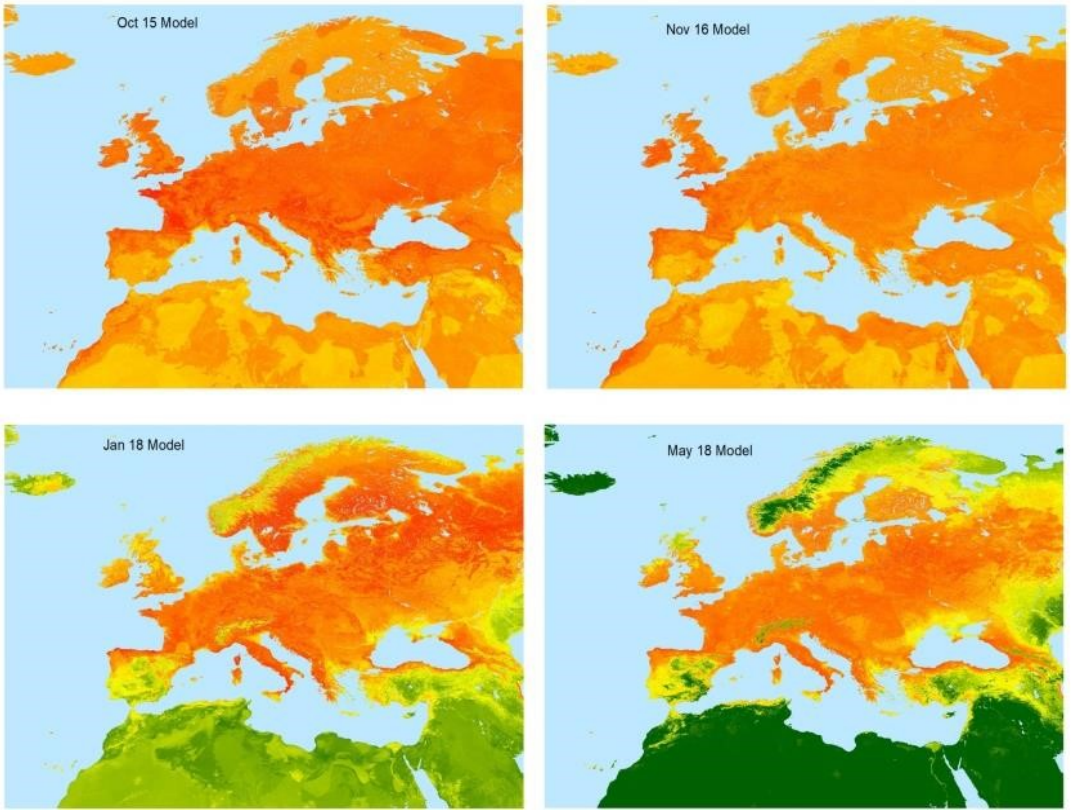
Getting the results out: Customisation and translation

WHAT IS GOOD ENOUGH: IS SOMETHING BETTER THAN NOTHING?

Evolution of changing midge models

Do we provide a model for Oct 15, or refuse

Did we know how wrong it would be?



Species:
7 *Culicoides* spp.

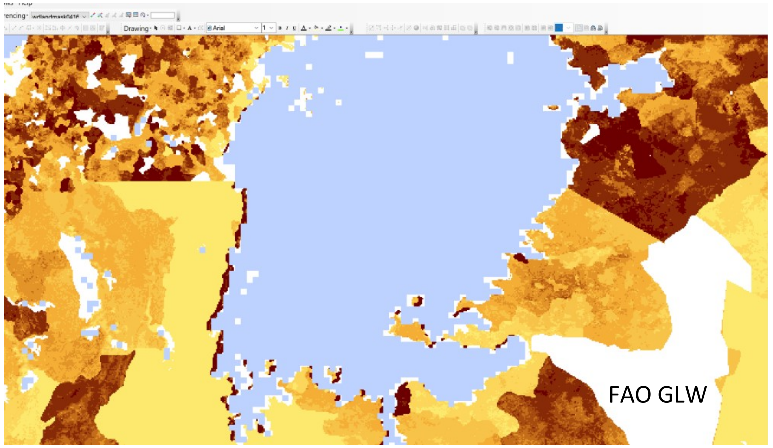
chiopterus
dewulfi
kingi
lupicaris
imicola
newsteadii
obsoletus
pulicaris
punctatus
scoticus

VectorNet

AIM COST Training School, Cyprus, January 2020

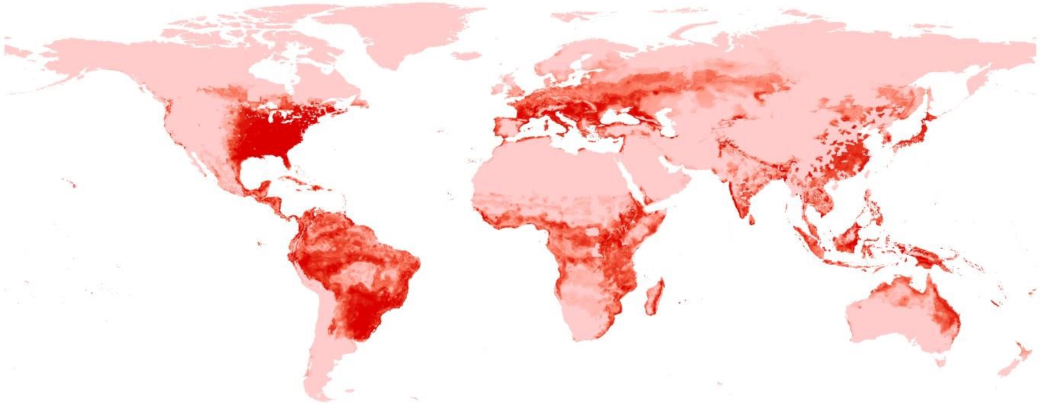
IS SOMETHING BETTER THAN NOTHING?

Global maps will have howlers: cattle



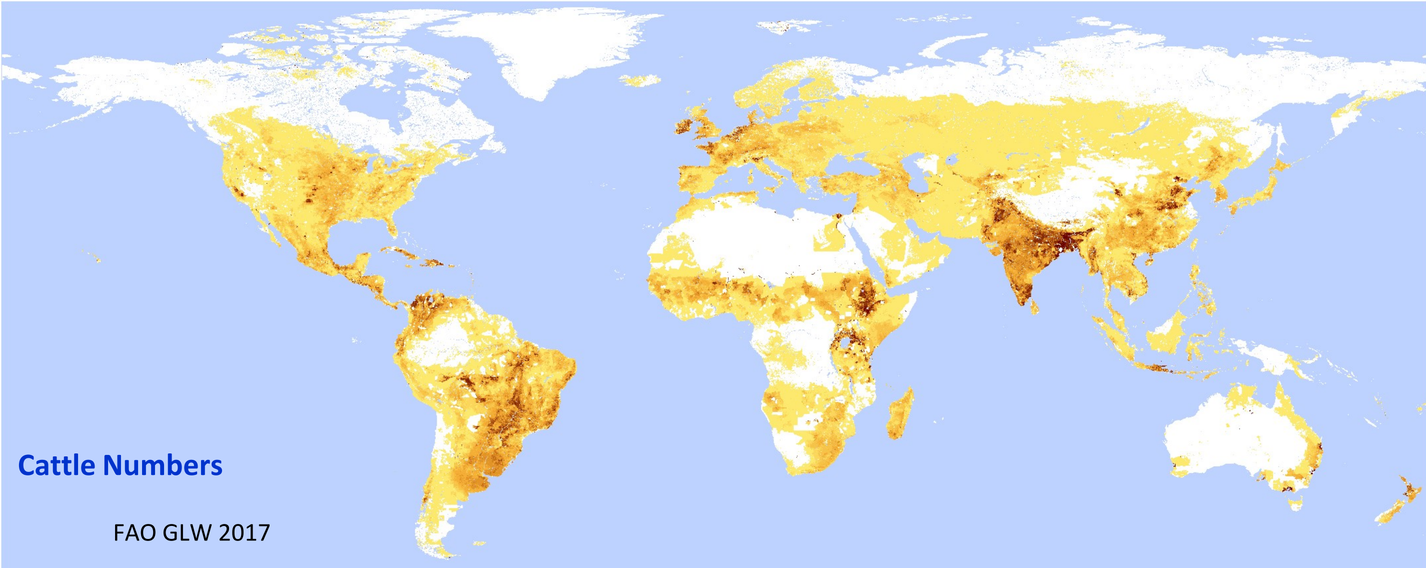
NB Cryptic tags can be useful to prove ownership!

Projection logic arguable, models misleading: Tiger Mosquito, Yellow Fever mosquito

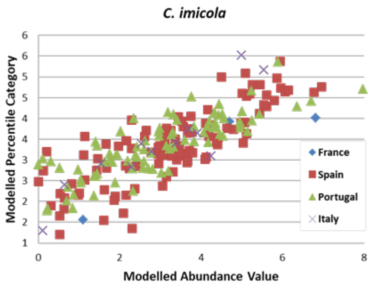


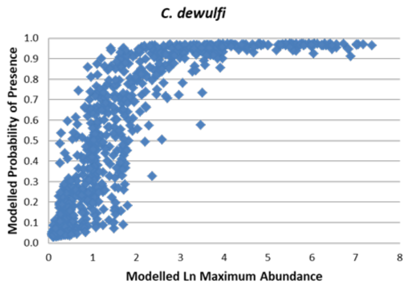
IS SOMETHING BETTER THAN NOTHING?

Presence absence or abundance?



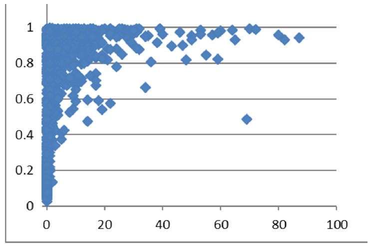
European Midges



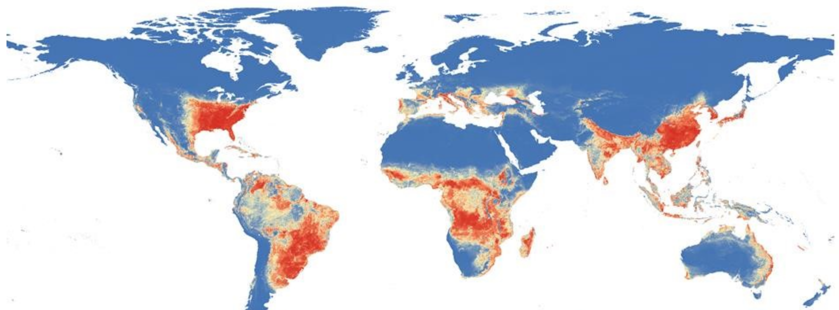


Wint et al, ECDC Technical Opinion, Dec 2018

Tick



Tiger Mosquito probability



Kraemer et al, Nature 2015

RISK OF PATHOGEN/VECTOR NOT JUST ABOUT PATHOGEN/VECTOR

Need information about:

Disease, Host, Vector

e.g. Dengue, ECF

Where there are the vectors AND

Where the vectors are infected AND

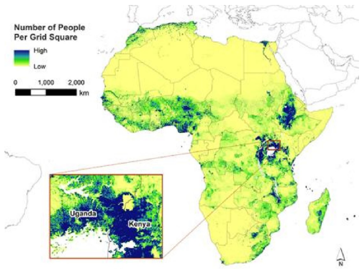
Where there are there are hosts (bovids, people)

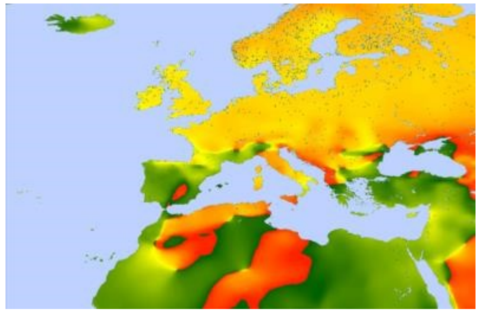
(Suitability later)

Covariates

Environmental, demographic, economic,…..

Customised



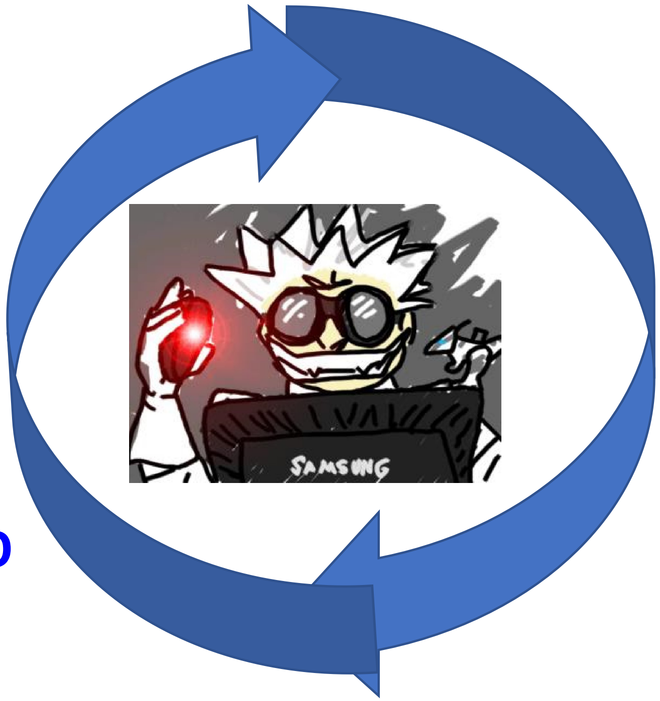


NOT JUST ABOUT MODELLING EITHER

GETTING THE DATA

SETTING THE OBJECTIVES

MAKING THEM WORK TOGETHER



DOING THE MODELLING

GETTING THE DATA OUT TO THE USERS

USING THE MODELS

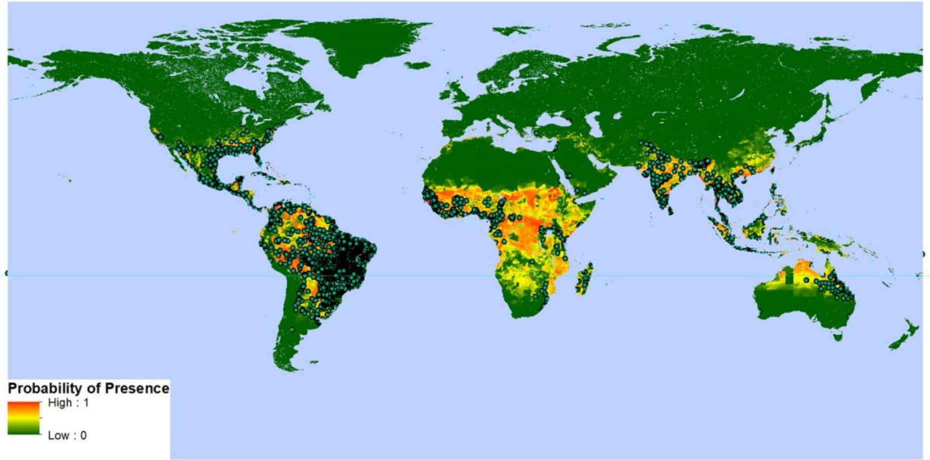
MODELING OFTEN THE EASY BIT

THE STEPS NEEDED: FINDING THE DATA

DATA TYPES: THE TARGETS = DISEASE, HOST, VECTOR

- **Most difficult stage:**
- **Often restricted data (wildlife, disease)**
- **Inevitably patchy, or inconsistent**
- **Often poor quality (zeros!)**
- **Large Scale models rely on LITERATURE and NETWORKS and data SHARING**

VectorNet



AIM
Aedes Invasive Mosquitoes

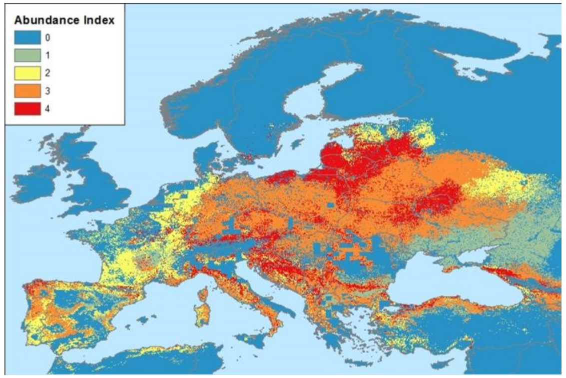
cost
EUROPEAN COOPERATION IN SCIENCE & TECHNOLOGY

THE STEPS NEEDED: FINDING THE DATA SHARING THE DATA

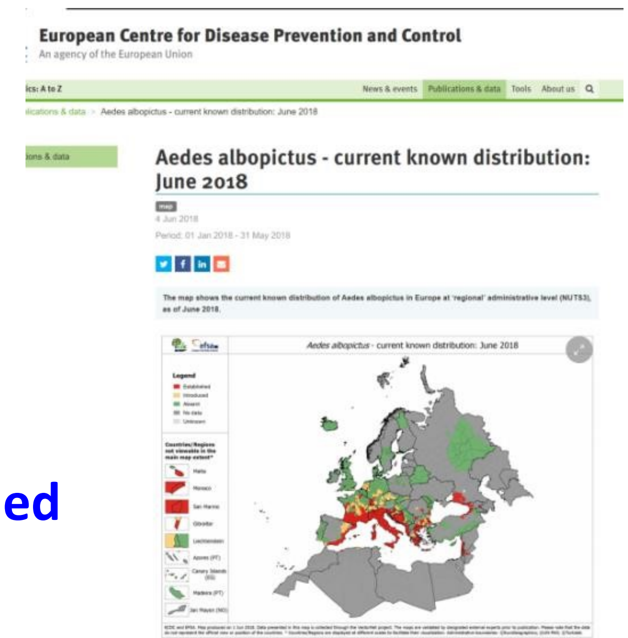
What is needed to facilitate data collection

What data owners can do

Networks
Data sharing
Anonymised
Summary aggregations
Transboundary



NEED PROTECTION OF OWNERSHIP



What data analysts can do

Share results & processed versions
Not one way flow! Culture change needed
Offer to adapt analysis for owners
Open data archives to owners
Explain outputs properly

AIM Aedes Invasive Mosquitoes
cost EUROPEAN COOPERATION IN SCIENCE & TECHNOLOGY

THE STEPS NEEDED: SHARING THE DATA
WHY SHARE YOUR DATA FOR MODELLING

PROS OF SHARING

- **Personal:**
  - It fulfills my contractual obligations
  - I get access to other data/outputs in return
- **Public Health/Societal:**
  - My data can be used to do things I cant to create PH/academic outputs .
  - It allows my data to be used at larger scales and have wider relevance

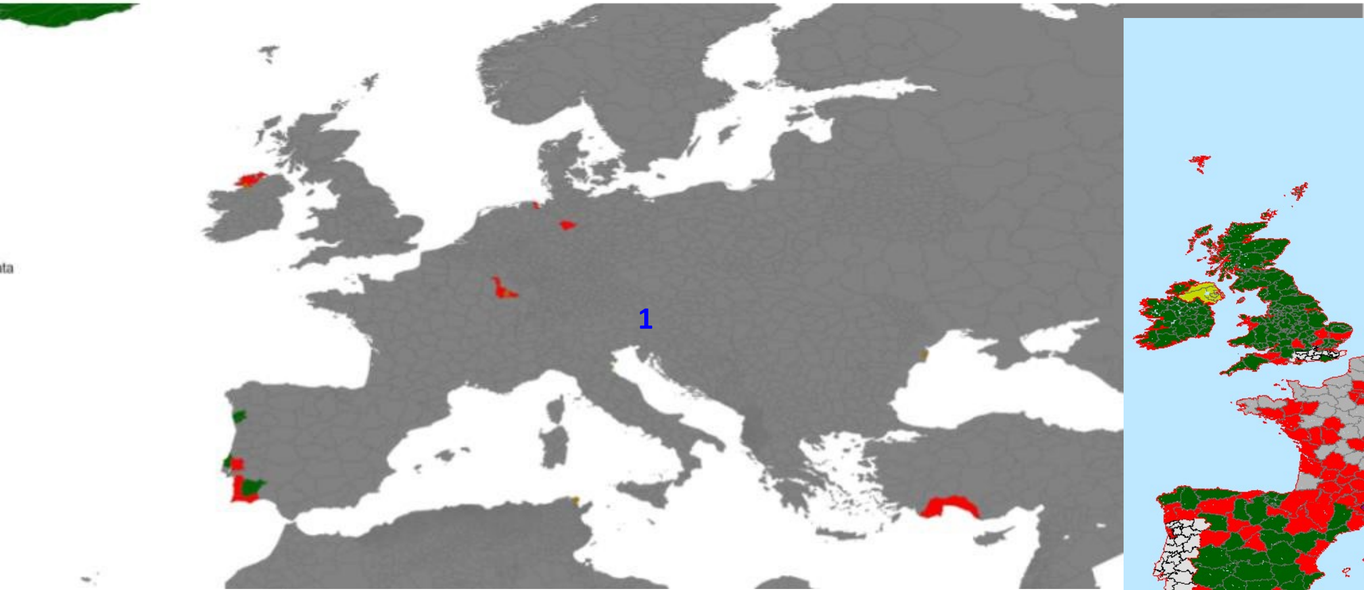
CONS OF SHARING

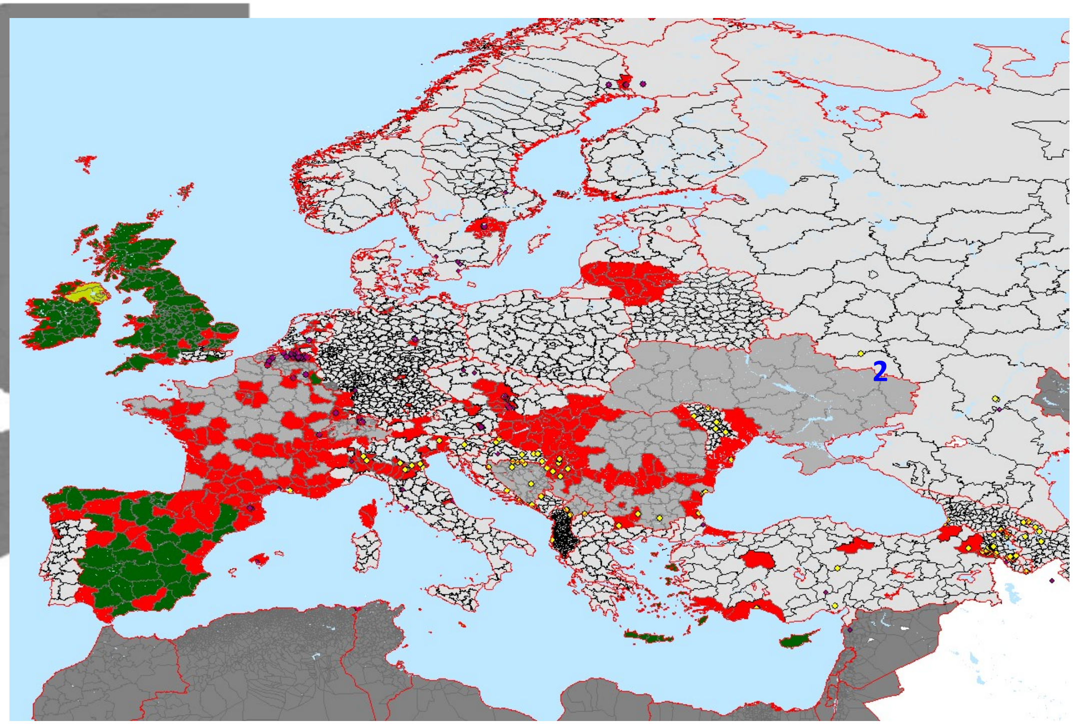
- **Personal:**
  - Someone else gets the benefit of my work before I do
  - I don’t get enough acknowledgement
  - I don’t get additional funding because I have shared my data
  - I don’t control the quality of what is done
  - Someone else is always hassling me to hurry up
- **Public Health/Societal:**

ENHANCING THE DATA

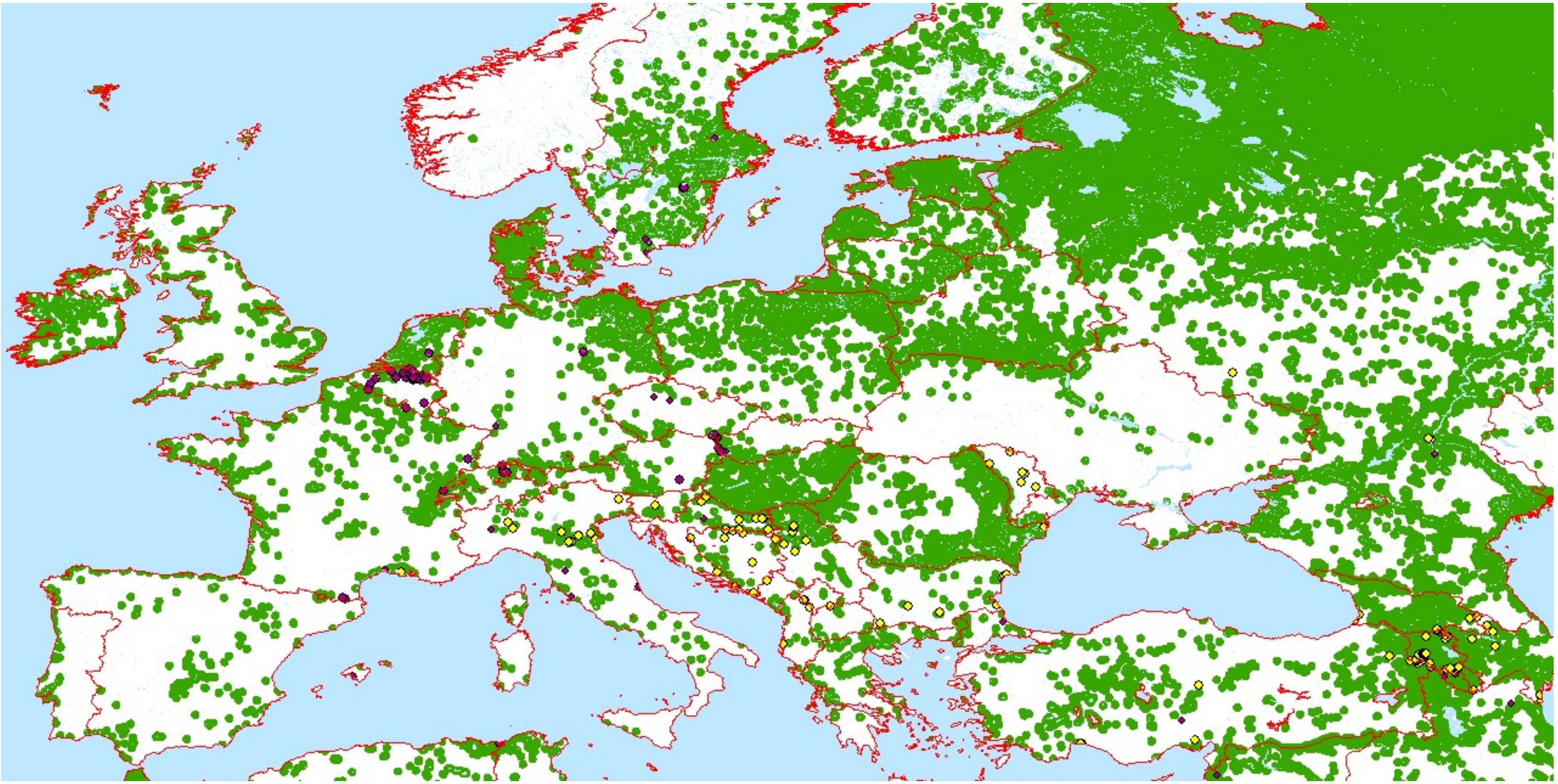
Sometimes it's a hopeless case, others we can 'cheat'

Can we define absences?





SUITABILITY



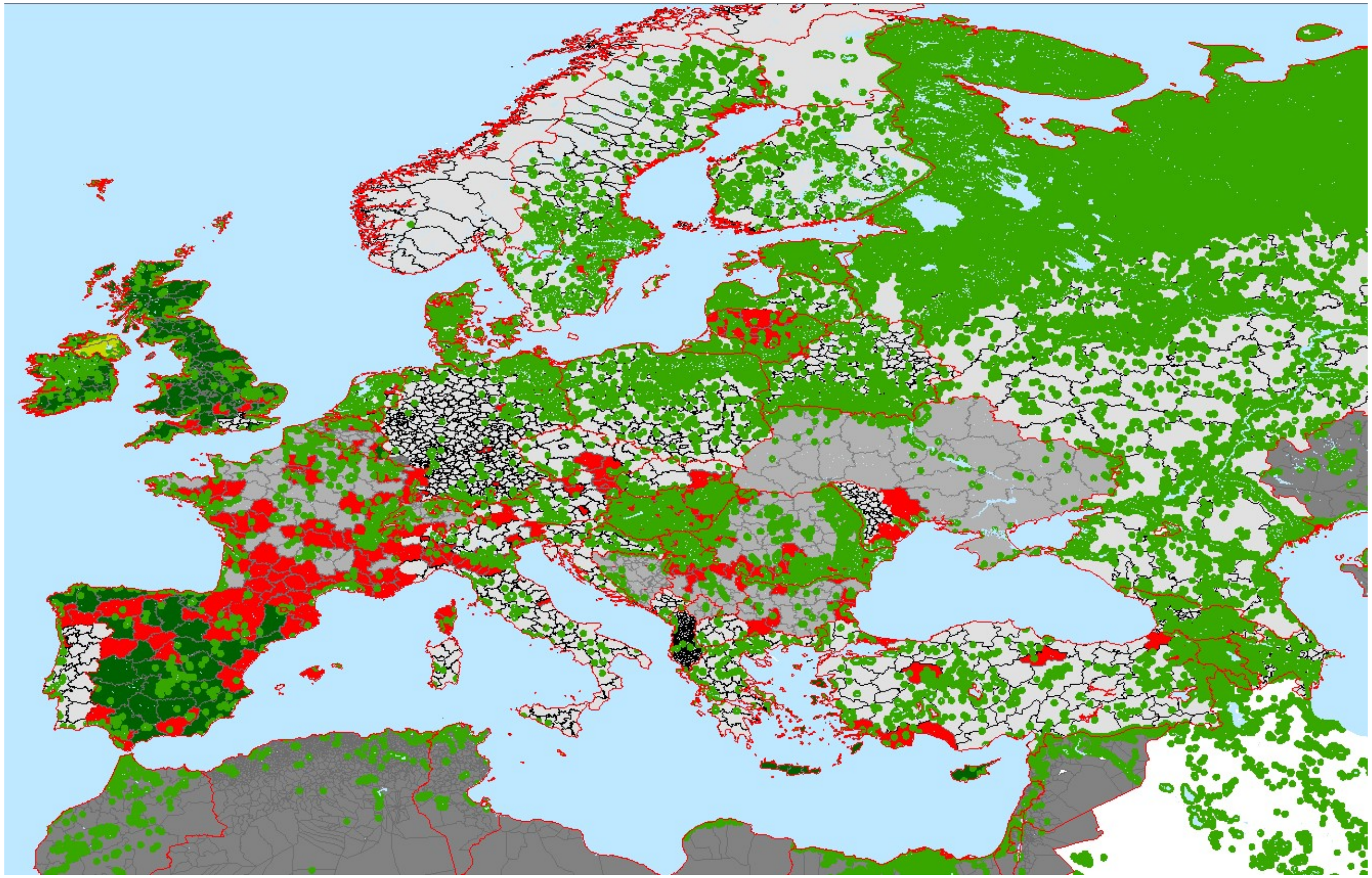
Suitability defined by land cover, environmental limits and dispersal capability

Unsuitable land identifiable

Absences can be assigned to unsuitable areas

Note (pleasing) correspondence between suitability and observations

USING SUITABILITY



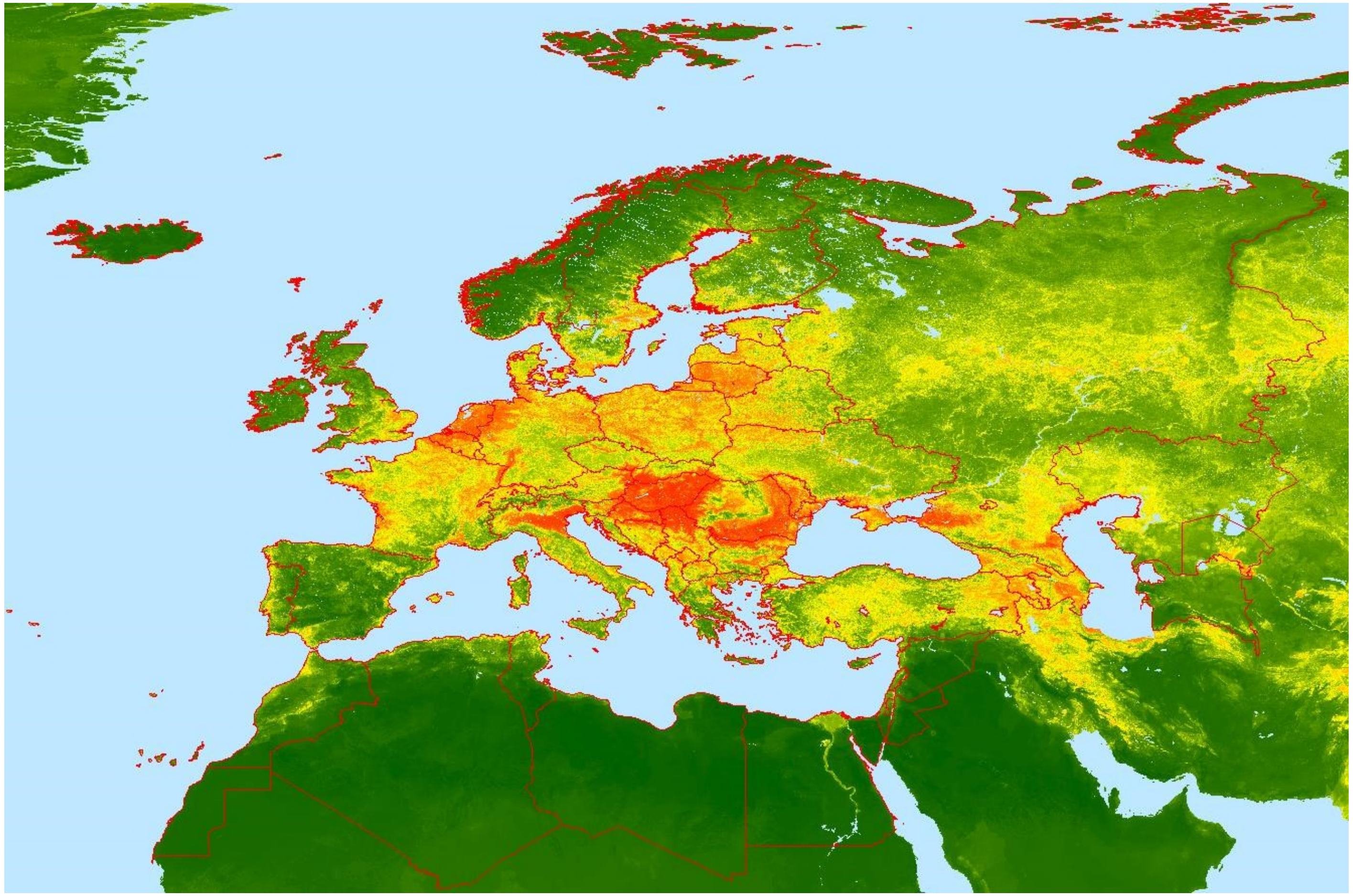
Absences can be assigned to unsuitable areas in unknown polygons,

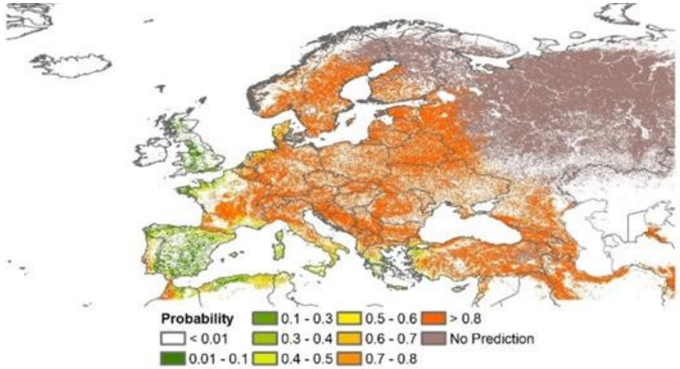
in all absence polygons,

and in unsuitable parts of presence polygons

SHOULD THEN BE BALANCED WITH PRESENCES
NB number or density??

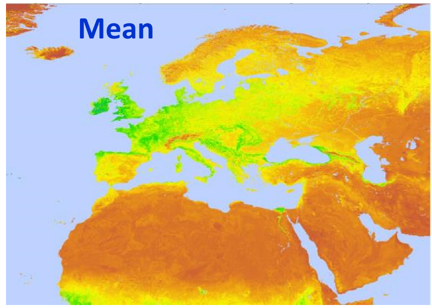
USING SUITABILITY





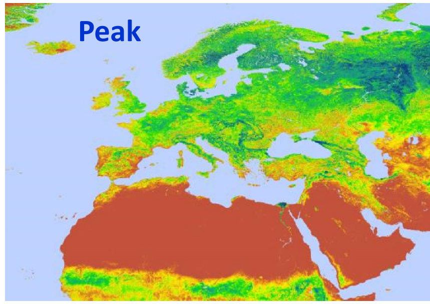
Big difference

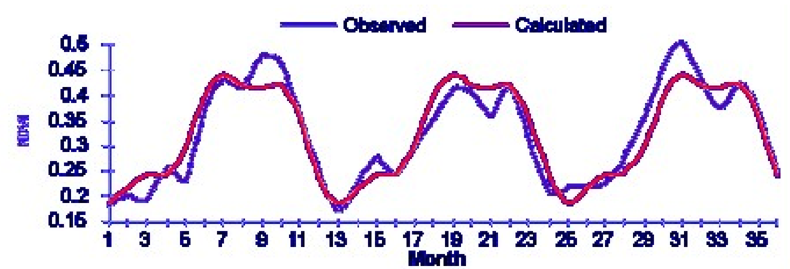
THE STEPS NEEDED: FINDING THE DATA COVARIATES: DATA REDUCTION

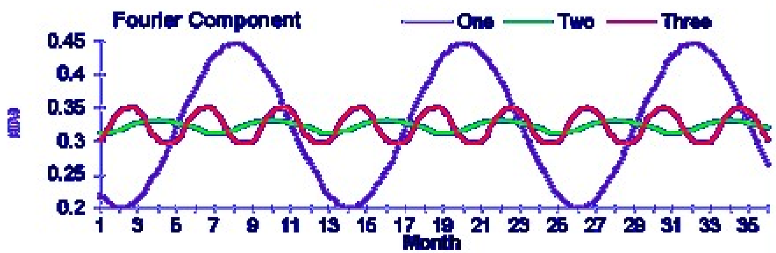


Turn large datasets to meaningful (manageable) parameters

Temporal Fourier Processing Imagery (2000-2016)





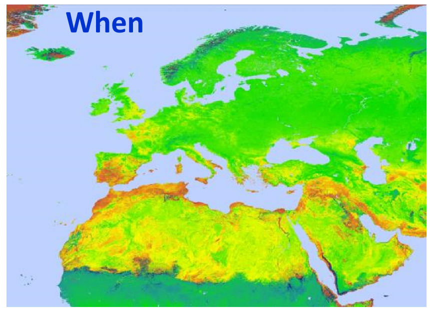


Decomposes irregular time series into regular components annual, bi annual, tri annual …….

Amplitude = Importance of each component
Here: annual>biannual>triannual

Mean = levels

Phase = timing of seasonal peaks



Provides aspects of seasonality as well as smoothed levels

AIM Aedes Invasive Mosquitoes

Funded by the Horizon 2020 Framework Programme of the European Union

cost EUROPEAN COOPERATION IN SCIENCE & TECHNOLOGY

FINDING THE DATA
COVARIATES: NEW VARIABLES

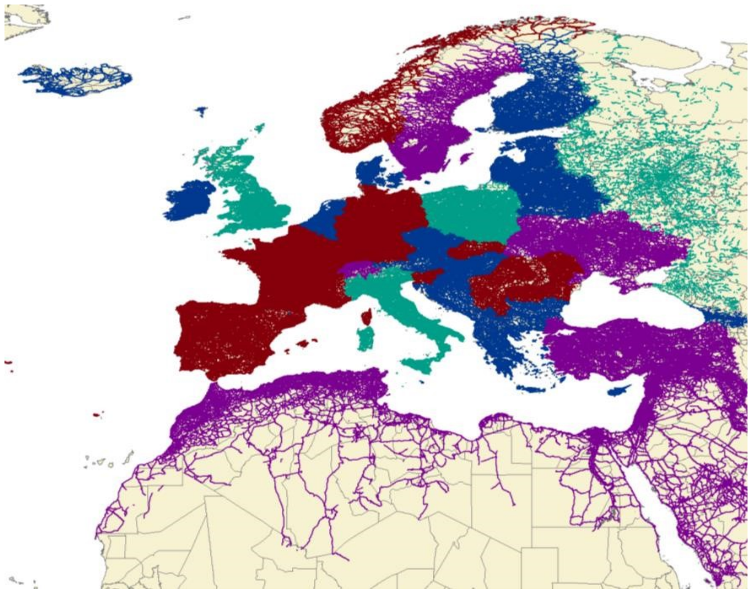
Accessibility:

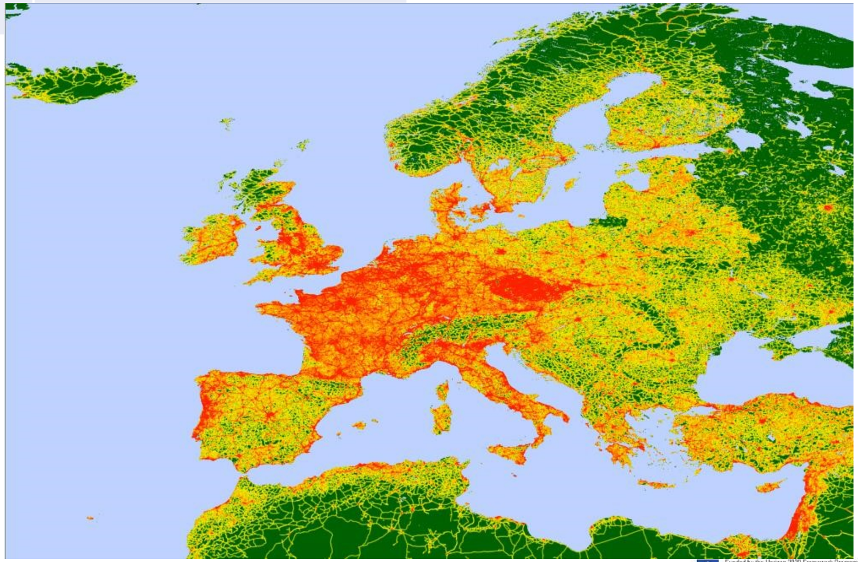
Measured as road density.

Could also be distance to…

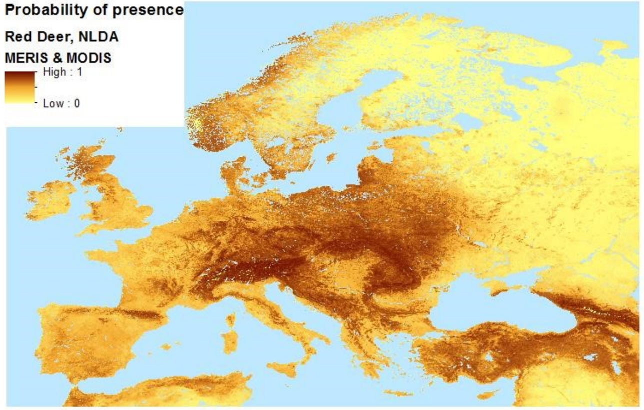
Global first surprisingly

Road Category	Filename	Approx No. Features, shape file size
All classified (motorway, trunk, primary, secondary, tertiary) with link roads	Eumenardsclassifiedwlinks	5.75M;1.7GB
Major (Primary, trunk, Motorway) and link roads	Eumenahighmajorwlinknov17	1.95m; 450MB
Secondary and link roads	Eumenardssecondarywlinknov17	1.53m: 450MB
Tertiary and link roads	Eumenardstertiarywlinknov17	2.3m: 820MB
Unclassified	Eumenardsunlcassifiednov17	2.5m; 1.4GB
Residential and 'living_streets' with link roads	Eumenardsresidlivingwlinknov17	13m; 2.6GB

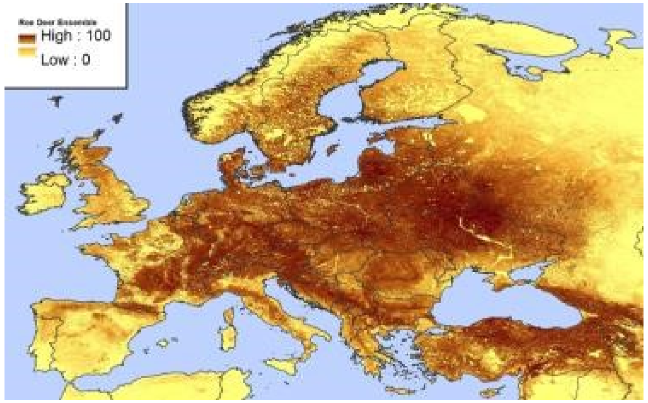


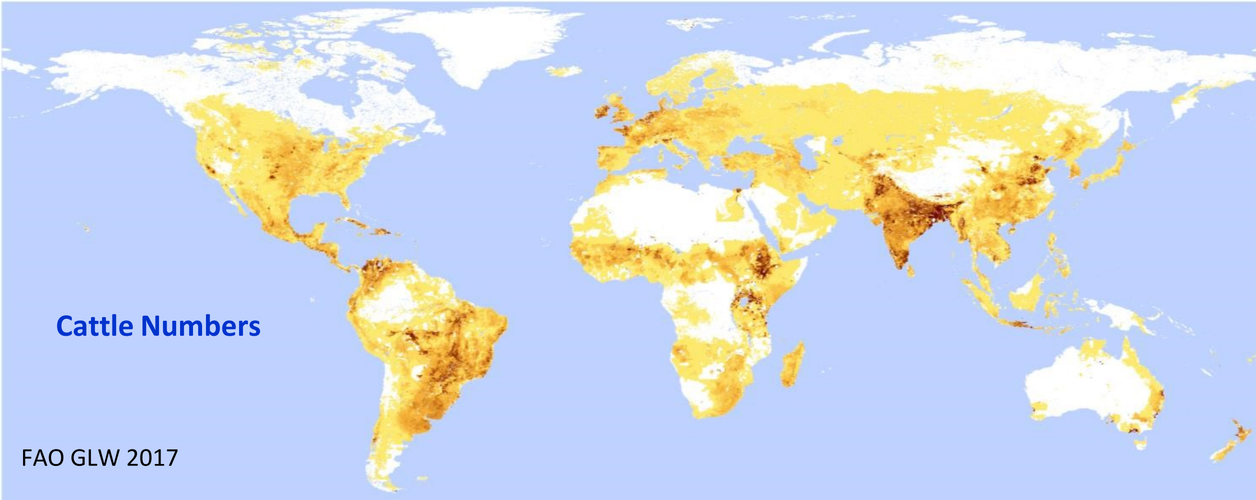


THE STEPS NEEDED: FINDING THE DATA



May also need to model drivers like hosts:
Deer for ticks
Cattle as denominator





THE STEPS NEEDED: FINDING THE DATA COVARIATE ARCHIVES

Don't re-invent the Wheel, you could get it wrong

Go to archives.... Standardised data

Google Earth Engine

SPATIAL DATA ARCHIVE

http://www.palebludata.com

DATA

Several Global Dataset Firsts

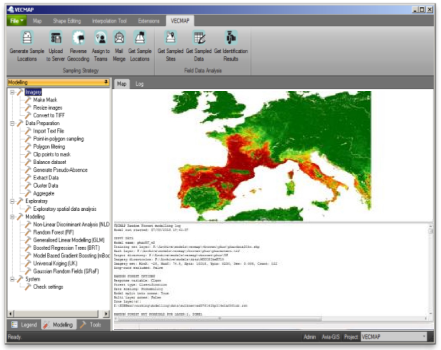
DOING THE MODELLING

Modelling methods:

- **Process Based:**
  - **Rare (difficult) at global level**
- **Stochastic:**
  - **Many methodologies**
  - **Spatial, Spatio Temporal**
- **Networks, Machine Learning**
- **If to be global (regional) standard,**
  - **People have more faith in the well established techniques**

BOOSTED REGRESSION TREES
RANDOM FOREST
MAXENT,
BAYESIAN, and many others

VECMAP

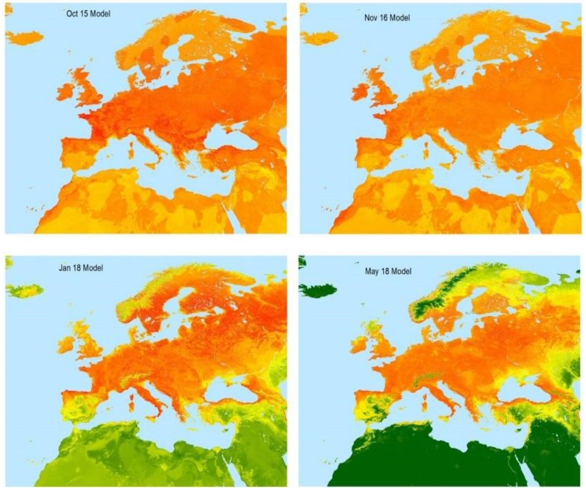


Funded by the Horizon 2020 Framework Programme of the European Union

AIM Aedes Invasive Mosquitoes

cost EUROPEAN COOPERATION IN SCIENCE & TECHNOLOGY

DOING THE MODELLING



Input data is primary determinant of model quality :

- **GARBAGE IN GARBAGE OUT**
- **Accuracy of data**
- **Non standardised sampling**
- **Mismatching data types**
- **Clustering**
- **Degree of extrapolation**

Models mostly pattern matching not identification of drivers

Exaggerated expectations:

- **85% accurate is good => 15% wrong**
- **models often not accurate at the pixel level**

MORE IMPORTANT TO GET DATA RIGHT: MANY METHODS WILL WORK

ALSO IMPORTANT TO MODEL USEFUL VARIABLES AND PRODUCE USEFUL OUTPUTS

Funded by the Horizon 2020 Framework Programme of the European Union
AIM Aedes Invasive Mosquitoes
cost EUROPEAN COOPERATION IN SCIENCE & TECHNOLOGY

DOING THE MODELLING

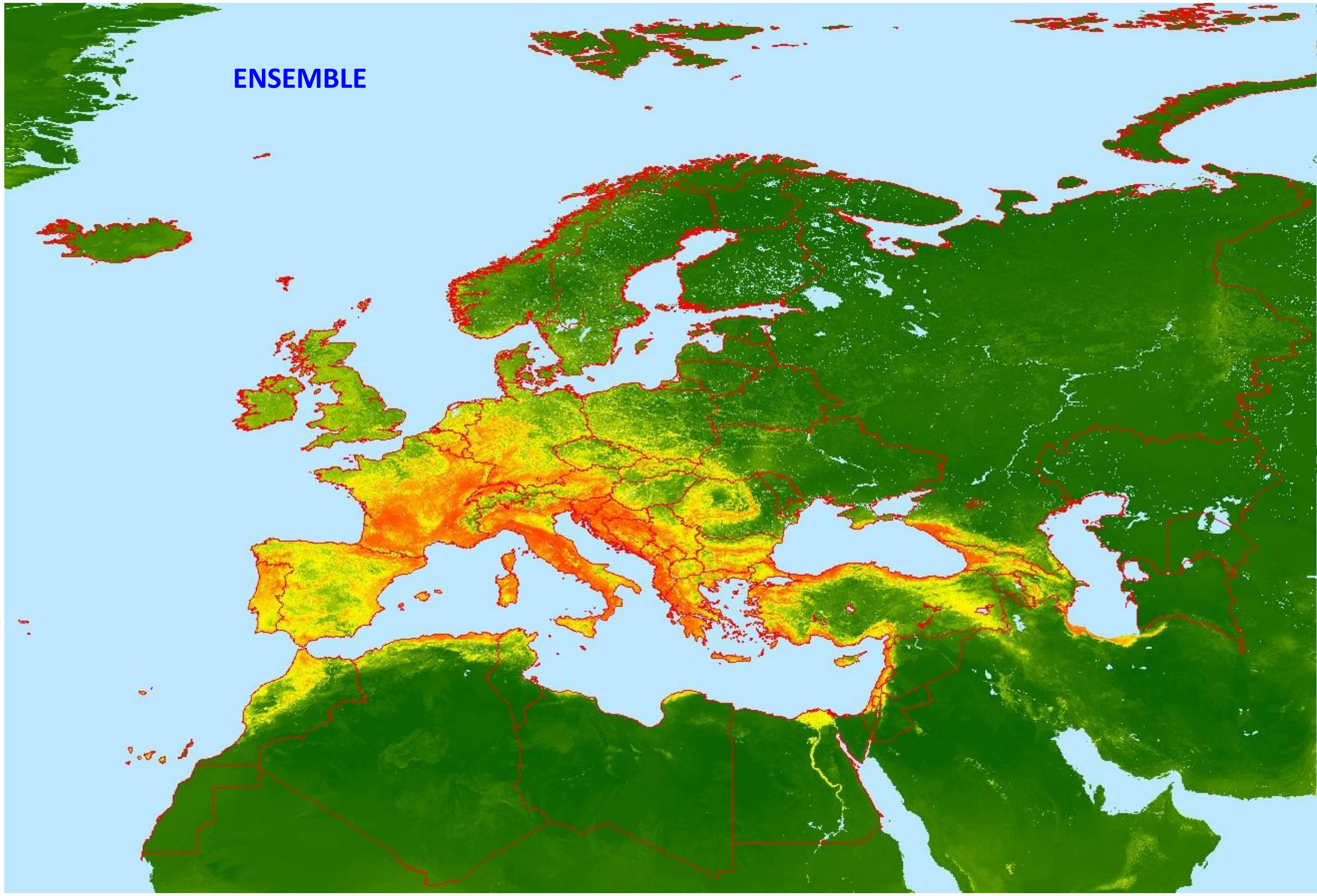
As have said earlier, most spatial modelling methods based on using sample to extrapolate.

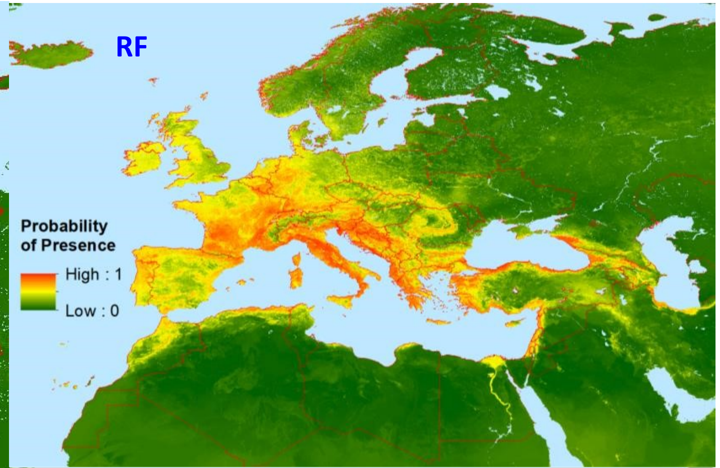
Different methods have different quirks. BRT tends to overfit, RF tends to smooth, but does categories well. BAYESIAN is very computer intensive, MAXENT purports not to need Absences. GLMM Logistic bias for presence, NLDA for absence. HURDLE combines PA and abundance modelling etc etc.

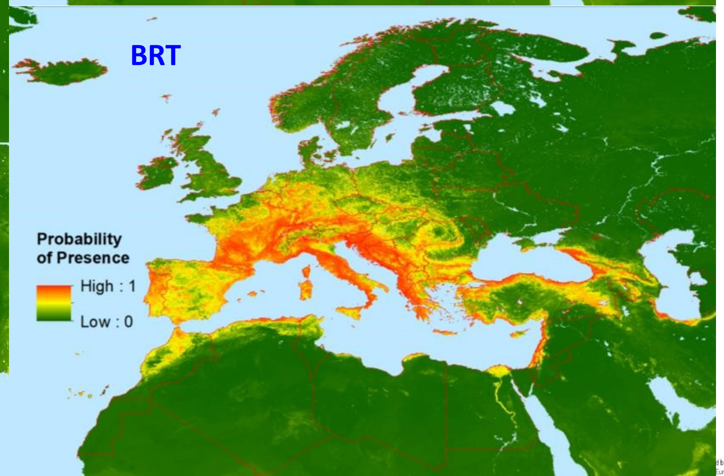
So get different resulst with same input data (like GCM projections !)

=> => strong case for ensembling methods, as well as doing replicate models.

DOING THE MODELLING: ENSEMBLING







AIM
Aedes Invasive Mosquitoes

cost
EUROPEAN COOPERATION
IN SCIENCE & TECHNOLOGY

DOING THE MODELLING: OUTPUTS

Decisions for Model output types

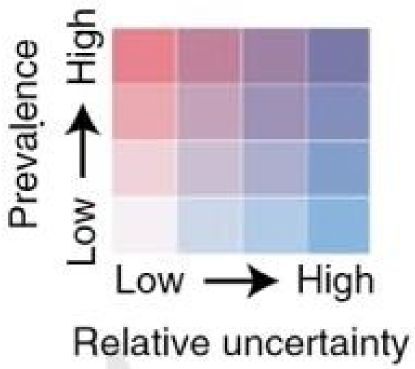
Depends on User, easy to get driven by "have tools will model"

Other things to consider:

- Resolution (depends on extent)
  - country
  - admin
  - pixel
- Update frequency
- Snapshot
- Projection

Levels of Uncertainty:

- Ensembles, Replicates
- Covariate errors
- Modelling errors
- Diminishing returns??



AIM
Aedes Invasive Mosquitoes

cost
EUROPEAN COOPERATION IN SCIENCE & TECHNOLOGY