

“— INNOVATIVE as a **pig** in the Orchard—”



GROUP 2

AZIENDA AGRICOLA
LUCREZIA SRL



POLITECNICO
MILANO 1863

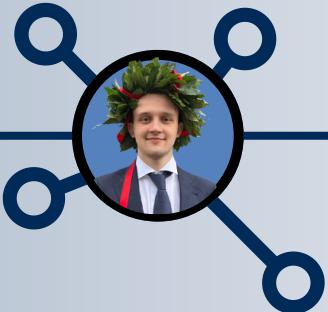


Prof. Matteo Matteucci
Prof. Filippo Maria Renga





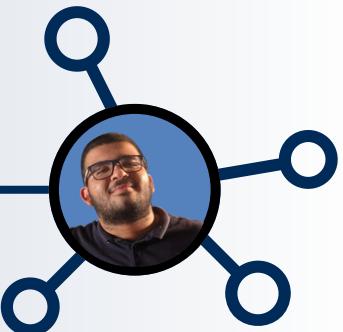
Manuela Mittica
10643012



Edoardo Talignani
10667788

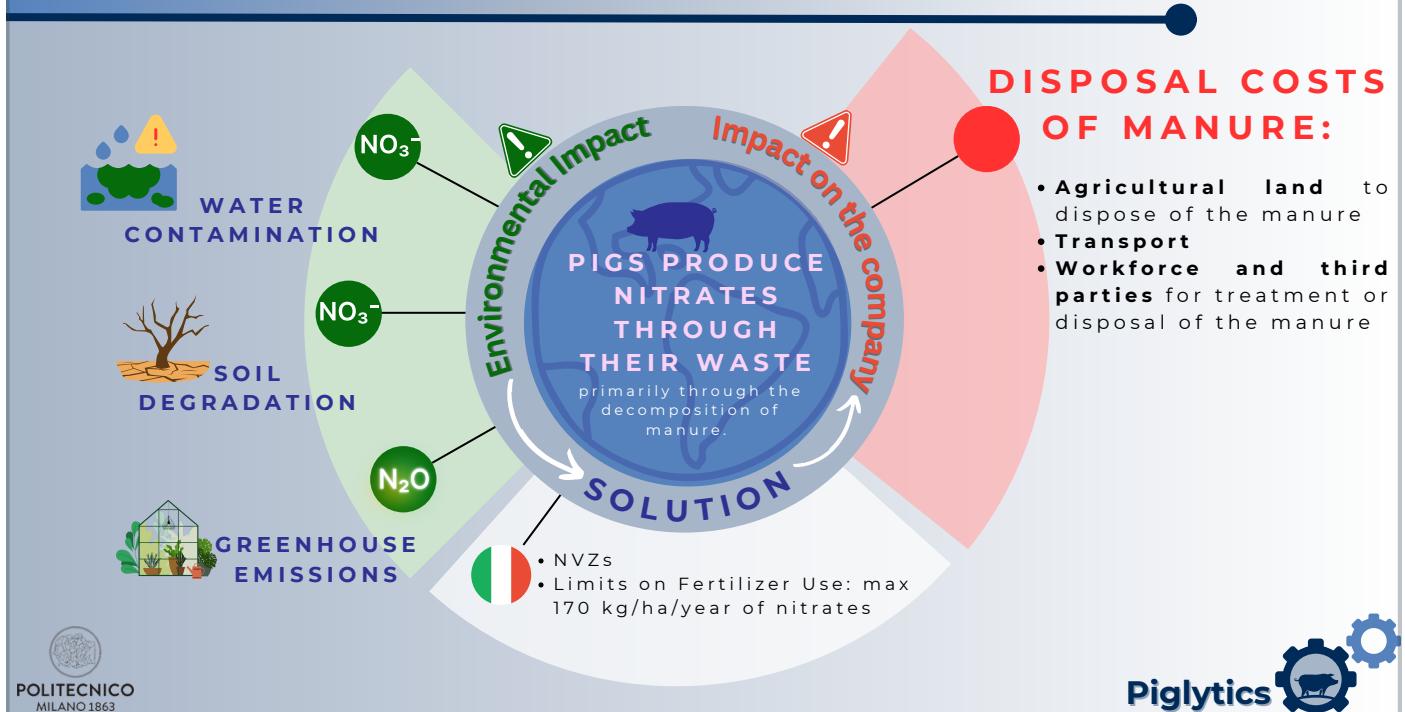


Bahram Hedayati
10870276

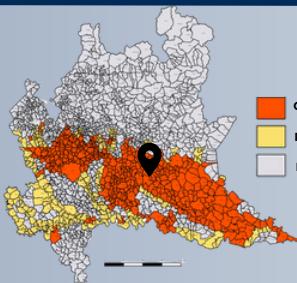


Ahmed Yassine
Hassairi
11074503

Nitrate Pollution from Pig Waste



KEY ASSUMPTIONS



Completely vulnerable
Partially vulnerable
Non-vulnerable

~67%
of the agricultural
land in Po Valley is
classified as NVZ
and the

Offlaga
is the Company
location,
surrounded by
NVZs

Company is
distributing
ALL manure in
NVZs

NVZs are increasing

WHAT IF the cost or limit of N/ha in NVZs changes?



REGULATORY PRESSURE

NVZs are likely to expand due to the regulatory pressure from EU (Nitrates Directive). Stricter regulations to reduce nitrate-related impacts.



REGIONAL ACTION PLAN (2024-2027)

Stricter regulations to reduce nitrate-related impacts.

BUT



STABLE LIMIT
The limit of 170 kg of nitrates per hectare is not being questioned.



STABLE PRICE
Although the significant expansion of NVZs, the price of 170€/ha is stable from the early 2000s.

ASSUMPTION

Price and limits for nitrates in NVZs will remain stable

Piglytics



Excess Nitrate Disposal in NVZs: costs and challenges

LAND OWNED:

50 Hectars

in Vulnerable
Zones

MAXIMUM CAPACITY:

**170 kg/ha of
nitrates**
for a
**total nitrates
disposable in
owned-lands
8500 kg**

EXCESS NITRATE DISPOSAL:

Total Nitrate
Production (2023):
20,724.8 kg
thus the

Excess Nitrates is
12,224.8 kg

TOTAL DISPOSAL COST

Disposal Cost per
Hectare
170 €/ha
(Laura Franchi)
Hectares Required
71.91ha

**Total Disposal Cost:
€12,224.8**



POLITECNICO
MILANO 1863

Piglytics



IS THERE A CHEAPER WAY?

INVESTMENT IN NITRO-DENITRO PLANT



POLITECNICO
MILANO 1863



Reduces the concentration of nitrates in manure by

at least 50%

Company would need much less lands for the dispose of manure

BUT
IS IT WORTH THE INVESTMENT?

Piglytics



Costs and Assumptions for Manure Management Plant



Evaluate the investment

We need to understand which costs the company would sustain in case it implements this investment.



Assumption: Negligible operational costs per year

- Self-sustaining system powered by solar energy.
- No need for advanced technology or constant monitoring.



Ideal case

The company doesn't need to recur to external lands in order to dispose of manure.



No more costs to sustain

To get rid of manure



POLITECNICO
MILANO 1863

Prof. Matteo Matteucci
Prof. Filippo Maria Renga



Costs and Assumptions for Manure Management Plant



THE PERFECT CASE DOESN'T EXIST

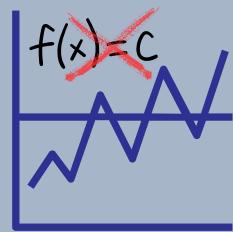
Distribution of manure varies significantly year by year

It is not possible to estimate the future production of nitrates based on the past.



Distribution of manure is not constant over time

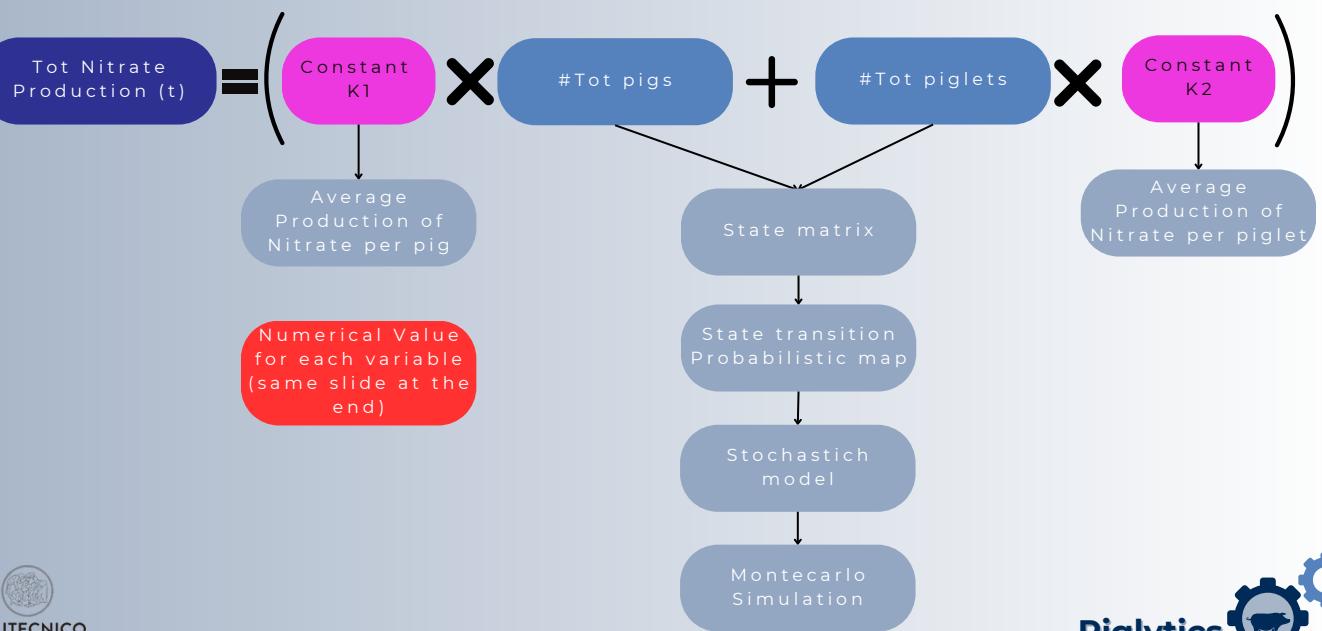
If production of manure is higher than the capacity of company-owned fields, the company still needs to rely on external lands.



OBJECTIVE: $f(x)$

Forecast the distribution of nitrates over time

The process



Data strategy: starting from a use case



Use Case:
evaluate the investment in "Nitro-Denitro" plant



DATA PROCUREMENT

Internal Data
(Structured)

Production

External Data
(Structured)

Weather

3Y of data (10/2022-10/2024)

Data strategy: starting from a use case

Merging and refining

- Merged external and internal Dataset
- Reconstructing missing information thorough data we have;
- Deleting unnecessary variables;
- Grouped from daily to weekly (ca. 156 weeks of data);
- Applied MA Filter to get rid of outliers



Use Data: ML and statistics

Business benefit: cost reduction in getting rid of Manure



Production Data: Event-based Datasets

Entrance

Entrance date of each sow, with age at entrance and birth date of the sow.

Birth

For each sow reports birth date of piglets, number of born alive and deads

Weaning

Reporting the Date, the kind of weaning and the number of weaned piglets for each sow.

Ecography

Date of coverage, date of ecography, result of ecography (pregnant yes/no)

Coverage

Coverage Date, number of attempts, date of birth or abortion.

Reform (exits)

Reason of exit, Date of reform, dead or sold.

Production Data: Event-based Datasets

Which are the
Hypothesis of the
models

Analysis of past Data: weekly population of pigs in last 3Y

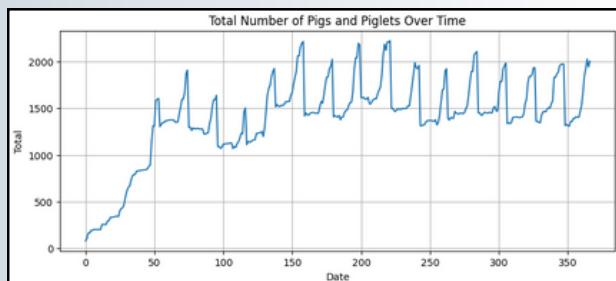
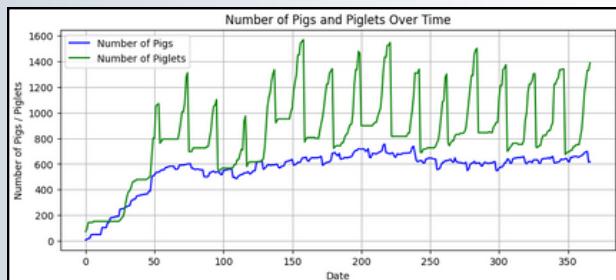
We first extract from the Dataset the weekly number of pigs and the weekly number of piglets.--> Presence of seasonality-->peaks every 25 weeks

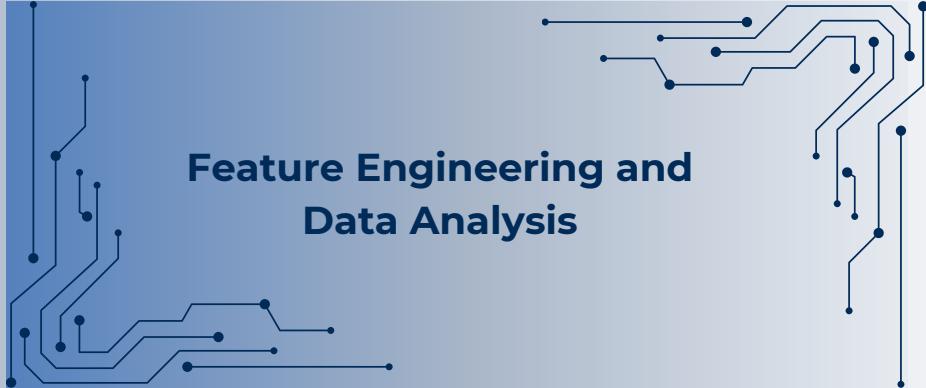
Sows: Number of initial sows + Numer of new entries - Number of exits. Number of initial sows are calculated as intialia exits which are not in datasets of entries and number of sows which are not in entries and not in exit (sows which are not indicated for errors)

Piglets extracted from dataset of born alive - sold piglets (weaning)

Then we sum the 2 functions together to get the total number of pigs and piglets.

As a first step we estimated the Nv "Born Alive" Parameter.
We then applied a Mean Filter, to decrease the variation of data.





Feature Engineering and Data Analysis



Prof. Matteo Matteucci
Prof. Filippo Maria Renga



**Fertility rate per every insemination and then do average
(function of figliata and weather)**

**Survival rate of Piglets per sow and then do average (average
number of alive newborn as a function of figliata and weather)**

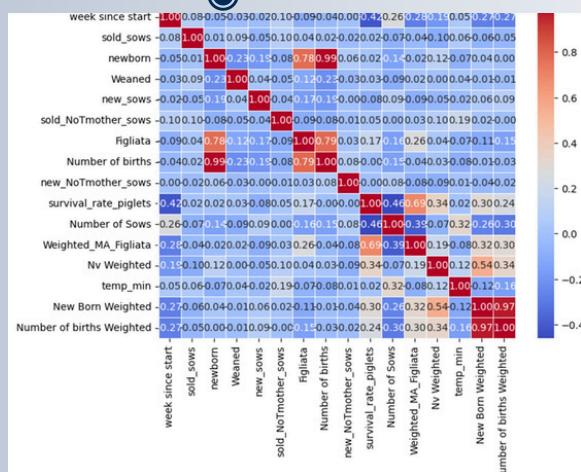
Average duration of pregnancy: 115 weeks

**Time between having birth and inseminated again: take the
distribution**

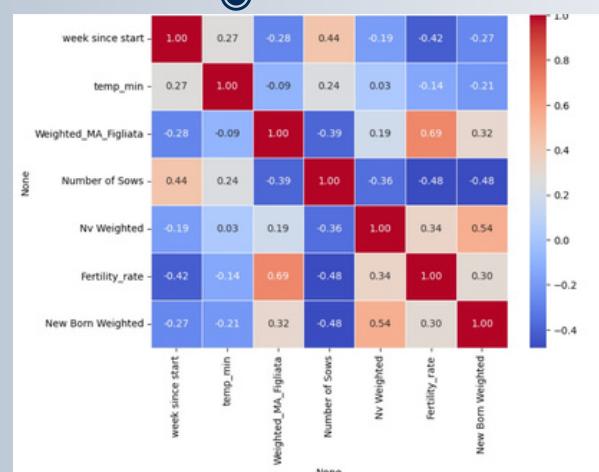
Prof. Matteo Matteucci
Prof. Filippo Maria Renga



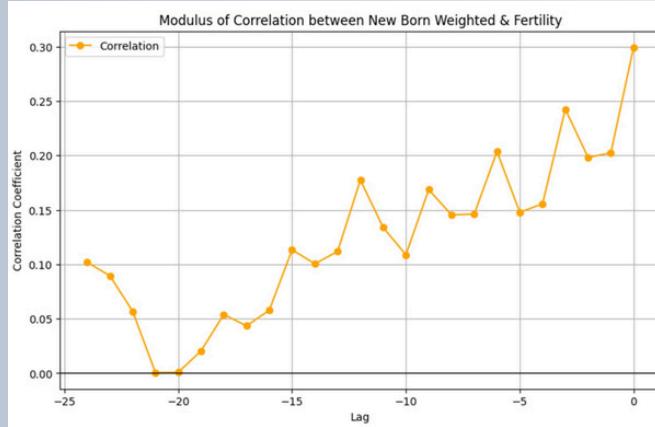
Correlation



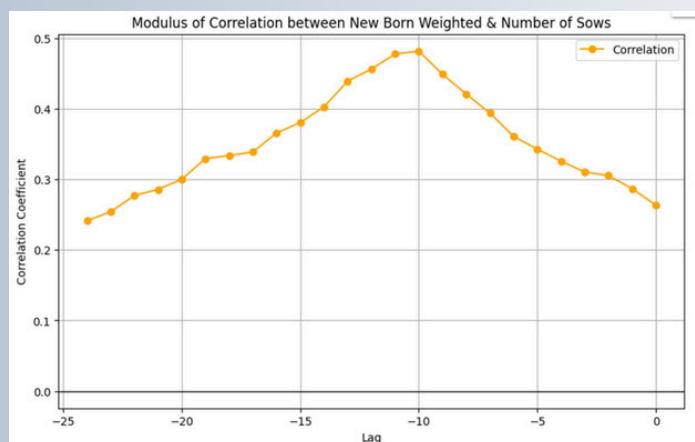
Correlation Matrix with shift



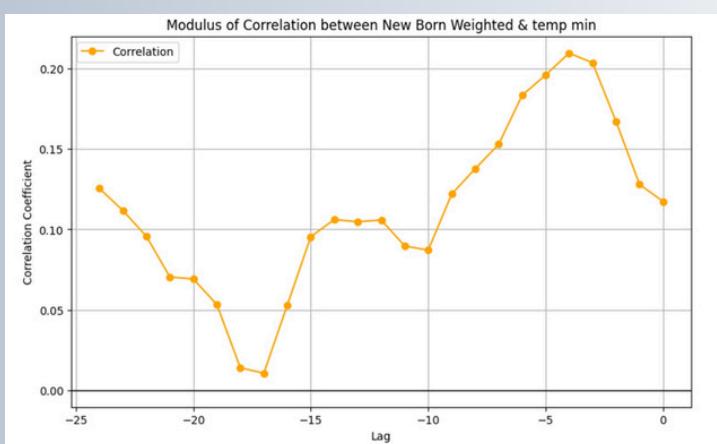
Correlation



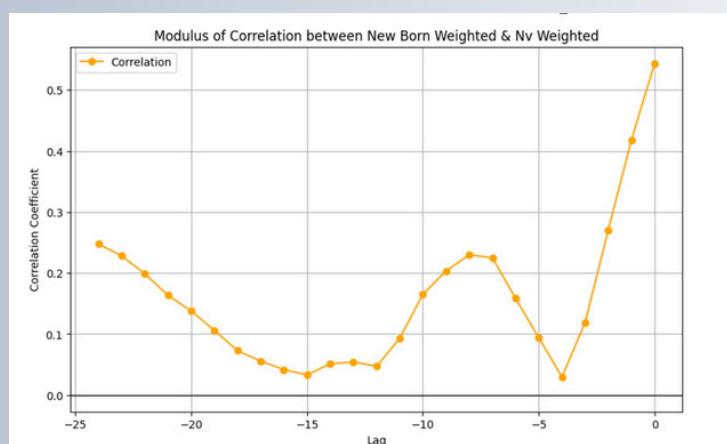
Correlation



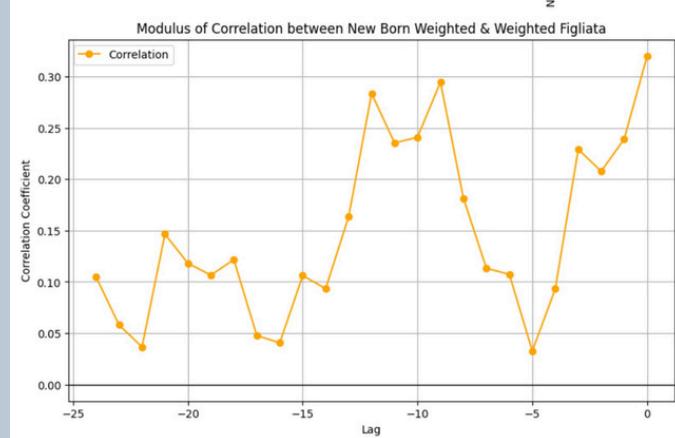
Correlation



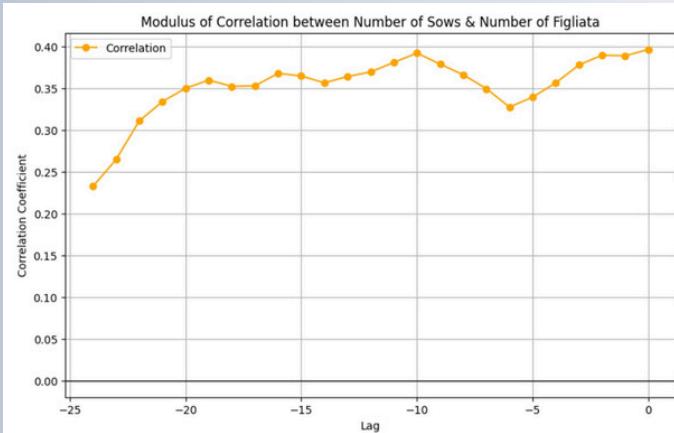
Correlation



Correlation



Correlation





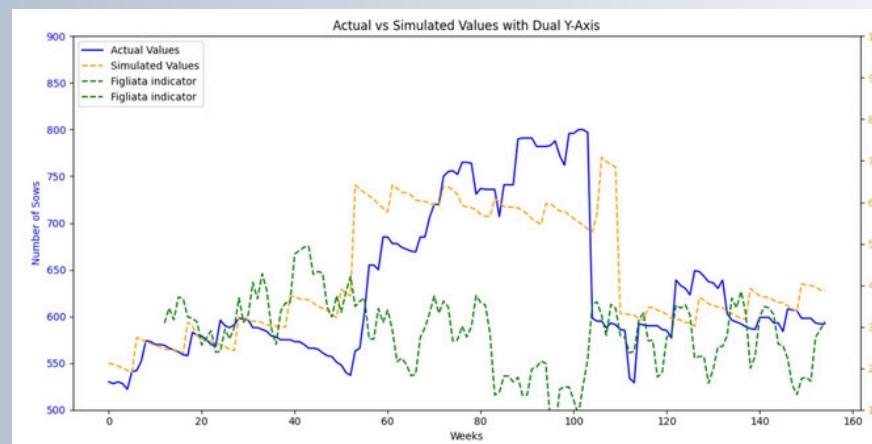
Forecasting

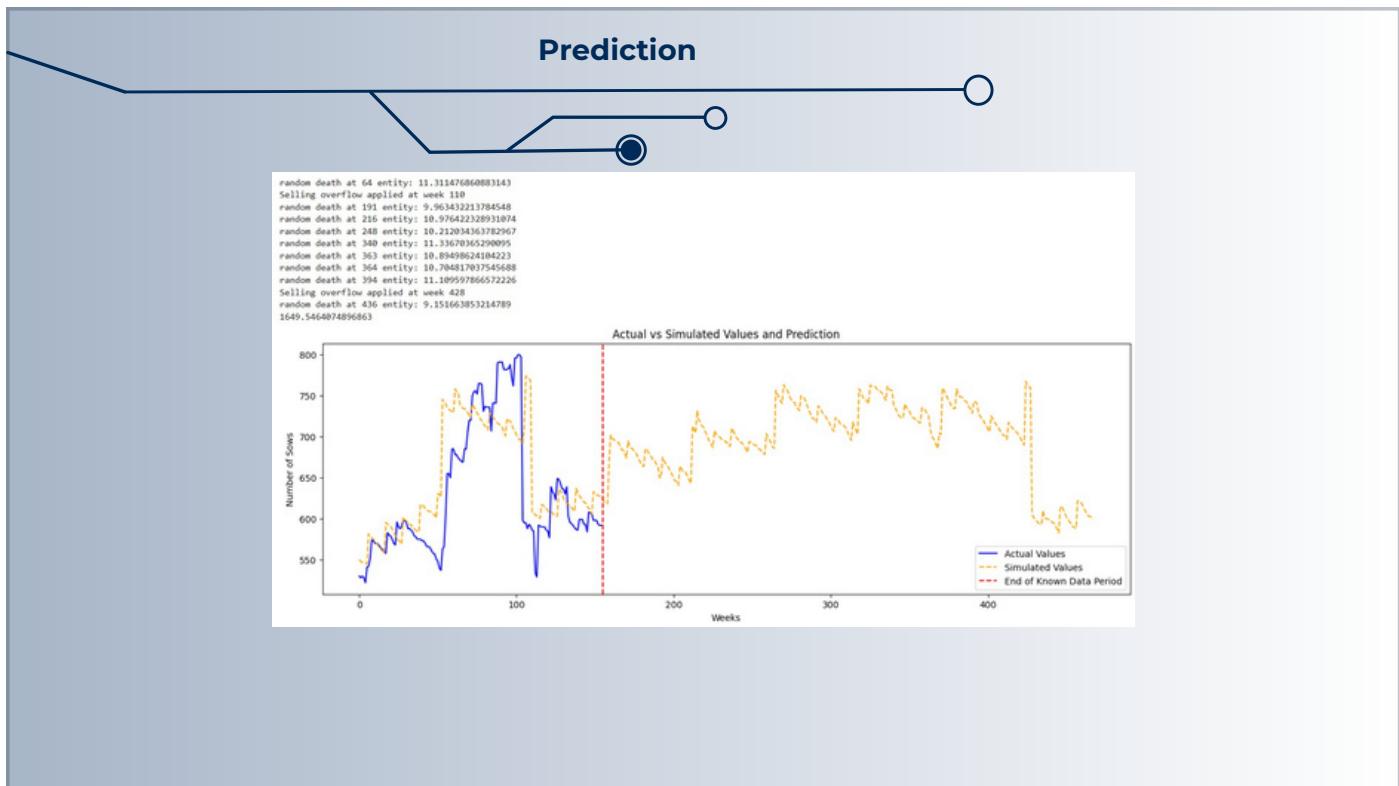


Prof. Matteo Matteucci
Prof. Filippo Maria Renga



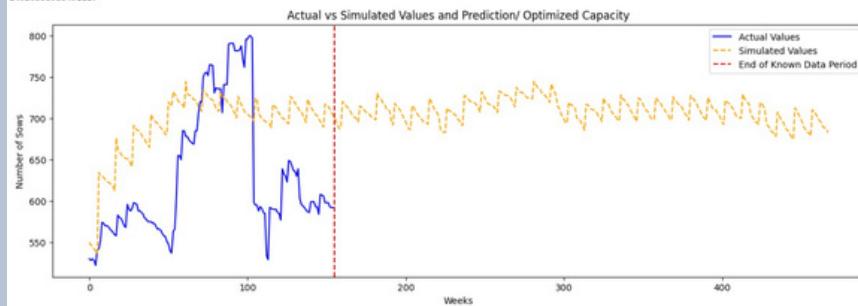
Actual vs Simulated Values

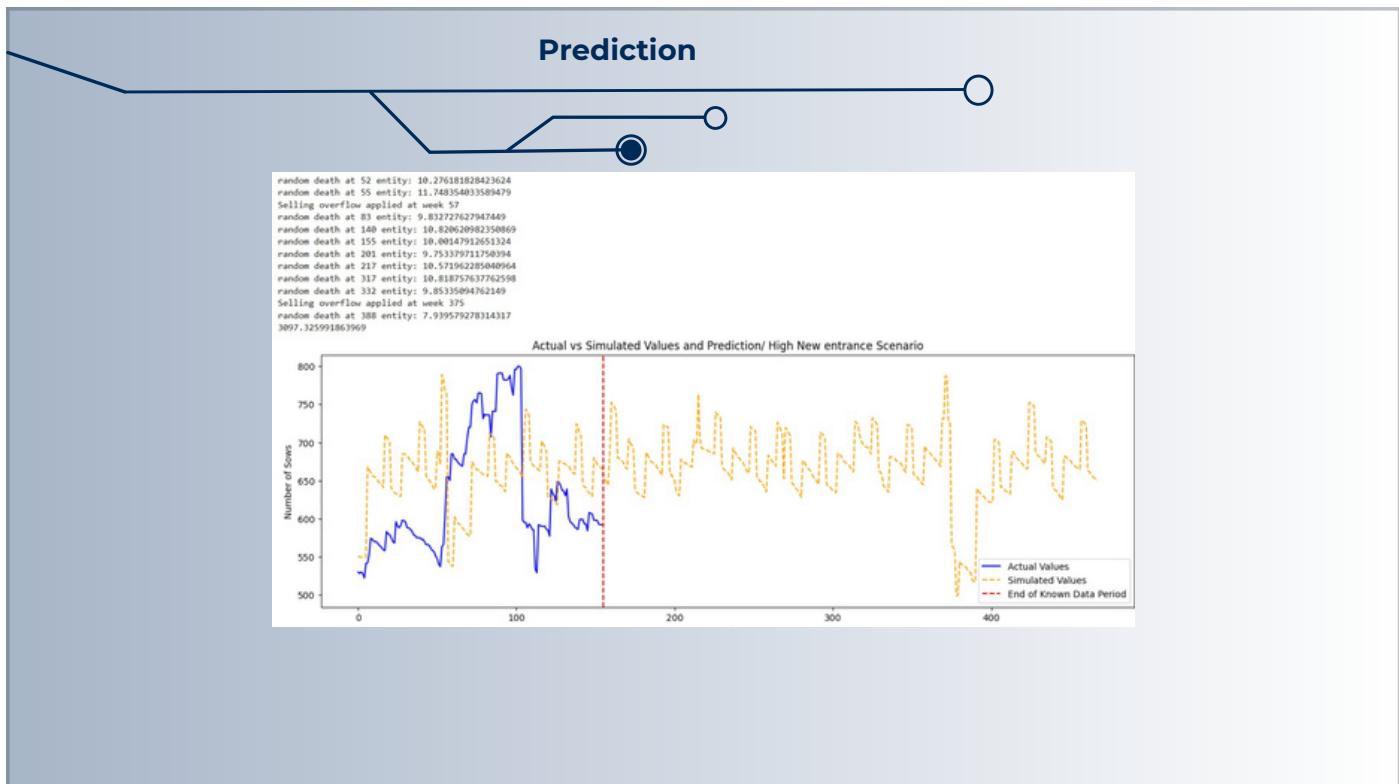




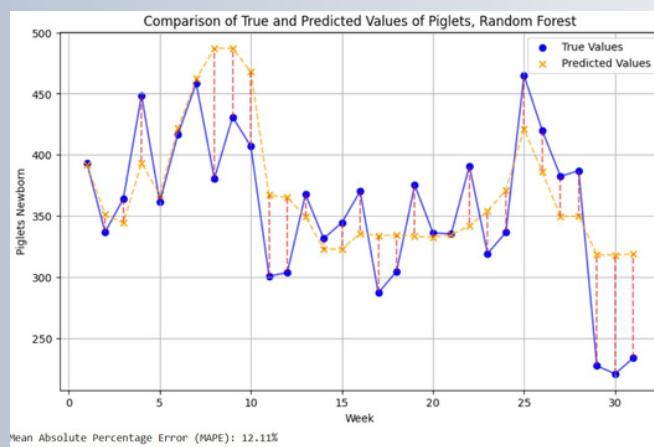
Prediction

```
random death at 18 entity: 10.14755035648444  
random death at 62 entity: 11.1705033439746  
random death at 106 entity: 10.46055576489135  
random death at 171 entity: 10.436032294545894  
random death at 191 entity: 10.626003654882258  
random death at 222 entity: 10.52026305717815  
random death at 296 entity: 10.9124946084372172  
random death at 388 entity: 10.69588565793545  
random death at 418 entity: 10.763816342814903  
random death at 428 entity: 10.65757142990187  
1492.35898349117
```

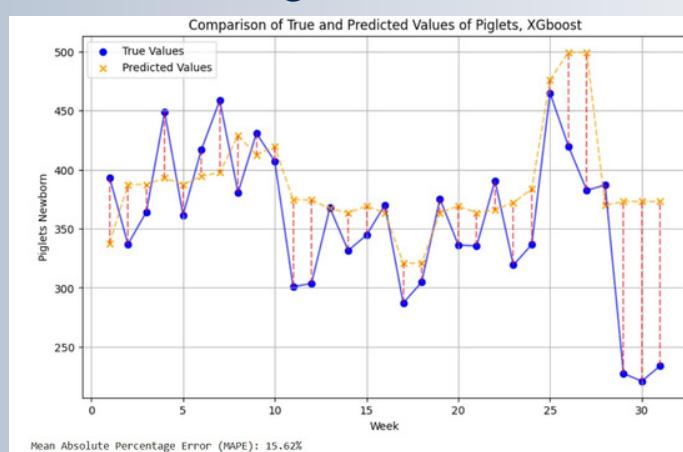




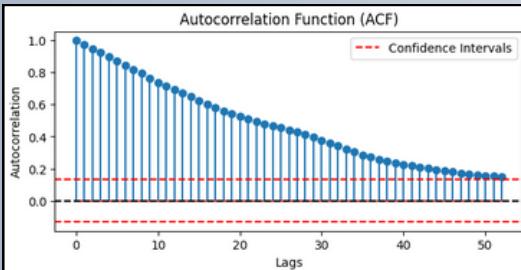
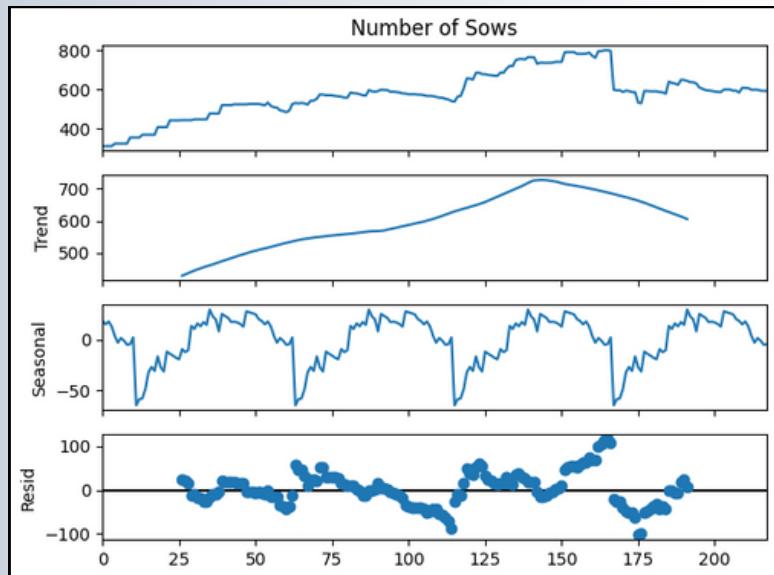
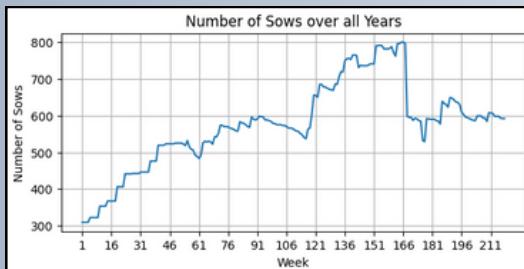
Prediction



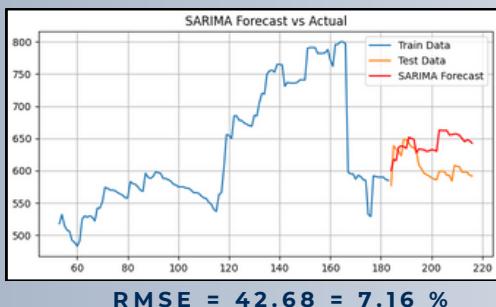
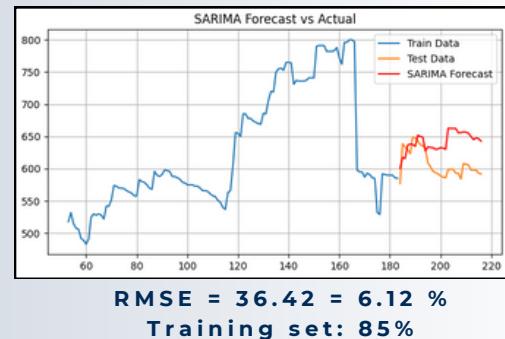
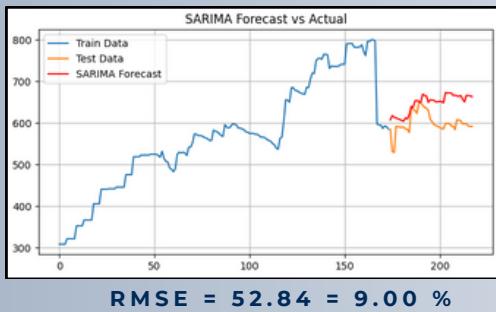
Prediction



Check for Seasonality

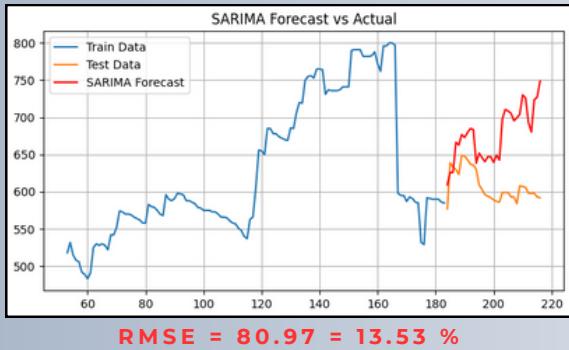


SARIMA Forecasting



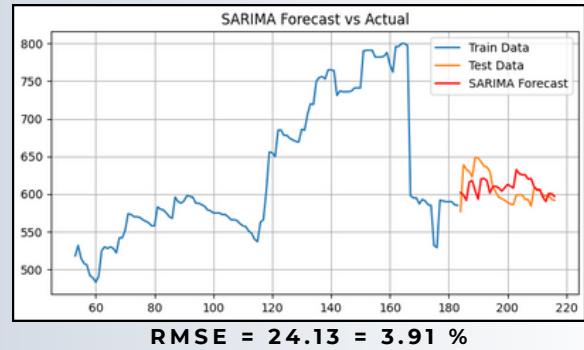
MeanAgeSows: 0.78
Figliata: 0.30
Number of births: 0.27
Weighted_MA_Figliata: 0.47

SARIMA Forecasting Augmented by Features



Augmented with features:

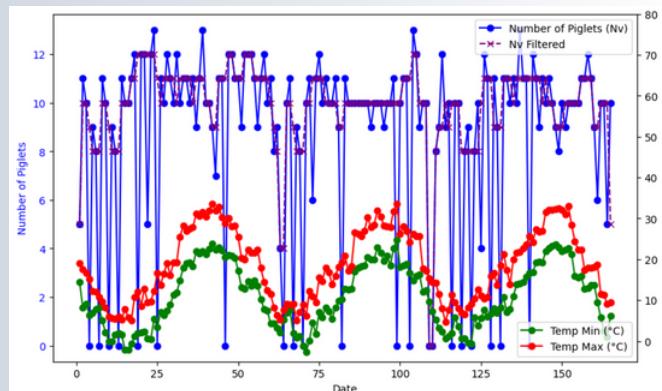
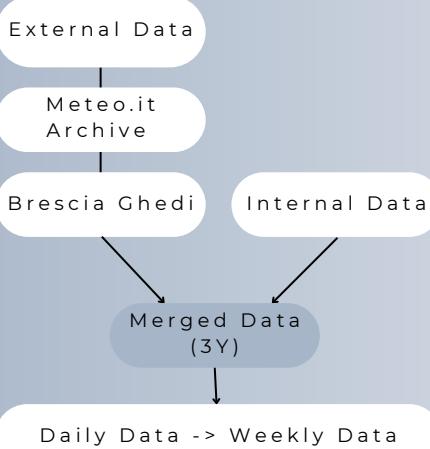
'MeanAgeSows',
'Figliata',
'Number of births'



Augmented with feature:

'MeanAgeSows'

Integration of production Dataset with external dataset: Temperature and Humidity Data

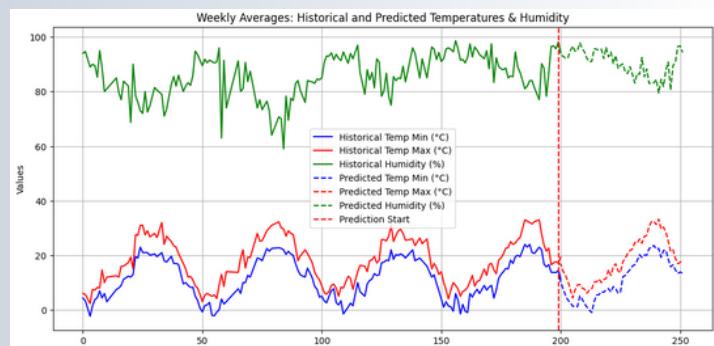


As a first step we estimated the Nv "Born Alive" Parameter.
We then applied a Median Filter, to decrease the variation of data.

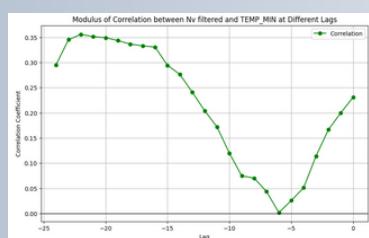
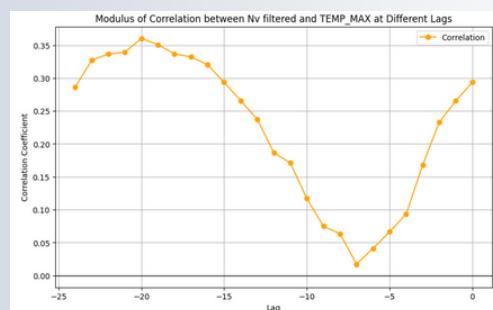
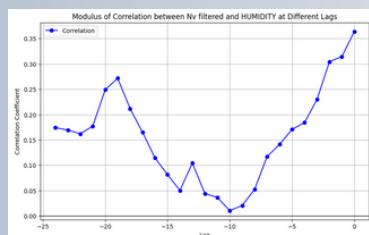
Prediction of next year Temperature and Humidity:

Random Forest

With Random Forest we forecasted Temperature and Humidity for next year.



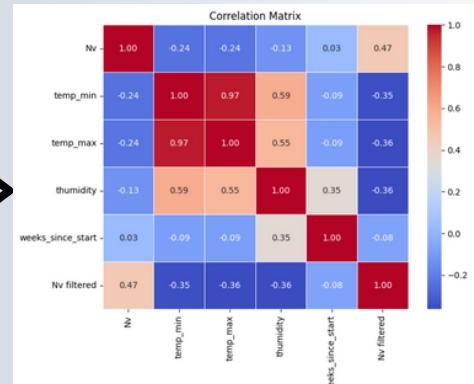
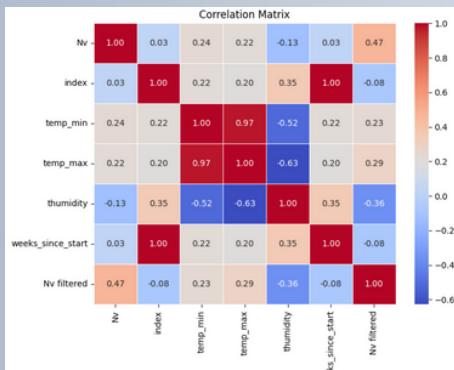
Correlation between Born Alive and several parameters with different lags



Heatmap with and without lags

Without Lags

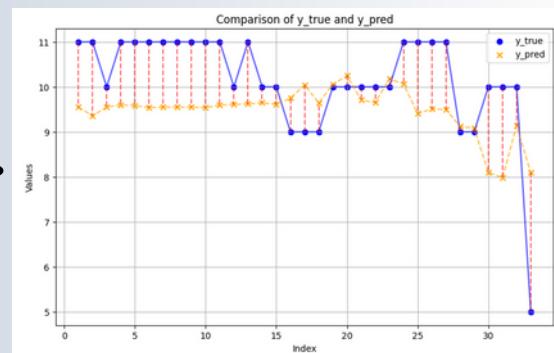
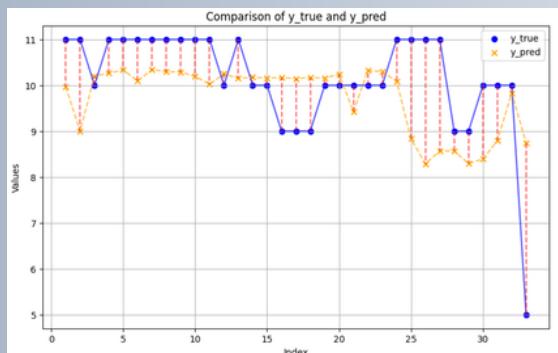
With Lags



Random Forest (test) with and without lags

Without Lags

With Lags



Born Alive Predicted (Next year)

With Lags



Prof. Matteo Matteucci
Prof. Filippo Maria Renga



MSE (With and Without lags)

Without Lags

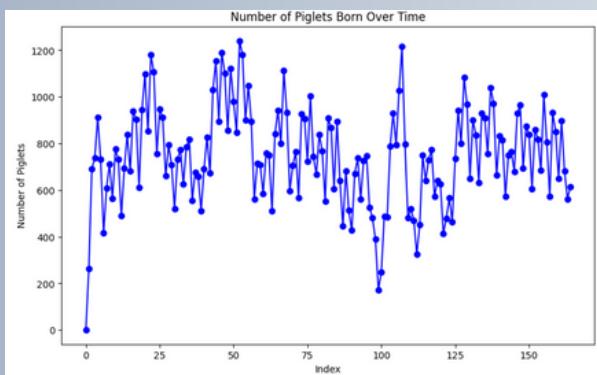
With Lags

Validation MSE: 4.68
Test MSE: 1.61
av. Error % = 12.51%

Validation MSE: 8.16
Test MSE: 1.58
av. Error % = 12.38%

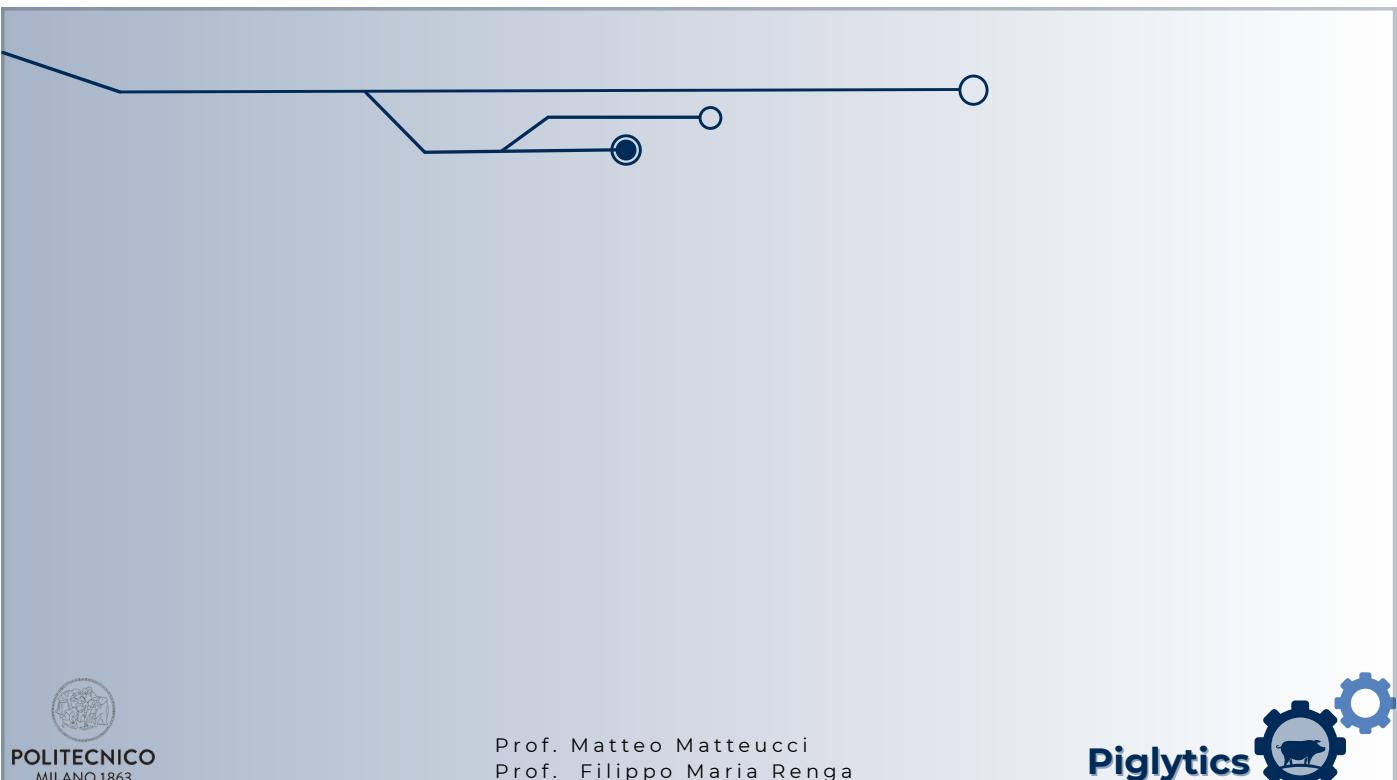
For the random forest we selected 80% of the data for the prediction while the remaining 20% is for test, selected randomly with a random pattern. we tested 20 random oatterns 1:20 and we averaging all the results

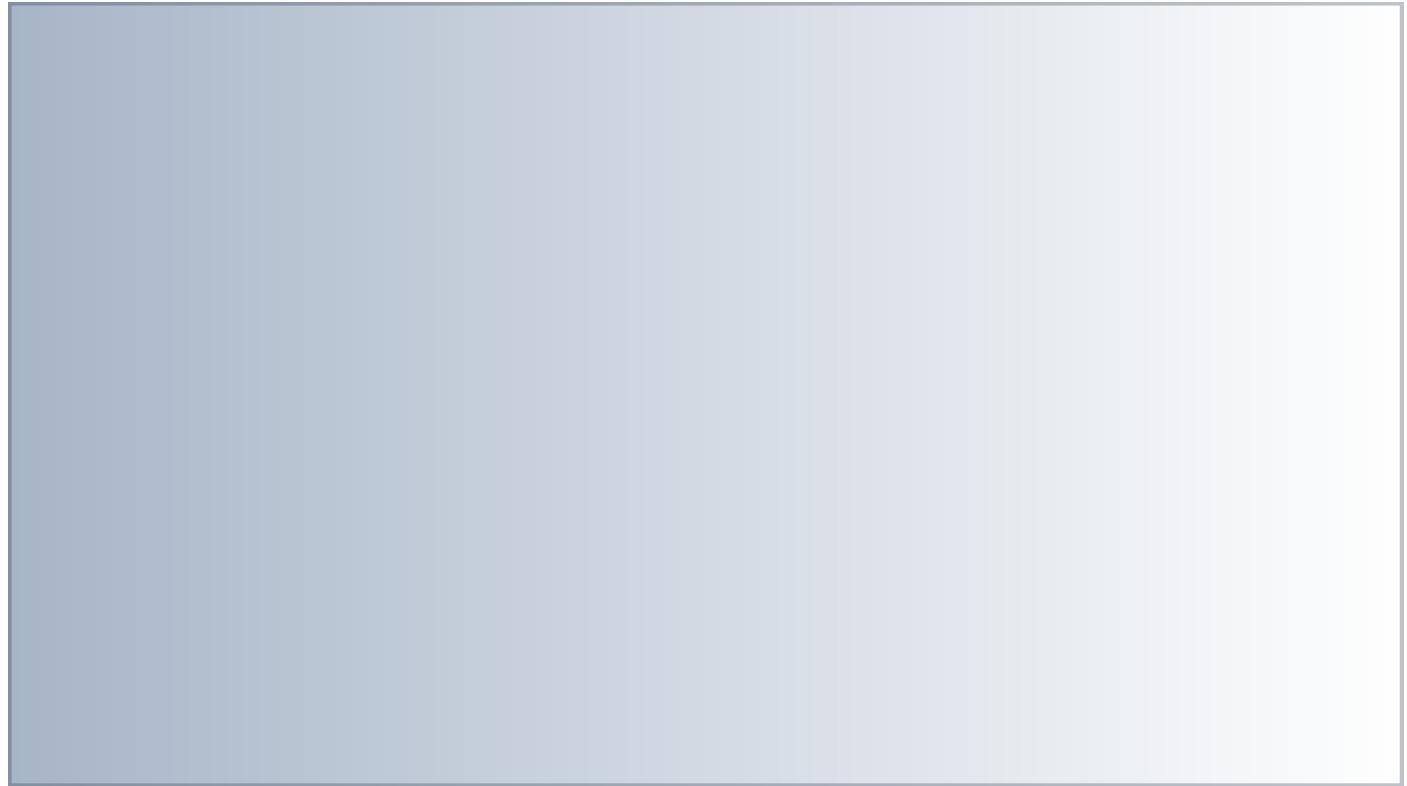
Number of piglets born



Correlation Matrix

	Number of Piglets	Nv	temp_min	temp_max	humidity	weeks_since_start	Nv filtered
Number of Piglets	1.00	0.35	0.03	0.04	-0.12	-0.04	0.33
Nv	0.35	1.00	0.24	0.22	-0.13	0.03	0.47
temp_min	0.03	0.24	1.00	0.97	-0.52	0.22	0.23
temp_max	-0.04	0.22	0.97	1.00	-0.63	0.20	0.29
humidity	-0.12	-0.13	-0.52	-0.63	1.00	0.35	-0.36
weeks_since_start	-0.04	0.03	0.22	0.20	0.35	1.00	-0.08
Nv filtered	0.33	0.47	0.23	0.29	-0.36	-0.08	1.00





Production Dataset

	Entrance Date	ID	Code	Birth Date	Age at entrance
0	2020-10-01	1063	1063	01/04/2020	183.0
1	2020-10-01	1066	1066	01/04/2020	183.0
2	2020-10-01	4635	4635	01/04/2020	183.0
3	2020-10-01	4691	4691	01/04/2020	183.0
4	2020-10-01	4829	4829	01/04/2020	183.0
...
1083	2024-10-02	24026	IT24026	28/03/2024	188.0
1084	2024-10-02	24051	IT24051	28/03/2024	188.0
1085	2024-10-02	24079	IT24079	28/03/2024	188.0
1086	2024-10-02	24231	IT24231	09/02/2024	236.0
1087	2024-10-02	24237	IT24237	28/03/2024	188.0

1088 rows x 5 columns

Entrance

	ID	Ecografy Date	Category	Birth Order	Result	Coverage Date	Nro
0	35127	2021-10-28	D1-ARA	4.0	Vuota	25/06/2021	1
1	0820	2021-12-16	E1-ARA	2.0	Vuota	18/11/2021	1
2	52938	2022-01-07	E3-ROS	6.0	Vuota	31/12/2021	1
3	0375F	2022-02-09	E4-BIA	4.0	Vuota	18/01/2022	1
4	2349	2022-02-11	E3-ROS	1.0	Vuota	04/01/2022	1
...
2716	7485	2024-11-07	G2-VER	1.0	Gravidia	02/10/2024	2
2717	7727	2024-11-07	G2-VER	1.0	Gravidia	11/10/2024	2
2718	D102	2024-11-07	G2-VER	4.0	Gravidia	07/10/2024	1
2719	6036	2024-11-07	G2-VER	2.0	Gravidia	10/10/2024	2
2720	020905	2024-11-07	G2-VER	10.0	Gravidia	08/10/2024	2

2721 rows x 7 columns

Birth

	Birth Date	ID	Birth Order	Category	Gestion	Duration	Nv	Mu	Ag	To	Dead	Remained	Alive
0	2021-10-09	49762	6.0	D1-ARA		110.0	7	0	13	2	13	0	5
1	2021-10-11	55998	4.0	D1-ARA		112.0	15	2	2	2	3	0	12
2	2021-10-11	0923	2.0	D1-ARA		112.0	18	1	1	4	2	0	13
3	2021-10-13	020905	2.0	D1-ARA		114.0	18	0	0	4	0	0	14
4	2021-10-13	020916	2.0	D1-ARA		114.0	12	0	7	0	3	0	16
...
2985	2024-11-27	D69	4.0	F6-GIAL		114.0	9	0	5	0	7	0	7
2986	2024-11-27	14830	1.0	F6-GIAL		116.0	7	0	0	0	0	0	7
2987	2024-11-27	11118	2.0	F6-GIAL		114.0	11	0	3	0	1	0	13
2988	2024-11-28	16596	5.0	F6-GIAL		115.0	14	0	0	0	0	0	14
2989	2024-11-28	12618	2.0	F6-GIAL		115.0	10	0	0	0	0	0	10

2990 rows x 12 columns

Weaning

	ID	Category	Birth Order	Weaning Date	Tipo	Nr	Ag	To	Morti	Svezzati	Age
0	56767	F3-ROS	11.0	2024-10-24	Svezzamento adottivo	2.0	9.0	0.0	0.0	9.0	21.0
1	020911	E7-BLU	7.0	2024-08-22	Svezzamento adottivo	2.0	14.0	0.0	0.0	14.0	31.0
2	6738	F3-ROS	1.0	2024-10-24	Svezzamento adottivo	2.0	13.0	0.0	0.0	13.0	29.0
3	2679	E6-GIA	2.0	2024-08-01	Svezzamento adottivo	2.0	1.0	1.0	0.0	0.0	24.0
4	2559	E7-BLU	6.0	2024-08-22	Svezzamento adottivo	2.0	14.0	0.0	0.0	14.0	37.0
...
2999	182	D1-ARA	5.0	2023-11-23	Svezzamento normale	1.0	2.0	0.0	4.0	9.0	36.0
3000	19	D1-ARA	5.0	2023-11-23	Svezzamento normale	1.0	1.0	0.0	2.0	11.0	38.0
3001	2015	D1-ARA	2.0	2023-11-23	Svezzamento normale	1.0	0.0	2.0	2.0	11.0	41.0
3002	14	D1-ARA	6.0	2023-11-23	Svezzamento normale	1.0	0.0	0.0	2.0	10.0	36.0
3003	11501	F6-GIAL	1.0	2024-11-27	Svezzamento normale	1.0	0.0	0.0	0.0	0.0	6.0

3004 rows x 11 columns

Production Dataset

Reforms

ID	Riforma Date	Riforma Reason	Category	Column6	Birth Order	Age at Riforma	Stadio\ngestaz.	Event	Exit Mode	IUSR
0	1 2024-10-21	18 - Scrofa morta	F4-BIA	6.0	NaN	47.0	Vuota	Svezzamento	Morti	0.0
1	3 2023-12-12	99 - Altre cause	C1-ARA	5.0	NaN	35.0	Sconosciuto	Copertura	Vendita (macello)	320.0
2	4 2024-03-27	12 - Caduta produttiva, età	E4-BIA	6.0	NaN	41.0	Vuota	Ecografia Vuota	Vendita (macello)	258.0
3	6 2024-02-06	99 - Altre cause	D3-RO	5.0	NaN	38.0	Vuota	Ecografia Vuota	Vendita (macello)	226.0
4	8 2024-02-06	99 - Altre cause	D4-BIA	6.0	NaN	37.0	Vuota	Ecografia Vuota	Vendita (macello)	164.0
...
774	Z177 2023-10-27	03 - Risultati parto scarsi	D3-RO	NaN	5	34.0	Sconosciuto	Copertura	Vendita (macello)	209.0
775	Z27 2023-01-21	18 - Scrofa morta	B1-ARA	NaN	4	25.0	Vuota	Svezzamento	Morti	1.0
776	Z5764 2023-11-05	08 - Malattie respiratorie	D1-ARA	NaN	7	42.0	Sconosciuto	Copertura	Vendita (macello)	133.0
777	ZD5 2023-10-26	02 - Incidente al parto	C7-BLU	NaN	5	34.0	Sconosciuto	Copertura	Vendita (macello)	129.0
778	ZT15 2022-02-01	19 - Prollasso uterino	D6-GIA	NaN	2	17.0	Gravida	Copertura	Morti	124.0

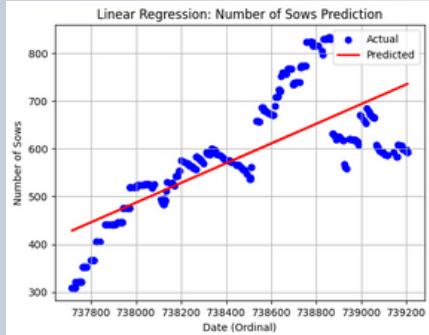
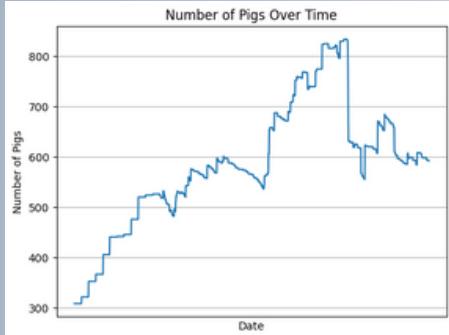
779 rows × 11 columns

The main challenge : The production dataset is event-based

The number of pigs and piglets on the farm must be computed according to the events

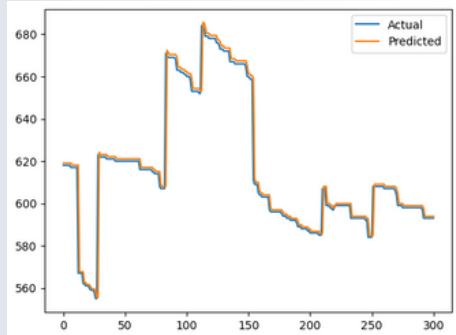
Prediction Models

Linear Regression



MSE = 7651.72

LSTM



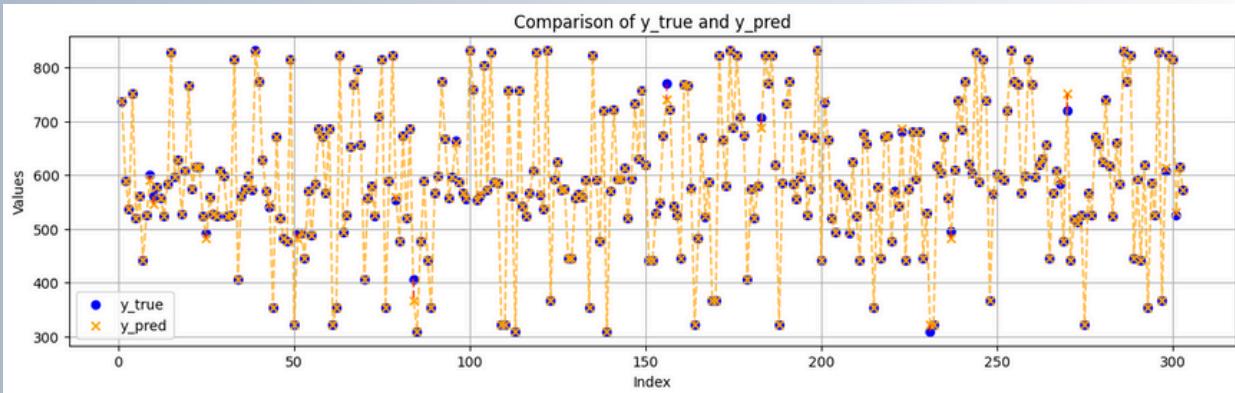
MSE = 7651.72

How about Polynomial?

4060.53

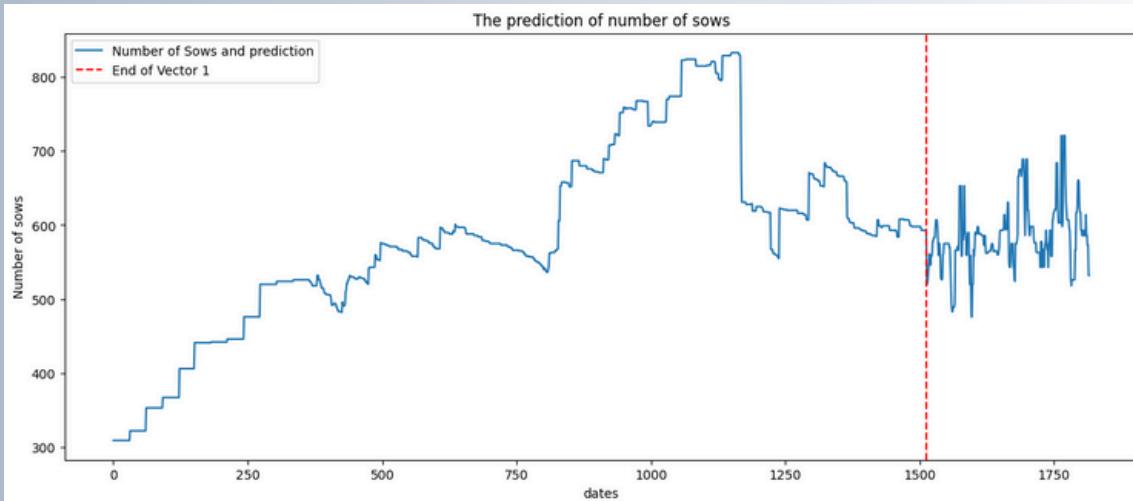
Prediction Models

Decision Tree Regression



MSE = 16.49

Doubts



circa 50 ettari a disposizione
costo al 2006 di asservimento secondo crpa
si aggirava a 170€ all'ettaro--> proiezione di
inflazione (circa 30% in più)
costo di distribuzione + manodopera = nelle
zone vulnerabili il costo medio era 250 euro
all'ettaro da rivalutare rispetto alla benzina
che prima era 0.72 e ora 1. Stima costo
totale.