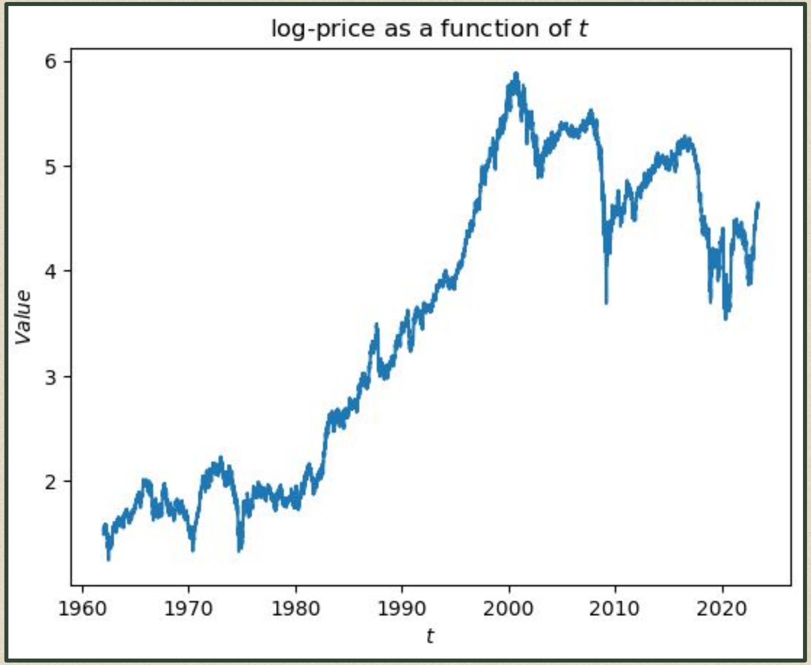
ABM Prediction

Marius Verdier, Pierre Tesio, Edward Lucyszyn
Paul Bastin, Mohamed Azahriou

What are the data?



Daily stock prices of General Electrics (from 1962 to 2023)

What is the model?

Majewski et al. type of model

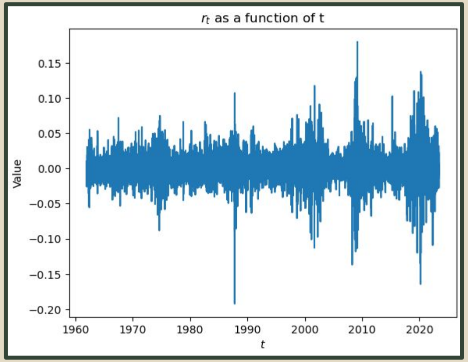
$$\begin{cases} p_{t+1} = p_t + \kappa_1 \left(v_t - p_t \right) + \beta \tanh\left(\gamma\, m_t \right) + \varepsilon_t \\ m_t = (1 - \alpha)\, m_{t-1} + \alpha \left(p_t - p_{t-1} \right) \\ v_t = (1 - \lambda)\, v_{t-1} + \lambda p_t \end{cases}$$

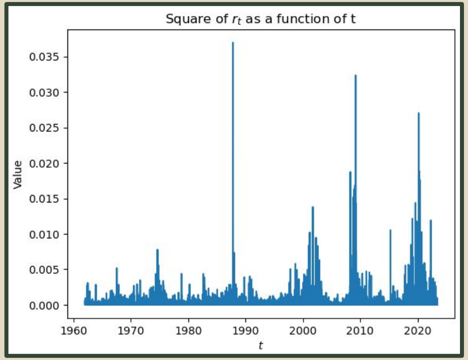
Where $\varepsilon_t \sim \mathcal{N}(0, \sigma(\varepsilon)^2)$

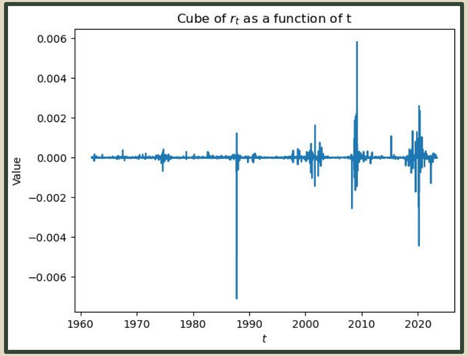
We need to estimate the parameters $\xi = (\kappa_1, \beta, \gamma, \alpha, \lambda, \sigma(\varepsilon))$

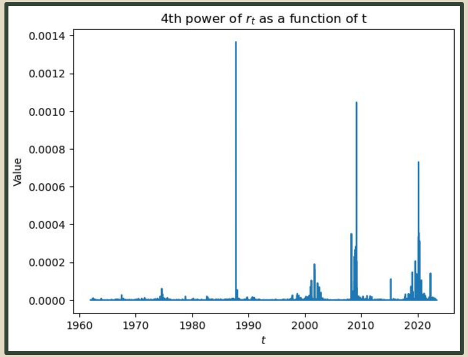
How to find the optimal parameters ?

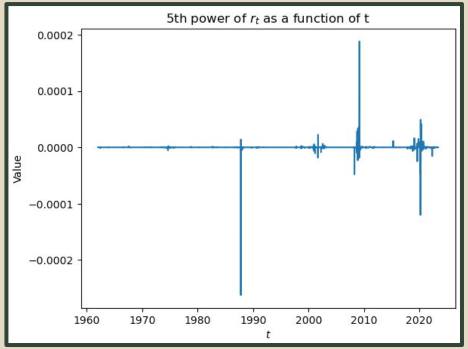
Consider $r, r^2, r^3, r^4, r^5, r^6$ the first six moments of data

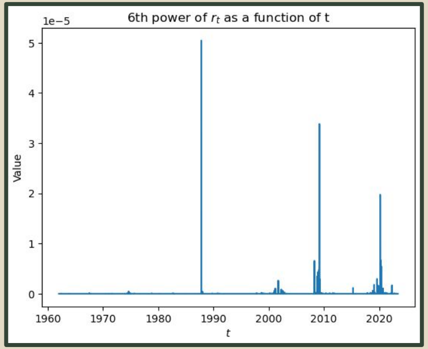












How to find the optimal parameters?

- Consider the function $f : (\kappa_1, \beta, \gamma, \alpha, \lambda, \sigma(\varepsilon)) \mapsto r_{ABM}, r^2_{ABM}, r^3_{ABM}, r^4_{ABM}, r^5_{ABM}, r^6_{ABM}$

  The function is stochastic !

- Consider the vector :

$$X = (\mathbb{E}(r) - \mathbb{E}(r_{ABM}), \mathbb{E}(r^2) - \mathbb{E}(r^2_{ABM}), \ldots, \mathbb{E}(r^6) - \mathbb{E}(r^6_{ABM}), ACF_r(1) - ACF_{r_{ABM}}(1), ACF_r(2) - ACF_{r_{ABM}}(2), ACF_r(3) - ACF_{r_{ABM}}(3))$$

- Minimize 100 times the function: $g : (\kappa_1, \beta, \gamma, \alpha, \lambda, \sigma(\varepsilon)) \mapsto \sum_{i=1}^{9} \sum_{j=1}^{9} X_j W^{(0)}_{i,j} X_i$

  with $W^{(0)} = I_9$

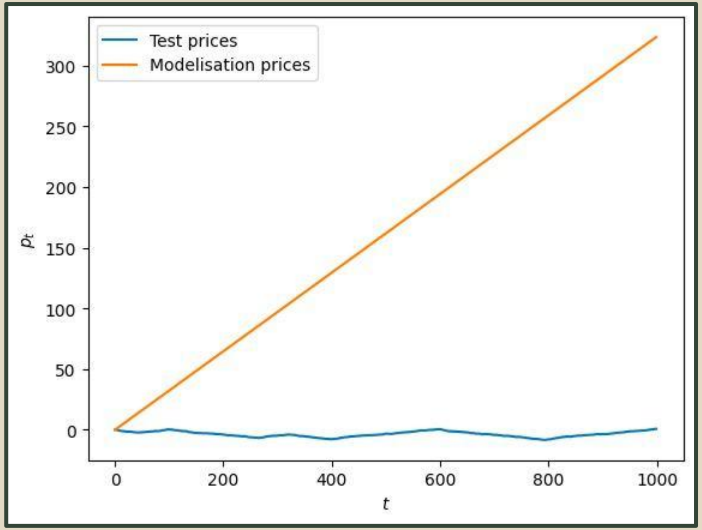
- To minimize, we use the function scipy.optimize.minimize with Powell method

- Calculate the mean of 100 minimisations : we call it $\xi^{(0)}$

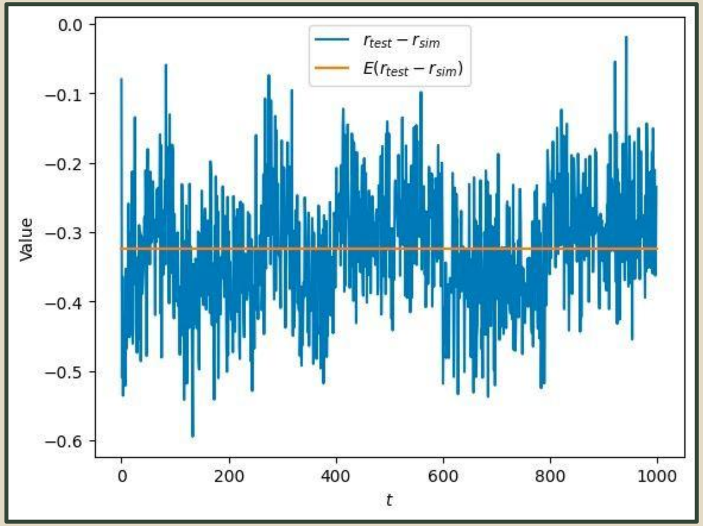
How to find the optimal parameters?

- For each minimized parameters array, we simulate 100 times the data, and calculate the 6 first powers of the returns. This creates an array of size 6x100x15000
- We calculate the mean on the third axis : we now have an array of size 6x100.
- We calculate the ACF, and take lags 1, 2, 3 : we get an array of size 3x100.
- We calculate the covariance matrix with the concatenation. The matrix is of size 9x9.
- We take the inverse of this matrix: $W^{(1)}$
- **With this new matrix, we calculate $\xi^{(1)}$ and we continue until the difference between two vectors is small enough.**

Results



Test prices
Modelisation prices
300
250
200
150
100
50
0
p_t
0
200
400
600
800
1000
t



$r_{test} - r_{sim}$
$E(r_{test} - r_{sim})$
0.0
−0.1
−0.2
−0.3
−0.4
−0.5
−0.6
Value
0
200
400
600
800
1000
t

Results

$$\xi_{test} = (\kappa = 0.08, \beta = 0.1, \gamma = 50, \alpha = 0.1, \lambda = 0.05, \sigma = 0.075)$$

$$\xi_{mode} = (\kappa = 0.12, \beta = 0.3, \gamma = 54, \alpha = 0.25, \lambda = 0.02, \sigma = 0.05)$$

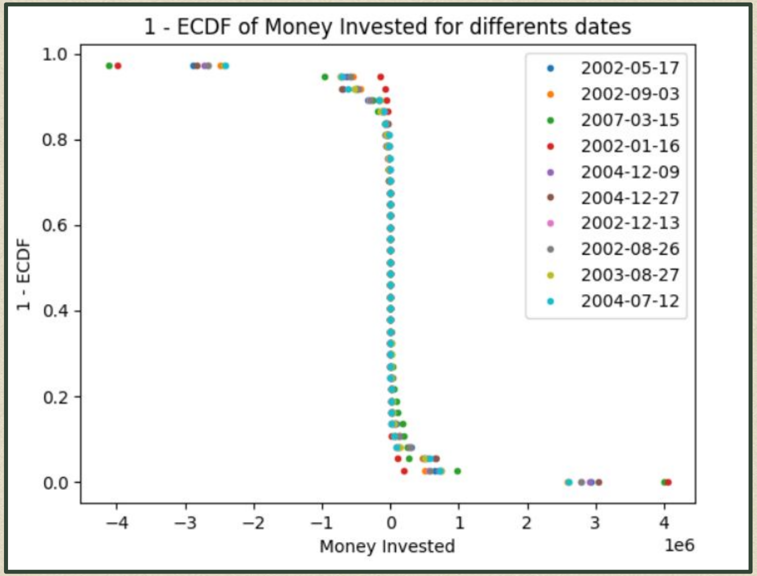
Because the calculus of m_t and v_t do not depend on noise, with the true data, we can predict the future values of m_t and v_t .

Investor behavior

We aim to classify investors in **two categories:**

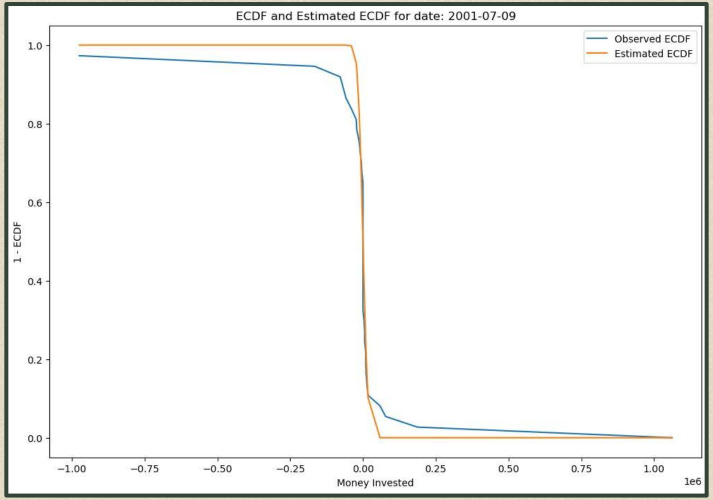
- trend followers
- contrarians

To do so, we plot $1 - ECDF(m)$ where each point represents an investor.



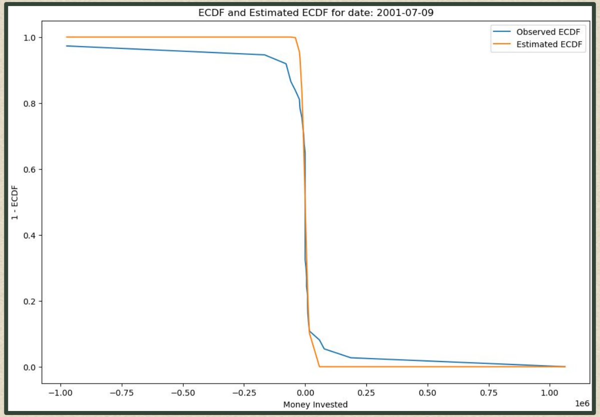
Investor behavior

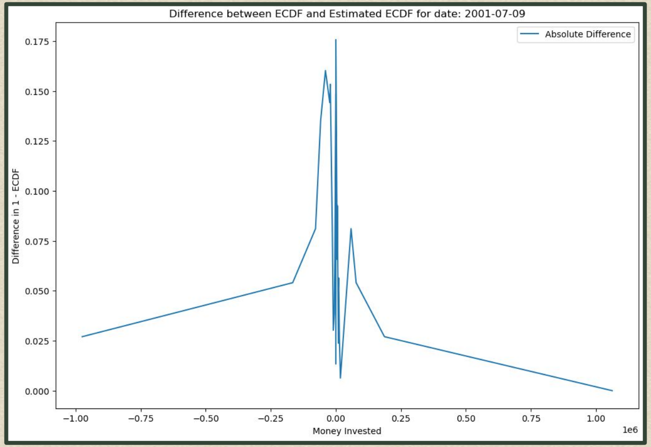
We then plot the 1-ECDF of the money invested:



Investor behavior

We then plot the 1-ECDF of the number of changes of invested volume:






```
message: Optimization terminated successfully.
success: True
 status: 0
    fun: 0.5452154714687333
      x: [ 1.380e+04]
    nit: 2
  direc: [[ 7.749e+02]]
   nfev: 45
```

We begin by try to fit our data with a **centered normal law** but approximation is not satisfying.

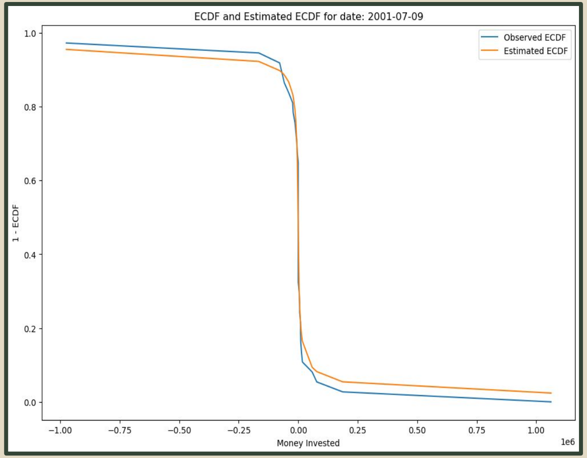
Investor behavior

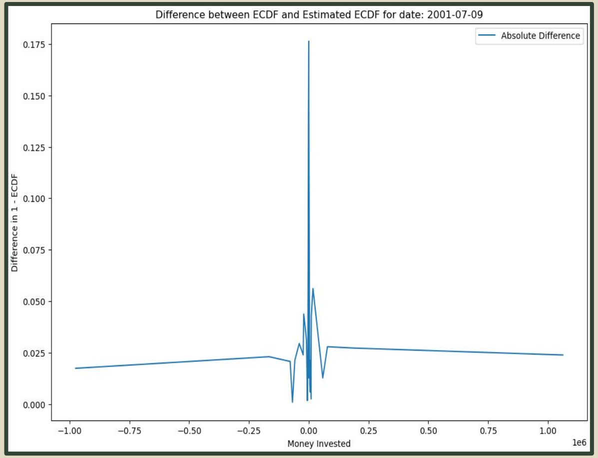
We will search our distribution in those which can be written under the form

$$\frac{\mathcal{T}(k) - \mu}{\sigma}$$

Investor behavior

And it's clearly better !

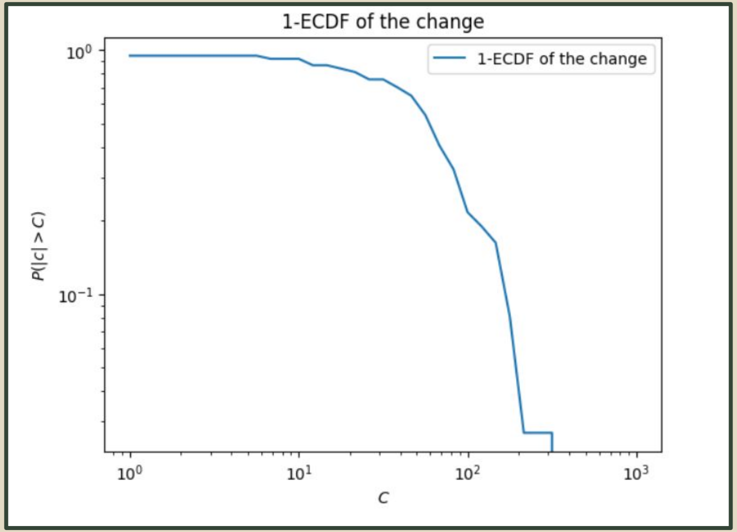





```
message: Optimization terminated successfully.
success: True
 status: 0
    fun: 0.40052282252885824
      x: [ 4.733e-01 -3.099e+02  4.496e+03]
    nit: 14
  direc: [[ 1.358e-03 -9.011e+00  8.230e+01]
          [ 0.000e+00  1.000e+00  0.000e+00]
          [-1.375e-03 -4.388e+00  3.319e+01]]
   nfev: 582
```

Investor behavior

We plot the 1-ECDF of the number of change



Investor behaviour

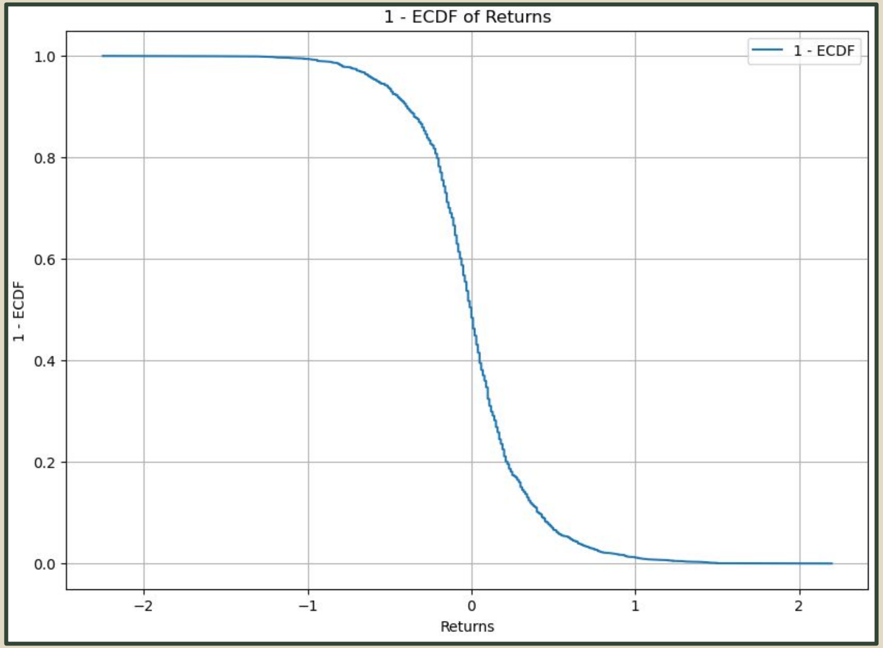
The CDF can be written under the form :

$$\alpha - e^{(x-\beta)}$$

or heavy-tailed with y = 4.35

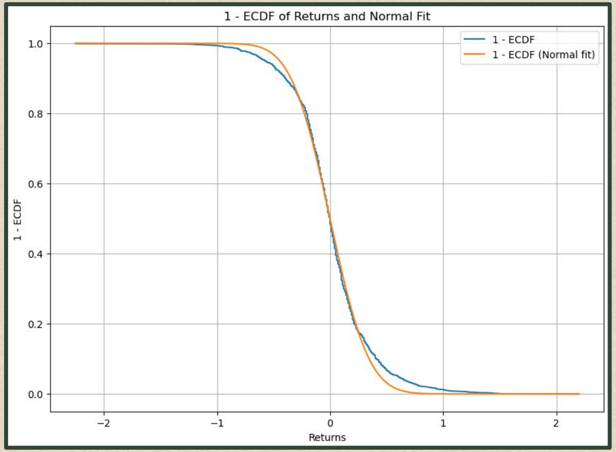
Investor behaviour

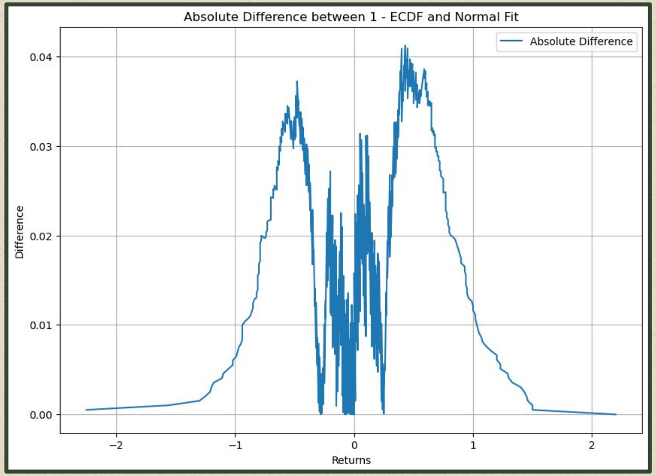
Bonus : we plot the 1-ECDF of the returns



Investor behaviour

We found a normal law !






```
 message: Optimization terminated successfully.
 success: True
  status: 0
     fun: 0.8719760579593364
       x: [ 2.698e-01]
     nit: 2
   direc: [[ 1.000e+00]]
    nfev: 28
[0.26983099]
```

Investor behavior

We then divide the investors between
mean-reverters, trend followers and neutral.

We calculate correlation between change of inventory and the same-day return for each investor.

If it is above 0, they are **trend followers**,
Below 0 they are **mean-reverters**,
Close to 0 we classify them as **neutral**.

```
investor_id
744      mean-reverting
2955            neutral
2963     mean-reverting
4066     mean-reverting
6018     mean-reverting
6198     mean-reverting
6447     mean-reverting
7319     mean-reverting
7936            neutral
9176     mean-reverting
9185     mean-reverting
9190     mean-reverting
9191     mean-reverting
10510    mean-reverting
10649           neutral
10710           neutral
11513           neutral
11872           neutral
14561           neutral
14774           neutral
15724    mean-reverting
15740    mean-reverting
15814           neutral
15936    mean-reverting
16433    mean-reverting
17837    mean-reverting
18159    mean-reverting
18169    mean-reverting
18178           neutral
18195           neutral
18228           neutral
18232           neutral
18317           neutral
18321    mean-reverting
18322           neutral
18333           neutral
18397    mean-reverting
dtype: object
mean-reverting    21
neutral           16
```

Investor behavior

We compute the probability of an investor changing from a class to another.

second_period	mean-reverting	neutral	trend-following
first_period			
mean-reverting	0.739130	0.217391	0.043478
neutral	0.571429	0.357143	0.071429

We finally determine which investors try to predict future price moves by **calculating the correlation between the change of inventory and future returns**.

neutral	26
mean-reverting	8
trend-following	3

Prediction with an ABM

- Take a model
- Compute the parameters of Majewski *et al.* that fits the most the model
- Compute the volatility and m_t with the last prices
- Predict the next value and repeat the process

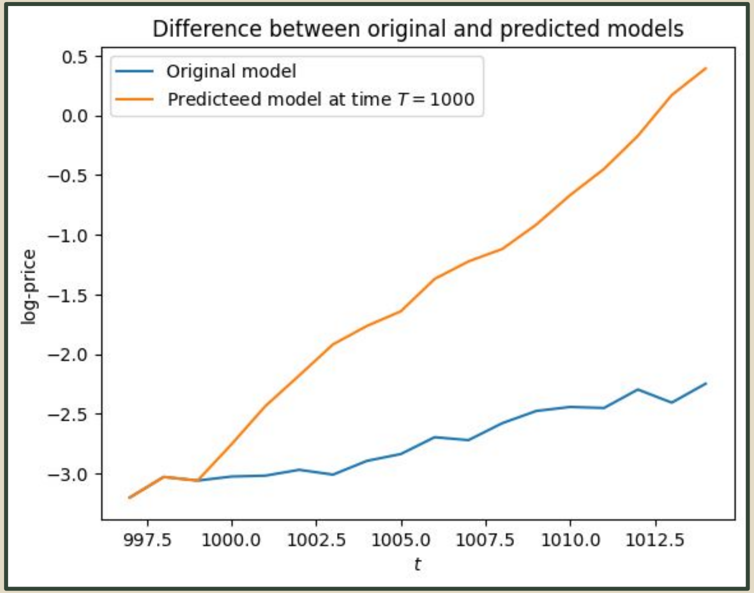
Also for a non stationary signal :

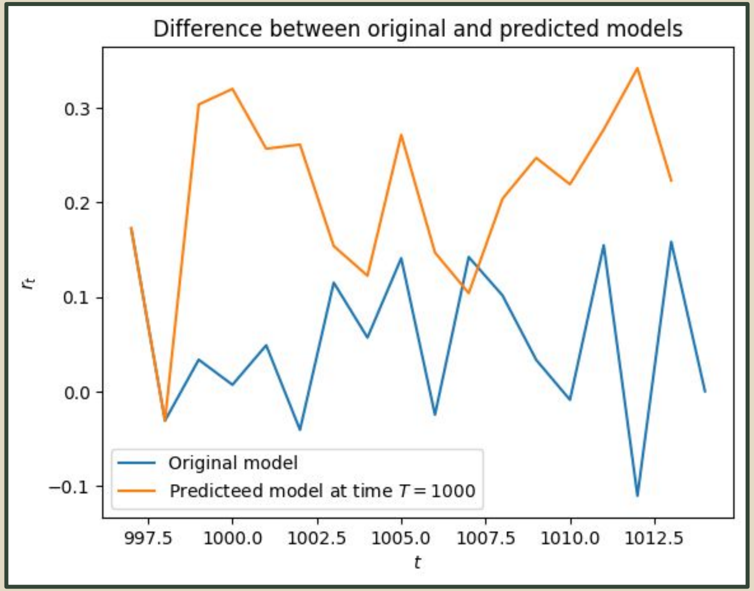
- Predict the n next values and repeat the process

Prediction with an ABM

Results for a model from *Majewski et al* type of model with the 1000 last values :

Parameters given : $\xi_{\text{test}} = (0.08, 0.1, 50, 0.1, 0.05, 0.075)$





Prediction with an ABM

Results :

	True for real model	False for real model
True for predicted model	11	4
False for predicted model	0	0

Precision = 0.7333333333333333

Hit ratio = 1.0

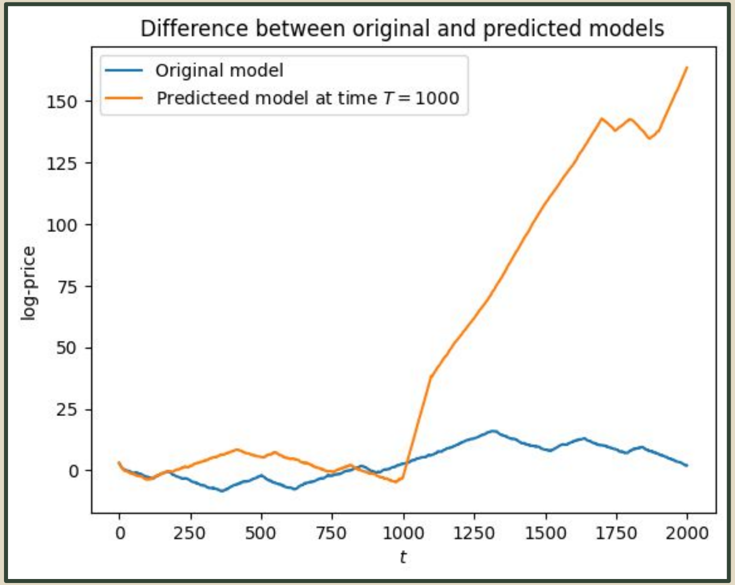
- Good precision and hit ratio
- Not enough prices estimated to compute a safe precision
- Took 9 minutes to compute 15 future values with only 4 iterations max for the GMM method

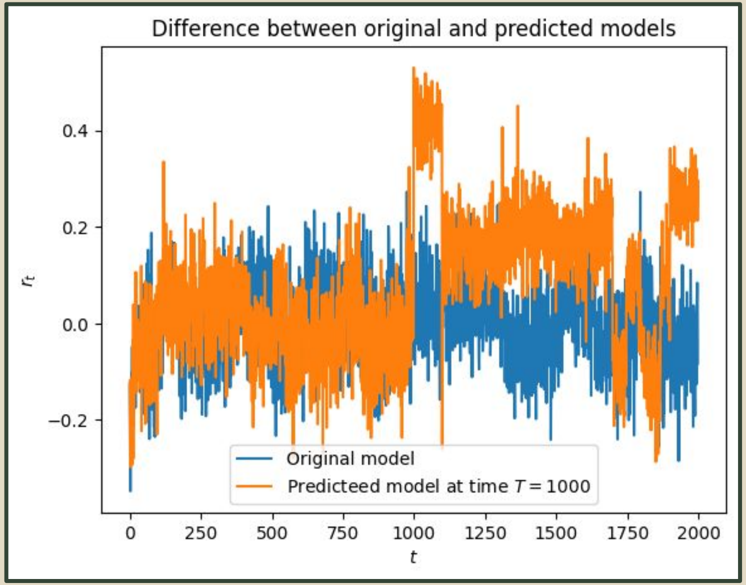
Can you we predict more parameters and do it faster ?

Prediction with an ABM

Results for a model from *Majewski et al* type of model with the 50 last values every 100 values :

Parameters given : $\xi_{\text{test}} = (0.08, 0.1, 50, 0.1, 0.05, 0.075)$





Prediction with an ABM

Results :

```
                              True for real model  False for real model
True for predicted model                      449                   428
False for predicted model                      53                    69
Precision = 0.5119726339794755
Hit ratio = 0.8667953667953668
```

- Took only 2 minutes to predict 1000 prices
- Less good precision but good balance between time and precision

Prediction with an ABM

Results for a model with stocks prices of General Electrics estimated with the last 7000 values :

	True for real model	False for real model
True for predicted model	3	2
False for predicted model	0	0

```
Precision = 0.6
Hit ratio = 1.0
```

- Took 18 minutes to predict only 5 prices
- Precision of 60% → good

Prediction with an ABM

What could have been done to improve the prediction ?

- More estimations of the precision with the prices of the stock exchange to have an idea of the precision for many values
- Take more time to compute with more iterations for the GMM method to see if it improves the precision
- Test on more data of different type (with crypto for example)
- Test if we can earn money with this method 💲

In conclusion

How easy is it to predict the market prices ?

- Financial markets are very complex : more parameters to catch market dynamic ?
- A huge amount of time and computational resources, as well as sophisticated models
- It is easier to understand better market movements, for example by studying investors behaviors

Thanks for your attention !
Tasks distribution:

- Pierre : part 1 & end of part 2
- Paul : part 1 & statistics analysis
- Mohamed : part 1 & beginning of part 2 & beginning of part 3
- Marius : part 1 & end of part 2 & beginning of part 3
- Edward : part 1 & beginning of part 2 & part 3