I. GO

# International PhD Excellence School **“Italo Gorini”**

Ph.D. School ITALO GORINI
ON LINE - September 04-09, 2020

Università degli Studi **Mediterranea** di Reggio Calabria

Università degli Studi di Napoli Federico II

associazione italiana gmee gruppo misure elettriche ed elettroniche

MMT Gruppo Nazionale

# What is measurement uncertainty? The case of COVID-19 outbreak forecasting

Nicola Giaquinto
nicola.giaquinto@poliba.it

Politecnico di Bari

# Agenda

# Brief introduction

Forecasting COVID-19: how can we approach the matter?

# COVID-19 forecasting: a HUGE amount of work has been already done

Source: https://www.cdc.gov/coronavirus/2019-ncov/covid-data/forecasting-us.html
on August 31th, 2020

- This week CDC received forecasts of national COVID-19 deaths over the next 4 weeks from 35 modeling groups. Of the 35 groups, 33 provided forecasts for both new and total deaths and two provided forecasts for total deaths only.

# Forecast assumptions and forecast models

These modeling groups **make assumptions about how levels of social distancing will change in the future**:

- Columbia University (Model: Columbia)
- Google and Harvard School of Public Health (Model: Google-HSPH)
- Georgia Institute of Technology, Center for Health and Humanitarian Systems (Model: GT-CHHS)
- Institute of Health Metrics and Evaluation (Model: IHME)
- John Burant (Model: JCB)
- Johns Hopkins University, Infectious Disease Dynamics Lab (Model: JHU_IDD)
- Notre Dame University (Model: NotreDame-FRED)
- Predictive Science Inc. (Model: PSI)
- University of California, Los Angeles (Model: UCLA)
- Youyang Gu (COVID-Projections) (Model: YYG)

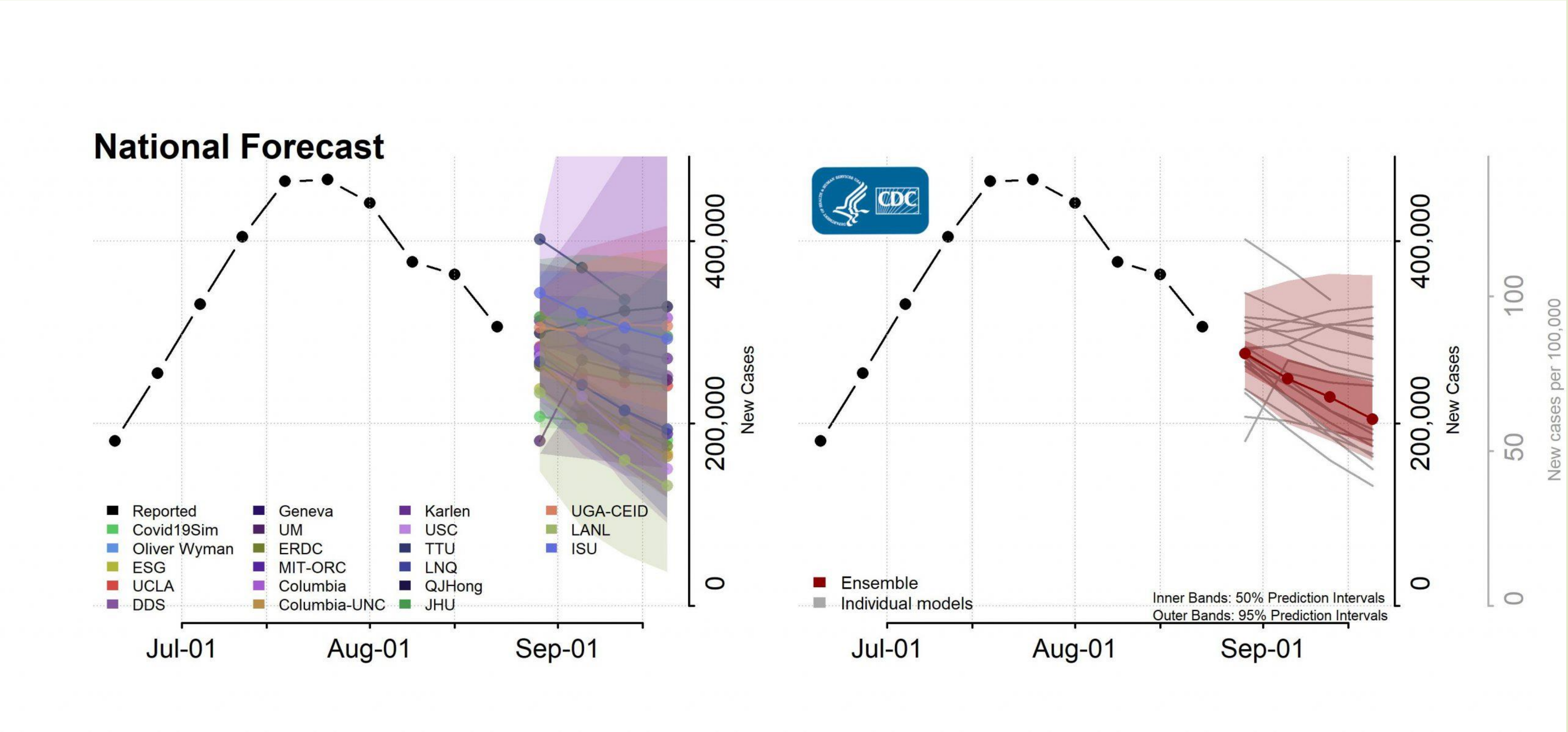
These modeling groups **assume that existing social distancing measures will continue through the projected four-week time period**:

- Carnegie Mellon Delphi Group (Model: CMU)
- Columbia University and University of North Carolina (Model: Columbia-UNC)
- Covid-19 Simulator Consortium (Model: Covid19Sim)
- … and other 26 research groups and models

From https://www.cdc.gov/coronavirus/2019-ncov/covid-data/forecasting-us.html
on August 31th, 2020

(Death forecasts)

# Example of forecasts provided to the general public




From https://www.cdc.gov/coronavirus/2019-ncov/cases-updates/forecasts-cases.html
(Cases forecasts, 31st August 2020)

# What we will NOT discuss

- State of the art
  - unpractical and out of the scope of this talk
- Quality of data (a big problem)
  - number of “positives”: is it reliable? Does it reflect the number of “infected”?
  - is there a count of the new hospitalized each day?
  - how many deaths “for” COVID, how many “with” COVID?
  - etc.
- Which data are meaningful?
  - before summer, **hospitalized seemed more important than positives**
  - after summer, **we are more interested in positives, even without symptoms**
  - number of positives? number of deaths? positives/swabs ratio?
- Other issues
  - predicting the effect of social distancing measures
  - long-term forecasting coming from opinions (more or less authoritative)

# What we will discuss

- Forecasting future data **in general**
  - positives, hospitalized, deaths, etc.
- Forecasting **by time series analysis**
  - from simple fitting to dynamic models
- **Uncertainty analysis**
  - from practice, to theoretical consideration, back to practice
- Same MATLAB code
  - we will try to be a little "practical" and "operational"
- The specific case study is the **cumulative number of positives in Italy** (in the following, some screenshots are in Italian)

# An early forecast

«There will be 14 000 positives at the end of the next week»

# The whistle blown by prof. De Nicolao, March 02, 2020

il Giornale.it

Coronavirus, l'algoritmo che allarma: "Entro domenica 14mila contagi"

msn notizie

Coronavirus, la previsione degli algoritmi: 14 mila contagi entro l'8 marzo

yahoo! notizie

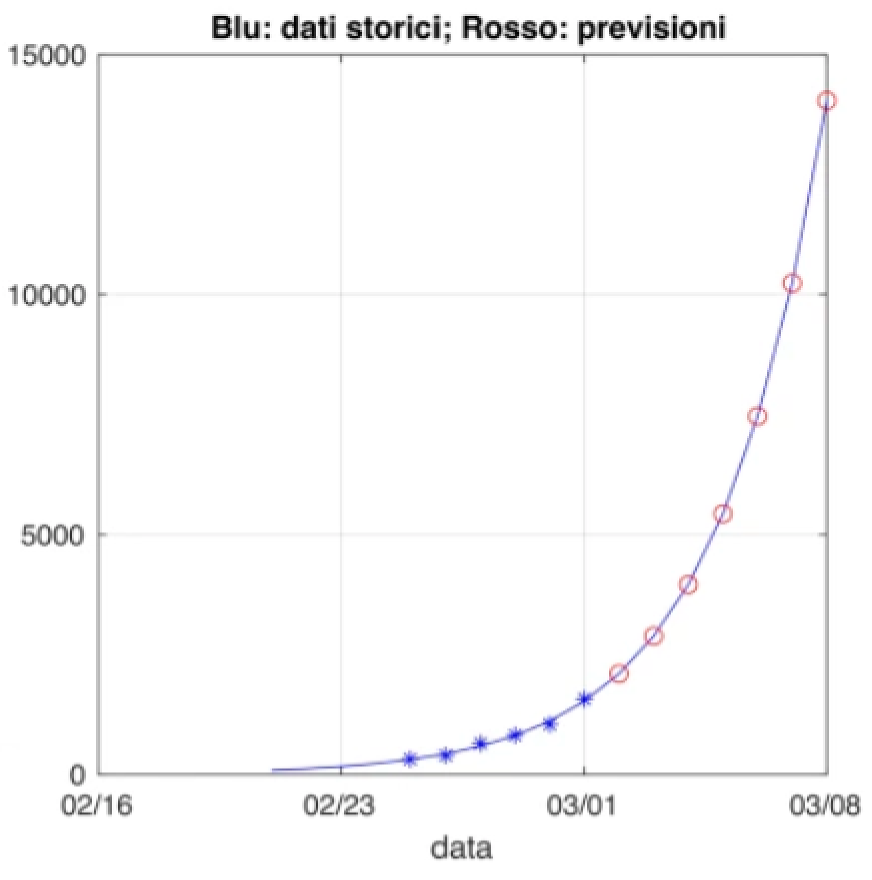
**Coronavirus, la previsione degli algoritmi: 14 mila contagi entro l'8 marzo**

**Coronavirus, 14mila contagi entro domenica: non lo prevede un virologo, ma gli algoritmi di un ingegnere informatico**

Business Insider Italia 6 marzo 2020

Il numero studiato dalla commissione di esperti formata dall'esecutivo è lo stesso calcolato dal professor Giuseppe De Nicolao dell'Università di Pavia dove insegna identificazione dei modelli e analisi dei dati. Foto di MIGUEL MEDINA/

# De Nicolao's model, and numbers




$N_{positivi}(t) = 1.535\ e^{0,3163\ t}$

$N_{positivi}(7) = 1.535 \times 1,372^{t} = 14.047$

more at:
https://statisticallearningtheory.wordpress.com/2020/03/02/previsione-della-crescita-esponenziale-dei-covid19-positivi-in-italia-lombardia-veneto-ed-e-romagna/

short link:
https://bit.ly/denicolao20200302

Nicola Giaquinto – Politecnico di Bari

# Let's compute this in MATLAB: model and parameters estimation…

$x(t) = a\exp(b\,t)$ **data model**

$y(t) = x + e$ observed data

$x_{log}(t) = \log(x) = \log(a) + b\,t = c + b\,t =$

$= A(t)\theta$ log of data model

$y_{log} = \log(y)$ log of observed data

$[\hat{c}, \hat{b}]^T = \hat{\theta} = (A^T A)^{-1} A^T y_{log}$ for $t = t_P$

OLS estimated param

$\hat{a} = \exp(\hat{\theta}(1))$ final estimated $a$

$\hat{b} = \hat{\theta}(2)$ final estimated $b$

## Data

```
y = [322, 400, 650, 821, 1049, 1577]'; % (past) data
N = length(y); % number of data samples
t_P = (-N+1:0)'; % past time intants (conventional)
```

## Regression

```
ylog = log(y); % log(y) = log(a) + b*t (straight line)
A_P = [ones(size(t_P)), t_P]; % regressor's matrix for past instants
theta_hat = A_P\ylog; % OLS regression: theta_hat = [log(a)_hat, b_hat];
```

## Estimated parameters

```
a_hat = exp(theta_hat(1)) % estimate of a
b_hat = theta_hat(2) % estimate of b
```

```
a_hat =          1534.7

b_hat =         0.31627
```

# … and fitting (past)+ forecasting (future)

$\hat{x}_{log} = \hat{c} + \hat{b}\, t =$

$= A(t)\hat{\theta}$ OLS estimated log of data

$\hat{x} = \exp(x_{log})$ final estimated $x$

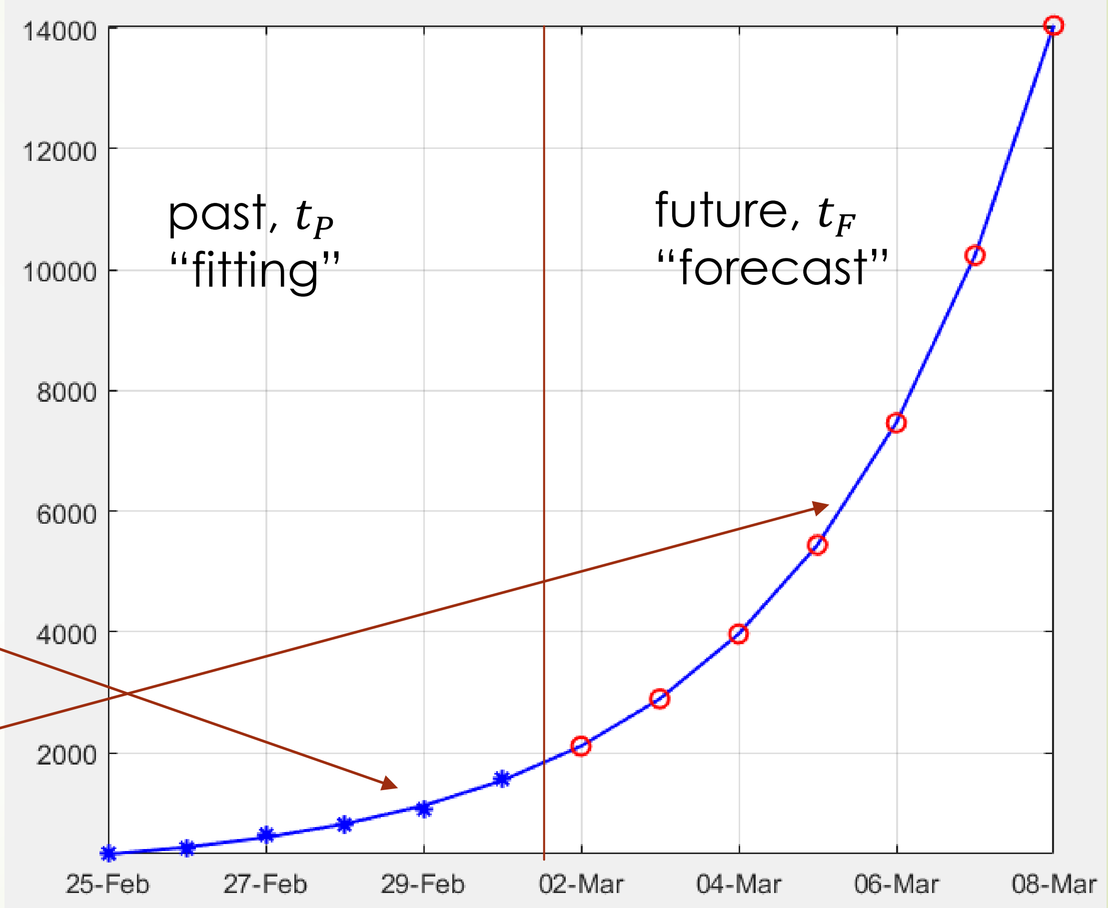
## Fitting (past)

```
xlog_hat = A_P*theta_hat; % estimate of log(x)
x_hat = exp(xlog_hat); % estimate of x
```

## Forecast (future)

```
t_F = (1:7)'; % future time instants
A_F = [ones(size(t_F)), t_F]; % regr. matrix for future instants
xlog_hat_F = A_F*theta_hat; % estimate of log(x)
x_hat_F = exp(xlog_hat_F);
x_end = x_hat_F(end)
```

```
x_end =          14045
```




$N_{positivi}(7) = 1.535 \times 1{,}372^{t} = 14.047$

# De Nicolao's concerns

> AVVISO IMPORTANTE (7/3/2020): i numeri dei soggetti positivi analizzati e predetti in questo post sono hanno cessato di essere affidabili da diversi giorni per due ragioni: (i) il cambio di

from https://bit.ly/denicolao20200302

- De Nicolao is a scientist. Not surprisingly, he has been the first to point out issues in his own forecast
- He published a note about the **unexpected difference between the forecast and the actual data**, and a discussion of the possible causes (you may want to read them!)
- But, before such discussion, a metrologist wants to know:
  - **What was the uncertainty of the forecast?**
- The uncertainty, indeed, basically **quantifies the difference that could be expected between forecasts and actual data**

# New MATLAB code: uncertainty computation (1/4)

Uncertainty in OLS estimates

$x_{log} = A\theta$ (log of) data model

$y_{log} = \log(y)$ (log) of observed data

$\hat{\theta} = (A^T A)^{-1} A^T y_{log}$ OLS estimated param

$\hat{x}_{log} = A\hat{\theta}$ OLS fitting

$u = u(y_{log}) = \frac{1}{\nu}\sum(y_{log} - \hat{x}_{log})^2$ input standard uncertainty ($\nu$ degrees of freedom)

$k = F_T^{-1}\left(\frac{1+cp}{2}, \nu\right)$ Student's t coverage factor

## Degrees of freedom and coverage factor

```
nu = length(t) - length(theta_hat); % degrees of freedom
cp = 0.95; % coverage probability
k = tinv((1+cp)/2,nu) % coverage factor
```

## Type A uncertainty of $y_{\log}$

Hypothesis: homoscedastic errors

```
u = sqrt(sum((ylog - xlog_hat).^2)/nu)
% standard uncertainty of ylog with nu d.o.f.
```

# Uncertainty computation (2/4)

Uncertainty in OLS estimates

| | |
|---|---|
| $\hat{\theta} = (A^T A)^{-1} A^T y_{log}$ | OLS estimated param |
| $\hat{x}_{log} = A\hat{\theta}$ | OLS fitting |
| $\hat{x}_{\log_F} = A_F\hat{\theta}$ | OLS forecast |
| $Q = (A^T A)^{-1}$ | matrix for uncertainty of $\hat{\theta}$ |
| $P = AQA^T$ | matrix for uncertainty of $\hat{x}_{log}$ |
| $S^2(\hat{\theta}) = u^2 Q$ | covariance matrix of $\hat{\theta}$ |
| $S^2(\hat{P}) = u^2 P$ | covariance matrix of $\hat{x}_{log}$ |

## Matrices for uncertainty evalutation in OLS method

```
Q = inv(A_P'*A_P); % for unc of PARAMETERS (theta)
P_P = A_P*Q*A_P'; % for unc in the FITTING (past)
P_F = A_F*Q*A_F'; % for unc in the FORECAST (future)
```

## Covariance matrices of $\hat{\theta}$ , $\hat{x}_{\log}$ , $\hat{x}_{\log_F}$

```
S2_theta = u^2 * Q; % for PARAMETERS
S2_xlog = u^2 * P_P; % for FITTING
S2_xlog_F = u^2 * P_F; % for FORECAST
```

# Uncertainty computation (3/4)

Uncertainty in OLS estimates

$u(\hat{\theta}) = \sqrt{diag(S^2(\hat{\theta}))}$ std uncertainty for $\hat{\theta}$

$u(\hat{x}_{log}) = \sqrt{diag(S^2(\hat{x}_{log}))}$

std uncertainty for $\hat{x}_{log}$

$U = k\,u$ expanded uncertainty

## Uncertainty of $\hat{\theta}$

```
u_theta = sqrt(diag(S2_theta)); % standard uncertainty
U_theta = k*u_theta; % expanded uncertainty

table(theta_hat, u_theta, U_theta)
```

## Uncertainty of $\widehat{x_{\log}}$ (fitting)

```
u_xlog = sqrt(diag(S2_xlog)); % standard uncertainty
U_xlog = k*u_xlog; % expanded uncertainty
```

## Uncertainty of $\widehat{x_{\log_F}}$ (forecast)

```
u_xlog_F = sqrt(diag(S2_xlog_F)); % standard uncertainty
U_xlog_F = k*u_xlog_F; % expanded uncertainty
```

# Uncertainty computation (4/4)

Propagation from $\hat{x}_{log}$ to $\hat{x} = \exp(\hat{x}_{log})$

$$u(\hat{x}) = u(\hat{x}_{log}) \exp(\hat{x}_{log}) = u(\hat{x}_{log})\, \hat{x}$$

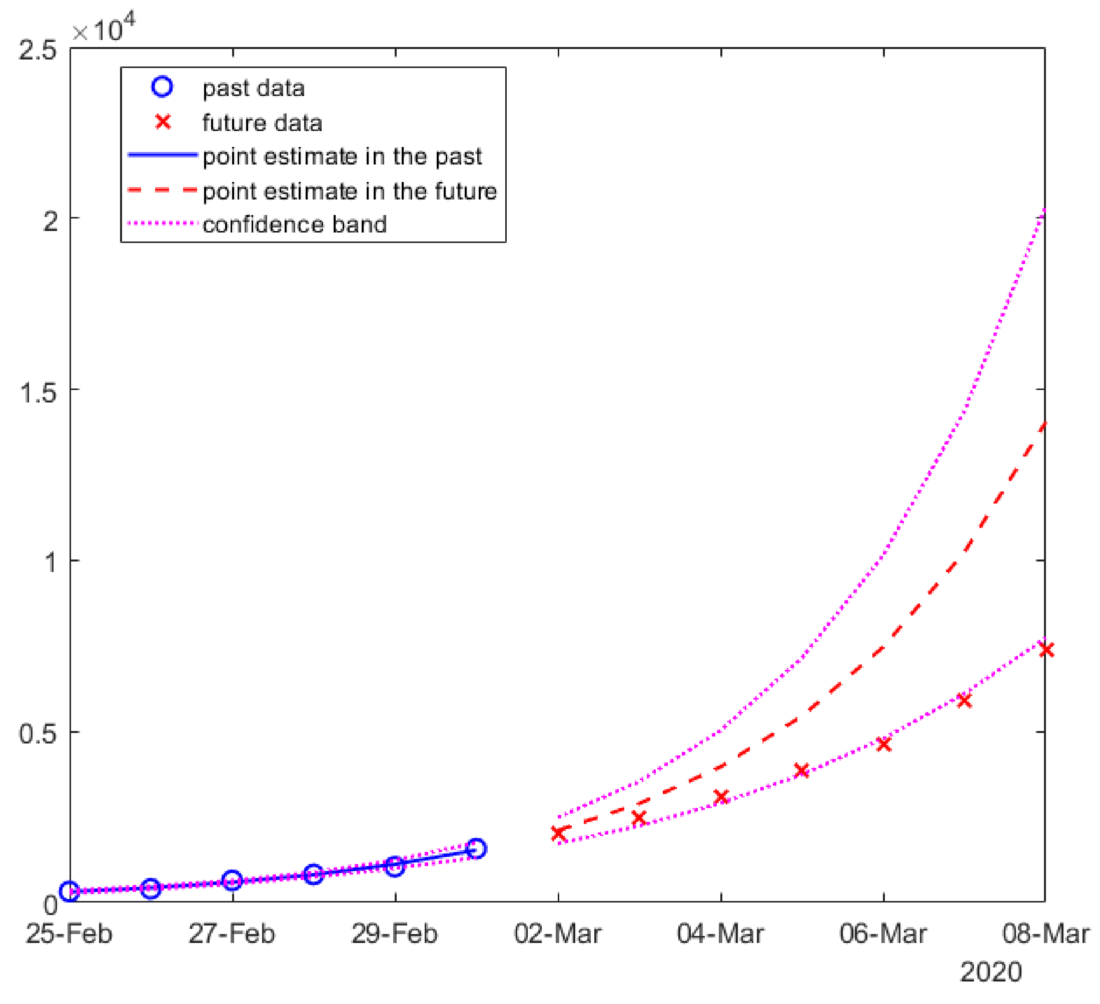
std uncertainty for $\hat{x}$

In this case, the same relation holds for the expanded uncertainty.

## Uncertainty of $\hat{x}$ in the past and in the future

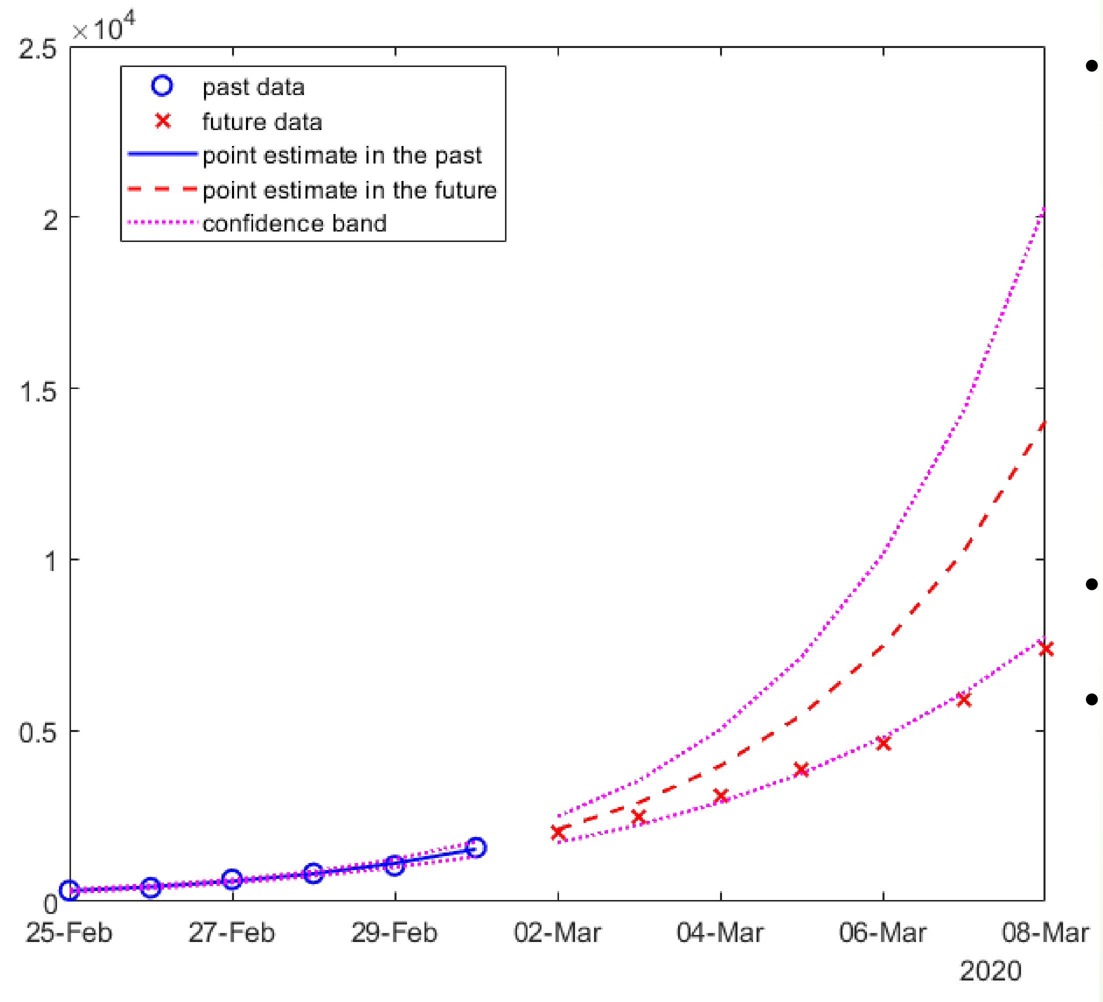
```
x_hat = exp(xlog_hat); % transformation
U_x = U_xlog.*x_hat; % uncertainty propagation
xL = x_hat - U_x; % uncertainty band (confidence band)
xU = x_hat + U_x;

x_hat_F = exp(xlog_hat_F);
U_x_F = U_xlog_F.*x_hat_F;
x_FL = x_hat_F - U_x_F;
x_FU = x_hat_F + U_x_F;
```

# Results of uncertainty computation



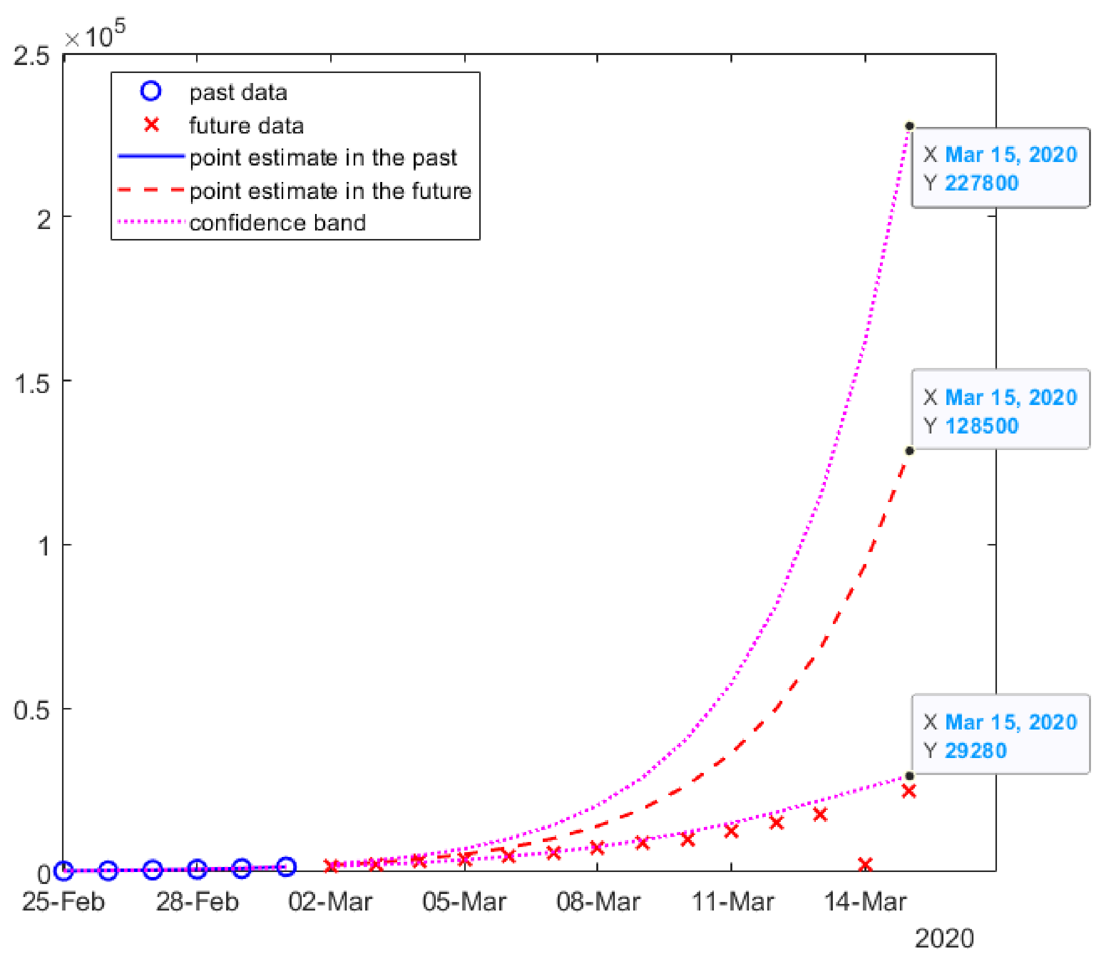

- The forecast is different from actual data, but is **not so incompatible with the uncertainty**
- It follows (almost) the lower edge of the confidence band

# So, we are satisfied with this forecast?




- Honestly, not much. But can we blame De Nicolao for the forecast?
  - The exponential model was reasonable according to epidemiology
  - It was also in accordance with data: $R^2 > 0.99$
  - Read also De Nicolao's notes
  - My opinion: at that time, the point was **to blow the whistle**
- We can construct a better model… but in hindsight
- The evolution of COVID-19 has taught us that:
  - **1. we must learn from actual experience**
  - **2. reliable forecasting is very, very difficult**

# Important information from uncertainty computation

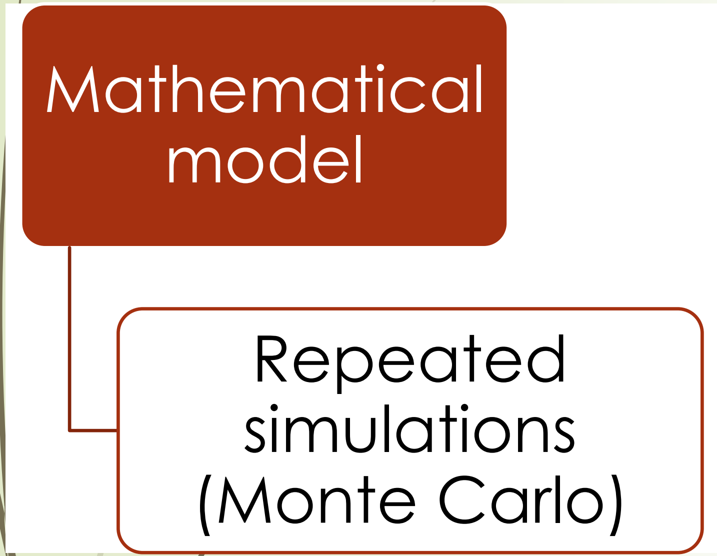



- Forecasting **many weeks ahead** with this information appears of little practical value, **due to increasing uncertainty**
- After **two weeks**, the model predicts an interval spanning one order of magnitude. Try with three or four weeks!
- Another model may have lower uncertainty. But only experience (**repeated forecasts**) can validate the modeling
- Repeated forecasts and uncertainty are, of course, strictly linked. Let us see this better

# Assessing our first forecast: uncertainty and long-run success rate

A simple (and short) technical check

# Checking the uncertainty band of our COVID-19 forecast using Monte Carlo method



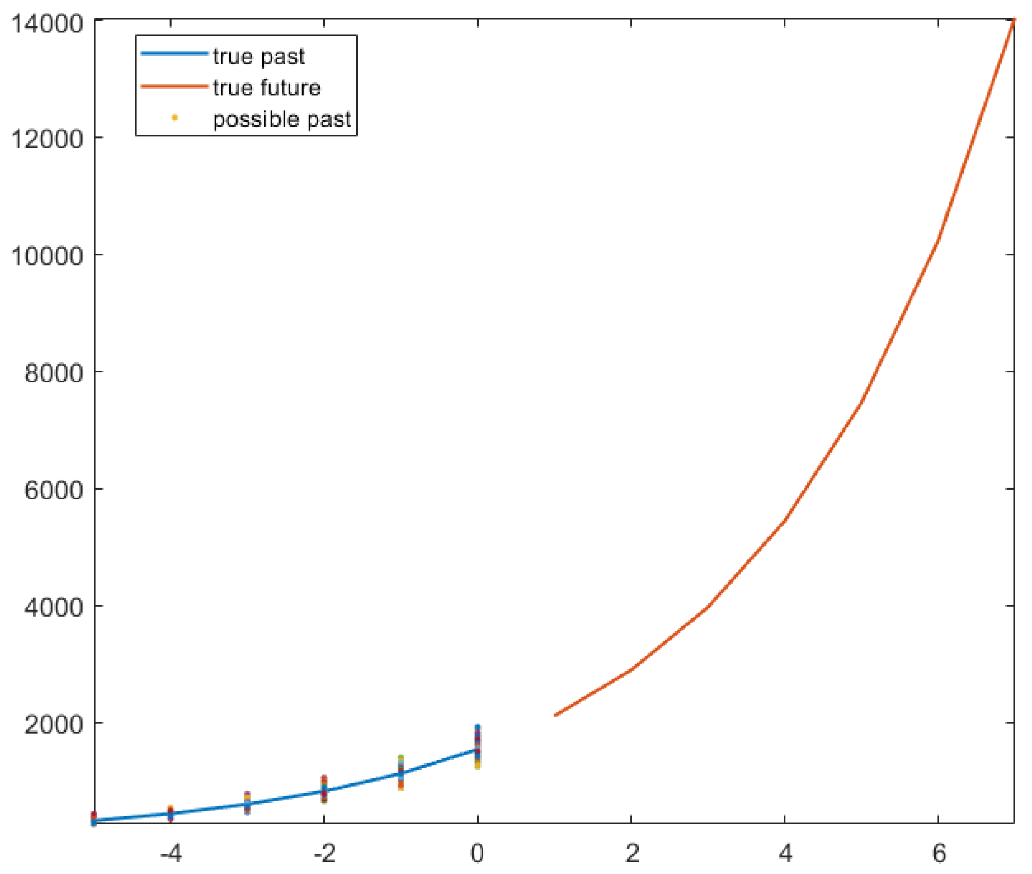

- We check the computed uncertainty band with repeated **simulated** forecasts (Monte Carlo method)
- We will compute the **long-run success rate (LRSR)**, and will compare it with the coverage probability (95%)
- Again, we present a Matlab implementation (quite easy and short)

# Nsim simulated measurements (past observations)

## Monte Carlo simulation

### Nsim simulated past measurements

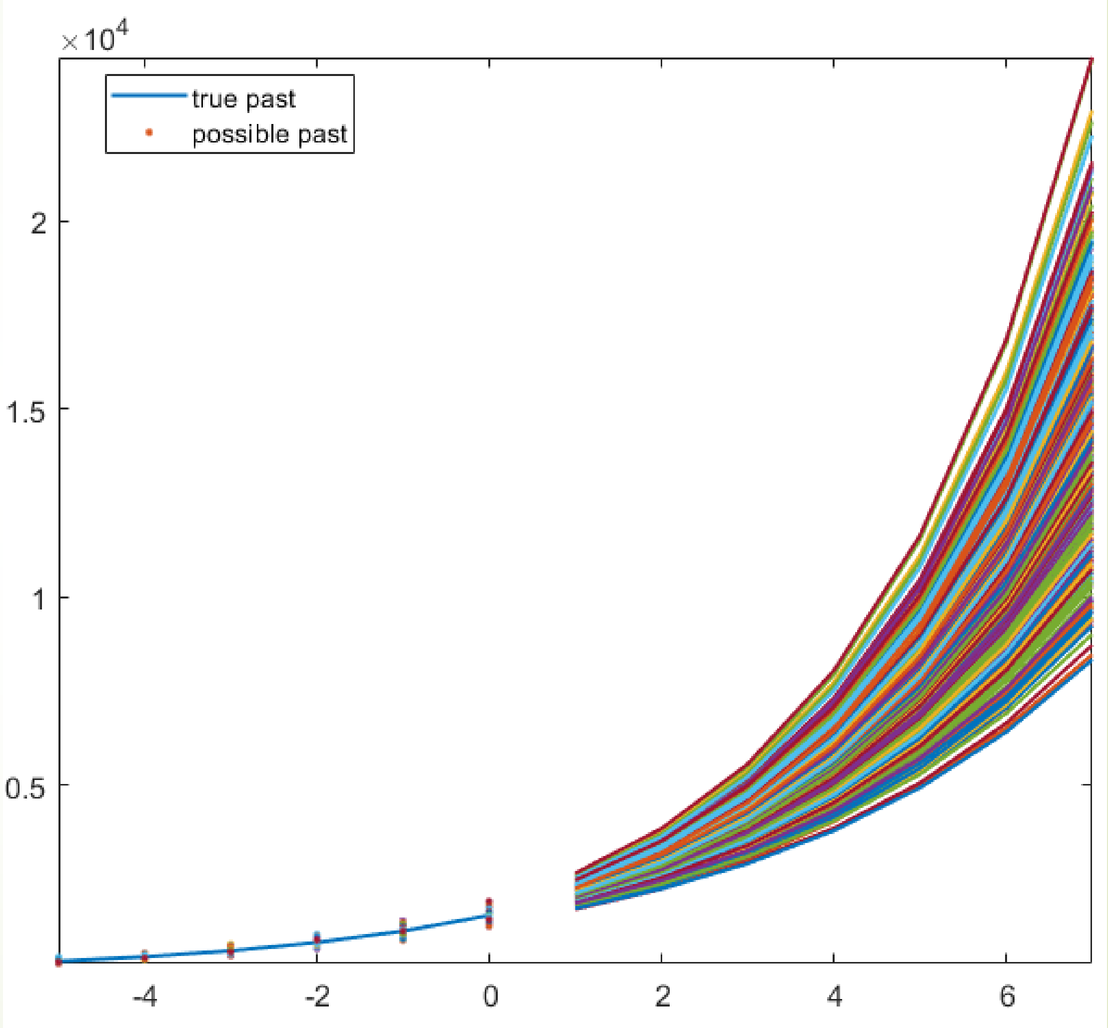
```
Nsim = 10000;

theta = theta_hat; % true parameters (fictional)
xlog = A_P*theta; % true log of signal (fictional)
xlog_F = A_F*theta; % true log of signal, future (fictional)

Xlog = repmat(xlog,[1,Nsim]); % Nsim copies of xlog
E = u*randn(size(Xlog)); % Nsim error sequences
Ylog = Xlog + E; % Nsim log of simulated measurements

x = exp(xlog); % true signal, past
x_F = exp(xlog_F); % true signal, future
Y = exp(Ylog); % Nsim simulated past measurements
```



# Nsim simulated forecasts

## Nsim simulated fittings and forecasts

```
Theta_hat = A_P\Ylog;
Xlog_hat = A_P*Theta_hat;
Xlog_hat_F = A_F*Theta_hat;
X_hat = exp(Xlog_hat);
X_hat_F = exp(Xlog_hat_F);
```



# LRSR of the simulated forecasts (computed intervals)

**Nsim type A standard uncertainty evaluations (at t = 7)**

```
u_sim = sqrt(sum((Ylog - Xlog_hat).^2)/nu); % Nsim type A stand. unc. u(y)
u_sim_Xlog_F_end = sqrt(u_sim.^2 * P_F(end,end)); % Nsim stand. unc. u(xlog_F(end))
```

**Nsim expanded uncertainty evaluations (at t = 7)**

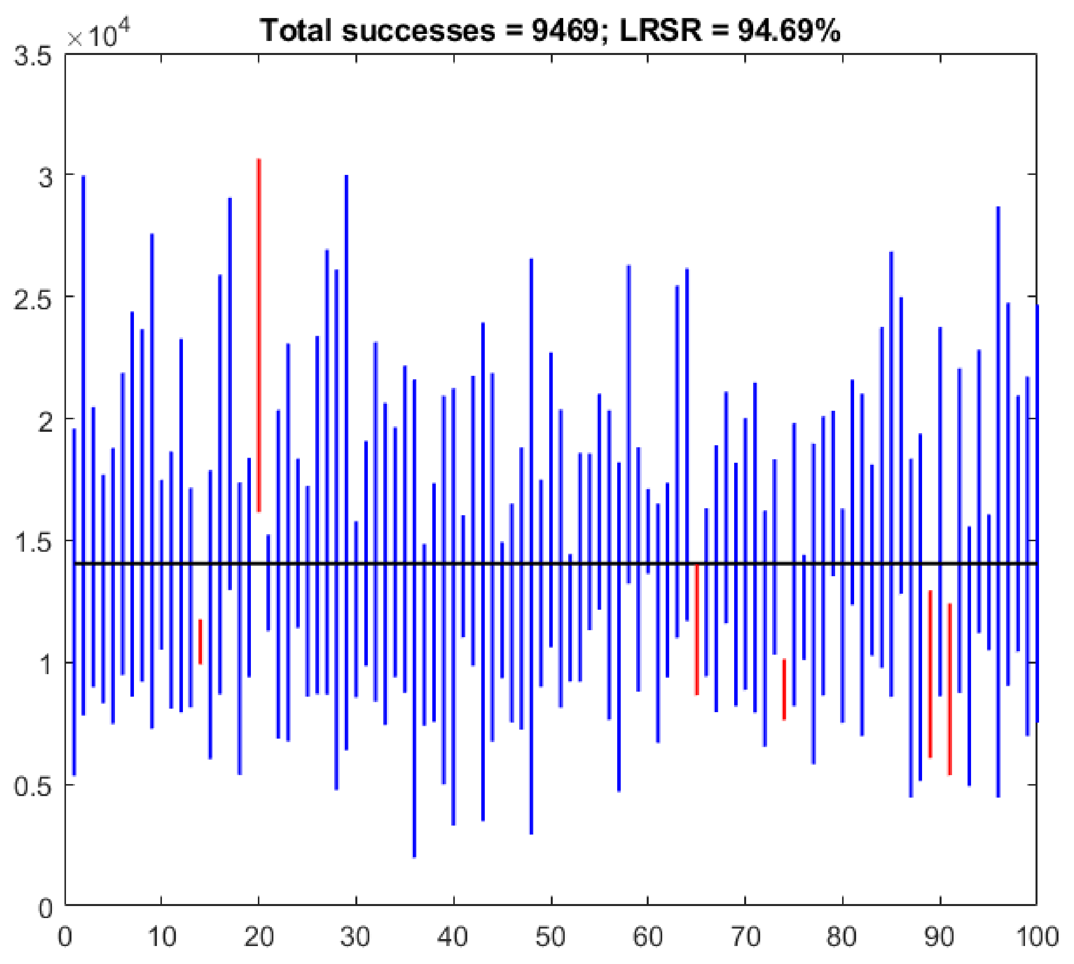
```
U_sim_Xlog_F_end = k*u_sim_Xlog_F_end; % Nsim expand. unc. U(xlog_F(end))
X_hat_F_end = X_hat_F(end,:); % Nsim forecasts at t = end = 7
U_sim_X_F_end = U_sim_Xlog_F_end.*X_hat_F_end; % Nsim expand. unc. U(x_F(end))
```

**Nsim uncertainty intervals (at t = 7)**

```
X_F_end_L = X_hat_F_end - U_sim_X_F_end;
X_F_end_H = X_hat_F_end + U_sim_X_F_end;
```

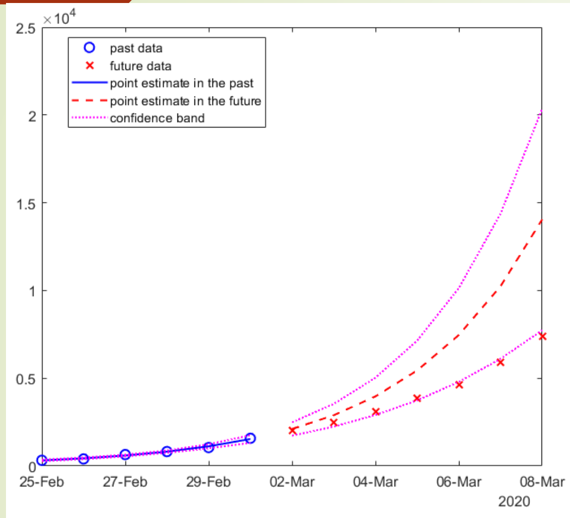
**LRSR: how many uncertainty intervals catch the "true" future?**

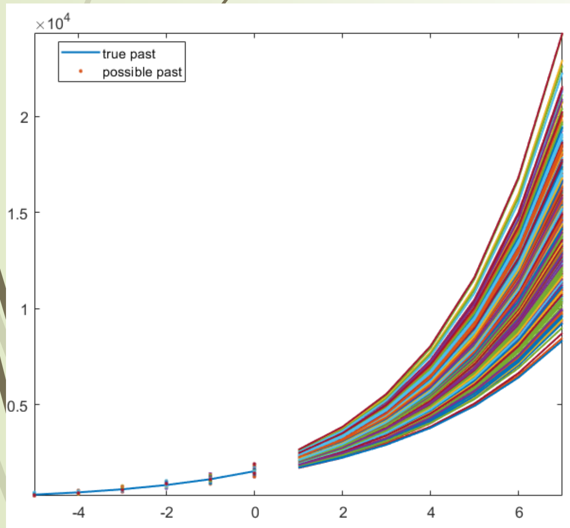
```
LRSR = sum((X_F_end_L <= x_F(end)) & (x_F(end) <= X_F_end_H))/Nsim
```



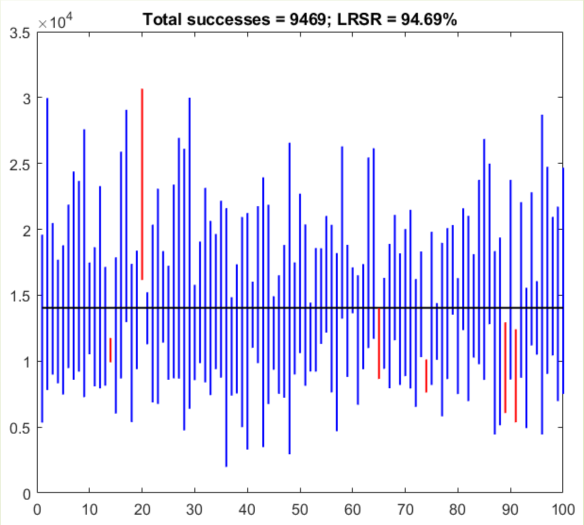

**The interval catches the (simulated) true value** in 95% of “experiments”

# Check is ok. But what is its value?




1. Is this check a "frequentist" and "old" approach to probability and uncertainty?
2. Is this possible only for the used forecasting method (a *frequentist estimate*)?
3. What if a non-frequentist method does not pass this kind of check?
4. Finally: what have we validated, actually?



# Uncertainty and long-run success rate: different opinions

Yes, this Section may be boring

# The question, in the words of prof. Ignacio Lira

IOP PUBLISHING METROLOGIA

Metrologia **46** (2009) 616–618 doi:10.1088/0026-1394/46/6/002

**On the meaning of coverage probabilities**

**I Lira**

It has been argued that the probability associated with a coverage interval […] should be about equal to the proportion of independent intervals reckoned over time that contain the measurand (Willink 2006 *Metrologia* **43** L39–42, Hall 2008 *Metrologia* **45** L5–8, Possolo *et al* 2009 *Metrologia* **46** L1–7)

The opposite point of view maintains that a coverage probability can only be interpreted as the degree of belief in the quoted interval containing the unknown value of the measurand (Lira 2008 *Metrologia* **45** L21–3).

# One of Lira's examples (in short)...

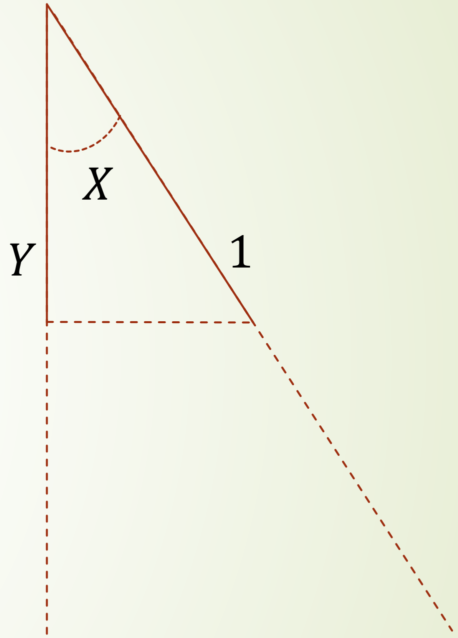
**2. The cosine of an angle**

It is desired to gauge the projection $Y$ of a line of known unit length onto another line. The measurement model is

$$Y = \cos X, \qquad (1)$$

where $X$ is the inexactly known angle between the two lines.

Suppose that the only information about $X$ is that it is contained in the interval $[-\pi/4, \pi/4]$. Since no preferences for any value $x$ in this interval are pre-established the uniform probability distribution $f_X(x) = 2/\pi$ should be adopted.



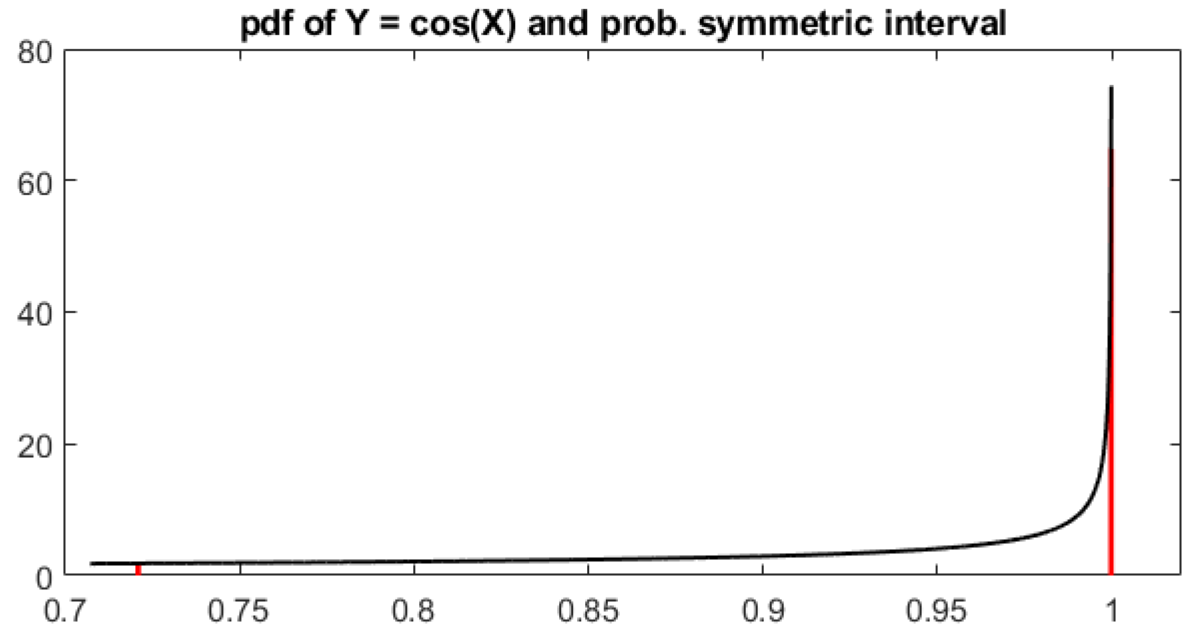

- $X$ unknown, $X \in [-\pi/4, \pi/4]$
- Give an interval estimate of $Y$ with a 95% coverage probability

# Lira's conclusions

An interval $[a, b]$ with coverage probability $p$ is:

$$a = \sin\frac{\pi}{4}(1+q) \qquad \text{and} \qquad b = \sin\frac{\pi}{4}(1+p+q).$$

with $q$ such that $0 < p + q \;\; 1$.




"Since in this example no actual measurements are involved, there is no randomness. […] So **it is not possible to imagine generating a large number of independent intervals** […] For $p$ = 0.95, the interval [0.7209, 0.9998] is the only [probabilistically symmetric] one; consequently **its 'long-run success rate' is meaningless**."

"[…] when a coverage interval summarizes the resulting state of knowledge, **the coverage probability should not be interpreted as a relative frequency** of successful intervals in a large series of imagined or simulated intervals, but **as the degree of belief of the party** that quoted the single actual interval in its being successful"

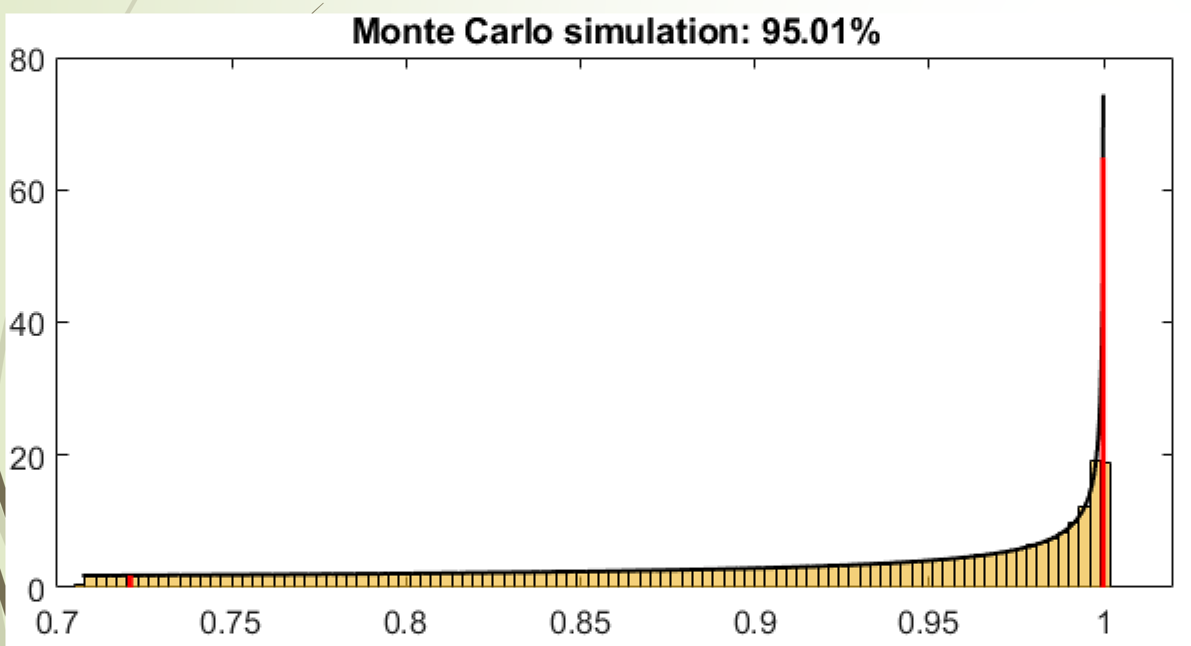
# A respectful observation (in MATLAB…)

**Interval**

```
p = 0.95;
q = 0.025;
a = sin(pi/4*(1+q))
b = sin(pi/4*(1+p+q))
```

**Random X, Y and LRSR**

```
Nsim = 1e6;
X = unifrnd(-pi/4,pi/4,[1,Nsim]);
Y = cos(X);
LRSR = sum((Y>=a & Y<=b))/Nsim
```




- Prof. Lira is a great Bayesian statistician and metrologist, but… *why we cannot compute the LRSR of his interval?*
- **Wherever there is a probability computation, a Monte Carlo simulation is possible**
  - This has nothing to do with the interpretation of probability (“frequency” or “degree of belief”)
- Lira’s computations are correct, but not so simple**: a mistake could have been possible** (it happens!)
- LRSR computation is simple, and validates his analytic solution

# I use LRSR to assess uncertainty, when practically possible

**In general:**

- If you don't, I don't think you are wrong.
- If you think my LRSR calculation are someway wrong, please write me a note
- If you think LRSR calculation should be legally banned… what can I say… write to the Government

**In the case of COVID-19 forecasts:**

- We will be happier if **actual future numbers fall in our uncertainty band in 95% of cases**
- People will be happier if we state: "**this forecast will be correct in 95 cases out of 100**", instead of: "**my/our degree of belief is 95%**"
- Of course, **in practice** this is all but simple! We must not confuse simulation with reality…

# A mistake to avoid: starting a cycle of harsh arguing (about LRSR, or whatever)

On the meaning of coverage probabilities

I Lira

Published 7 October 2009 • 2009 BIPM and IOP Publishing Ltd

Metrologia, Volume 46, Number 6

COMMENT

Probability, belief and success rate: comments on 'On the meaning of coverage probabilities'

R Willink

Published 12 May 2010 • 2010 BIPM & IOP Publishing Ltd

Metrologia, Volume 47, Number 3

REPLY

Reply to 'Probability, belief and success rate: comments on "On the meaning of coverage probabilities"'

I Lira

Published 12 May 2010 • 2010 BIPM & IOP Publishing Ltd

Metrologia, Volume 47, Number 3

and so on, endlessly…

Se il Covid fa litigare gli scienziati

*di Luciano Fassari*

***A volte per questioni formali, altre più di sostanza. M scienziati questa emergenza è stata una scintilla (an attizzato polemiche e botta e risposta anche molto ac scienza in tutta la penisola e a tutti i livelli***

**In Onda, scintille tra Giulio Sapelli e Andrea Crisanti: "Non sono uno statistico? Parli tu che hai visto l'epidemia in tv"**

- Lessons learned from COVID-19: **harsh arguments among scientists are harmful for themselves and the society**
- We send only a message:
  - **"we don't even agree among metrologists"** (or other categories of scientists)
- If the fight is about the uncertainty of Y = cos(X), it is not better

# So, uncertainty must be associated to a LRSR?

- My answer is: WHENEVER POSSIBLE (sometimes it is NOT PRACTICALLY POSSIBLE)
- Again: we must always keep distinct:
  - **reality** (and actual experiments)
  - **mathematical models**, aiming at mirroring the real phenomenon (and simulated experiments)

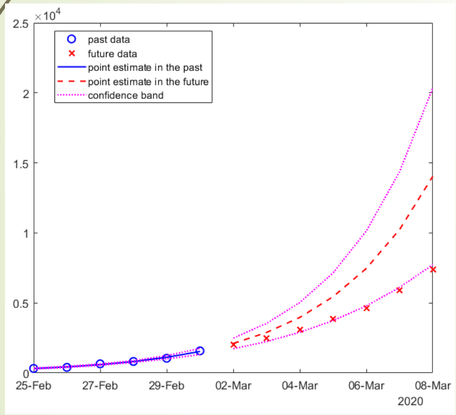
**Reality** → **Mathematical model**

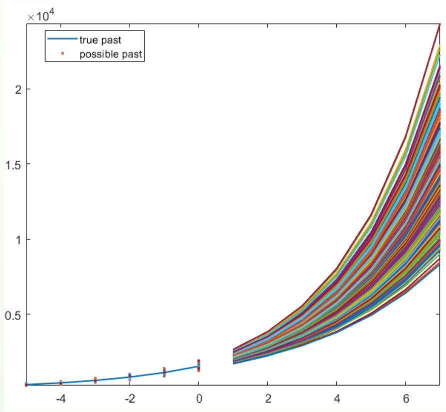
Reality:
- Repeated experiments (in some way)
- It checks / assesses experimentally the uncertainty
- We can expect probability = success rate QUITE LOOSELY

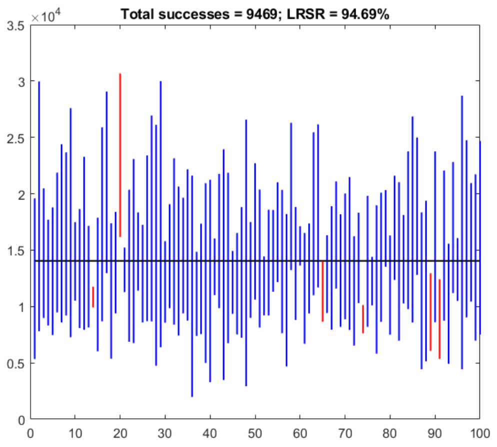
Mathematical model:
- Repeated simulations (Monte Carlo)
- It checks the uncertainty computation, NOT the correctness of the model
- We must have probability = success rate ALMOST EXACTLY

# Using Monte Carlo simulations

- Of course, they check only the correctness of uncertainty computations **within the model**
- They do NOT check that the mathematical model is satisfactory for the actual phenomenon.
- But are MC simulations and LRSR a "frequentist" device? What if we use e.g. **Bayesian methods**?




**Mathematical model**

- Repeated simulations (Monte Carlo)
- It checks the uncertainty computation, NOT the correctness of the model
- We must have probability = success rate

  ALMOST EXACTLY

# Monte Carlo simulations for Bayesian methods

**Mathematical model**

- Repeated simulations (Monte Carlo)
- It checks the uncertainty computation, NOT the correctness of the model
- We must have probability = success rate ALMOST EXACTLY

- Monte Carlo computation of LRSR is *always possible* (even if it may be unpractical)
- **Making this check does not mean being "frequentist" instead of "Bayesian" or whatever**
- See paper below

Metrologia

PAPER

## Interval estimations in metrology

G Mana[1] and C Palmisano[2]

Published 1 May 2014 • © 2014 BIPM & IOP Publishing Ltd

Metrologia, Volume 51, Number 3

This paper has examined interval estimation from both the Neyman and Bayesian viewpoints and investigated differences that are not always perceived. It has demonstrated a frequentist model of the coverage probability of Bayesian intervals, where a single interval is built for the same measured valued repeatedly sampled according to different

Both the Neyman and Bayesian approaches are correct, but confidence and credible intervals are solutions of different problems, namely (2) and (3). Hence, what is the best approach is an ill posed question. Whether to use one or the other to express the uncertainty of measurements depends on what problem we must solve and on decision theoretic considerations that are outside the scope of this paper. The following thoughts may supply some guidelines.

# What about reality?

**Reality**

- Repeated experiments (in some way)
- It checks / assesses experimentally the uncertainty
- We can expect probability = success rate QUITE LOOSELY

- Uncertainty can be consistent with the mathematical model and NOT with reality – **we will see exactly such a case with COVID-19 forecasting**
- Repeating **real** experiments?
- Real experiments are not always repeatable
  - repeatability is a typical "laboratory" case
- We cannot "repeat" **actual** COVID-19 observations and forecast at March 1, 2020!
- Likewise, we cannot "repeat", to check forecasting algorithms:
  - the past evolution of the Bitcoin (e.g. from 2009)
  - the past quotation in Stock exchange of Microsoft shares
  - the 2020 USA presidential elections
  - etc.

# “Repeating” experiments, in wide sense

**Reality**

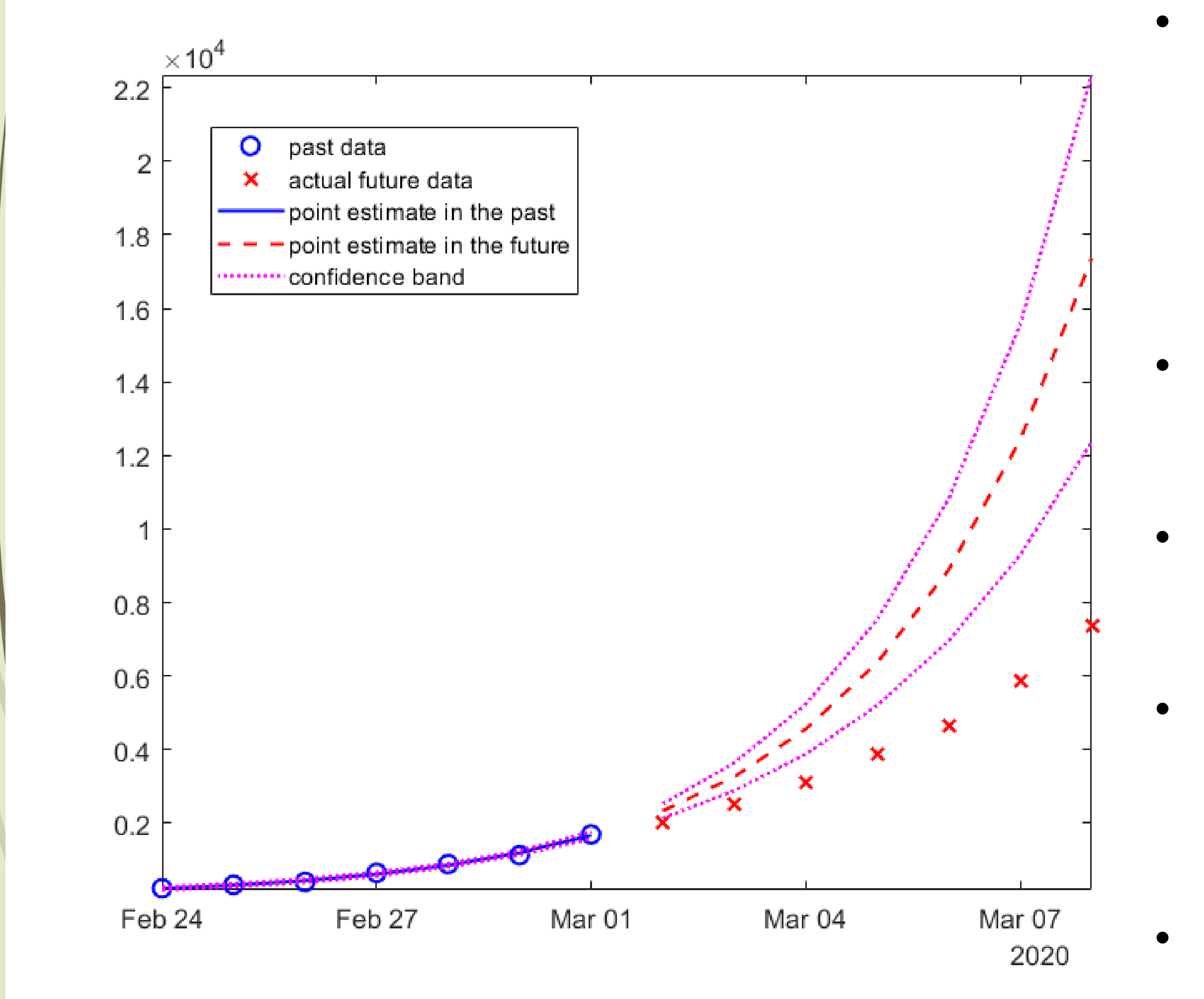
- Repeated experiments (in some way)
- It checks / assesses experimentally the uncertainty
- We can expect probability = success rate QUITE LOOSELY

- In a sense we can “repeat” actual experiments and evaluate the LRSR. What is the success rate:
  - of a financial advisor, human or artificial, *in the long run*? (“high-frequency trading”)
  - of a person/algorithm that predicts elections outcome, *in many different elections*?
  - **of our COVID-19 forecasting method, e.g. *day by day, week by week, etc.?***
- This not an exact procedure, like with mathematical models, but
  - it is commonsense
  - is **dutiful**

# Back to COVID-19 forecasting

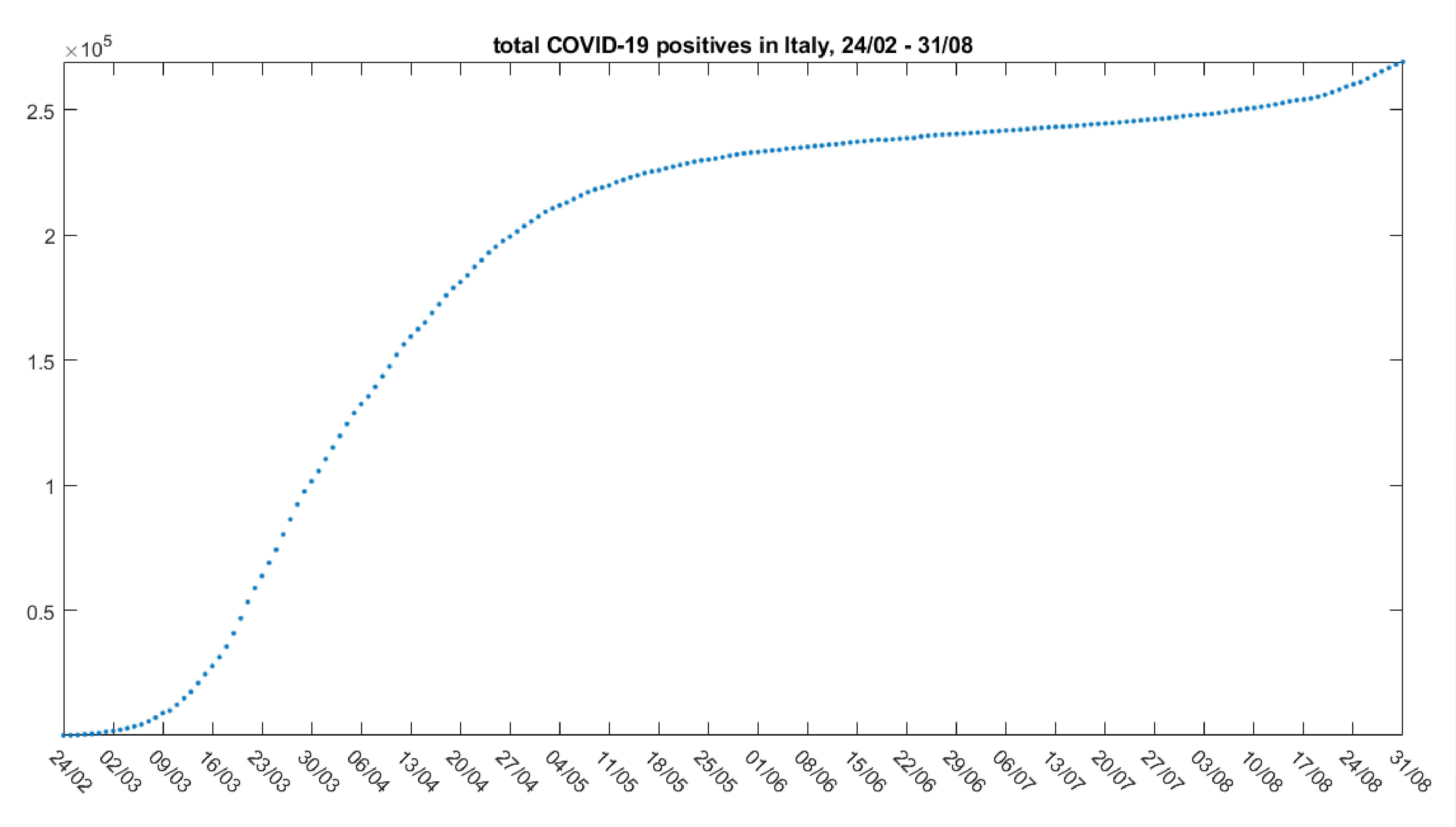
Forecasting from Feb 24 till today, with more flexible models

# Forecast with exponential model - corrected data




- The forecast differs from De Nicolao's one, because of **posterior correction of official data**
  - Feb 24 has been added
  - Numbers from Feb 25 to Mar 01 have been corrected, increasing them
- With these numbers, the exponential model is clearly NOT satisfactory
- This is a bit surprising, also because $R^2 = 0.9933$ in the fitting
- Note: "true" infected could well have been exponential, but we are working to forecast an available time series
- This is just one reason to try a non-exponential model

# How to model the COVID-19 evolution?




- Second reason is that the evolution, of course, can be exponential only “sometimes”
- Simple-form curves (exponential, logistic, etc.) are not suitable
- Which model could we use?

# A popular modeling/forecasting tool: ARIMA

ARIMA = auto-regressive integrated moving average

$ARIMA(p, d, q)$ model:

$$y'_t = c + \Phi_1 y'_{t-1} + \cdots + \Phi_p y'_{t-p} + \Theta_1 \epsilon_{t-1} + \cdots + \Theta_q \epsilon_{t-q} + \epsilon_t$$

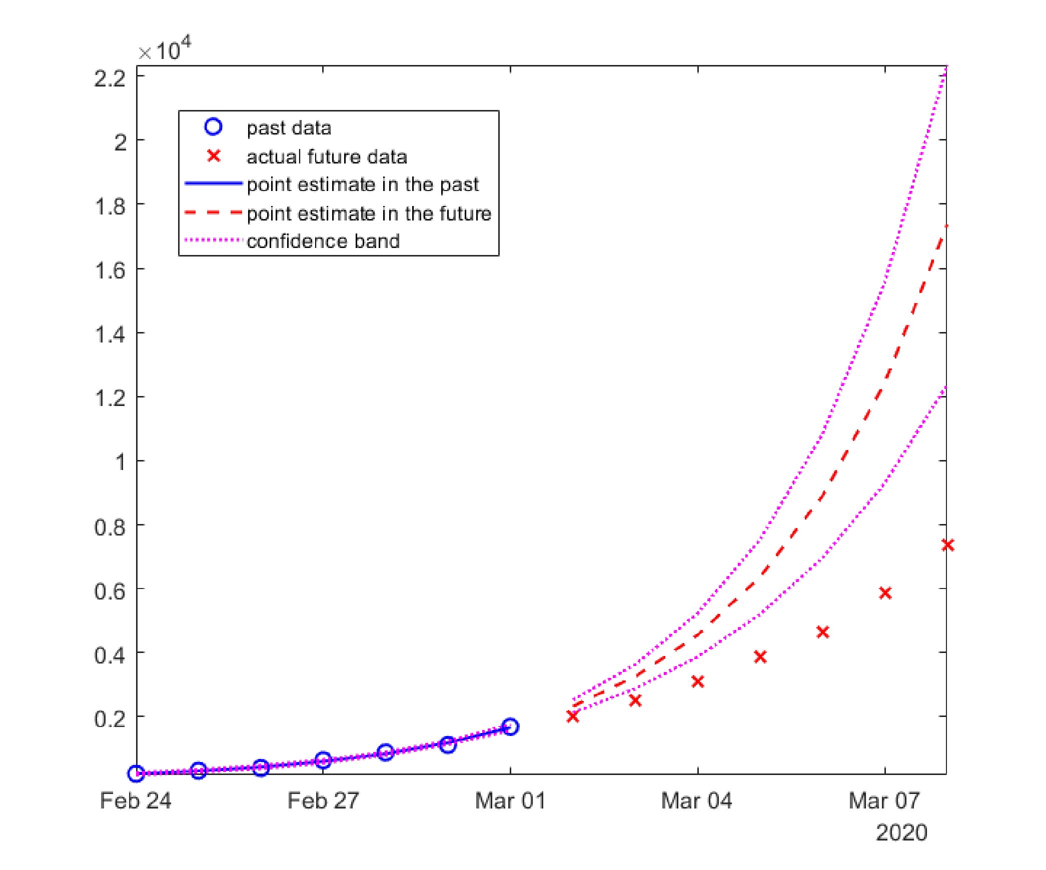
where:
$y_t$ = time series
$y'_t$ = d-times differenced series: $y'_t = diff(y_t, d)$
$\epsilon_t$ = white noise
$\Phi_1, \dots, \Phi_p$ = p coefficients, auto-regressive part
$\Theta_1, \dots, \Theta_q$ = q coefficients, moving average part

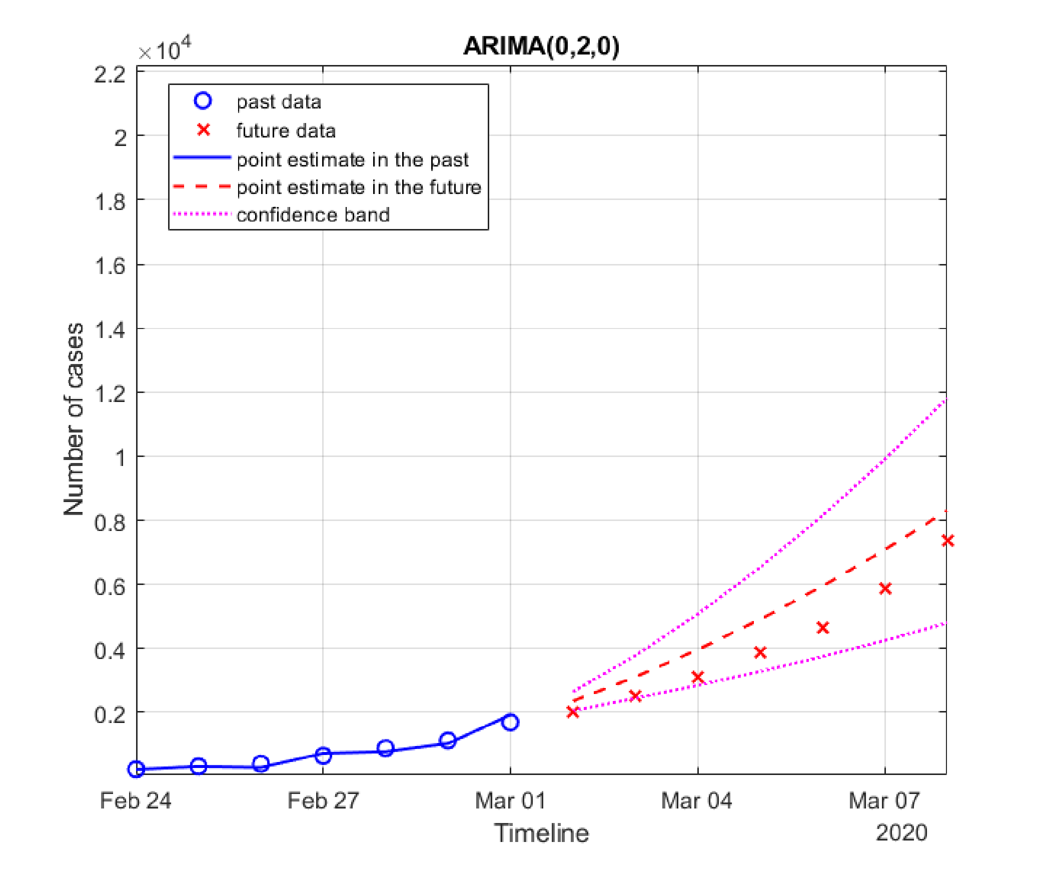
There is a quite big amount of theory on the subject. You may want to study something more about it!

We just focus on results of practical interest here

ARIMA modeling has been already used in many scientific works about COVID-19 evolution and forecasting

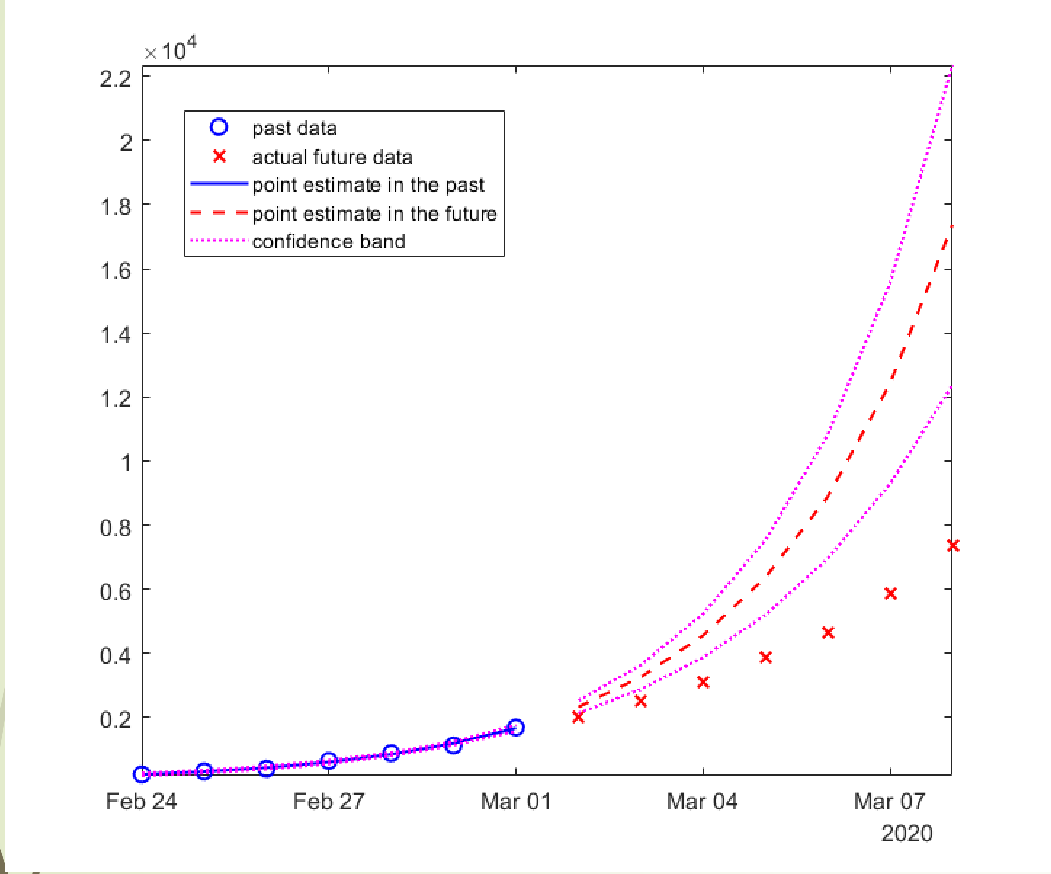
# Exponential vs. ARIMA(0,2,0)



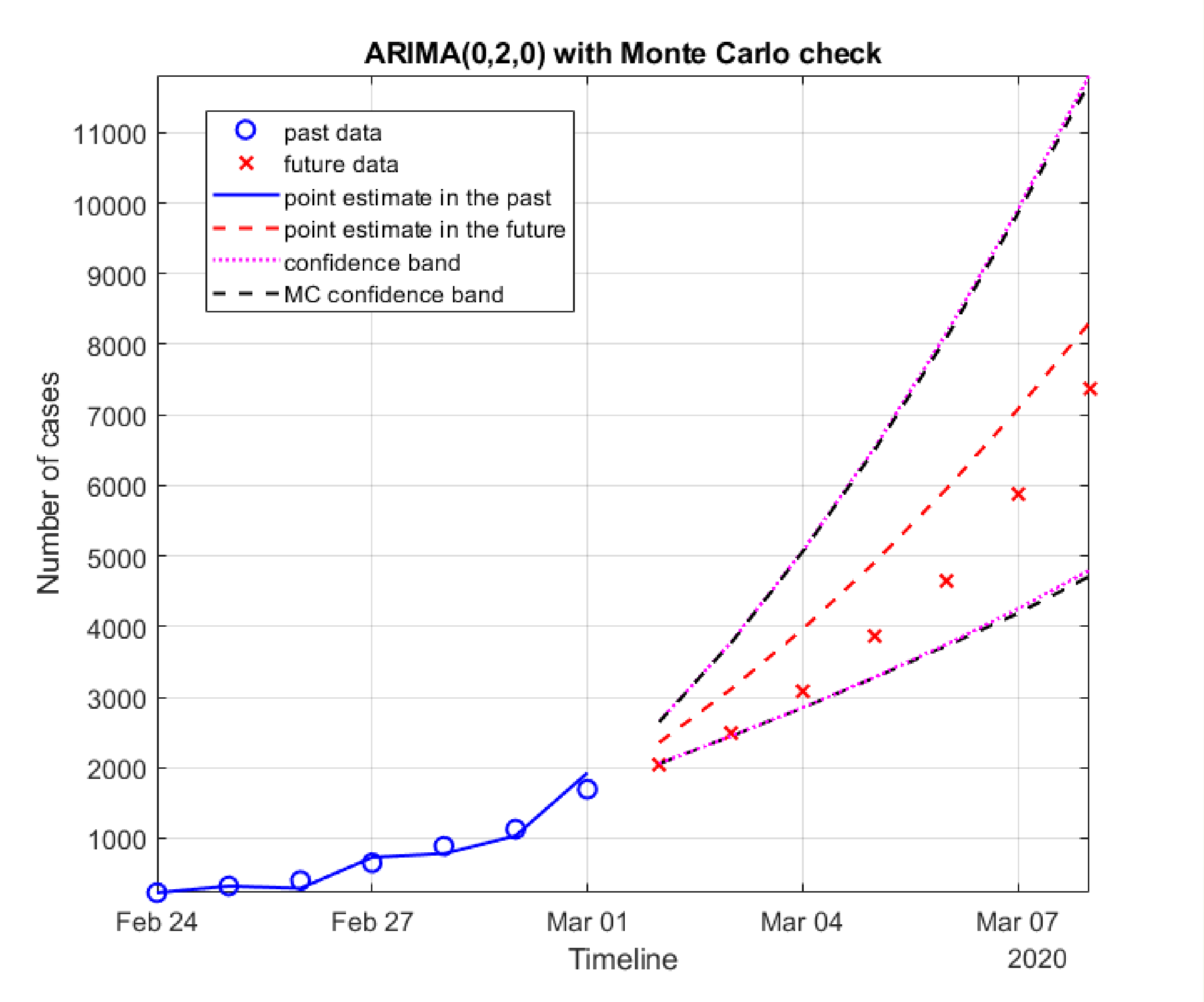

- ARIMA(0,2,0) is a very, very simple model. But its forecast here **is better**
- Besides, **the uncertainty band is narrower**: about 7000 vs. 10000
- But **pay attention** to the meaning of this “uncertainty”!

# Meaning of uncertainty with exponential model



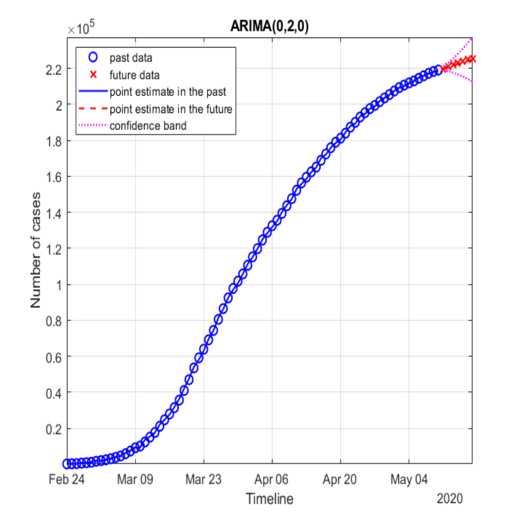

- The exponential model is justified by epidemiology
- With very good reasons we can state: "the evolution appears to be exponential"
- In this case, the computed uncertainty has also a clear meaning:
    - "If the evolution is exponential, this is the future, with this uncertainty, and this associated probability"
- (Of course in this case uncertainty computation becomes useless, because the model fails)

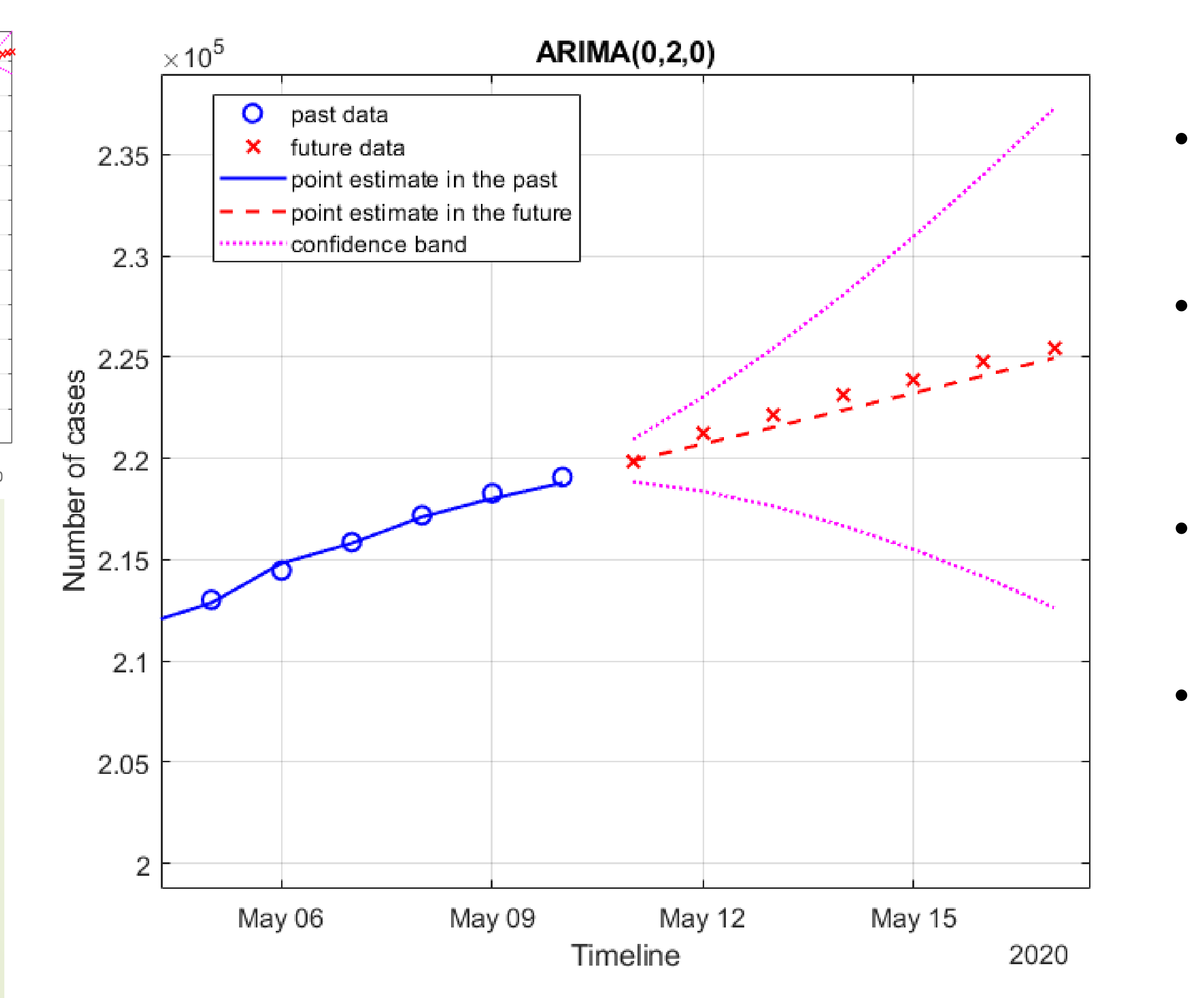
# Meaning of uncertainty with ARIMA model




- We have checked also the ARIMA model uncertainty with Monte Carlo simulation: it is "right". It appears realistic, too.
- But the computed uncertainty has ***little real meaning*** here
- The reason is that ARIMA model is NOT related with the epidemiologic mechanism (like the exponential model). Even if it works well in forecasting!
- This statement may appear surprising, but let's have a look at another forecast…

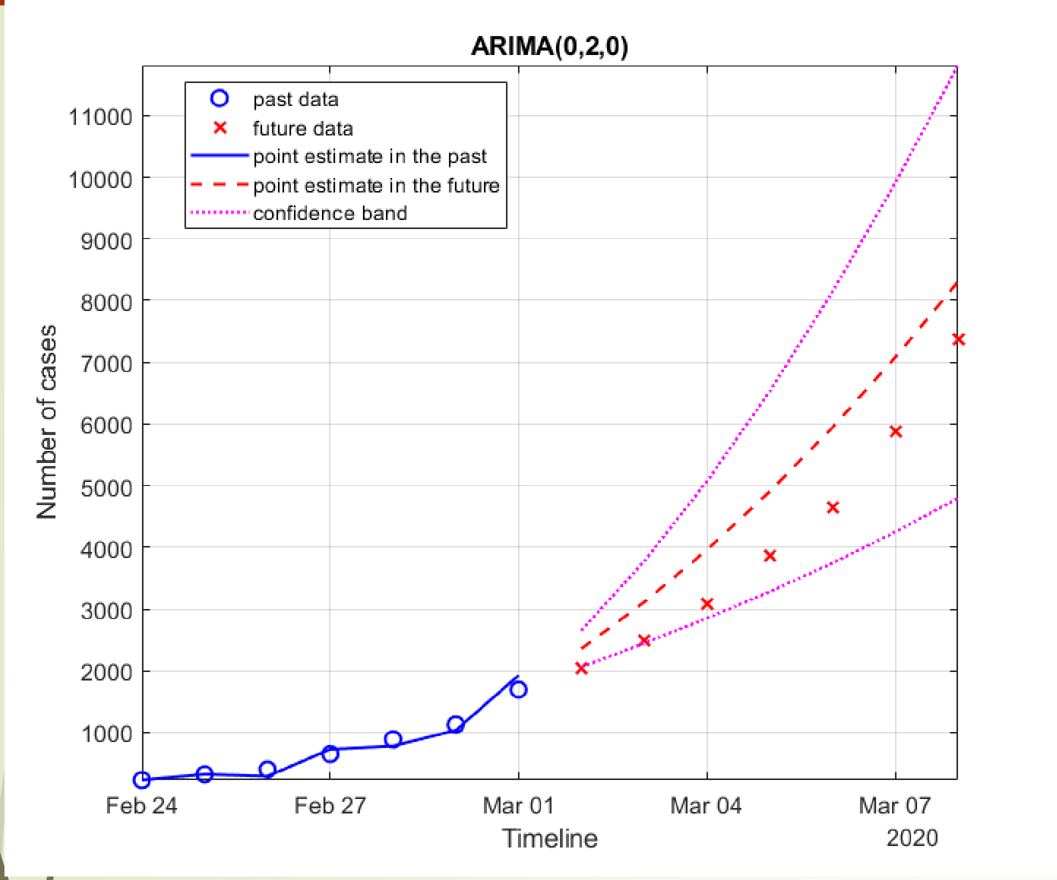
# Forecast and “uncertainty” at May 10th

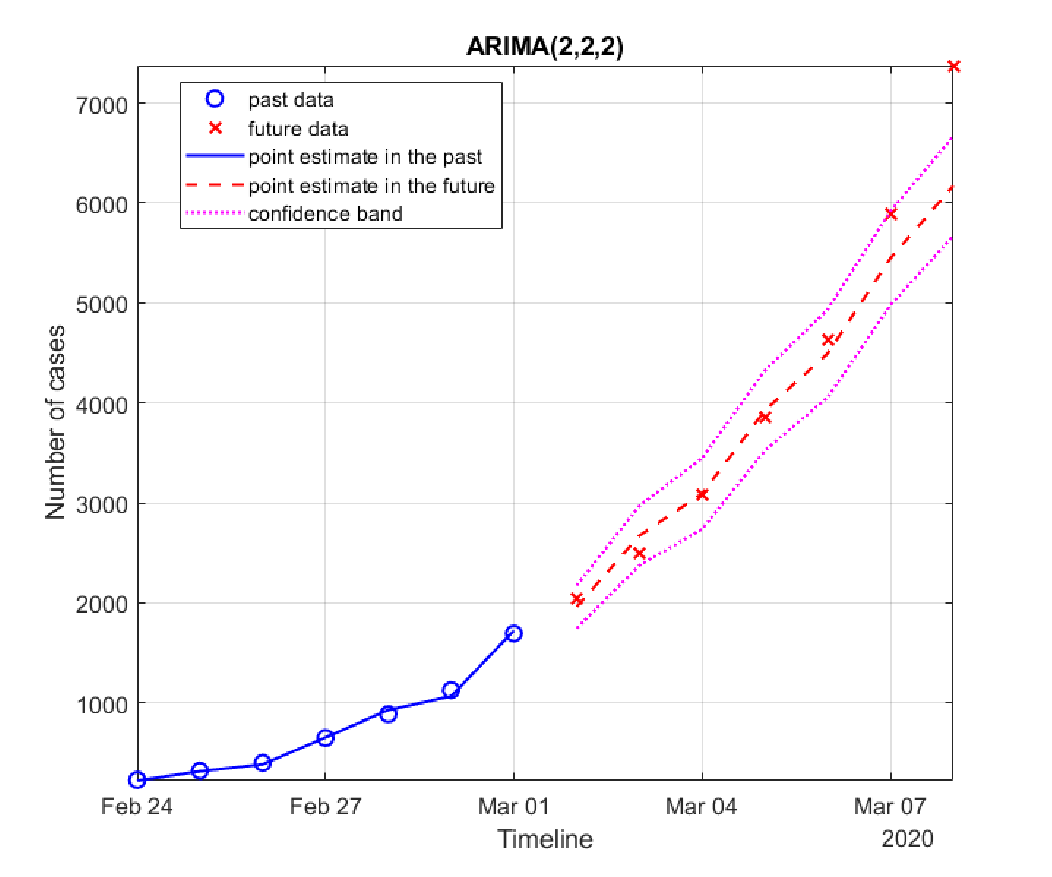



- The 1-week forecast of ARIMA(0,2,0) is nice
- Has the mathematically computed uncertainty a real meaning here?
- Can the cumulative number of cases **decrease**?
- An ARIMA process **has far more freedom than the epidemiological process**

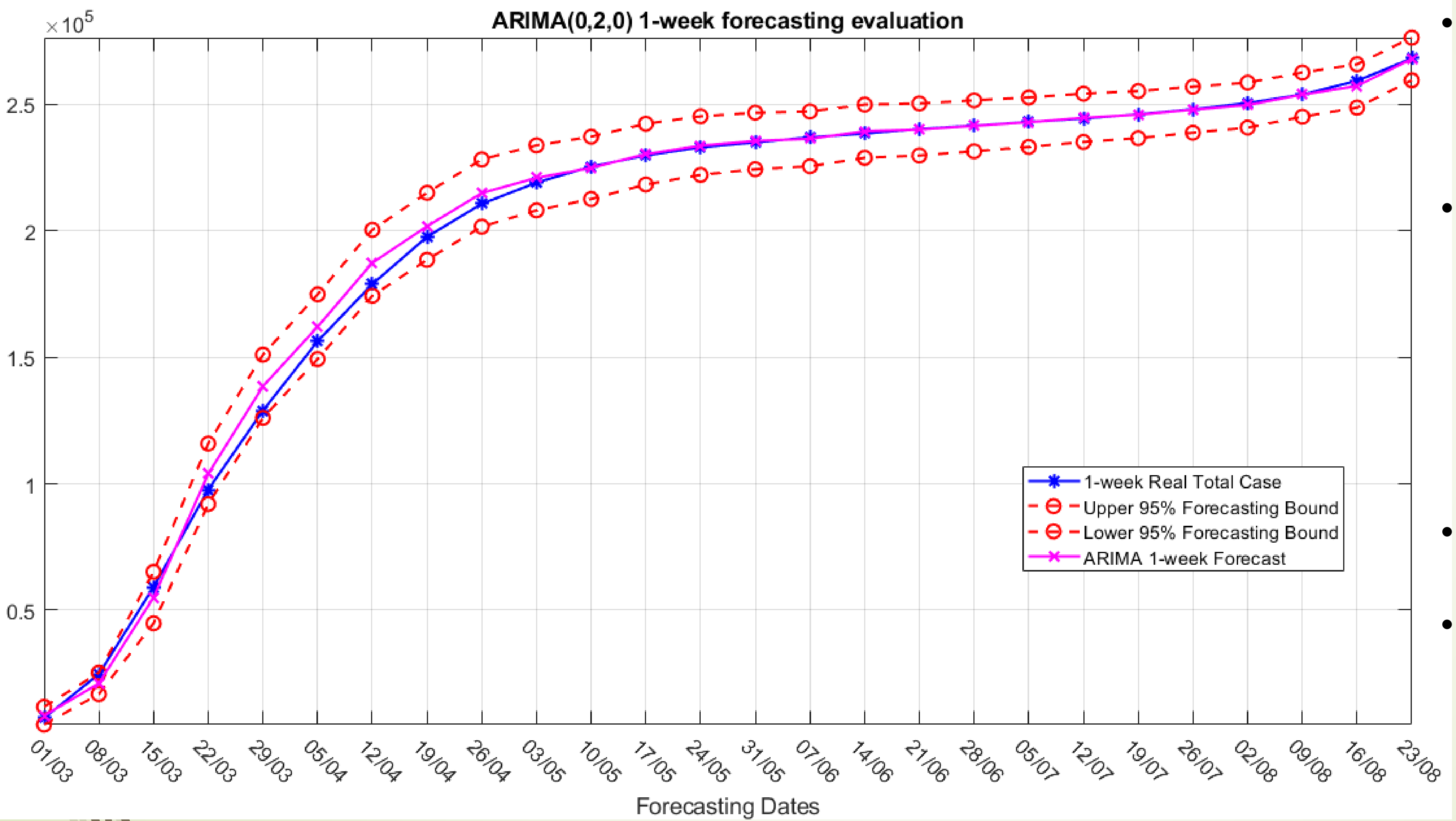
# ARIMA(0,2,0) vs. ARIMA(2,2,2)



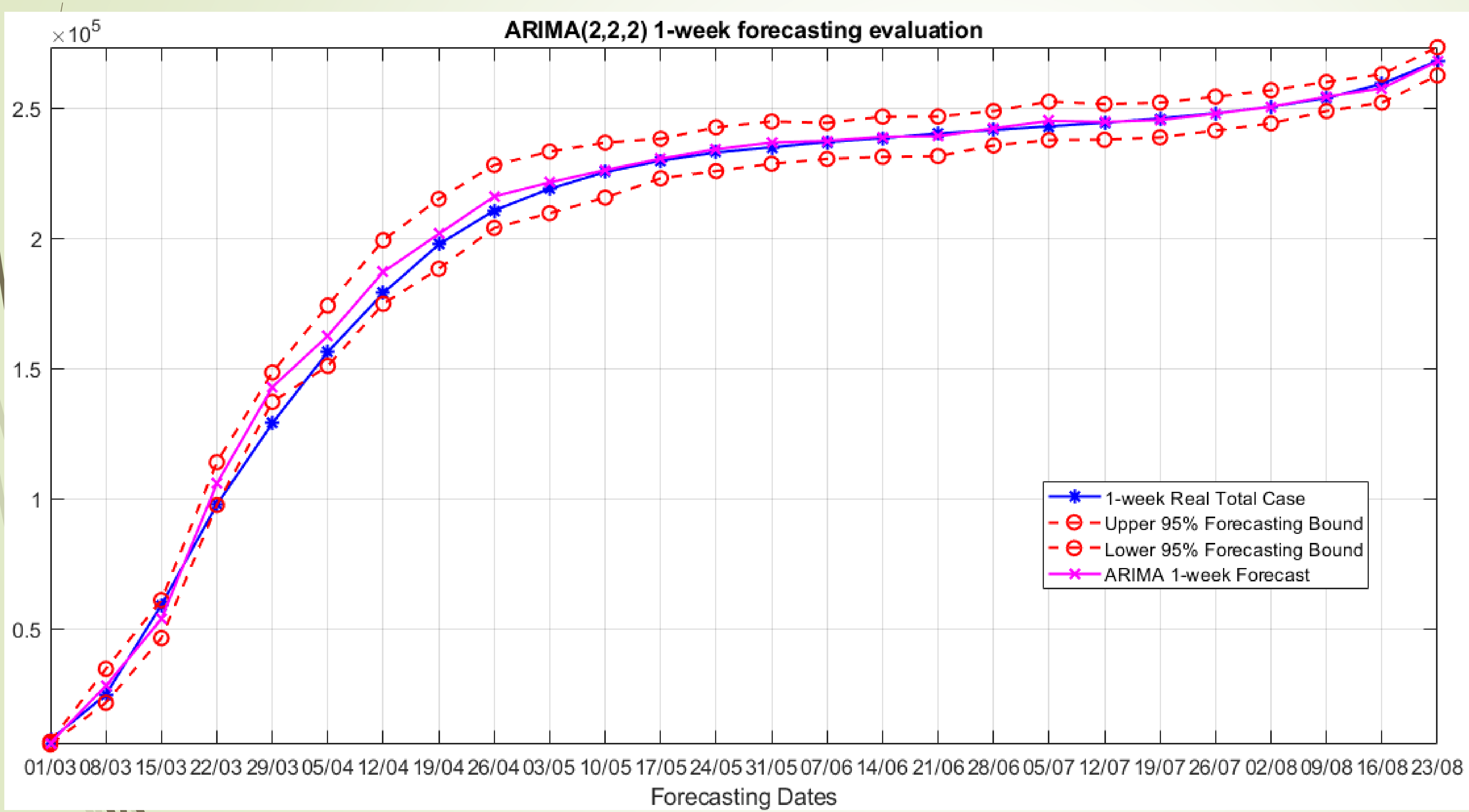

Uncertainties are quite different! Which is the “true” one? Note: forecasting ability is similar
We must check **actual performance of forecast methods in repeated, real forecasts**

# ARIMA(0,2,0) forecasts 1 week ahead




- In this graph, at 01/03 we see the forecast for 08/03 (1 week after), etc.
- The graph shows:
  - the *real* number of cases
  - the *forecasted* number of cases
  - the “uncertainty”
- The forecast appears nice
- Uncertainty band is “realistic” at the beginning, then **clearly too pessimistic**

# ARIMA(2,2,2) forecasts 1 week ahead




Quite similar results.

Maximum absolute error:

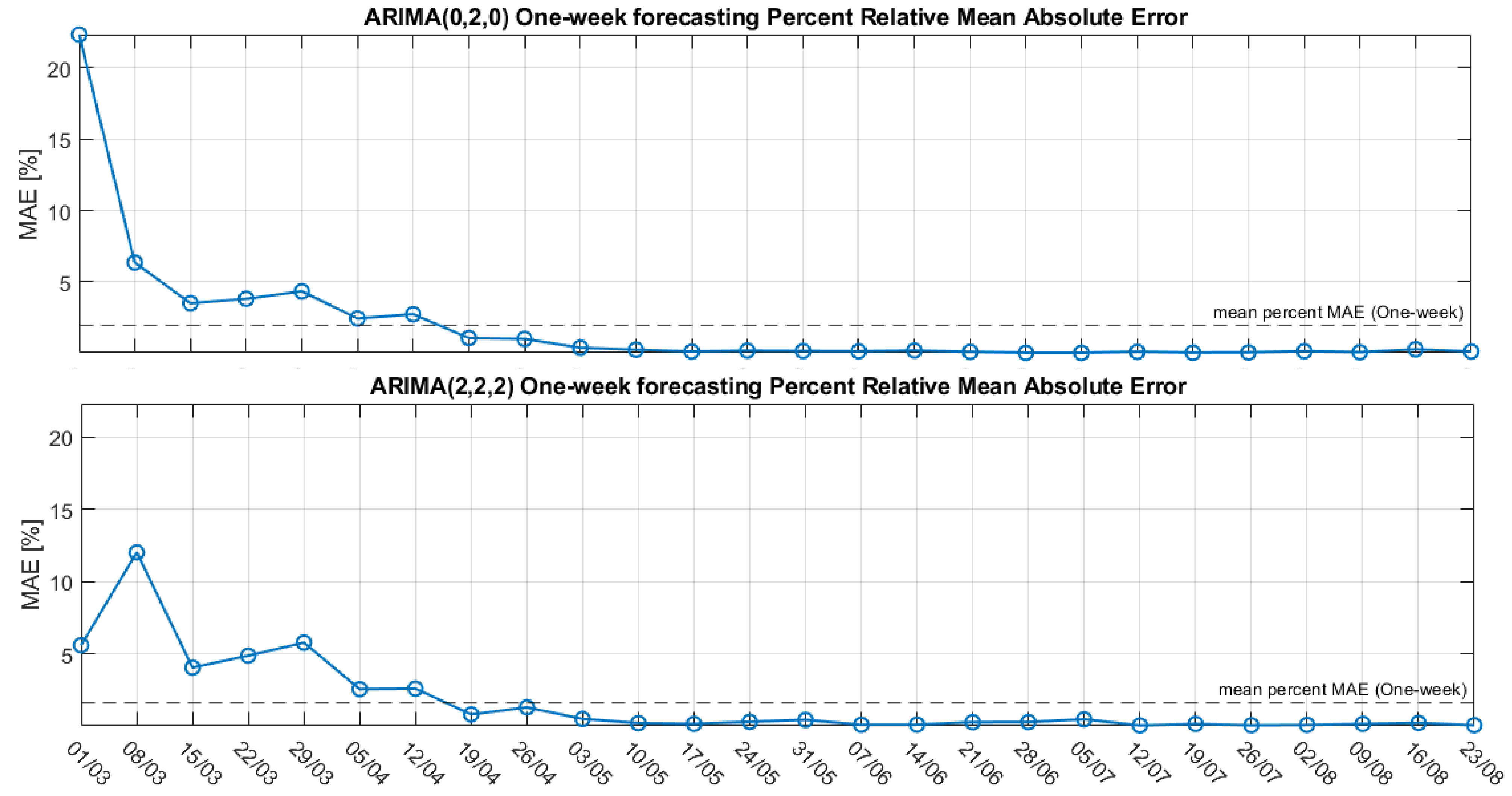
ARIMA(0,2,0)
**MaxAE = 28%**

ARIMA(2,2,2)
**MaxAE = 16%**

It is, in a sense, a **"worst-case" uncertainty**

Maximum error is too pessimistic to assess the whole forecast algorithm

# Forecast MAE, 1 week ahead (relative errors)

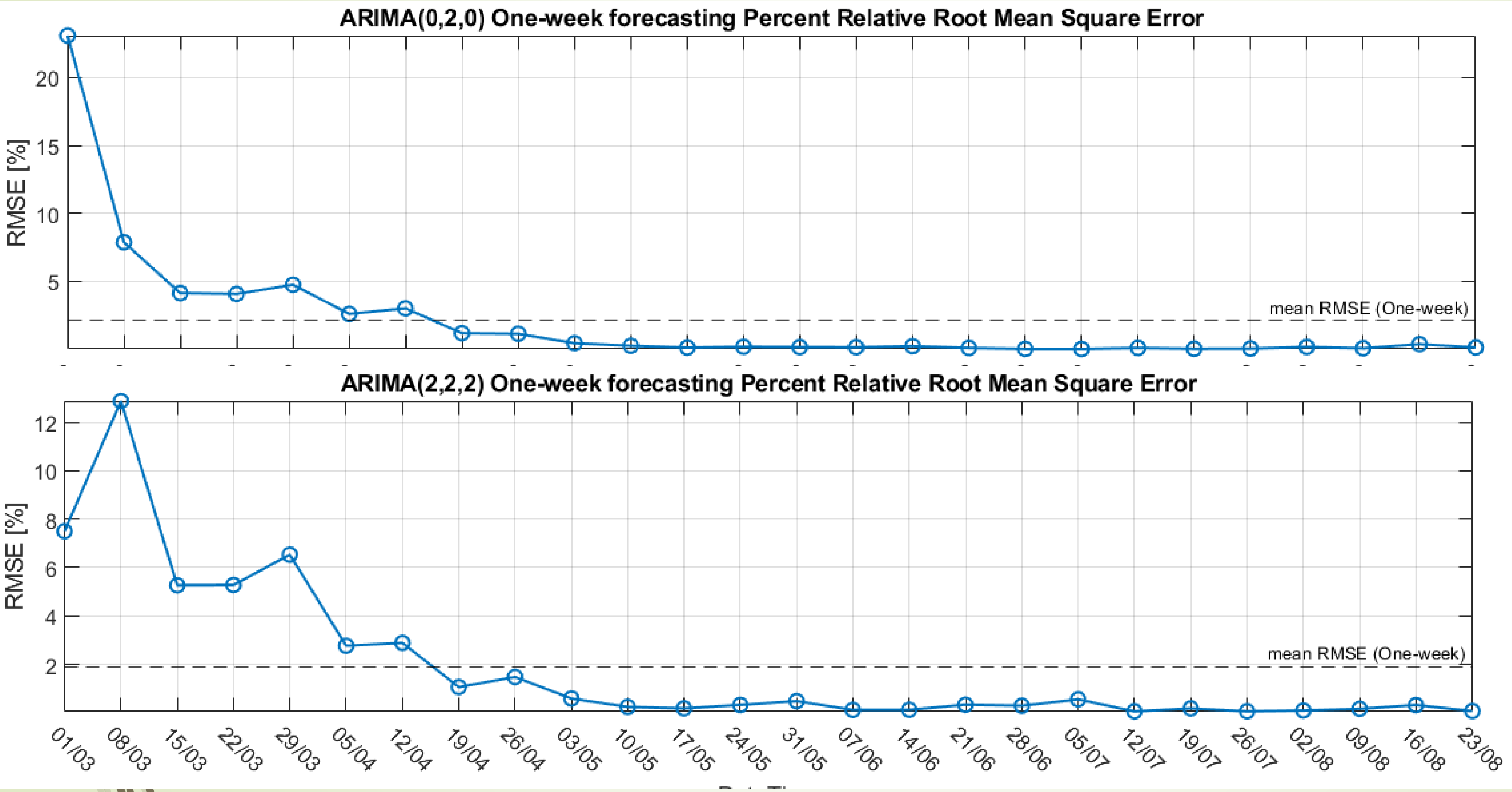



Graph reports the Mean Absolute Error (MAE), in each 1-week forecast

**It is an assessment of the forecasting method, different from the "theoretical" uncertainty**

mean % MAE:
**ARIMA(0,2,0): 1.9%**
**ARIMA(2,2,2): 1.7%**

# Forecast RMSE, 1 week ahead (relative errors)

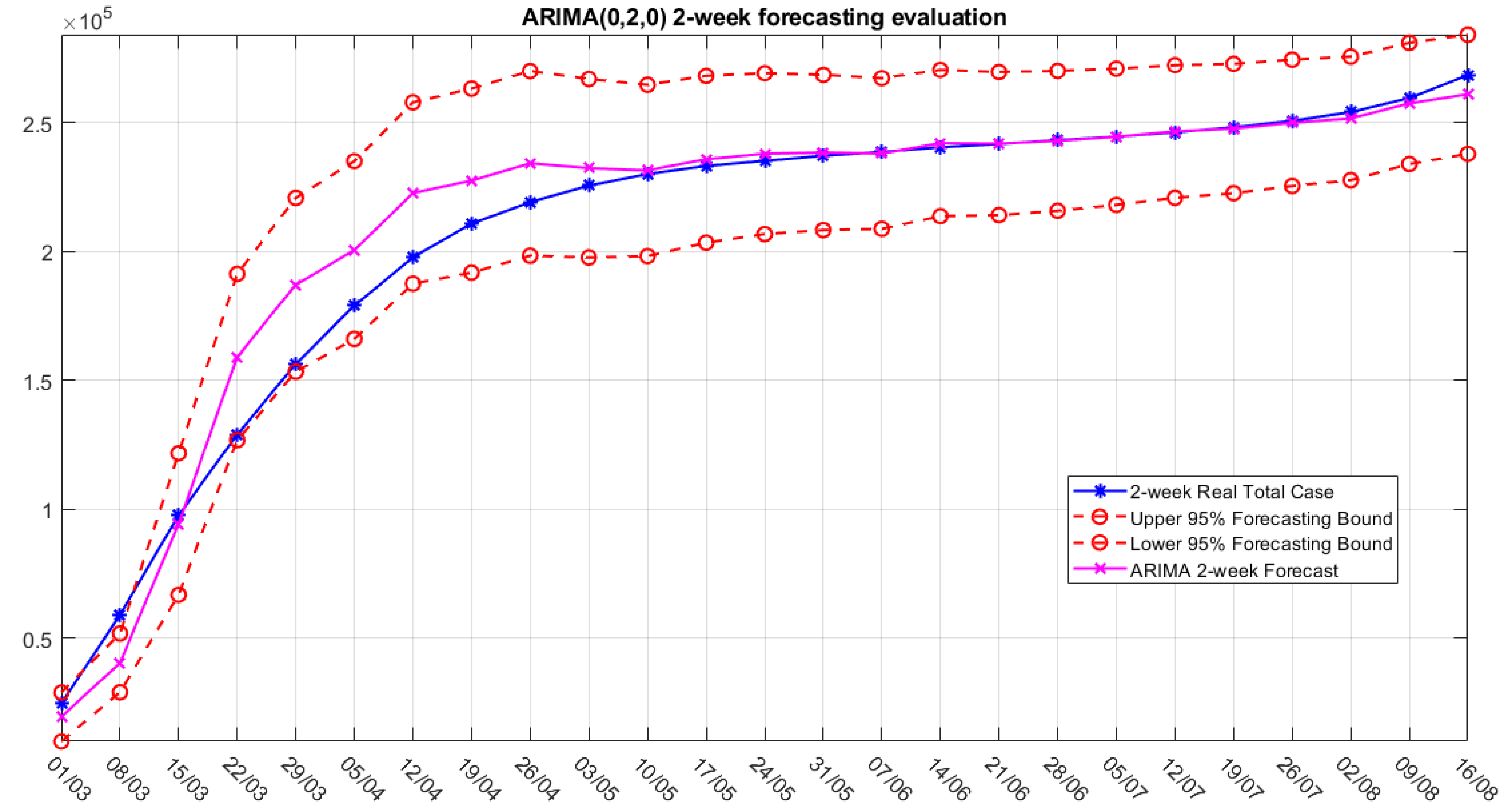



Graph reports the Root Mean Squared Error (RMSE), in each 1-week forecast

**It is another assessment of the forecasting method, in a way similar to a type A (experimental) standard uncertainty**

mean % RMSE:
**ARIMA(0,2,0): 2.1%**
**ARIMA(2,2,2): 1.9%**

# ARIMA(0,2,0) forecasts 2 weeks ahead




**01/03 = forecast for 15/03** (2 week after), etc.

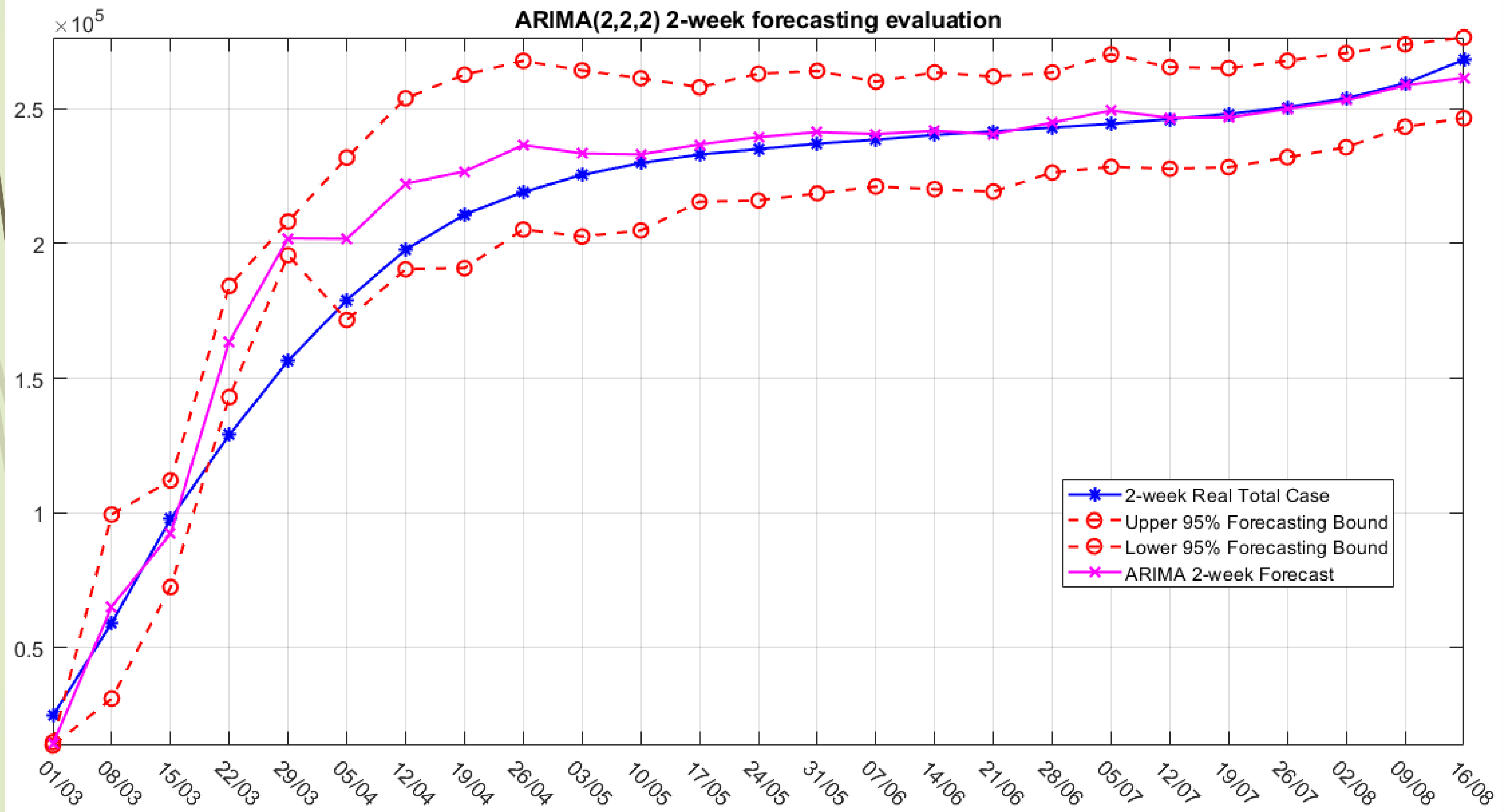
Forecast errors are, of course, larger

Again, **the "theoretical" uncertainty band becomes too pessimistic in the second part of the time axis**

Maximum absolute error:

MaxAE = 32%

# ARIMA(2,2,2) forecasts 2 weeks ahead

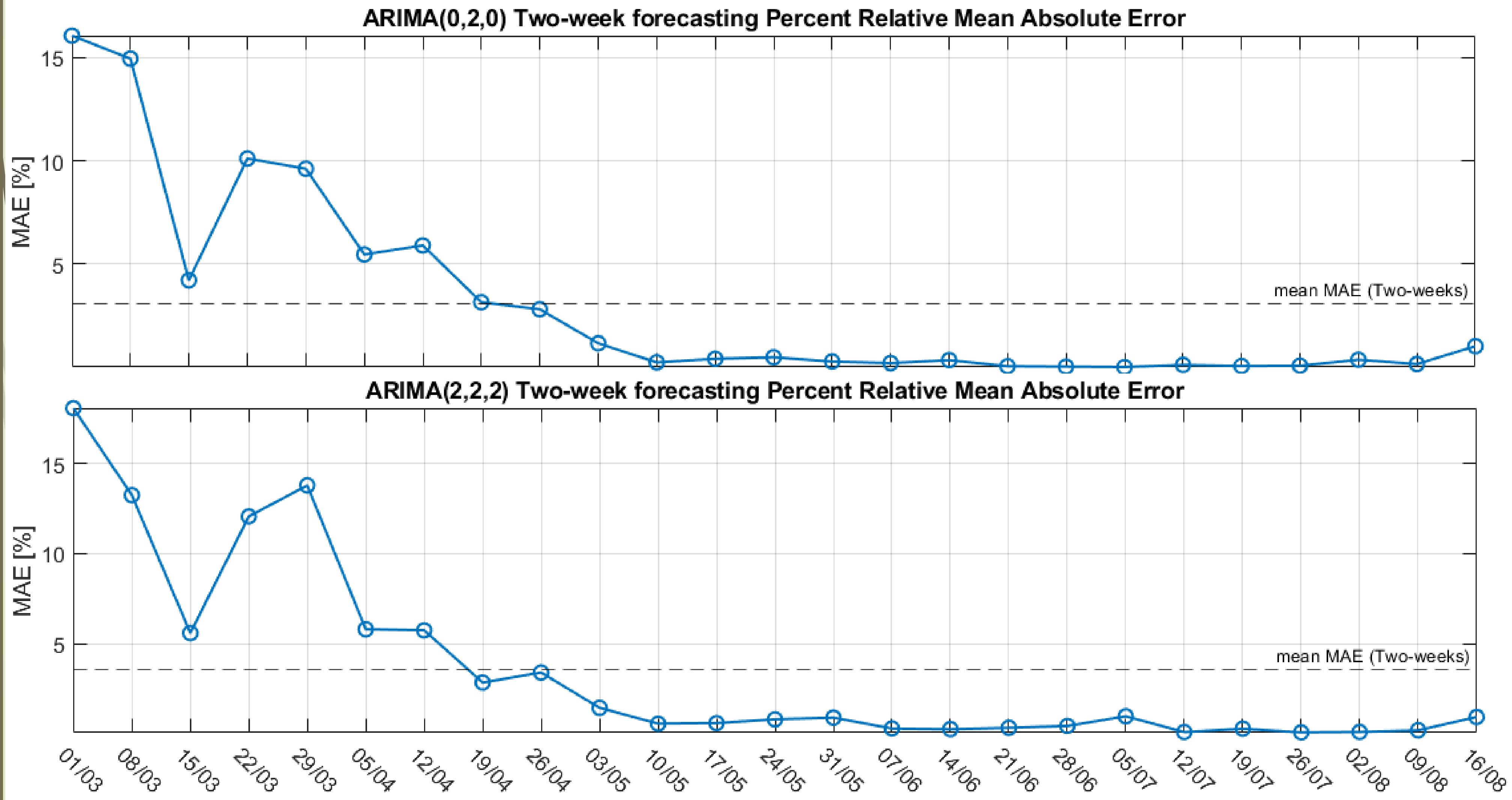



Maximum absolute error:

ARIMA(0,2,0)
**MaxAE = 32%**

ARIMA(2,2,2)
**MaxAE = 42%**

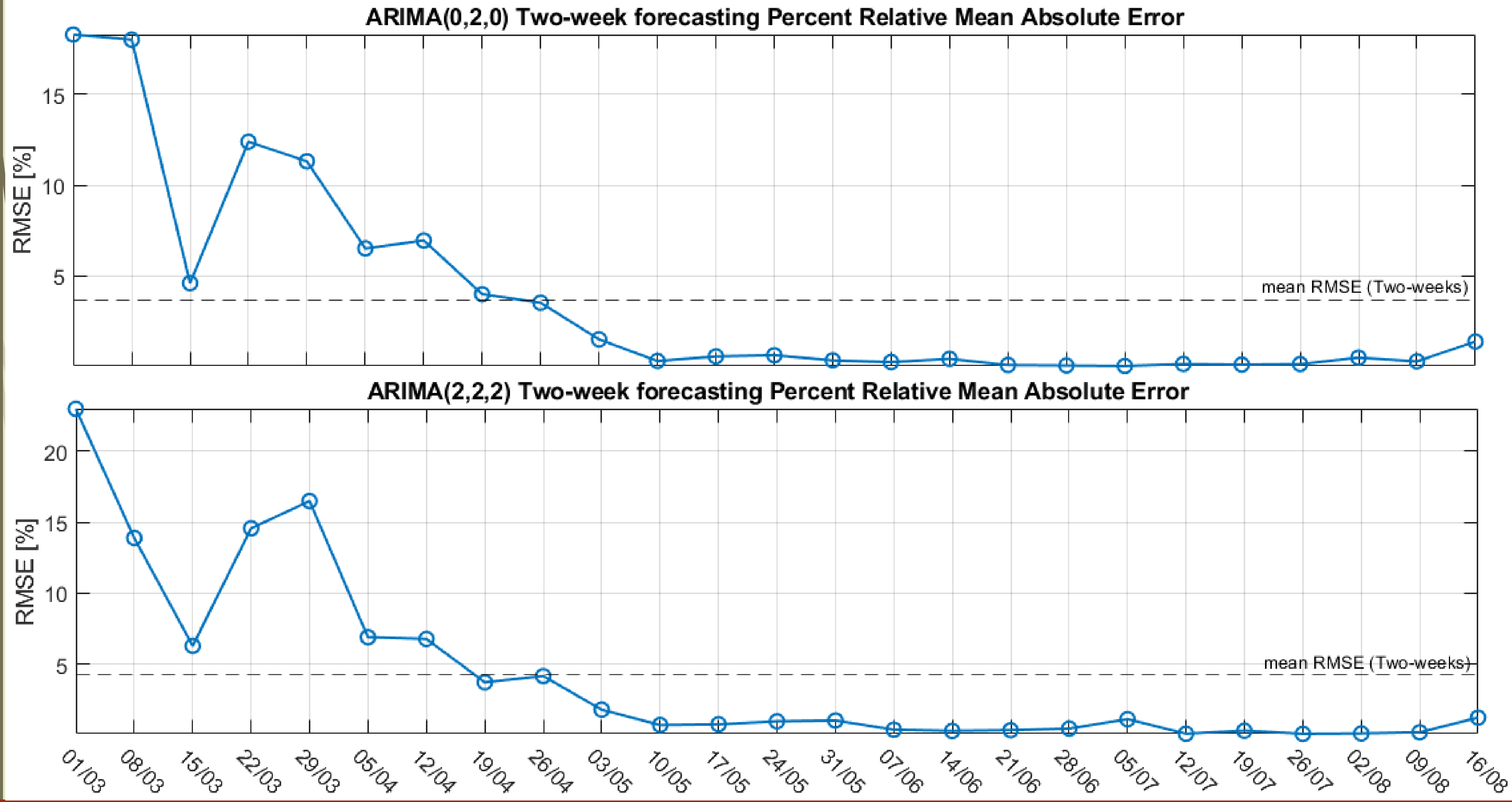
# Forecast MAE, 2 weeks ahead (relative errors)




mean % MAE:

**ARIMA(0,2,0): 3.1%**
**ARIMA(2,2,2): 3.6%**

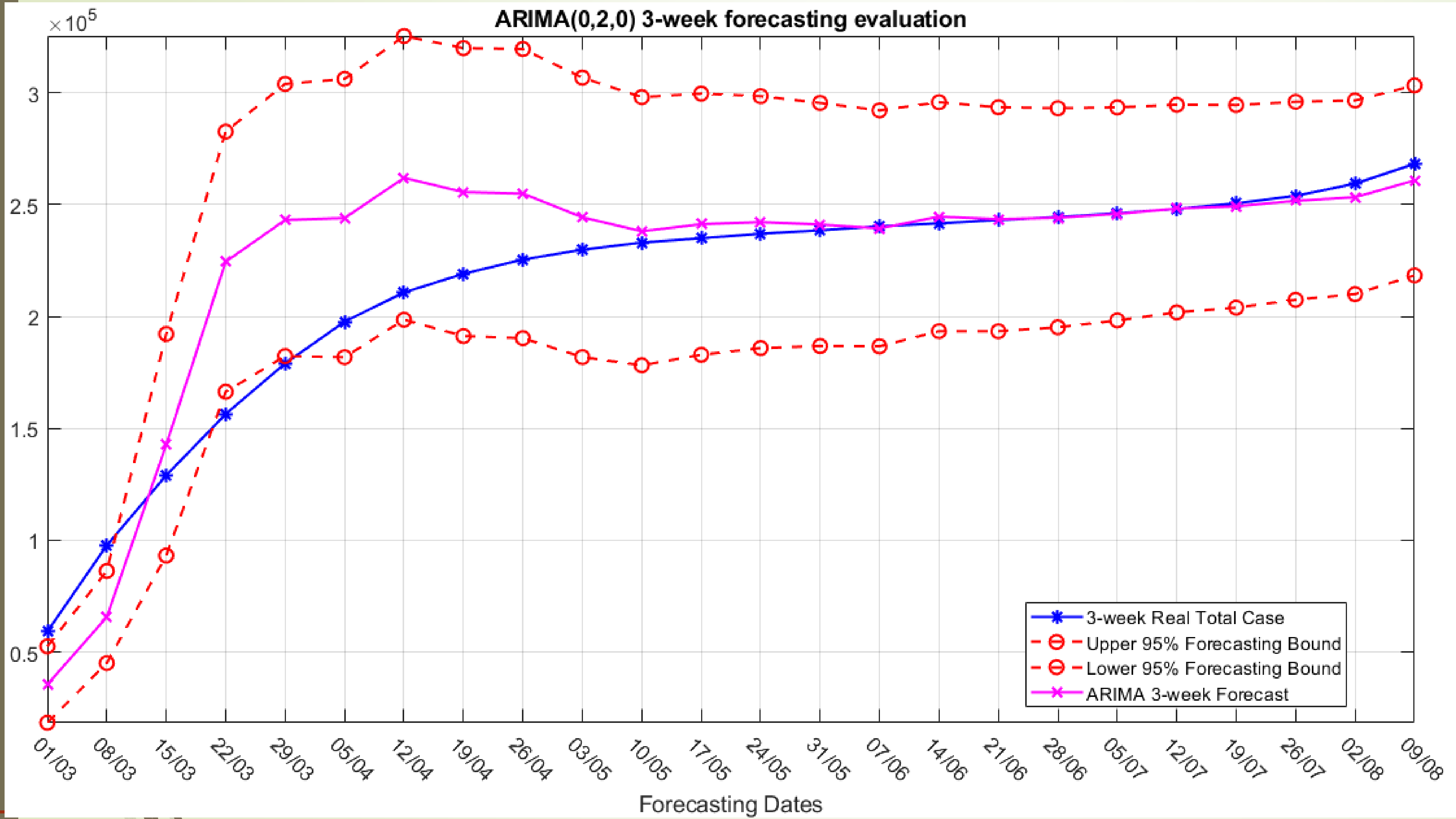
# Forecast RMSE, 2 weeks ahead (relative errors)




mean % RMSE:

**ARIMA(0,2,0): 3.7%**
**ARIMA(2,2,2): 4.2%**

# ARIMA(0,2,0) forecasts 3 weeks ahead




**01/03 = forecast for 22/03** (3 weeks after), etc.

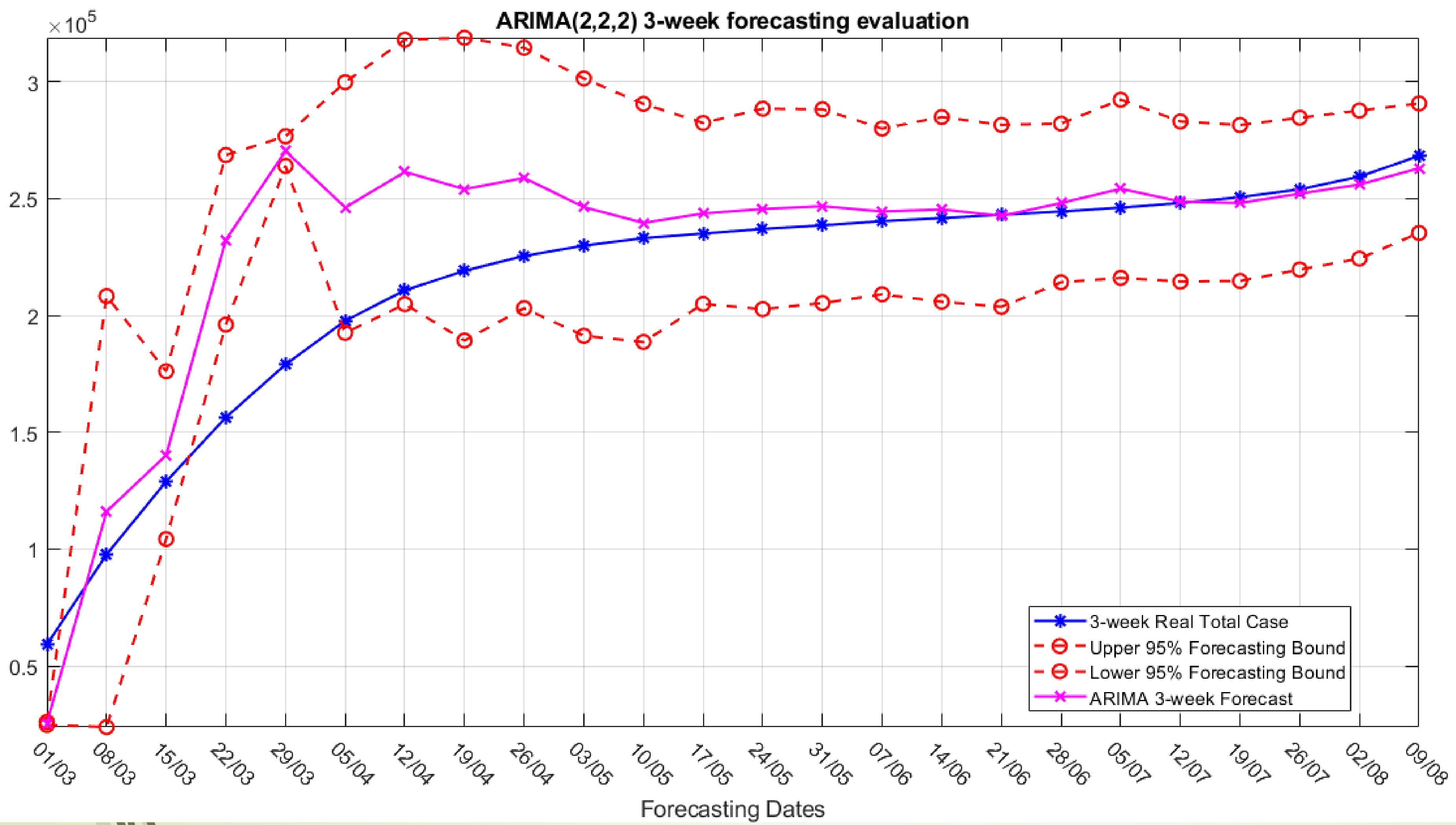
**Forecast errors are initially large**, then they become acceptable

Errors increase in August: (the algorithm was not informed about new social distancing conditions)

Maximum absolute error:

MaxAE = 44%

# ARIMA(2,2,2) forecasts 3 weeks ahead

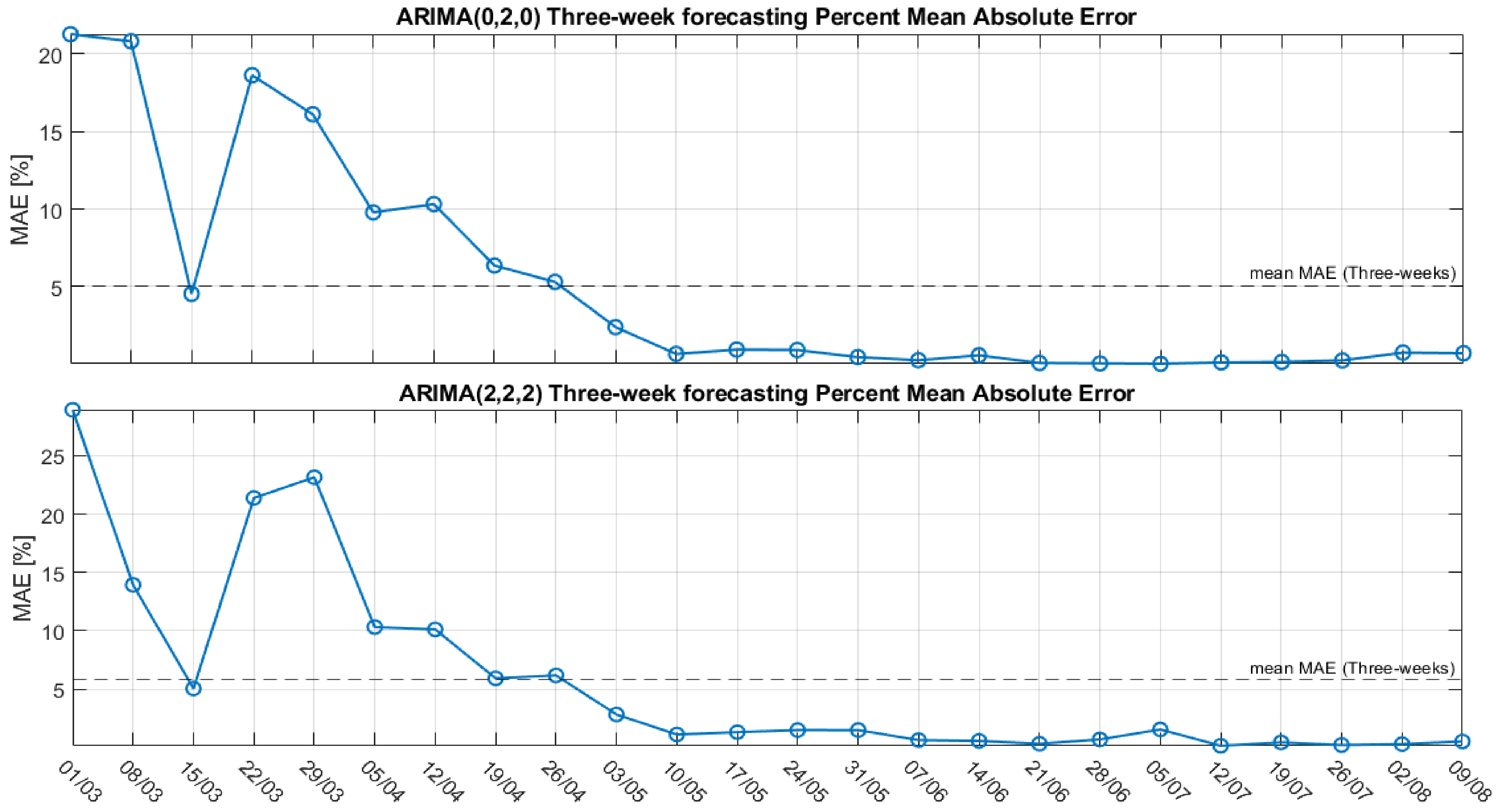



Maximum absolute error:

ARIMA(0,2,0)
**MaxAE = 44%**

ARIMA(2,2,2)
**MaxAE = 57%**

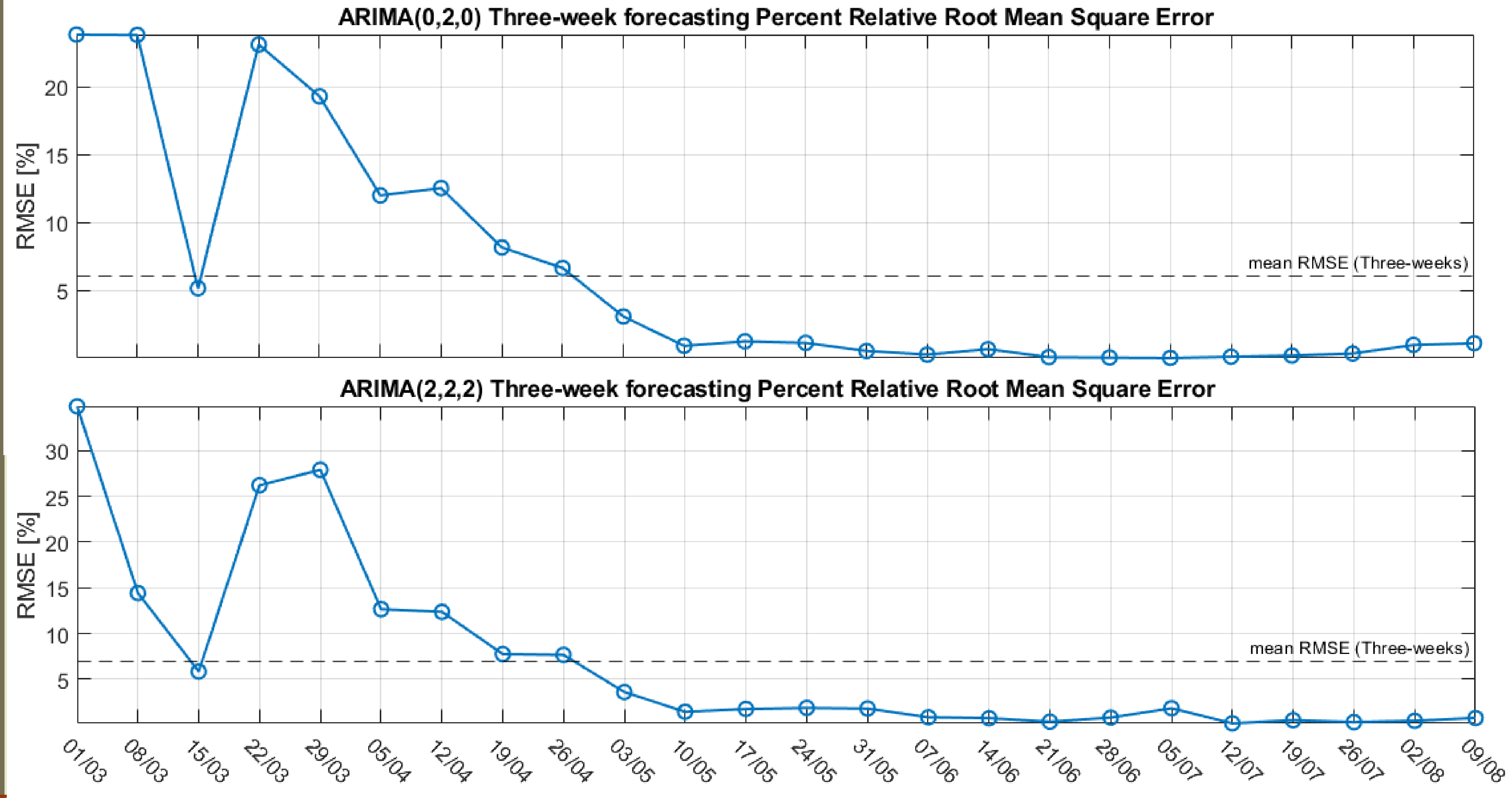
# Forecast MAE, 3 weeks ahead (relative errors)




mean % MAE:

**ARIMA(0,2,0): 5.1%**
**ARIMA(2,2,2): 5.8%**

# Forecast RMSE, 3 weeks ahead (relative errors)




mean % RMSE:

**ARIMA(0,2,0): 6.1%**
**ARIMA(2,2,2): 7.0%**

# Summing up…

In general terms:
ARIMA(0,2,0) seems better when there are slow variations
ARIMA(2,2,2) seems better when there are faster variations

MaxAE %

| Model | 1 week | 2 weeks | 3 weeks |
|---|---|---|---|
| ARIMA(0,2,0) | 28% | 32% | 44% |
| ARIMA(2,2,2) | 16% | 42% | 57% |

MAE %

| Model | 1 week | 2 weeks | 3 weeks |
|---|---|---|---|
| ARIMA(0,2,0) | 1.9% | 3.1% | 5.1% |
| ARIMA(2,2,2) | 1.7% | 3.6% | 5.8% |

RMSE %

| Model | 1 week | 2 weeks | 3 weeks |
|---|---|---|---|
| ARIMA(0,2,0) | 2.1% | 3.7% | 6.1% |
| ARIMA(2,2,2) | 1.9% | 4.2% | 7.0% |

# Conclusions

Avoiding a lengthy summary…

# Four points about uncertainty

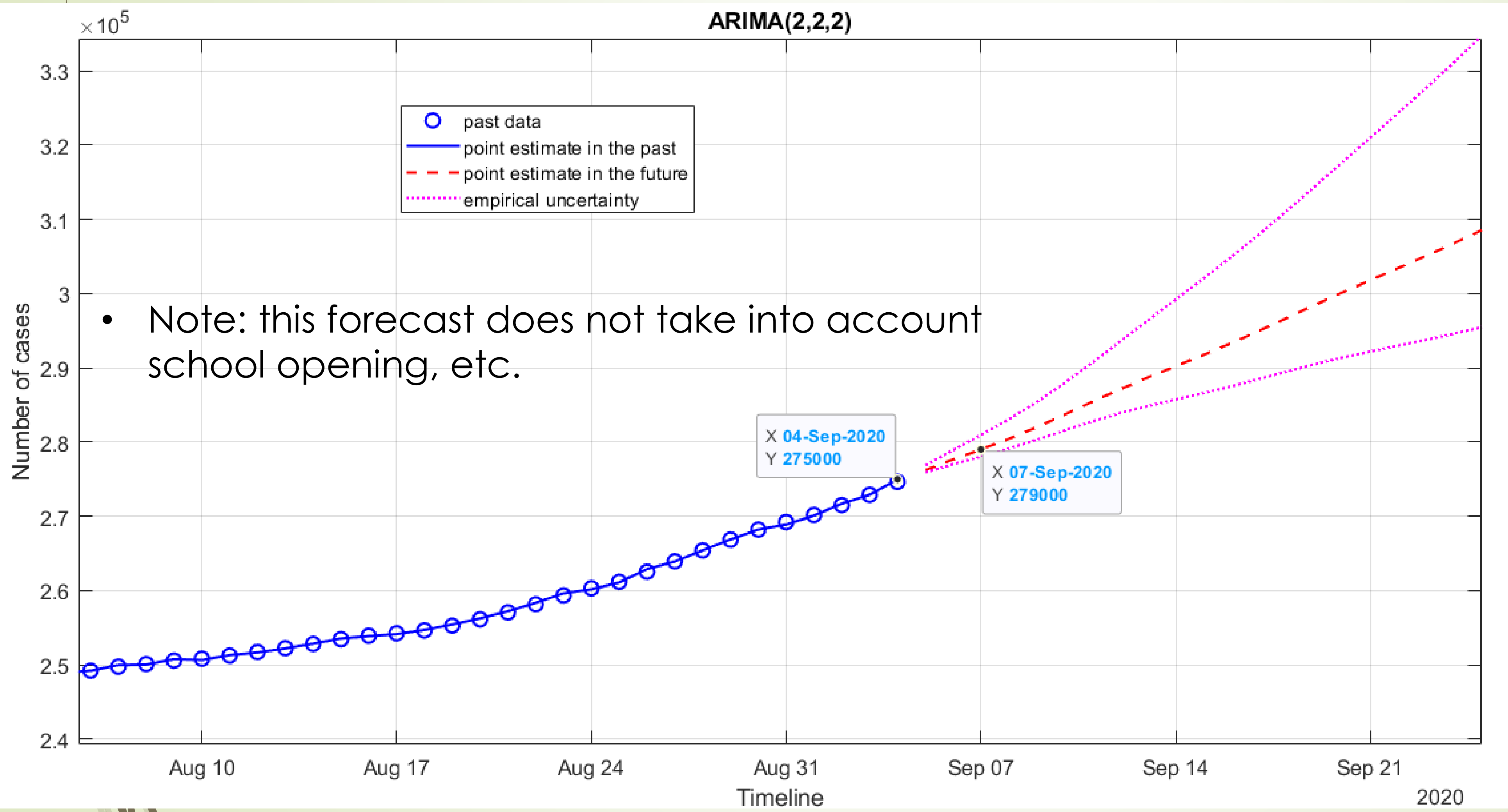
1. **Evaluate uncertainty**
   1. Forecasts / measurements without uncertainty: how much are they worth?
   2. There are **many possible ways**: don't be too dogmatic (but GUM is the key reference)

2. You may want to **check the uncertainty computation with MC simulation**
   1. It is possible and often convenient, not mandatory

3. Check uncertainty **against reality**, whenever possible
   1. If you have reference data, it is dutiful

4. Always check uncertainty **against common sense**
   1. Your mathematical model may mislead you

# What I would like to see in the news




- Note: this forecast does not take into account school opening, etc.

- **Reliable forecasts for the next week, with uncertainty**
- Tentative forecasts for the next two, three weeks
- No forecasts as mere opinions (if not "divination"), nor "posterior forecasts"
- Let's call for **more metrology in societal issues**!

# Thank you for your attention!