



# **Interbeef Technical Committee meeting**

Uppsala, Dec 6-7, 2012

# Created an incoming data checking program



Why a checking program?

A) Check the content in the file to prevent the programs to crash during the evaluation due to invalid characters or inconsistencies in the data files.

B) Useful if/when a database for Interbeef is developed.

Could use the same principle as for dairy [?] Program will be used by member organizations. A zipfile will be created IF both the parameter and performance file is correct. The zip file should then be uploaded into the database via a restricted area.

# Checks in the program



Checks each field in the file so that the file has the correct content.

- a) Integer in an integer field and so on.
- b) correct abbreviations is used in the field
- c) country in herdid = reporting country
- d) no empty spaces (in the field or within the filed)
- e) consistency checks between number of enviromental effects and number environmental effects actually given within the file and between the parameter file and performance file.

# Programs used in the evaluation



- Mix99
- python programs are used to create the infiles to mix99
- python programs are under Mercurial version control

# Data reception



- Countries did well in keeping the deadline
- Character fields were renumbered (herid, animalid), but renumbering was kept to a minimum to ensure traceability during the evaluation
- Extra renumbering was needed for outstanding large integer values.
- There is an limit to the maximum integer number MiX99 (Fortran) can read. On 32-bit platforms the data file can contain integer numbers in the range of  $\pm 2.147.483.648$ . Therefore we would kindly ask you to do the renumbering on your side. We will add this as an extra check in our checking program



# National evaluation statistics

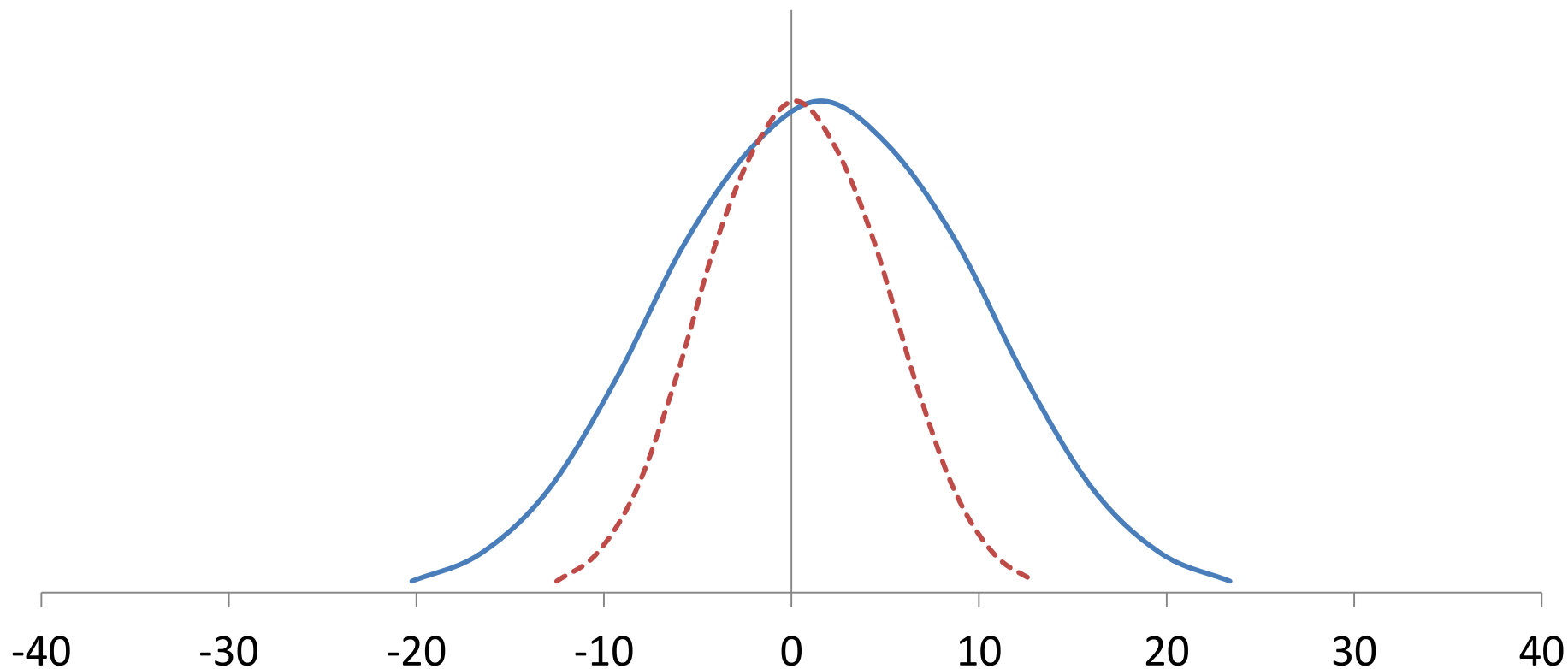
<b>Country:</b>	CZE	CZE	CZE	CZE	GBR	ESP	ESP	FRA	FRA	IRL	FIN	FIN	SWE	SWE
<b>Breed:</b>	LIM	LIM	CHA	CHA	LIM	LIM	LIM	CHA	LIM	ALL	CHA	LIM	CHA	LIM
<b>Effect:</b>	Direct	Maternal	Direct	Maternal	Direct	Direct	Maternal	Direct	Direct	Direct	Direct	Direct	Direct	Direct
<b>Mean EBV:</b>	1.91	-4.01	18.52	-4.64	11.47	1.57	0.20	2.64	4.50	0.51	-1.77	-0.50	3.42	6.48
<b>SD EBV:</b>	9.32	5.32	11.01	6.21	12.27	7.27	4.24	11.68	11.35	5.63	9.34	9.61	7.98	8.78
<b>Min EBV:</b>	-42.50	-36.60	-35.50	-39.20	-53.00	-35.07	-23.48	-64.30	-57.71	-53.70	-44.90	-44.00	-34.30	-34.60
<b>Max EBV:</b>	58.70	27.10	124.00	24.80	84.00	39.13	28.09	62.30	68.21	55.10	46.80	46.50	55.70	53.80
<b>Mean REL:</b>					0.35	23.15	16.28	0.32	0.40	0.24	0.36	0.42	0.58	0.59
<b>SD REL:</b>					0.18	13.22	9.66	0.12	0.11	0.16	0.15	0.16	0.11	0.10
<b>Min REL:</b>					0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.06	0.08
<b>Max REL:</b>					0.98	96.69	89.92	0.99	0.99	0.99	0.93	0.95	0.99	0.98



# National evaluation statistics



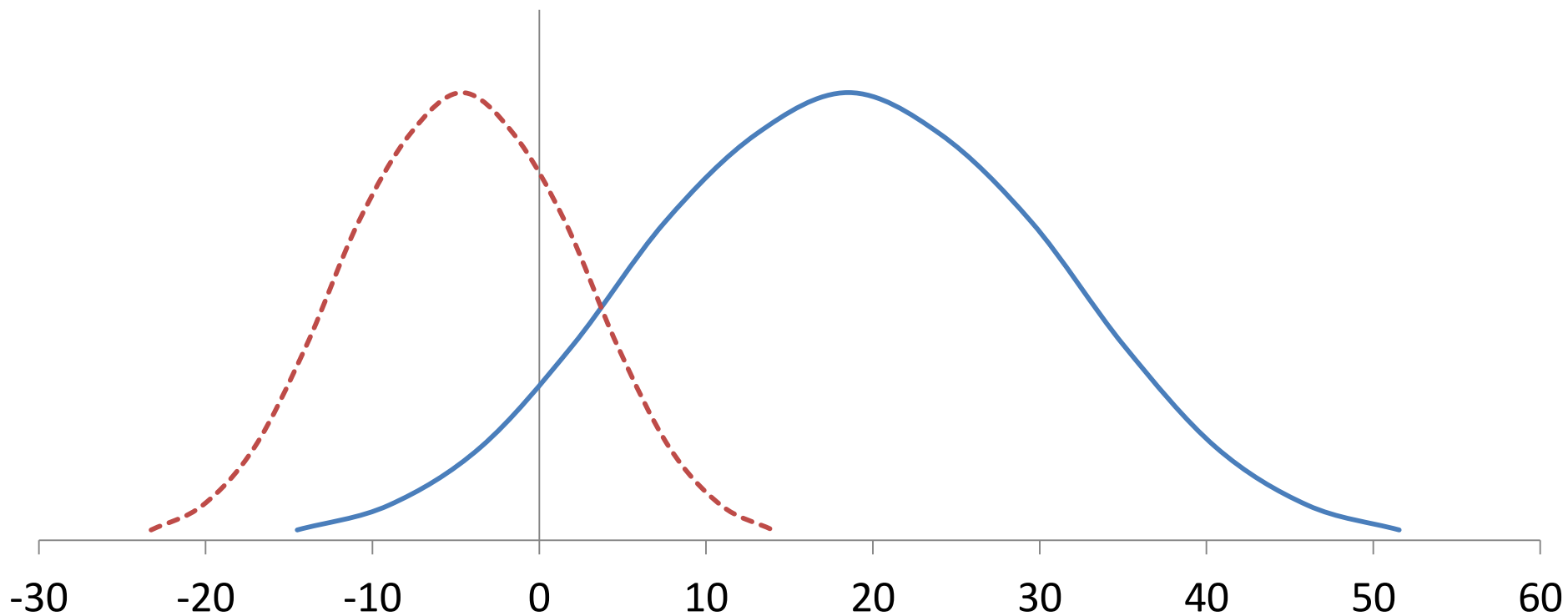
— ESP LIM Direct    - - - ESP LIM Maternal



# National evaluation statistics



— CZE CHA Direct    - - - CZE CHA Maternal





# Correlation matrix



- Mix99 uses variance-covariance matrix instead of correlations
- Previous run:
  - LIM: dnk, fra, gbr, irl, swe
  - CHA: irl, dnk, swe, fra
- Current run: irl, cze, dnk, swe, fra
  - LIM: **cze**, dnk, **fin**, fra, gbr, irl, swe, **esp**
  - CHA: irl, **cze**, dnk, swe, fra, **fin**

# Correlation matrix



- How did we calculate the variance components?
  - Adding all new countries together with the "previous" countries → matrix not positive definite.
- Therefore we started with two countries and made the matrix positive and estimated the variance components. Then we added one country at the time and reestimated the variance components until all countries were included
- Still need to re-run with same edits applied in the EBV estimation procedure





# CHA – (Co)variance matrix



	irl (1)	cze (2)	dnk(3)	swe(4)	fra (5)	fin (6)
irl (1)	502.44134	182.94663	353.19552	262.77831	293.86344	192.57205
cze (2)		163.96132	155.62158	56.19973	93.23760	117.68589
dnk(3)			260.59693	165.03197	176.30375	145.55582
swe(4)				184.42304	206.56169	94.76283
fra (5)					328.42673	101.67578
fin (6)						197.51698

# CHA – Correlation matrix



	irl (1)	cze (2)	dnk(3)	swe(4)	fra (5)	fin (6)
irl (1)		0.63740	0.97609	0.86326	0.72341	0.61129
cze (2)			0.75286	0.32319	0.40179	0.65396
dnk(3)				0.75279	0.60264	0.64157
swe(4)					0.83931	0.49651
fra (5)						0.39921



# Country of origin – LIM

(rel  $\geq$  0.5 in at least one scale = 433 422 animals)



Country	Frequency	Percent
BEL	1	0.0
CAN	11	0.0
CZE	59	0.0
DEU	9	0.0
DNK	4449	1.0
ESP	213	0.1
FIN	477	0.1
FRA	412007	95.1
GBR	1847	0.4
IRL	331	0.1
LUX	24	0.0
NLD	6	0.0
NZL	1	0.0
SWE	13984	3.2
USA	3	0.0

# All scales stats on EBV & Rel - SIM



✦ (rel  $\geq$  0.5 in at least one scale = 433 422 animals)

Variable	N	Mean	Std Dev	Sum	Minimum	Maximum
EBV_CZE	433422	4.47815	5.53446	1940928	-22.328	36.791
EBV_DNK	433422	8.19257	9.37235	3550840	-42.123	59.448
EBV_ESP	433422	4.13094	5.06546	1790440	-19.402	38.1
EBV_GBR	433422	4.38417	5.52926	1900194	-25.58	39.093
EBV_IRL	433422	8.76526	10.42439	3799058	-46.613	64.019
EBV_SWE	433422	8.64117	9.94854	3745274	-49.494	60.143
EBV_FRA	433422	8.99656	10.25483	3899306	-44.856	62.163
EBV_FIN	433422	5.55665	6.20697	2408375	-28.689	42.134
RCZE	433422	0.29805	0.05997	129183	0.04496	0.93078
RDNK	433422	0.45119	0.06818	195554	0.11475	0.98191
RESP	433422	0.19378	0.06371	83988	0.05121	0.96443
RGBR	433422	0.18913	0.06749	81975	0.05486	0.98061
RIRL	433422	0.4718	0.07027	204487	0.07029	0.97964
RSWE	433422	0.44296	0.07054	191990	0.05147	0.96877
RFRA	433422	0.55151	0.08631	239037	0.12906	0.99972
RFIN	433422	0.1819	0.05095	78838	0.04566	0.90134



# Country of origin – CHA

(rel  $\geq$  0.5 in at least one scale = 411 563 animals)



Country	Frequency	Percent
CAN	40	0.0
CZE	196	0.1
DEU	5	0.0
DNK	413	0.1
FIN	239	0.1
FRA	403108	98.0
GBR	22	0.0
HUN	1	0.0
IRL	2459	0.6
LUX	14	0.0
NOR	1	0.0
SVK	1	0.0
SWE	5063	1.2
USA	1	0.0

# All scales stats on EBV & Rel - CHA



✦ (rel  $\geq$  0.5 in at least one scale = 433 422 animals)

Variable	N	Mean	Std Dev	Sum	Minimum	Maximum
EBV_IRE	411563	7.74866	10.51796	3189061	-49.986	62.68
EBV_CZE	411563	2.93954	4.29352	1209806	-30.341	27.786
EBV_DNK	411563	4.89522	6.74896	2014692	-34.17	45.066
EBV_SWE	411563	5.07493	6.87563	2088653	-35.137	41.044
EBV_FRA	411563	7.89081	10.71679	3247565	-55.869	63.478
EBV_FIN	411563	3.1308	4.3917	1288520	-33.725	40.432
RIRE	411563	0.31986	0.05859	131641	0.18571	0.98549
RCZE	411563	0.13219	0.06621	54403	0.05076	0.93543
RDNK	411563	0.23867	0.05982	98227	0.17889	0.95509
RSWE	411563	0.40928	0.0657	168443	0.05092	0.97762
RFRA	411563	0.56667	0.09174	233219	0.07813	0.9994
RFIN	411563	0.11699	0.04715	48147	0.07813	0.86577

# Correlations - CHA

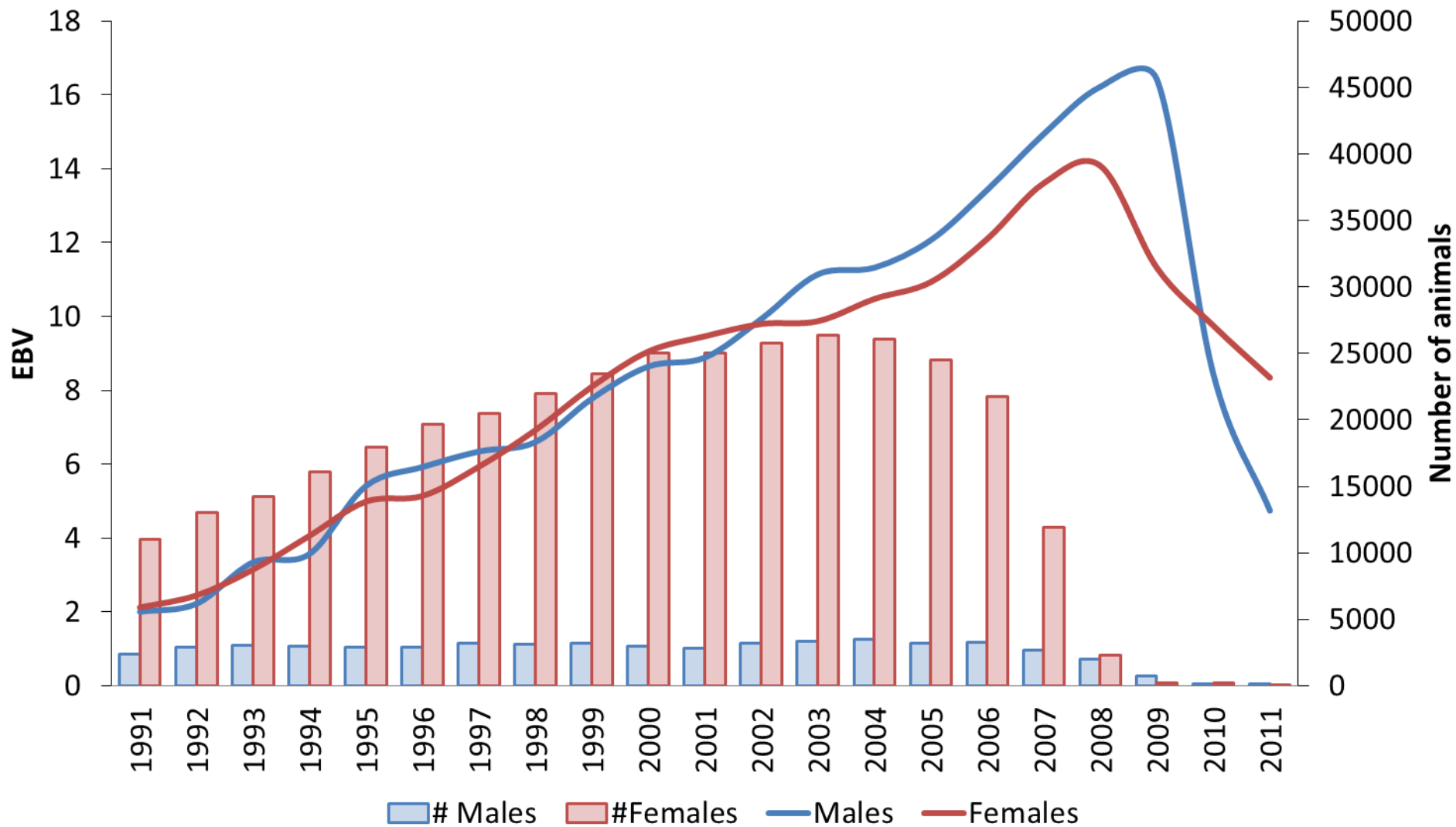


	EBV_IRL	EBV_CZE	EBV_DNK	EBV_SWE	EBV_FRA	EBV_FIN
EBV_IRL	1	0.93197	0.99384	0.98434	0.96852	0.95243
		<.0001	<.0001	<.0001	<.0001	<.0001
EBV_CZE	0.93197	1	0.96103	0.86284	0.8507	0.97379
	<.0001		<.0001	<.0001	<.0001	<.0001
EBV_DNK	0.99384	0.96103	1	0.95888	0.93756	0.96976
	<.0001	<.0001		<.0001	<.0001	<.0001
EBV_SWE	0.98434	0.86284	0.95888	1	0.99284	0.90307
	<.0001	<.0001	<.0001		<.0001	<.0001
EBV_FRA	0.96852	0.8507	0.93756	0.99284	1	0.88375
	<.0001	<.0001	<.0001	<.0001		<.0001
EBV_FIN	0.95243	0.97379	0.96976	0.90307	0.88375	1
	<.0001	<.0001	<.0001	<.0001	<.0001	



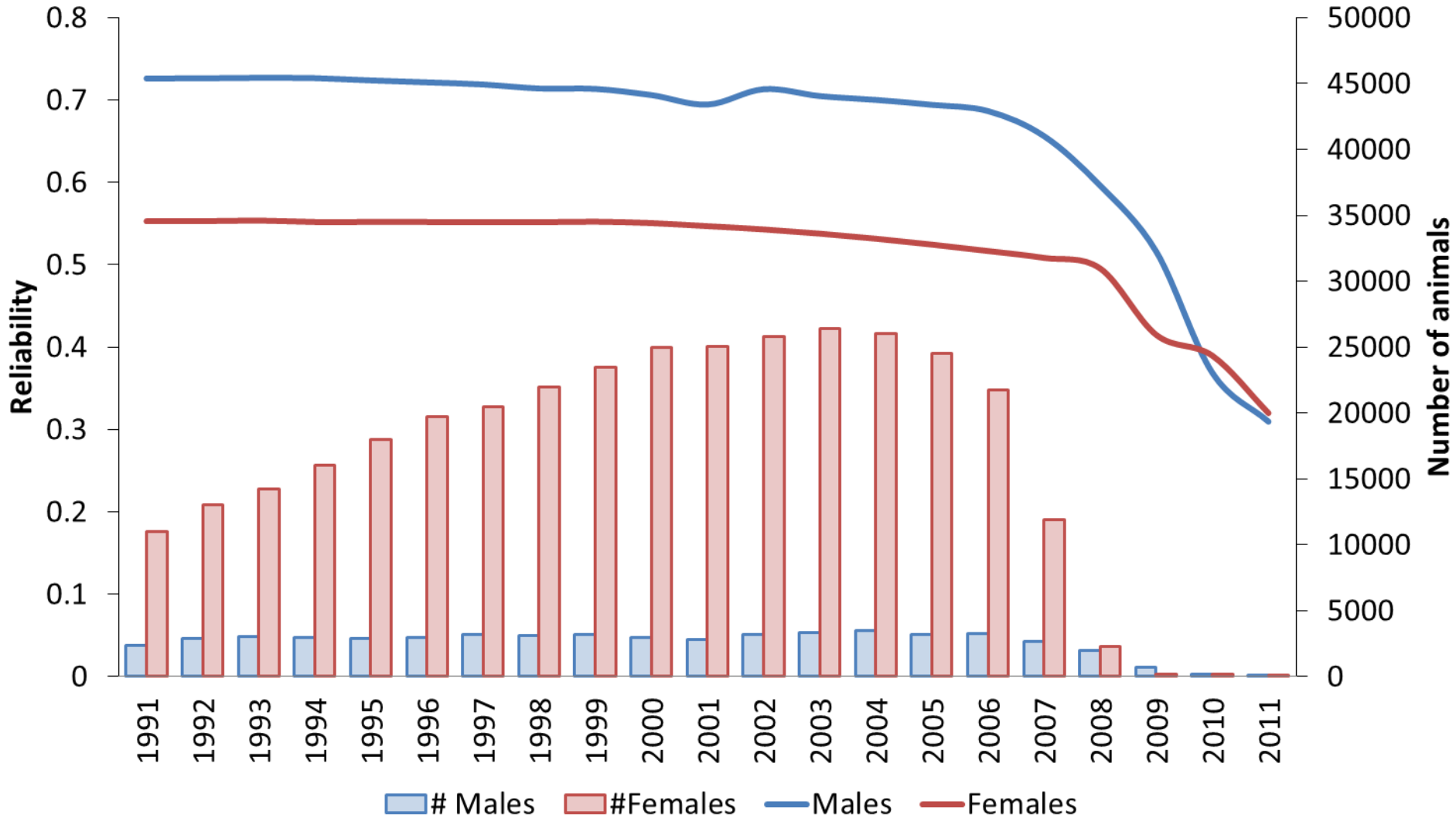
# Genetic trend for CHA in FRA scale - Model 1

(rel  $\geq 0.50$  in at least one scale)



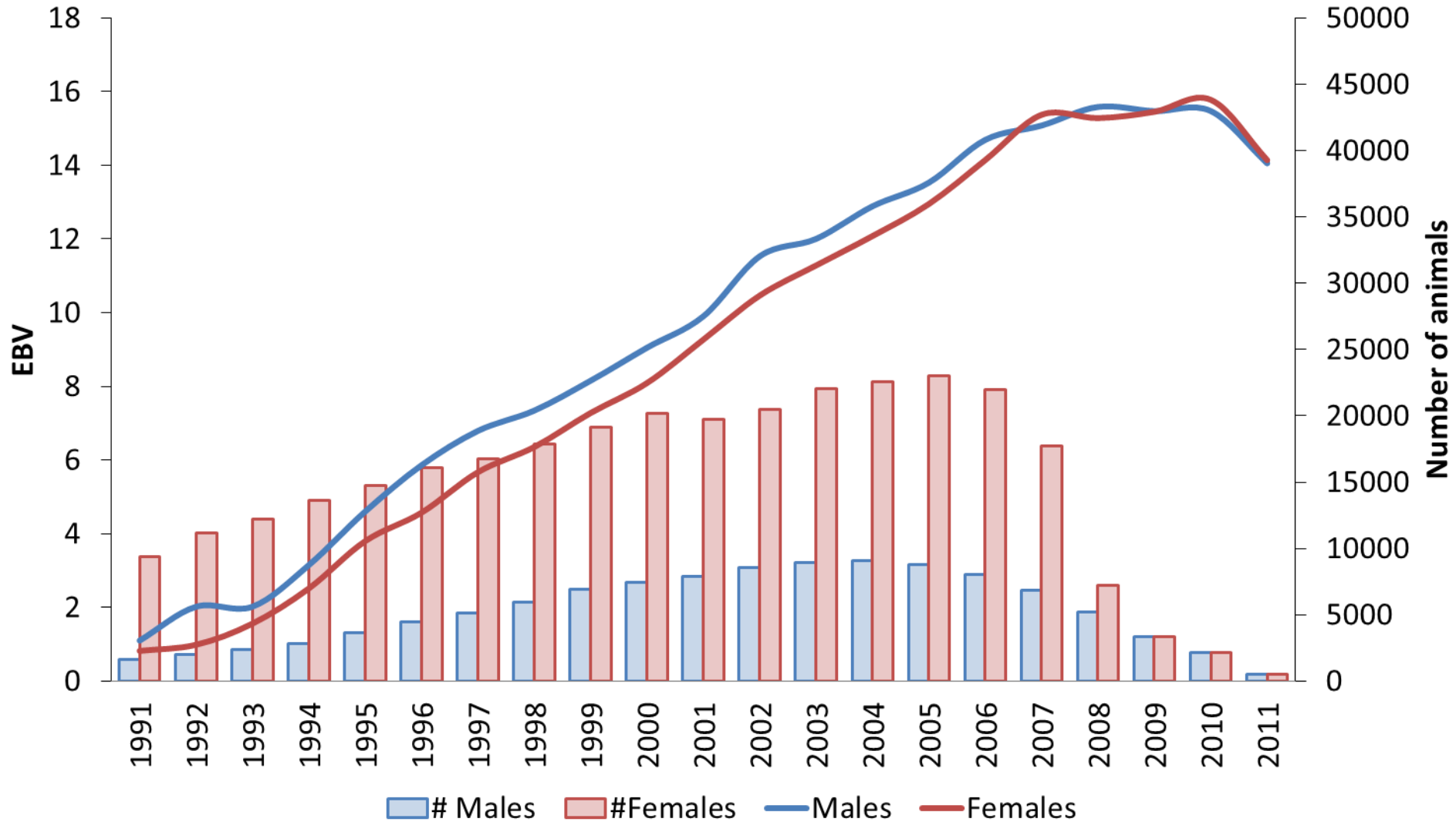
# Average reliability for CHA in FRA scale - Model 1

(rel  $\geq 0.50$  in at least one scale)



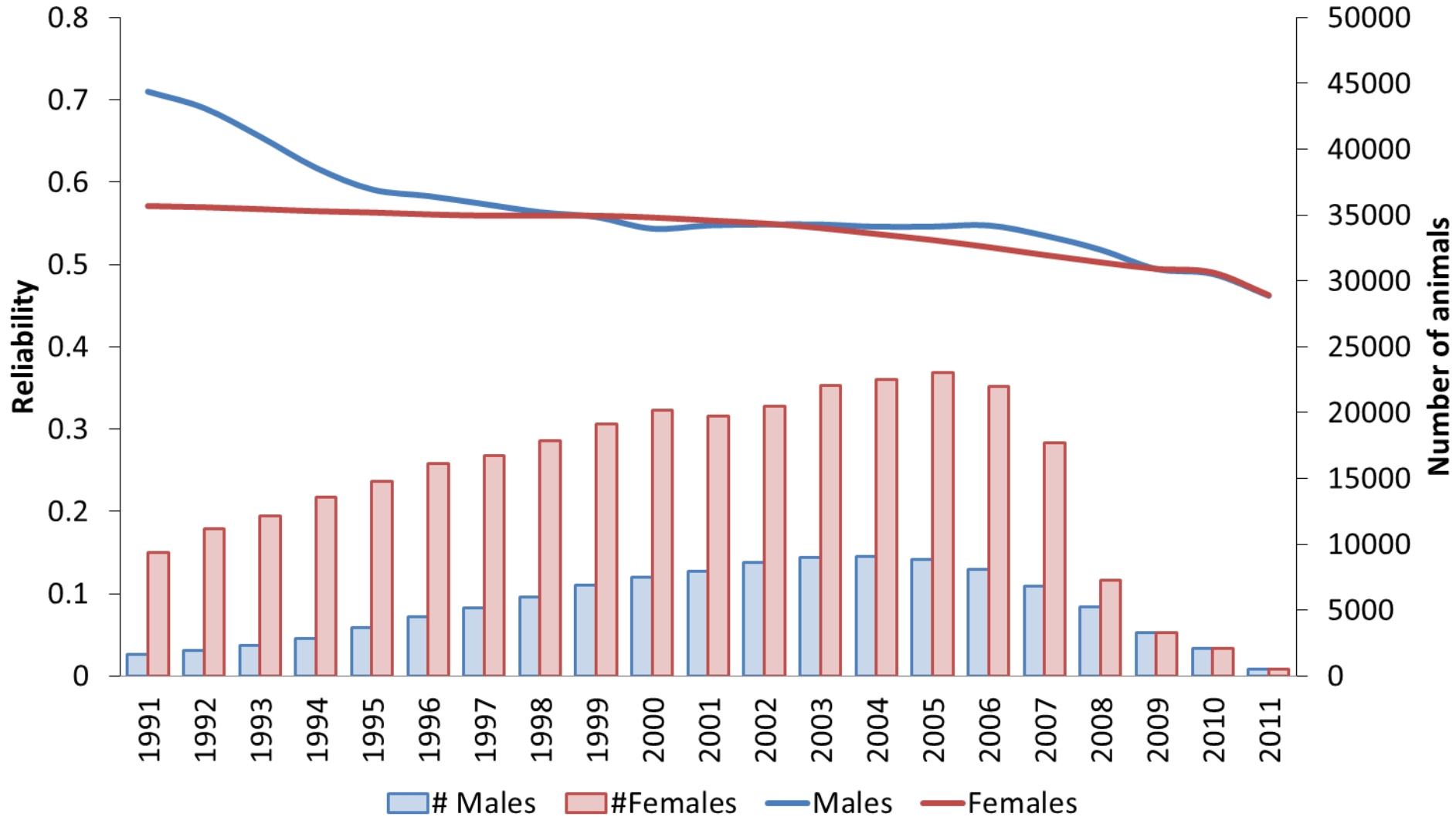
# Genetic trend for LIM in FRA scale - Model 1

(rel  $\geq 0.50$  in at least one scale)

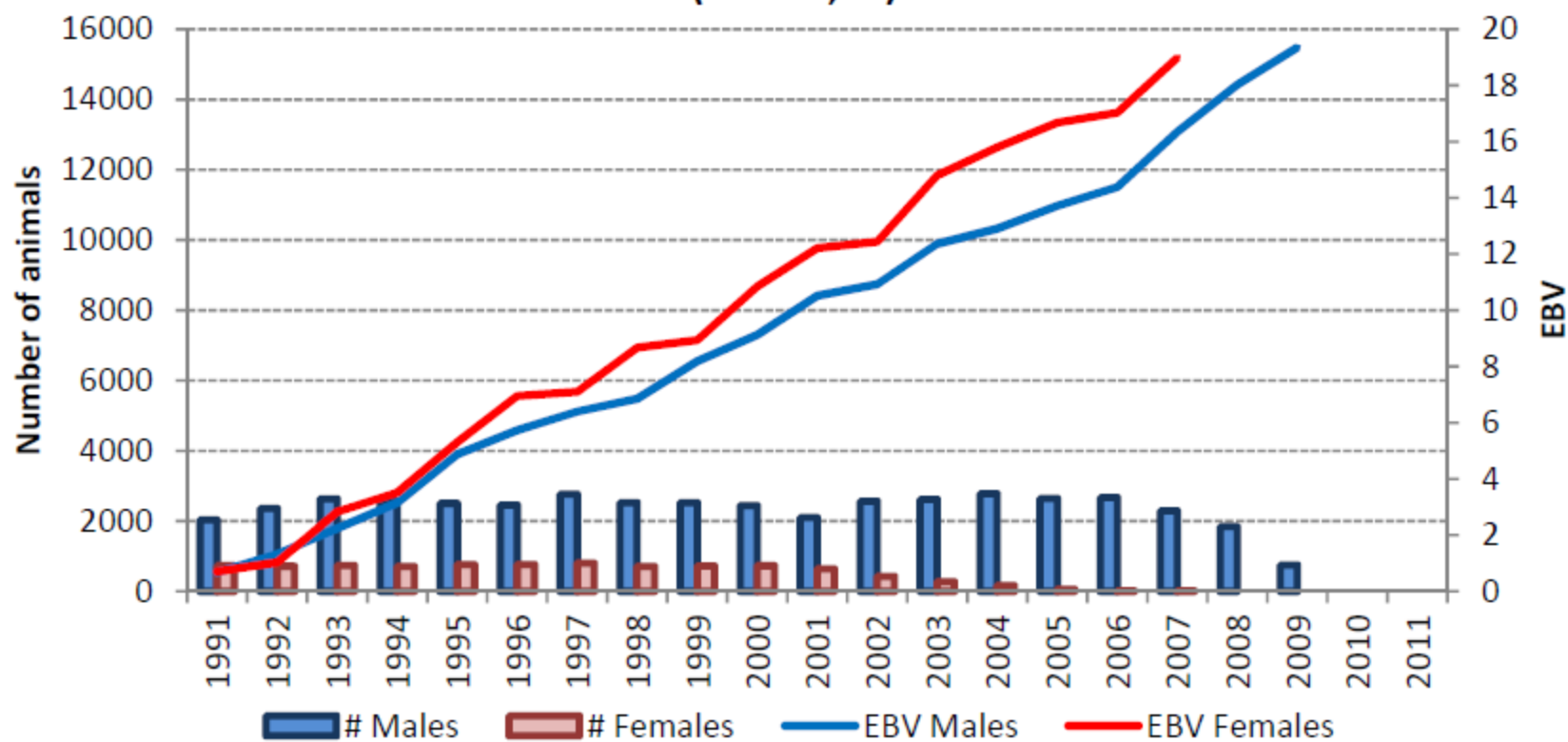


# Average reliability for LIM in FRA scale - Model 1

(rel  $\geq 0.50$  in at least one scale)

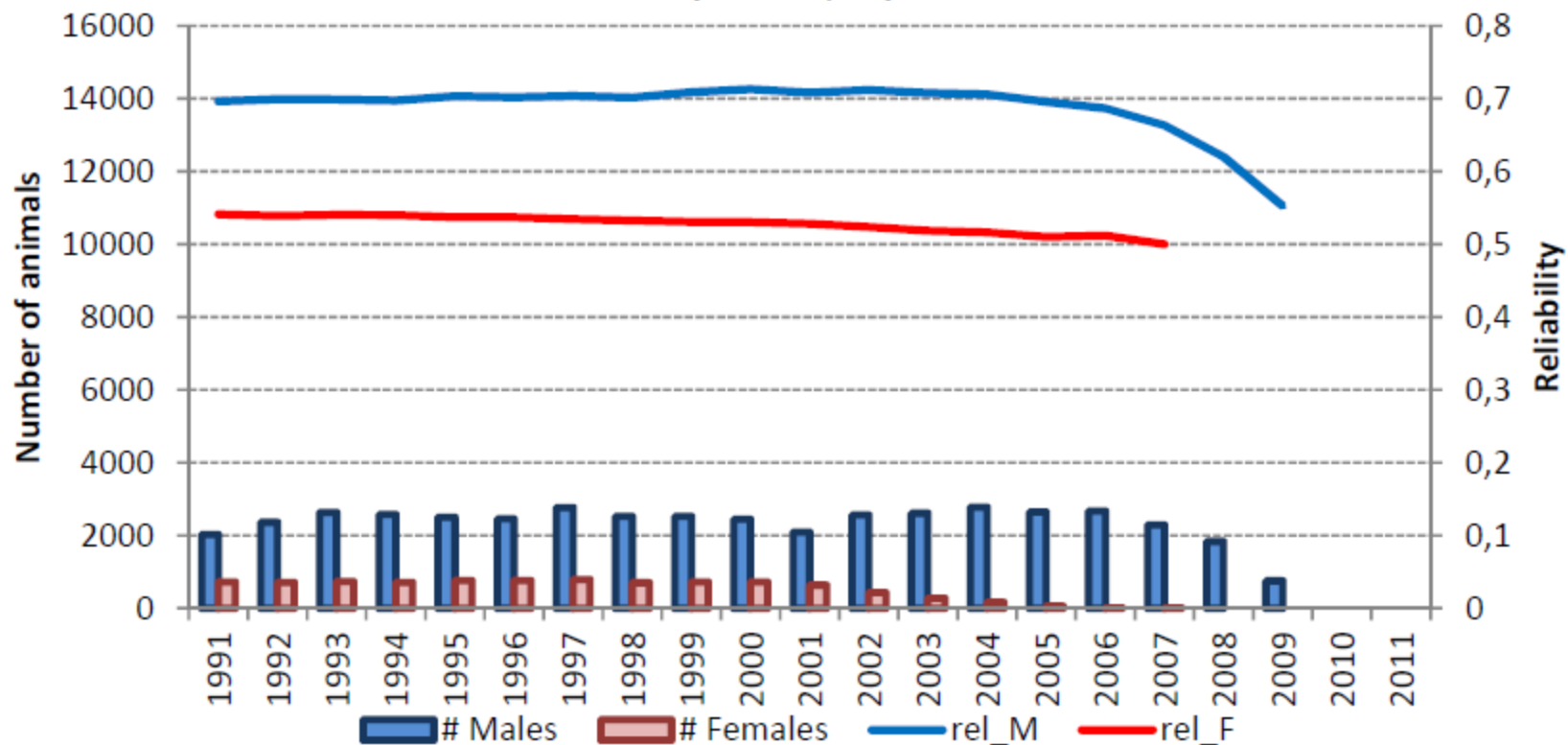


## Genetic trend for CHA in France (rel $\geq 0,50$ )



# Average reliability for CHA in France

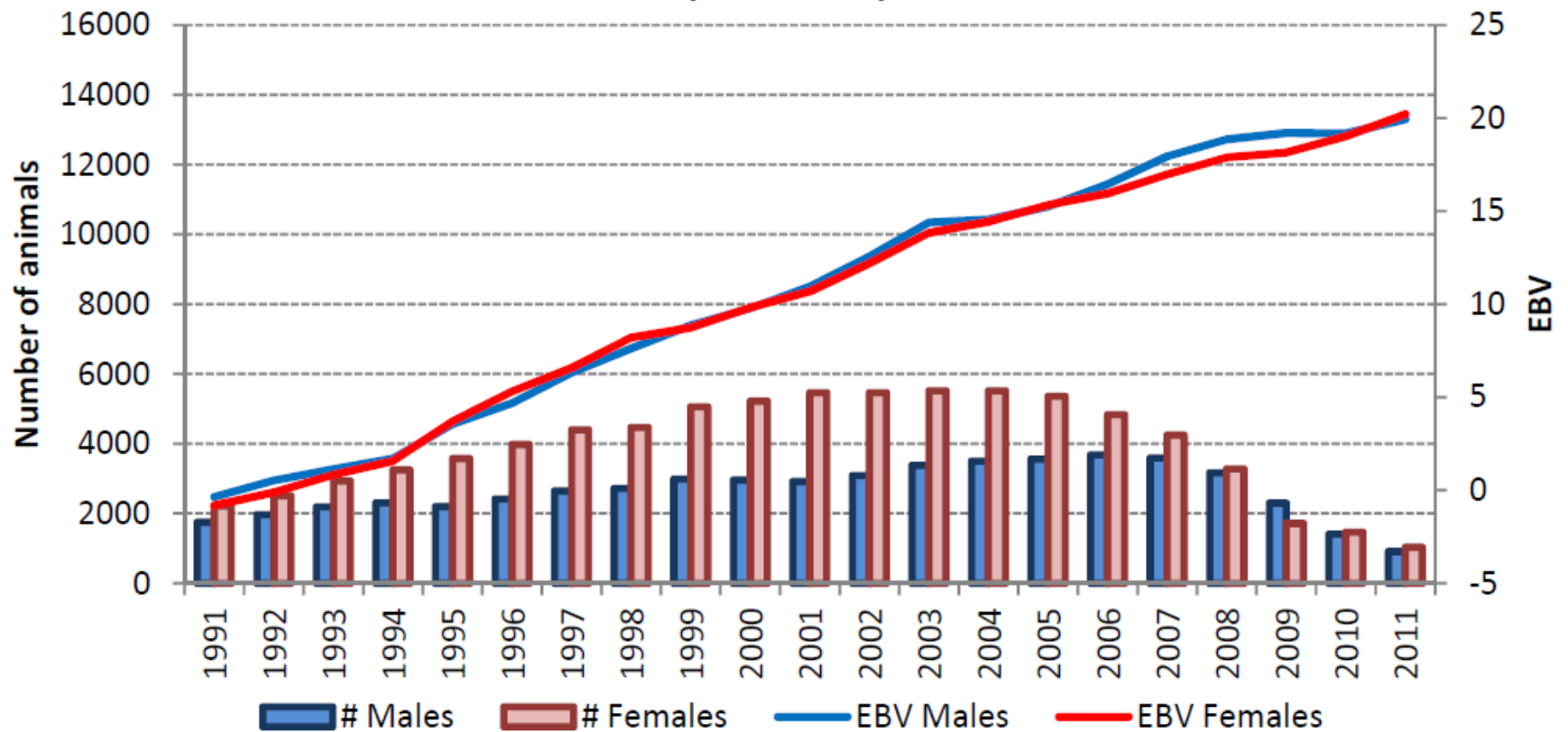
(rel  $\geq 0,50$ )





# Genetic trend for LIM in France

(rel  $\geq 0,50$ )



# Average reliability for LIM in France

(rel  $\geq 0,50$ )

