Regression model to forecast the structure of agriculture in Finland

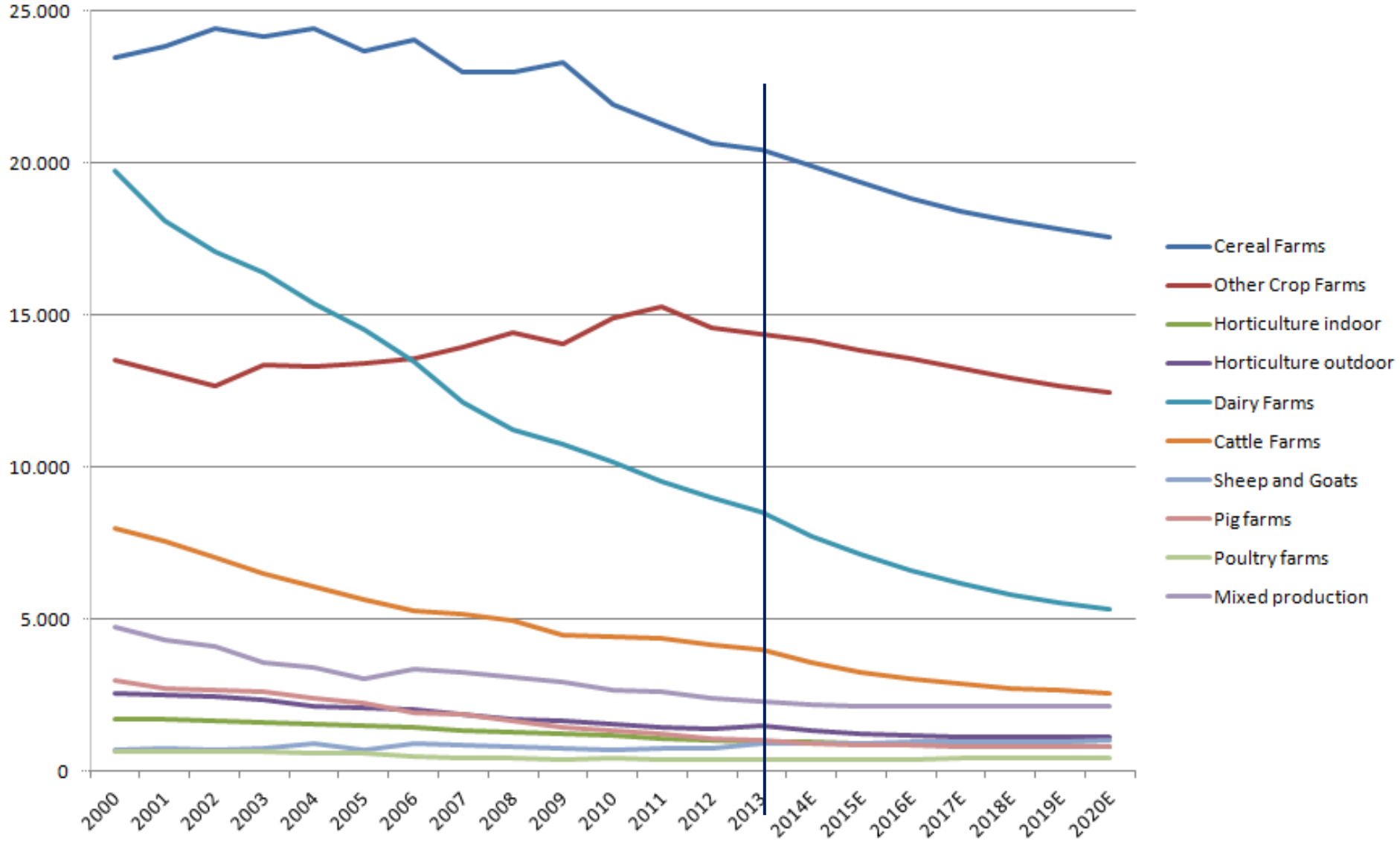
Arto Latukka & Jukka Tauriainen

MTT Agrifood Research Finland

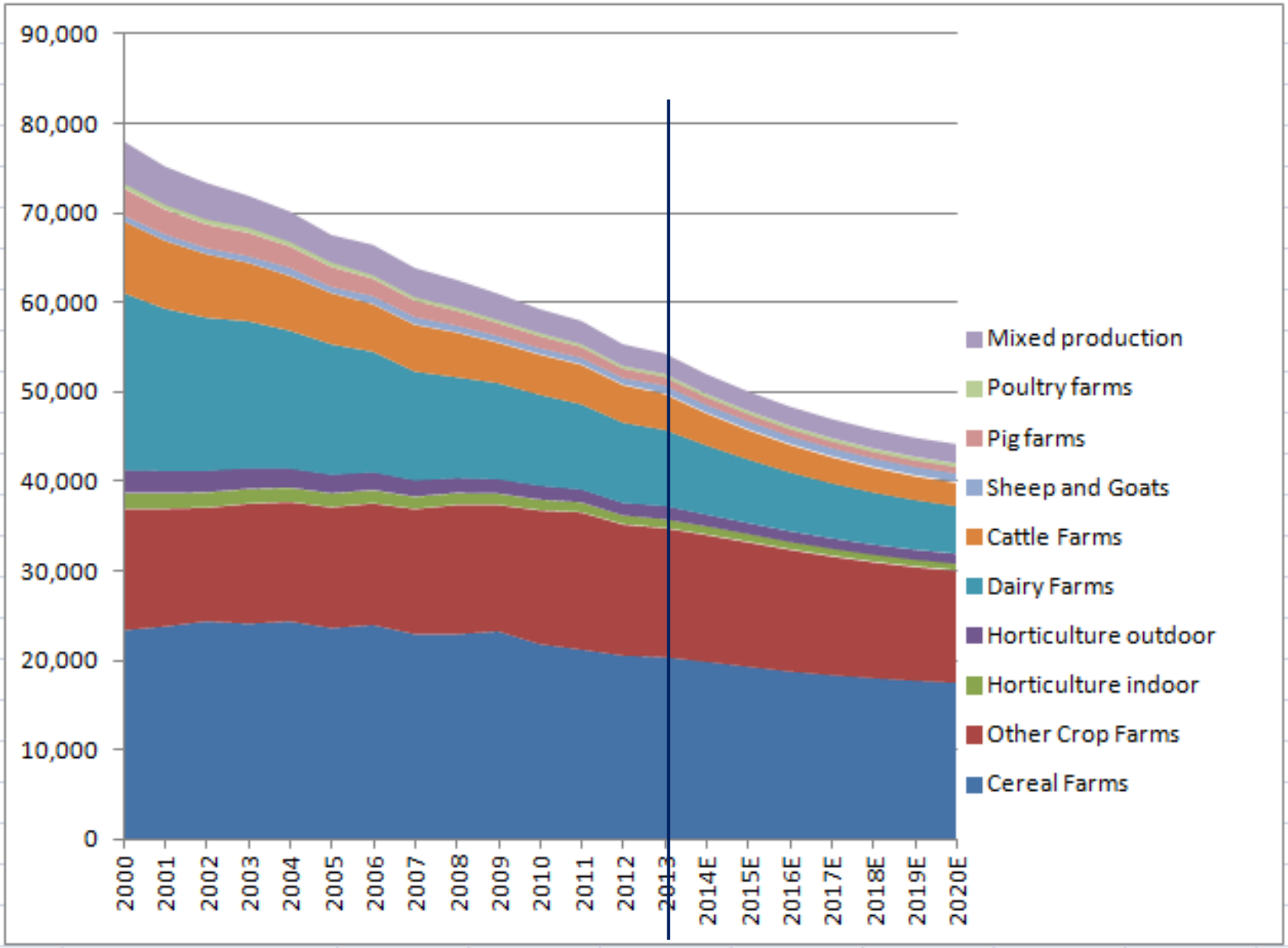
22nd PACIOLI –workshop

28th of September – 1st of October 2014
Castleknock Hotel, Dublin, Ireland

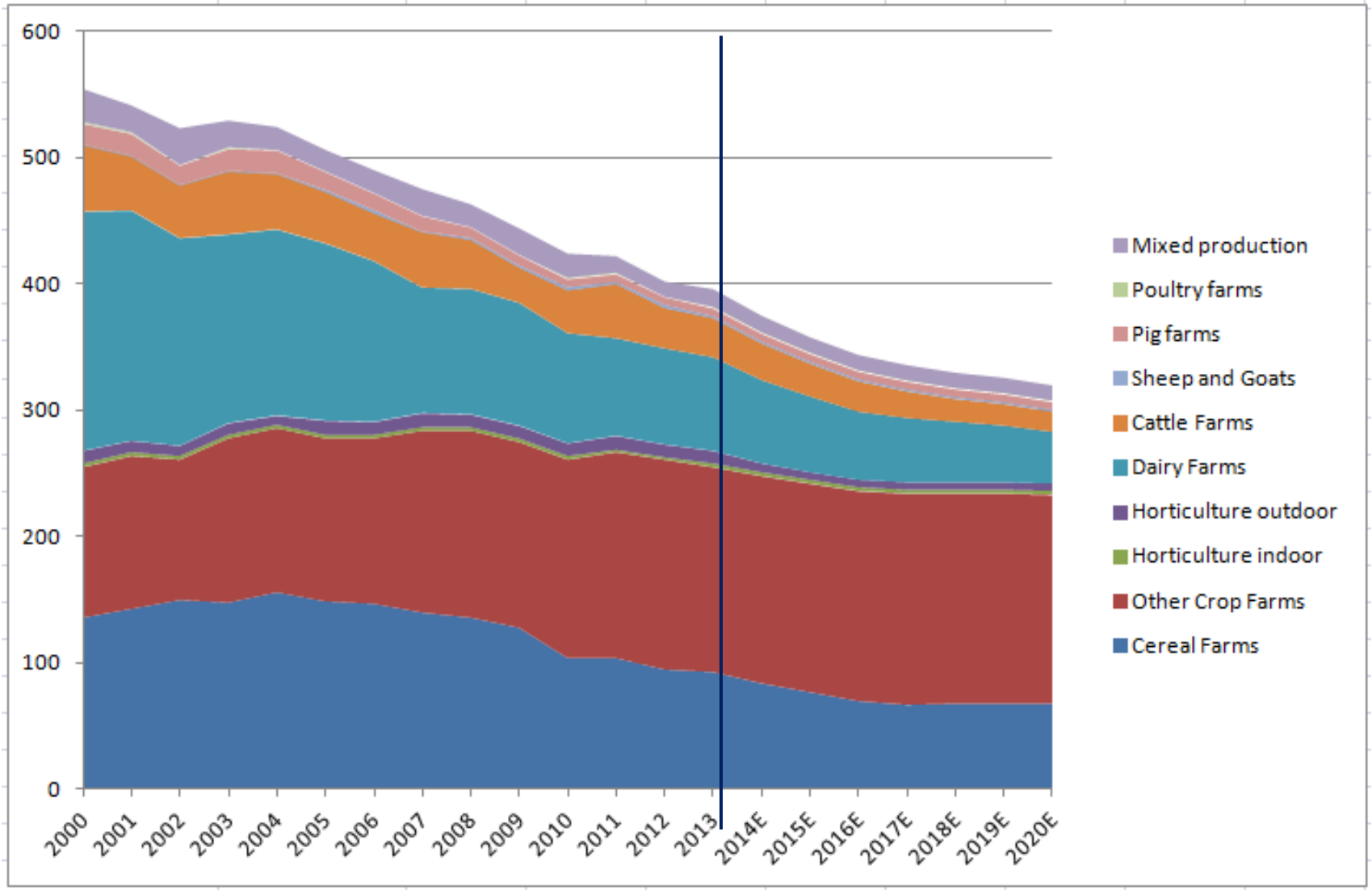
Results: Number of farms 2000-2020E in Finland



Results: number of farms in 2000-2020 in Finland



Number of farms in municipality ’xxx’



Forecasting model

- **All the farms from the farm register 2000 - 2013 are included (900 000 farms)**
- **SO-typology has been calculated for all of the farms (type of farm, size class)**
- **All the regional classifiers are included into the dataset**
- **In order to forecast the number of farms for year 2014, seven years time span has been taken from the dataset (2007-2013)**

Data for regression model

Farms has been classified yearly into groups based on six classifying factors (seven classifiers)

- ‘subsidy regions’ 7
- ’extension service centres’ 20
- ‘municipalities‘ 320
- ‘language (finnish, swedish)’ 2
- ‘production types’ 10
- ‘size class’ 14

There are yearly more than 12 million cells, but farms can be found in 12.000 different cells yearly (for example municipality belongs only to one subsidy region)

The number of farms in each cell/group is calculated yearly.

The regression model is based on the development of the number of farms in each cell/group

Regression model

Regression model has been constructed to find intercepts and regression coefficients for each combination of classifying factors (about 12 000 regression analysis)

proc reg by ‘subsidy region’ ’extension centre’ ‘municipality‘ ‘language’ ‘production type’ ‘size class’

model ‘Number of farms’ = ‘year’

The linear regression model

- The dependent variable is ‘Number of farms’
- The independent variable is ‘Year’

Problem: The change from last year to first forecasted year

- The change from the last year of real data into first year of forecasted results is critical (2013->2014)
- There might be large change (shift) of the number of the farms from 2013 to 2014E
- Solution: The regression lines of each class combination is shifted up or down so that the lines goes via the correct number of the farms in the last real year (so 2013)
- Intercepts of the regression lines are changed

Calculation of the numbers of farms to year 2014E (forecast)

- The dataset intercepts and regression coefficients of each combination of classifiers are included into the dataset

- Number of farms for 2014E for each class is calculated

  N=intercept + regression coefficient x year

Change the numbers of farms to integers

- Numbers of the farms calculated are not always integers
- -> ‘round’ or ‘floor’ the number of farms in each class to integers

if the regression coefficient (b1) is negative and the size class is small then ‘floor’; otherwise ‘round’

- if b1 <0 then if size class <7 then N=floor(N);else N=round(N);

if the regression coefficient is zero and the size class is small then ‘floor’; otherwise ‘round’

- if b1 =0 then if size class <7 then N=floor(N);else N=round(N);

if the regression coefficient is positive and the size class is large then ‘round’; otherwise ‘floor’

- if b1 >0 then if size class >9 then N=round(N);else N=floor(N);

Sometimes number of the farms can be negative

If the number of the farms (forecast) is negative in certain class

- The number of farms is changed to zero in that class
- The number of the farms in the next, larger size class is decreased by that negative amount (or as much as possible)
- And all the “negative” amount can’t be subtracted, then the number of farms in the next larger size class is subtracted

Based on the number of farms in each class, the farm dataset is expanded

- Expand the numbers of the farms in each class to farm data
- If there exists for example five farms in class, then there will be five rows in farm dataset
- In 2014E there will be 52 000 rows, so 52 000 farms

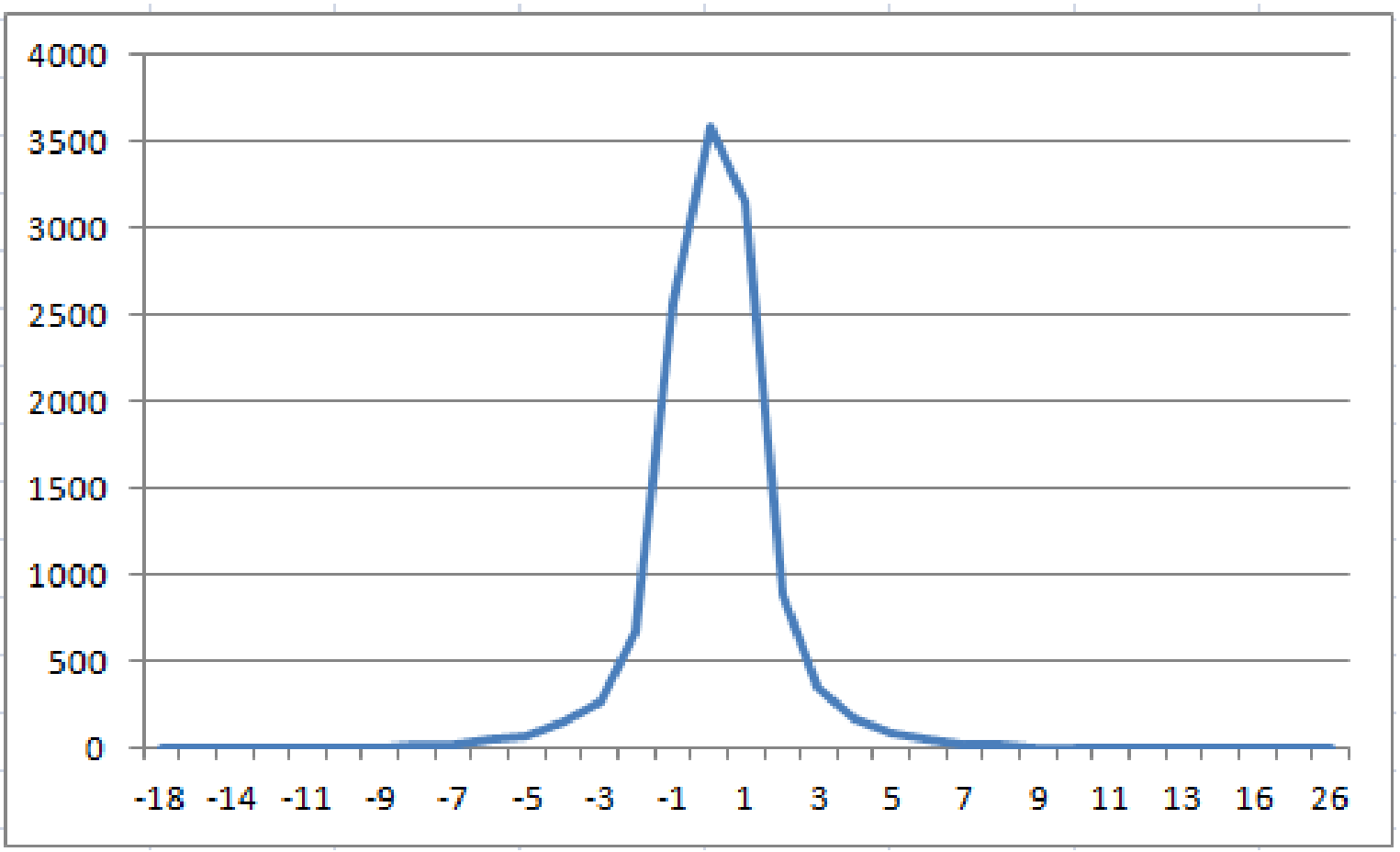
- The forecasts for year 2015E is done based on 2008-2014E
- The circle/loop is done for each year until all the forecasts have been calculated (until 2020E)
- About 84 000 regression analysis had to be done (12 000 regression for each 7 years, 2014-2020)

Testing of the forecasting model with year 2013

- Regression model was tested by forecasting the number of farms in 2013 with 2006-2012
- The difference is calculated between the real number of farms and forecasted number of farms
- 12147 regression analysis

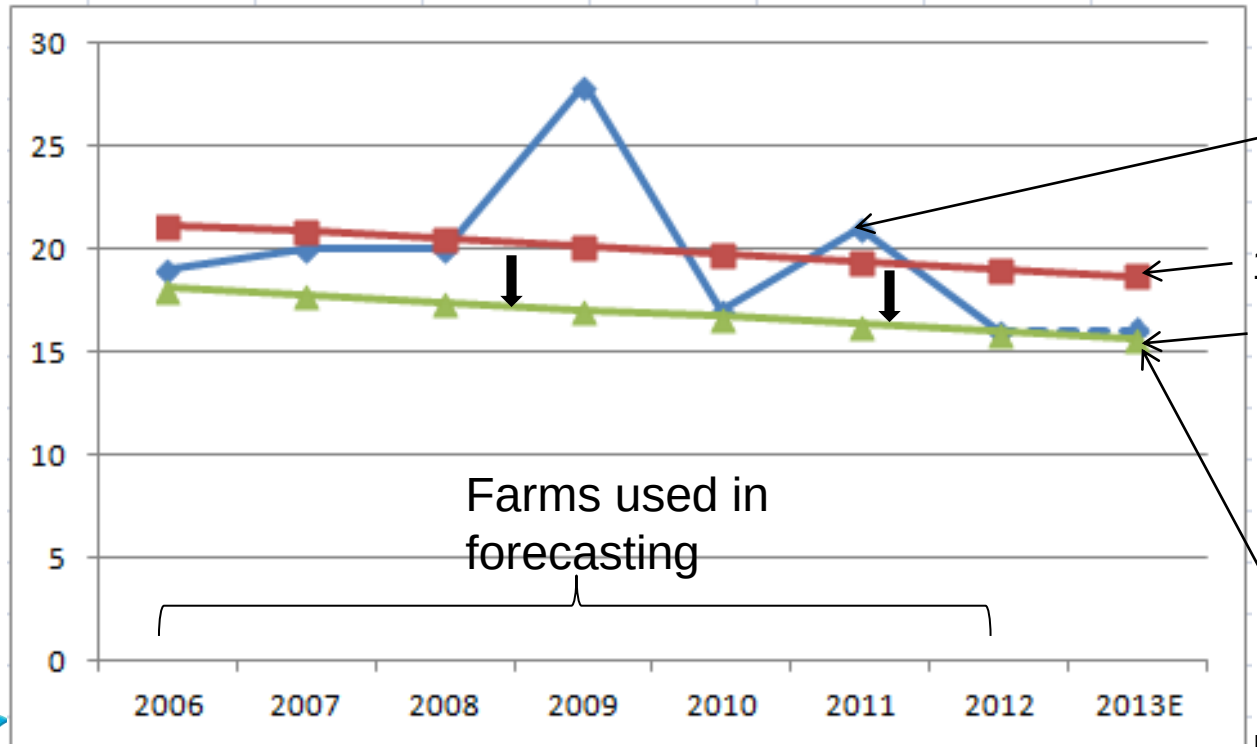
	Mean	Standard deviation
Difference (number of farms -forecast)	0.117	1.826
Absolute difference	1.179	1.400
Real number of farms	54,367	farms
Forecasted number of farms	52,946	farms
Difference	1,421	farms
Difference, %	2.614	%

Difference between the real and forecasted number of farms in 2013



Example of regression model with forecasting 2013

Subsidy region	‘C2’
Extension centre	‘Etelä-Pohjanmaa’
Language	‘Finnish’
Municipality	‘xxx’
Production type	’Ceral farms’
Size class	‘4’



Number of Farms in dataset, blue

18.71 forecast 2013, red

15.64 shifted, green

16 farms, after rounding = forecast 2013, result

Number of farms in real dataset is16 in 2013, blue

Forecasting error = 0

Language selection: Suomeksi | På svenska | In English

MTT ECONOMYDOCTOR www.mtt.fi/economydoctor

EU MEMBER STATES
EU Member State FADN-results

FADN Standard Results (SO)
Results of EU member countries 2004-2012e

Results updated 15.8.2014

FADN Advanced Results (SO)
Key ratios calculated of EU results by member countries 2004-2011

Results updated 15.8.2014

FADN Standard Results (SGM)
Results of EU member countries 1989-2009

Results updated 15.8.2014

FADN Advanced Results (SGM)
Key ratios calculated of EU results by member countries 1989-2009

Results updated 15.8.2014

FINLAND

Agriculture and horticulture
Results 2000-2012 updated 19.6.2014
Forecasts 2013E updated 11.8.2014

Reindeer farming
Results 2012/2013 updated 28.5.2014
Forecasts 2013/2014e updated 28.5.2014

Fur farming
Results of accounting year 2006

Forecast of Structural Development
Structure Development 2000-2020E
Results updated 8.5.2014

The Structural Development of Agriculture
Structure Development 2000-2012
Results updated 8.5.2014

Total Calculation
Results 2000-2012 updated 19.6.2014
Forecasts 2013E updated 11.8.2014

Forecast results can be found in EconomyDoctor - service

2000-2020E | Own criteria | info

- Subsidy Regions
- Regions
- Ely-centers
- Provinces
- FADN region
- ProAgria Centre
- Production types

FADN region

MTT Economydoctor. Forecast of Structural Development (mtt.fi/economydoctor). MTT EconomyDoctor Forecasting m

Number of farms	2020E	2019E	2018E	2017E	2016E	2015E	2014E	2013	2012	2011	2010	20
Etelä_Suomi	20.457	20.730	21.104	21.571	22.143	22.836	23.614	24.614	25.088	26.263	26.878	27.
Sisä_Suomi	8.815	9.009	9.267	9.593	9.950	10.371	10.856	11.351	11.528	12.026	12.303	12.
Pohjanmaa	9.360	9.547	9.764	10.023	10.336	10.711	11.156	11.386	11.551	12.109	12.440	12.
Pohjois_Suomi	5.613	5.685	5.762	5.875	6.019	6.217	6.443	7.016	7.225	7.614	7.693	7.
All the Farms	44.245	44.971	45.897	47.062	48.448	50.135	52.069	54.367	55.392	58.012	59.314	61.

Service Production: MTT Agrifood Research Finland | www.mtt.fi/english

