Data Assimilation

The Story of the Kalman People

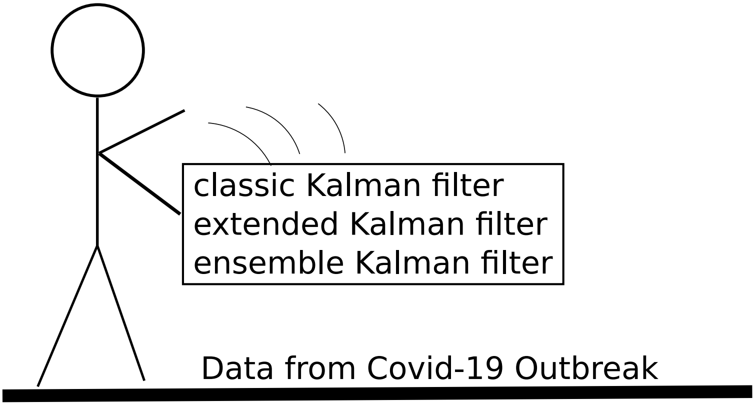
A. P. Braun, I. Kottlarz, K. H. Mok, N. Weiß

June 30, 2022

Content

Goals

The Vision of the Kalman People



Quick reminder:

- $x^f = \mathcal{M}(x^a)$
- $x^a = x^f + K[y - \mathcal{H}(x^f)]$

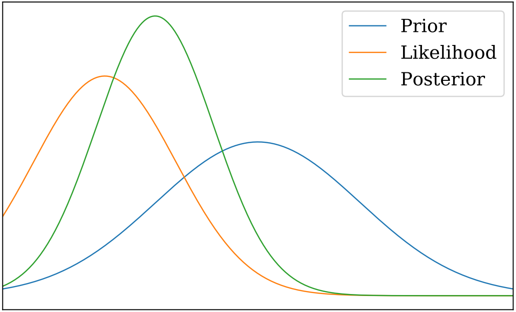


Figure: Bayes formula

The models we used

Models

The Lorenz9 model [RLS+98]

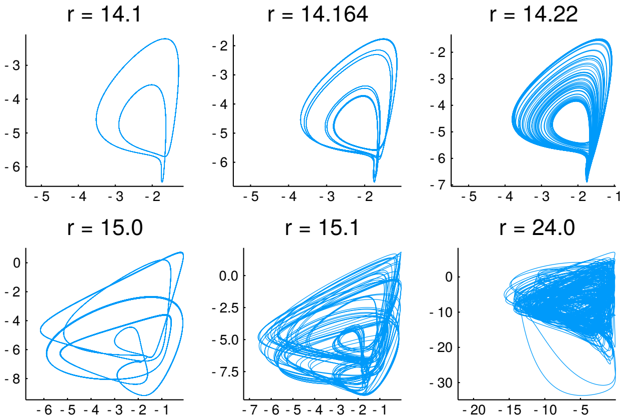
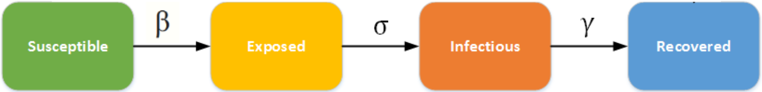


Figure: The C_6-C_7-plane where $t \in (2500, 3000]$, $u_0 = [0.01, 0, 0.01, 0, 0, 0, 0, 0, 0.01]$, $\sigma = a = 0.5$.

Models

The SEIR-Model

classic		our adjustment
$\dot{S} = -\beta \cdot S \cdot I / N$	$\rightarrow$	$\dot{S} = -\beta \cdot S \cdot E / N$
$\dot{E} = \beta \cdot S \cdot I / N - \sigma \cdot E$	$\rightarrow$	$\dot{E} = \beta \cdot S \cdot E / N - \sigma \cdot E$
$\dot{I} = \sigma \cdot E - \gamma \cdot I$		
$\dot{R} = \gamma \cdot I$		



Models

The SEIR-Model

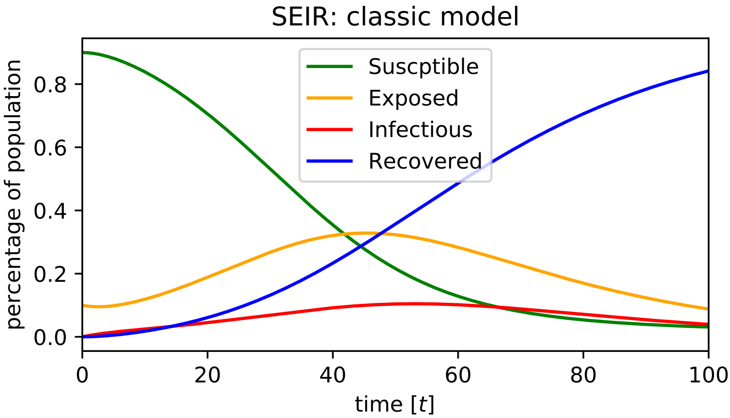


Figure: Plot of the SEIR-model with $\beta = 0.5$, $\sigma = 1/24$ and $\gamma = 1/4$.

Models

Our SEIR-Model

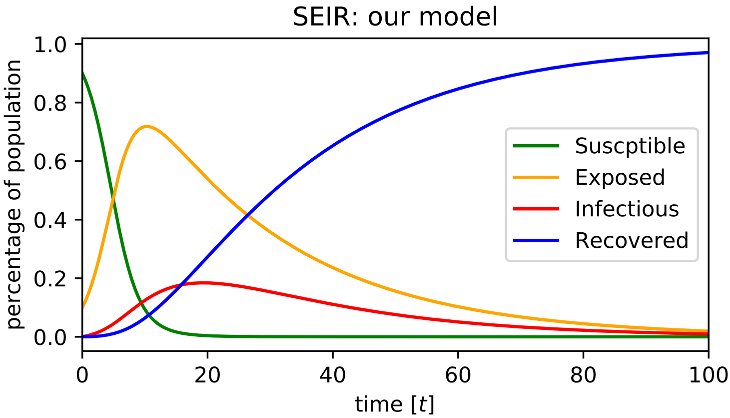


Figure: Plot of our SEIR-model with $\beta = 0.5$, $\sigma = 1/24$ and $\gamma = 1/4$.

Handling Real World Data

Handling Real World Data

Challenges

- Easily accessible source for data
  - $\rightsquigarrow$ github! [Uni, DXY]
- Estimation of covariance matrices
- Which models describe the observations?
- What model parameters fit the observations?

Handling Real World Data

Some examplary plots

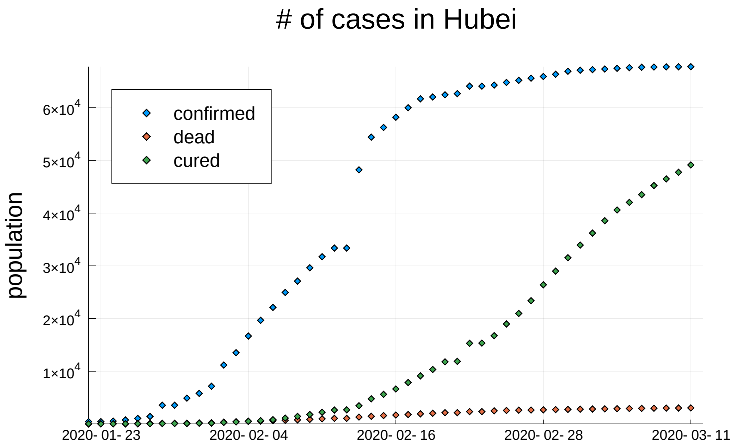


Figure: Covid-19 Cases in Hubei

Handling Real World Data

Some examplary plots

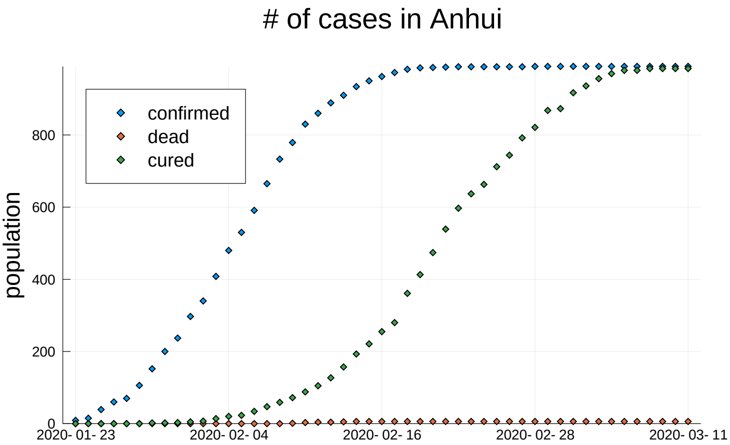
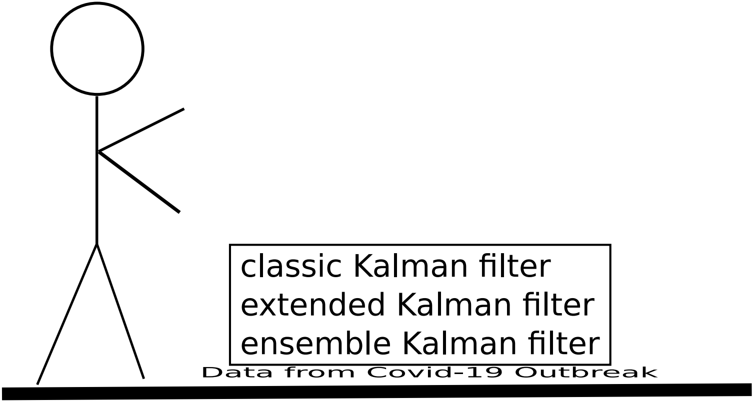


Figure: Covid-19 Cases in Anhui

Filters

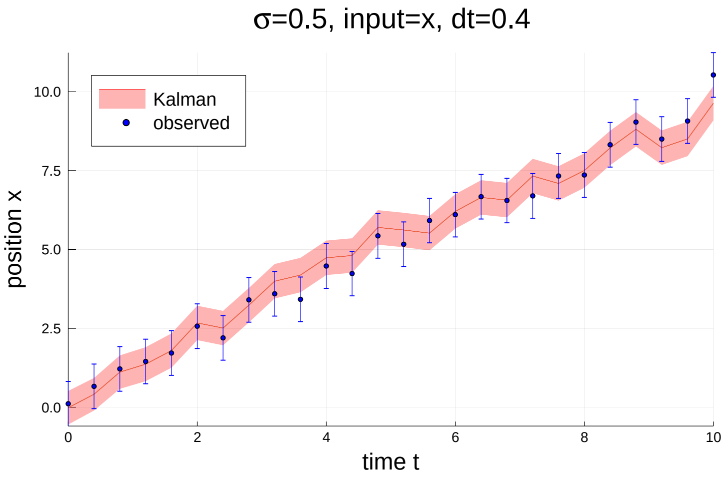
The Rise of the Kalman People



The classic Kalman filter (CKF)

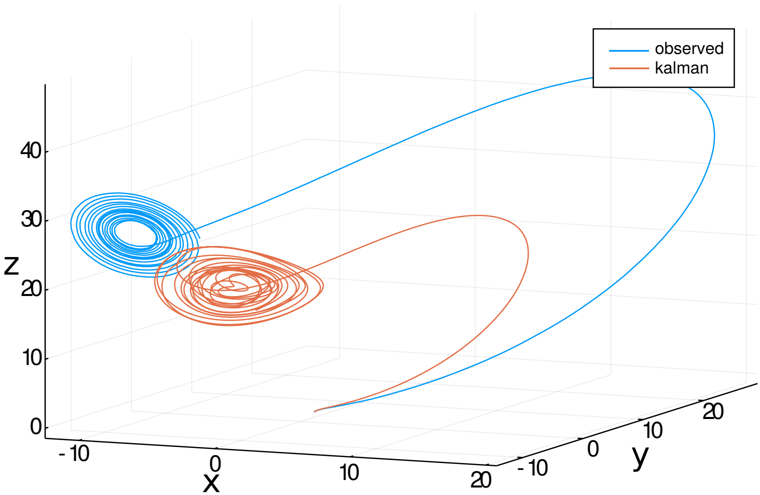
The classic Kalman filter (CKF)

Robot model



The classic Kalman filter (CKF)

lorenz model



The extended Kalman filter (ExKF)

The extended Kalman filter

Lorenz3 - chaotic - 3 variables

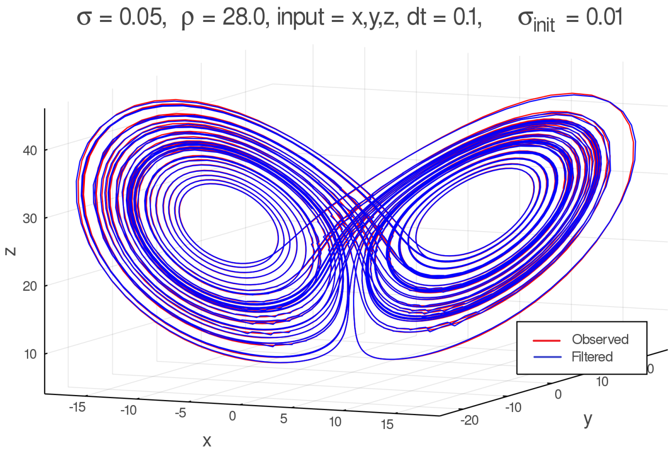


Figure: $t \in (150, 200]$

The extended Kalman filter

Lorenz3 - chaotic - 2 variables

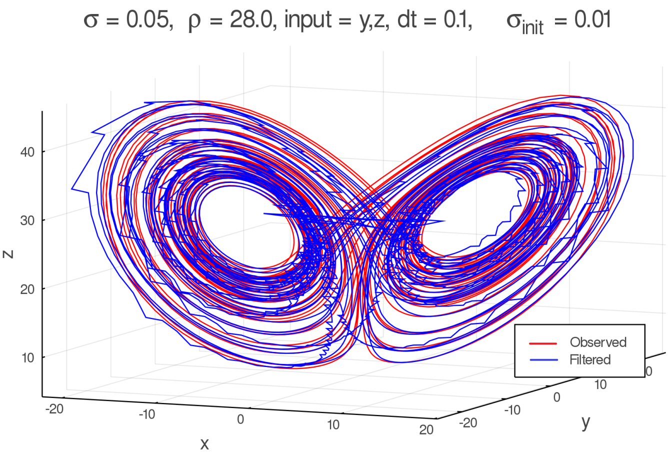


Figure: $t \in (150, 200]$

The extended Kalman filter

Lorenz3 - chaotic - 1 variable

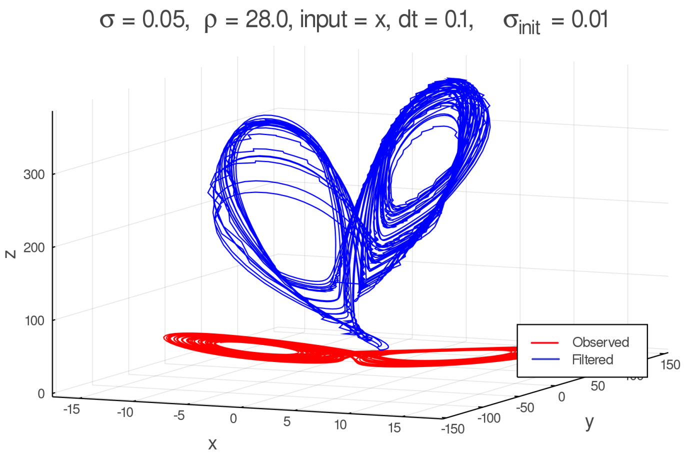


Figure: $t \in (150, 200]$

The extended Kalman filter

Lorenz3 - chaotic - 1 variable

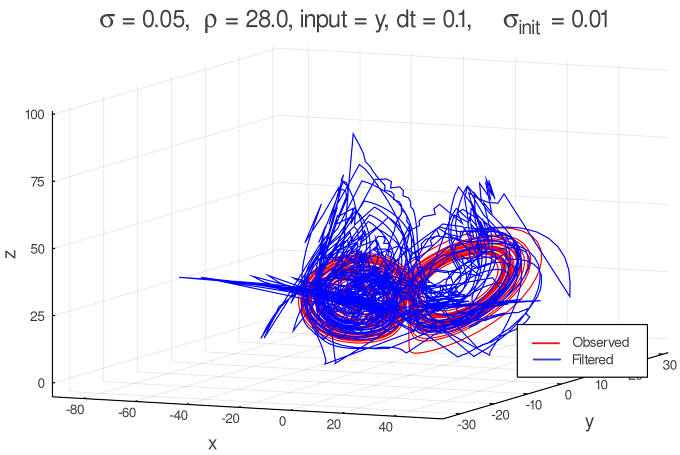


Figure: $t \in (150, 200]$

The extended Kalman filter

Lorenz9 - chaotic - 6 variables

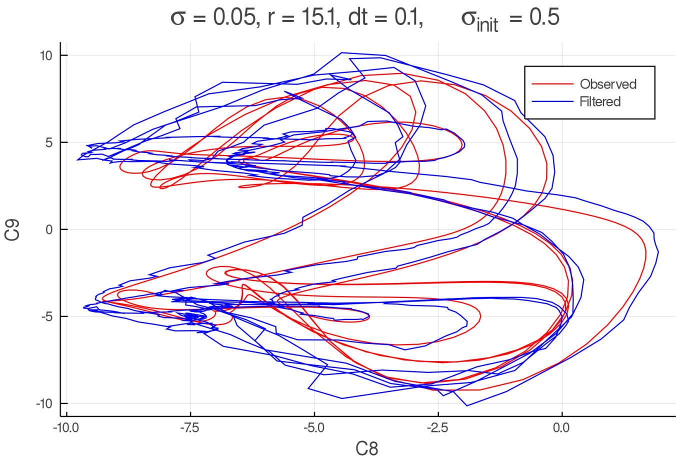


Figure: Input variables: C1-C6

The extended Kalman filter

Lorenz9 - chaotic - 6 variables

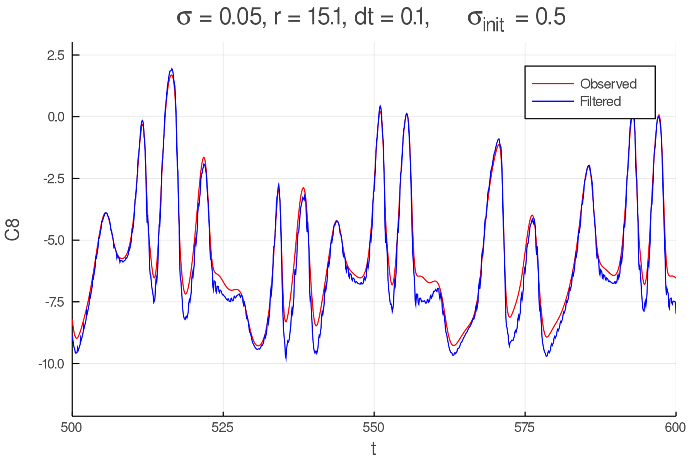


Figure: Input variables: C1-C6

The ensemble Kalman filter

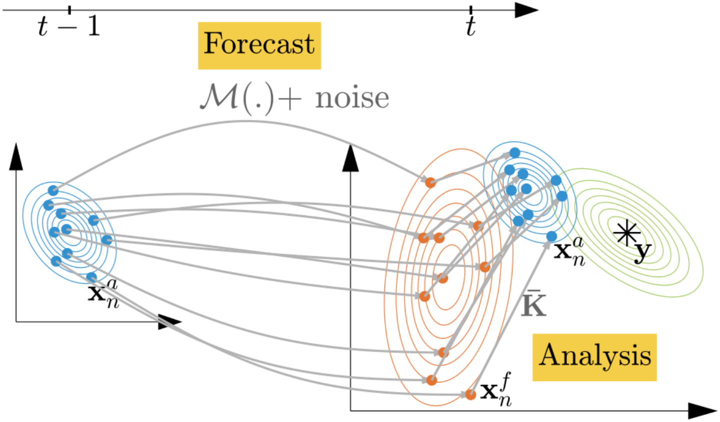
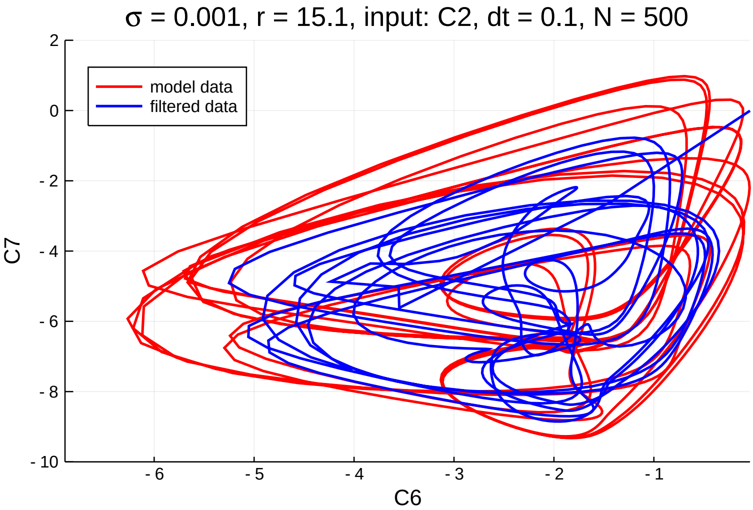


Figure: The principle of the Ensemble Kalman Filter. [Raa19]

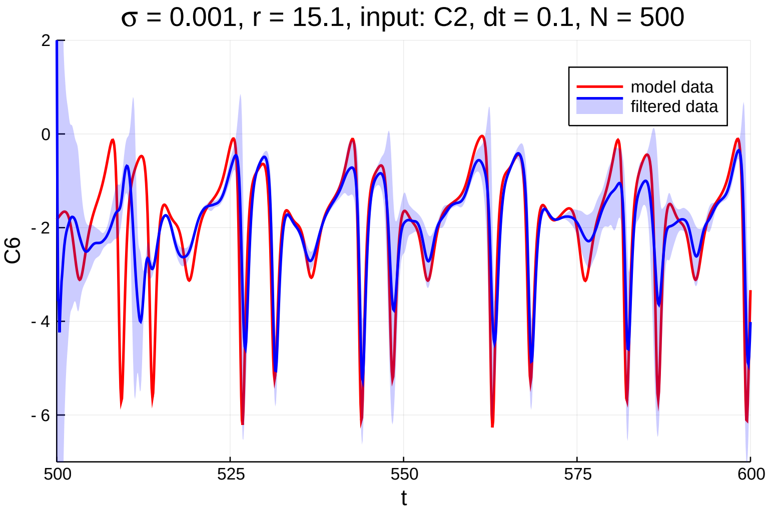
The ensemble Kalman filter

Lorenz9 - chaotic - one variable



The ensemble Kalman filter

Lorenz9 - chaotic - one variable



The ensemble Kalman filter

SEIR for Hubei - doesn't really match

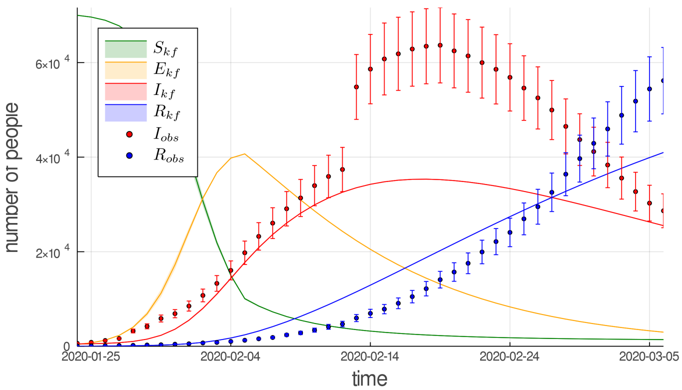
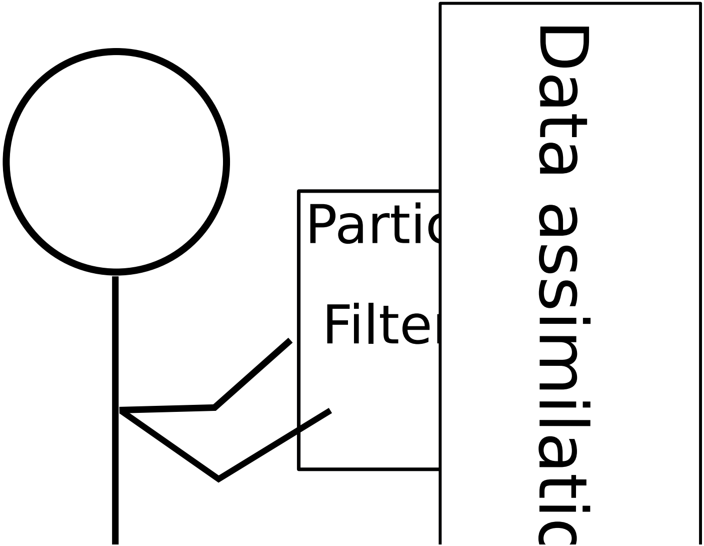


Figure: The data from [DXY] with the adapted SEIR for the province of Hubei, the guessed parameters were $\beta = 0.7(0.3 \text{ for } t > 14)$, $\sigma = 1/10$ and $\gamma = 1/24$. The initial value was $u_0 = [7000, 400, 500, 0]$

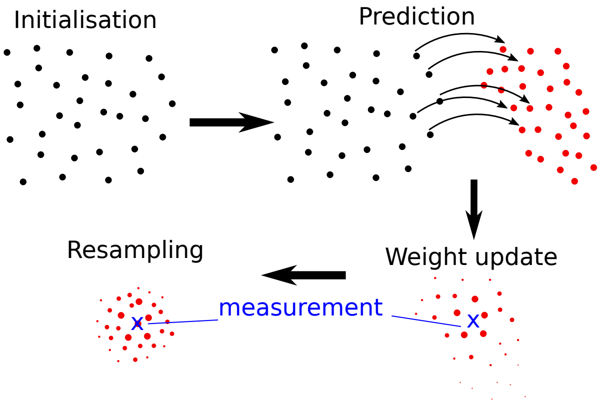
The Particle Filter

Another Love



The Particle Filter

Algorithm



The Particle Filter

Results - the best

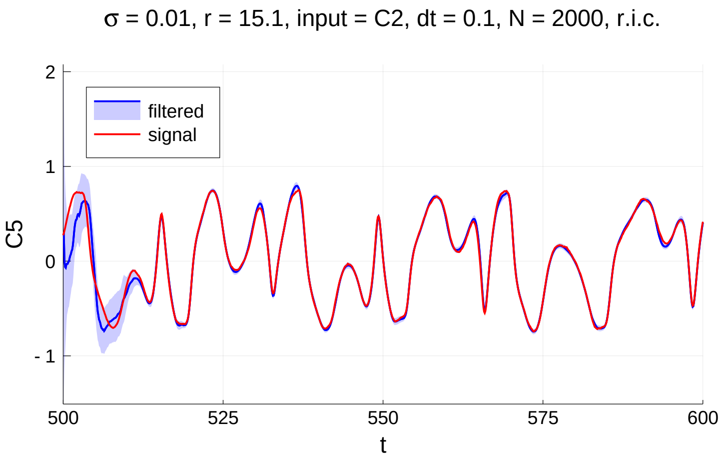


Figure: Particle Filter with unknown initial condition and large noise in chaotic regime. First rough adaptive resampling used.

The Particle Filter

Results - the worst

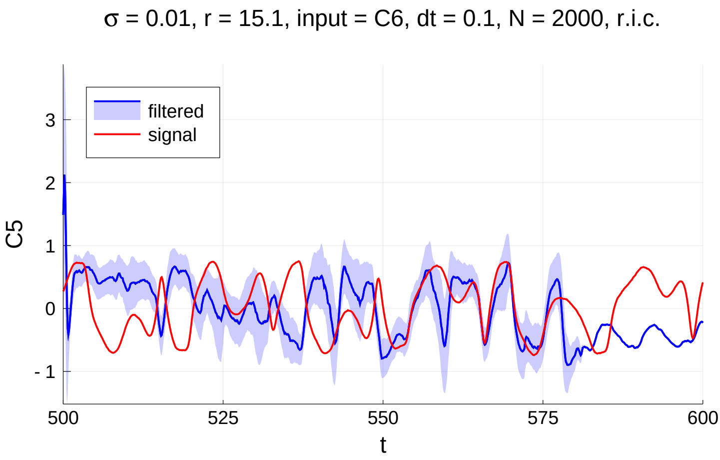
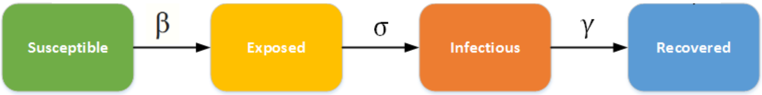


Figure: Particle Filter with unknown initial condition and large noise in chaotic regime. Same rough adaptive resampling used as before.

The Particle Filter
The SEIR-Model

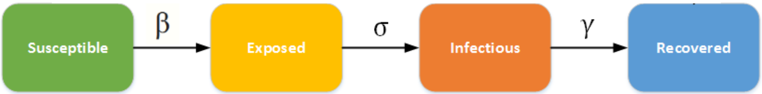
$$\dot{S} = -\beta \cdot S \cdot E/N$$
$$\dot{E} = \beta \cdot S \cdot E/N - \sigma \cdot E$$
$$\dot{I} = \sigma \cdot E - \gamma \cdot I$$
$$\dot{R} = \gamma \cdot I$$



The Particle Filter

The SEIR-Model

$$\dot{S} = -\beta \cdot S \cdot E/N$$
$$\dot{E} = \beta \cdot S \cdot E/N - \sigma \cdot E$$
$$\dot{I} = \sigma \cdot E - \gamma \cdot I$$
$$\dot{R} = \gamma \cdot I$$
$$\dot{\beta} = 0$$
$$\dot{\gamma} = 0$$



The Particle Filter

Parameter estimation of Real World Data

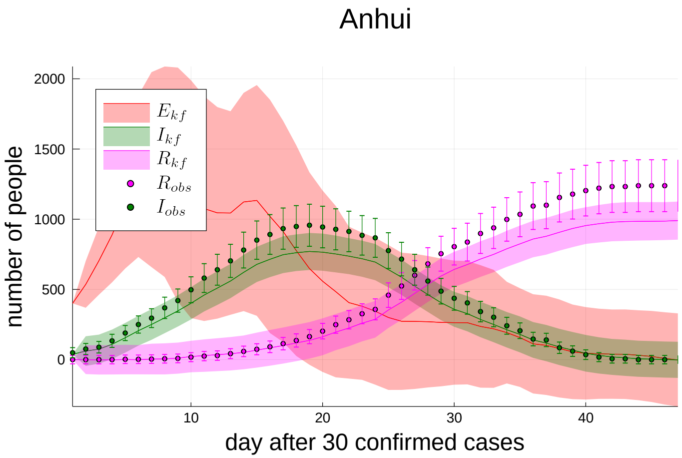


Figure: Plot of observed vs. Kalman-estimated data.

The Particle Filter

Parameter estimation of Real World Data

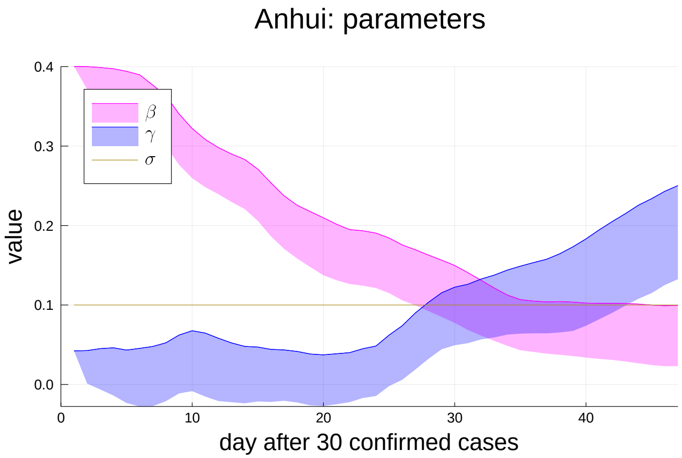
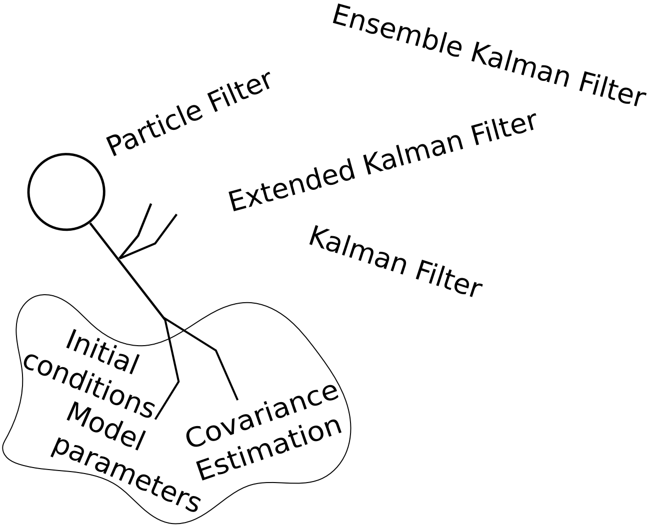


Figure: Plot of the Kalman-filtered parameters

Summary

The Curse of the Real World Data



Summary

Filters with unknown initial condition

	CKF	ExKF	EnKF	ParF
$r = 15.0$, $\#\text{obs} = 9$	$\times$	$\checkmark$	$\checkmark$	$\checkmark$
$r = 15.0$, $\#\text{obs} = 1$	$\times$	$\times$	$[\checkmark]$	$[\checkmark]$
$r = 15.1$, $\#\text{obs} = 9$	$\times$	$\checkmark$	$\checkmark$	$\checkmark$
$r = 15.1$, $\#\text{obs} = 1$	$\times$	$\times$	$[\checkmark]$	$[\checkmark]$

Table: Comparison of the implemented filters on grounds of performance with the Lorenz9-model. Here, the filters were not given the exact inital condition of the model.

Outlook

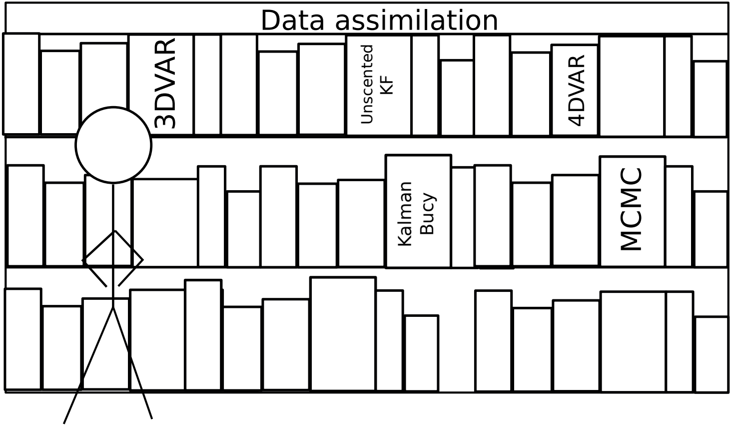
Outlook

Things to be improved or further explored

- EnKF: Runtime
  - $\rightsquigarrow$ EnKF slower than ParF for otherwise similar implementation
- ParF: Optimal noise in resampling
  - $\rightsquigarrow$ dependence on process noise? using multivariate distribution?
- Covariance estimation for sparse data

Outlook

The undiscovered Country



References

DXY git repository, https://github.com/BlankerL/DXY-COVID-19-Data/tree/master/csv, Accessed: June 30, 2022.

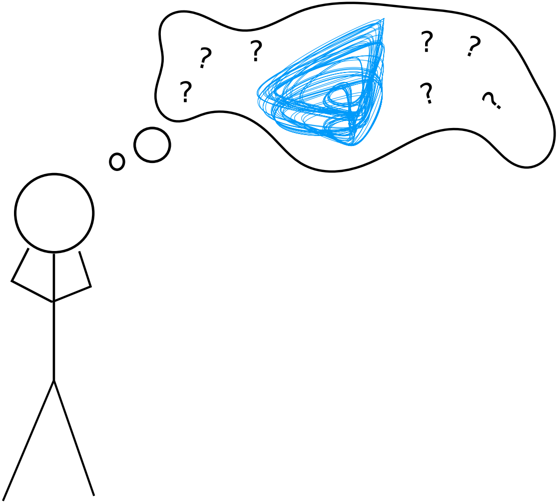
Patrick N. Raanes, *Introduction to data assimilation and the ensemble kalman filter*, 2019.

Peter Reiterer, Claudia Lainscsek, Ferdinand Schürrer, Christophe Letellier, and Jean Maquet, *A nine-dimensional lorenz system to study high-dimensional chaos*, Journal of Physics A: Mathematical and General (1998) **31** (1998), 7121–7139.

John Hopkins University, *John Hopkins CSSE git repository*, https://github.com/CSSEGISandData/COVID-19/tree/master/csse_covid_19_data/csse_covid_19_time_series, Accessed: June 30, 2022.

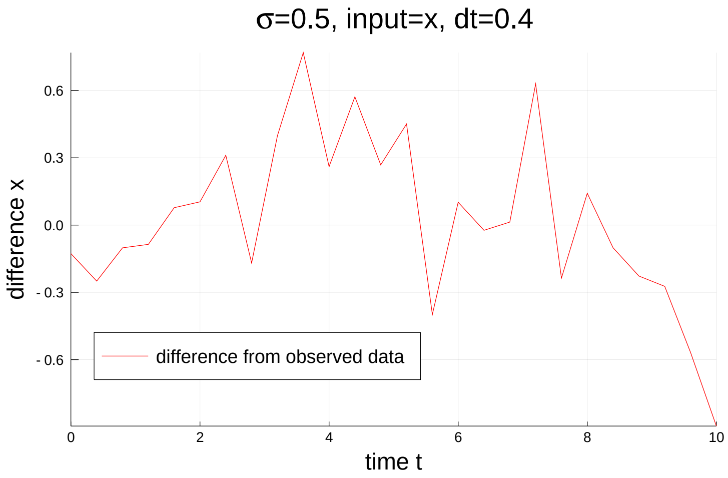
Questions?

The Assault on Reason

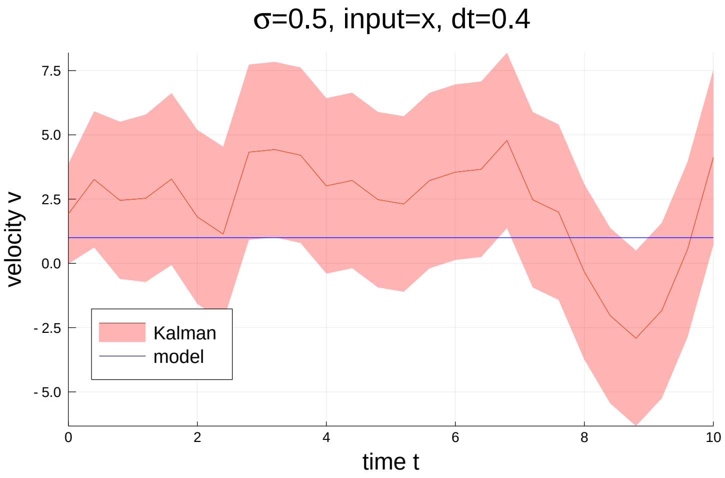


Backup slides

Robot model: difference of position x

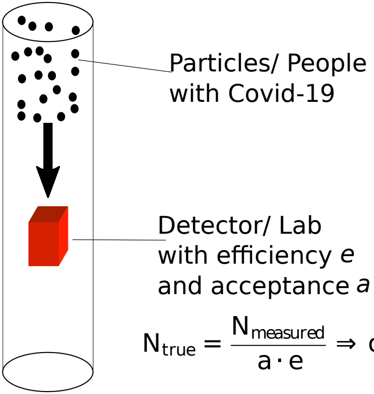


Robot model: velocity v



Handling Real World Data

Errors



$$N_{\text{true}} = \frac{N_{\text{measured}}}{a \cdot e} \Rightarrow \sigma_{N_{\text{true}}} = \sigma_{N_{\text{true}}}(\sigma_a, \sigma_e)$$

Figure: Efficiency *e* and acceptance *a* of the tests are estimated with error

Handling Real World Data

Some examplary plots

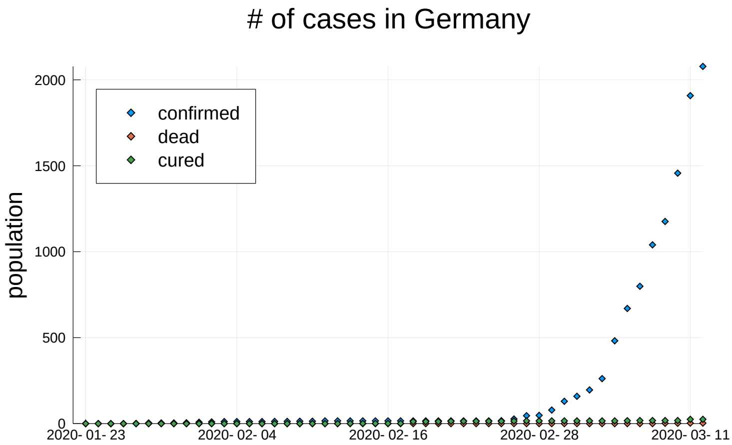


Figure: Covid-19 Cases in Germany

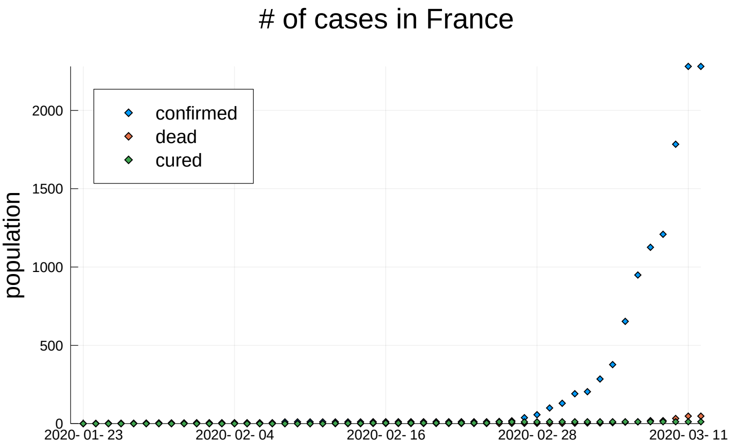


Figure: Covid-19 Cases in France

Handling Real World Data

Some examplary plots

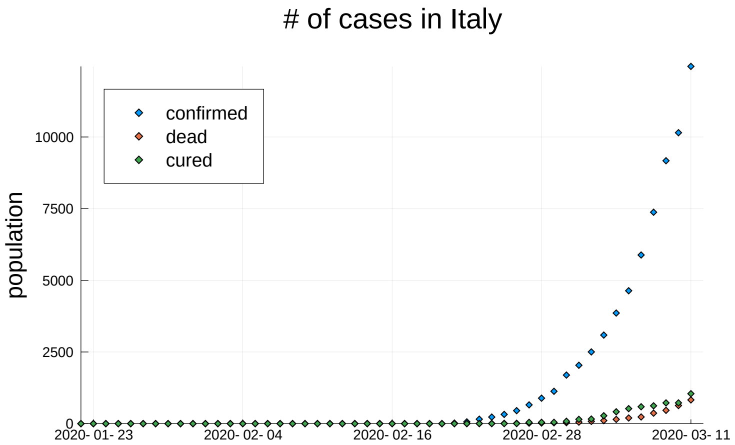


Figure: Covid-19 Cases in Italy

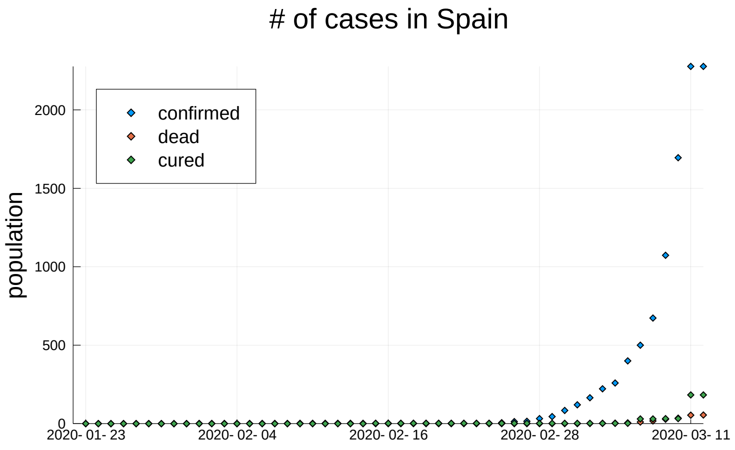


Figure: Covid-19 Cases in Spain

Handling Real World Data

Some examplary plots

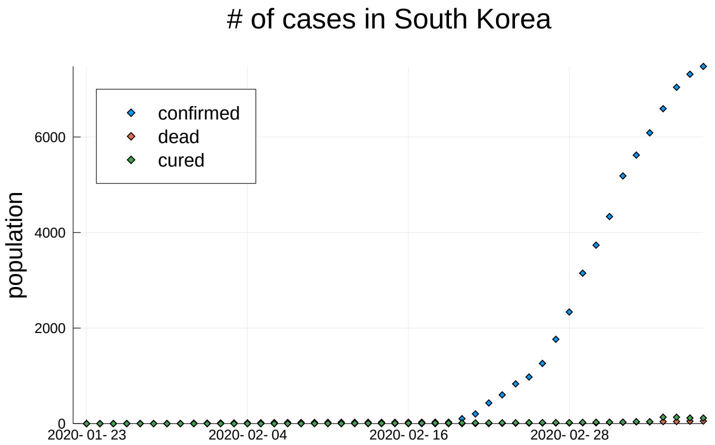
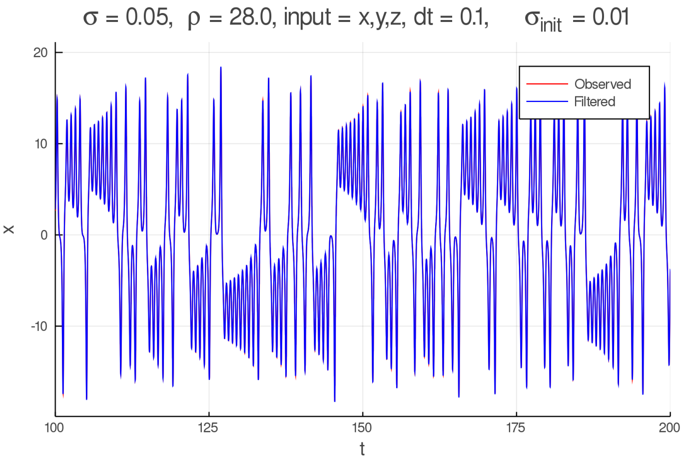


Figure: Covid-19 Cases in South Korea

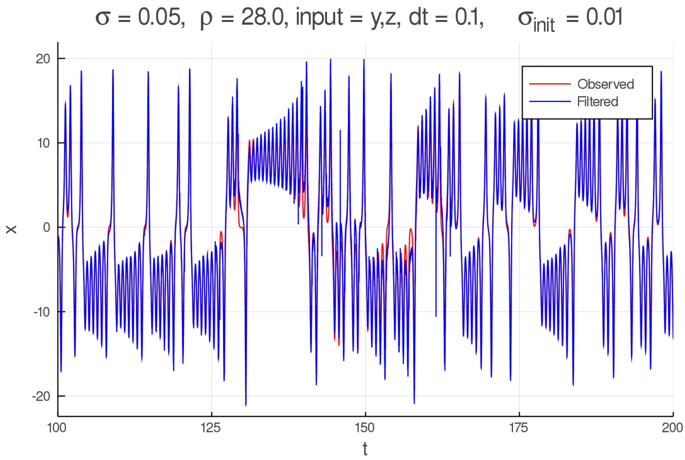
The extended Kalman filter

Lorenz3 - chaotic - 3 variables



The extended Kalman filter

Lorenz3 - chaotic - 2 variables



The extended Kalman filter

Lorenz3 - chaotic - 2 variables

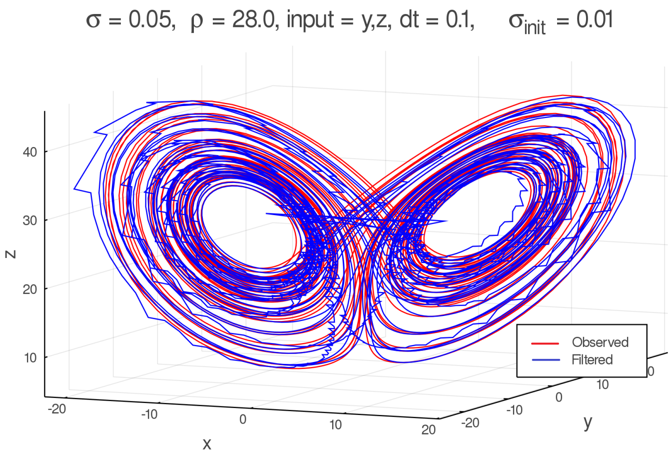
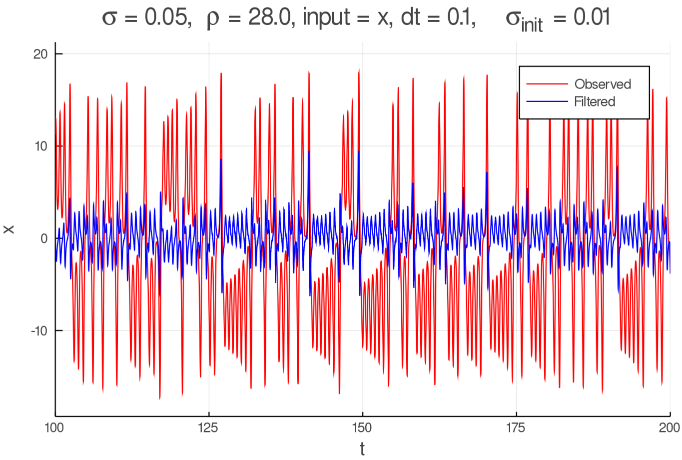


Figure: $t \in (150, 200]$

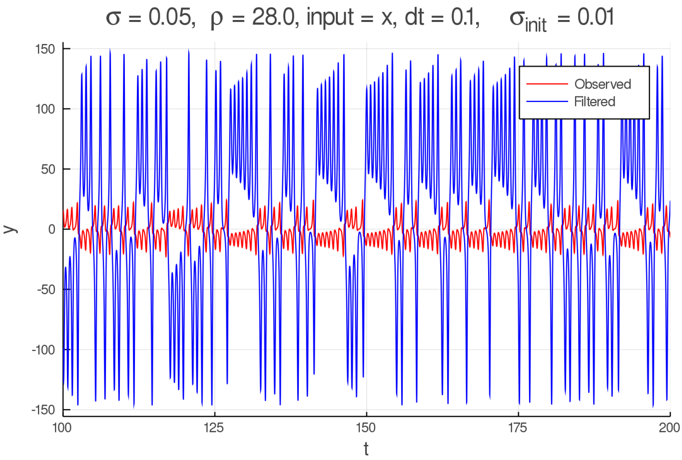
The extended Kalman filter

Lorenz3 - chaotic - 1 variable



The extended Kalman filter

Lorenz3 - chaotic - 1 variable



The extended Kalman filter

Lorenz9 - chaotic - 6 variables

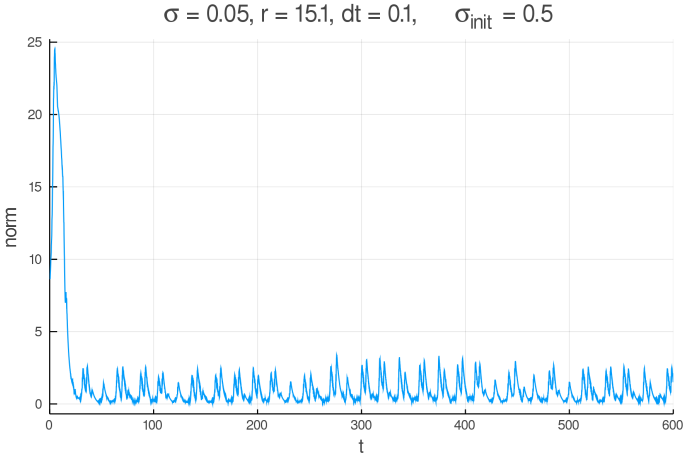


Figure: L_2-norm of discrepancies between observed variables and filtered estimation (Input variables: C1-C6)

The extended Kalman filter

Lorenz9 - chaotic

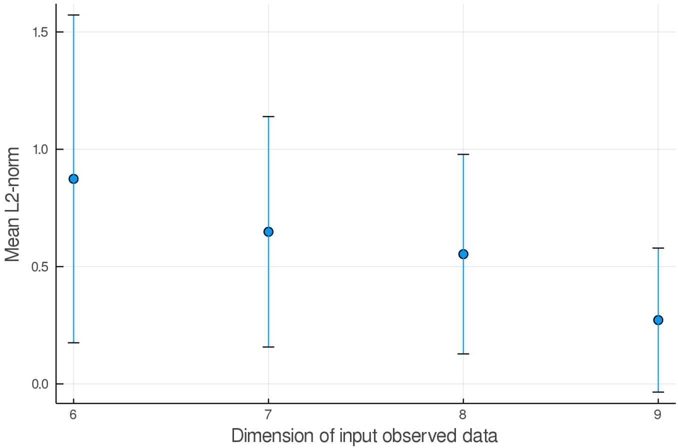


Figure: Mean L_2-norm of discrepancies between observed variables and filtered estimation for different dimensions of input observed data

The extended Kalman filter

Lorenz9 - chaotic

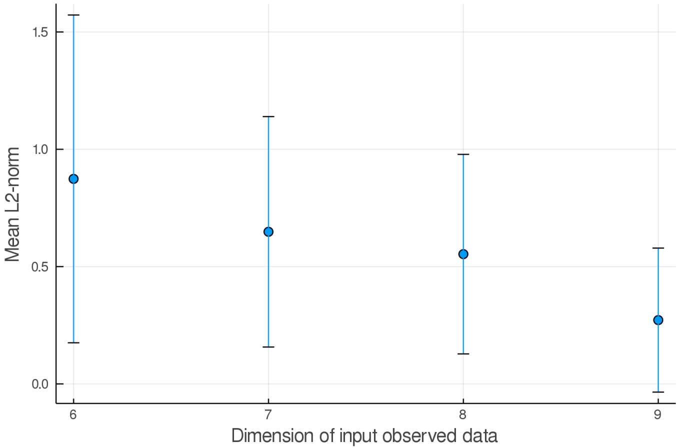


Figure: Mean L_2-norm of discrepancies between observed variables and filtered estimation for different dimensions of input observed data

The extended Kalman filter

Lorenz9 - forecast

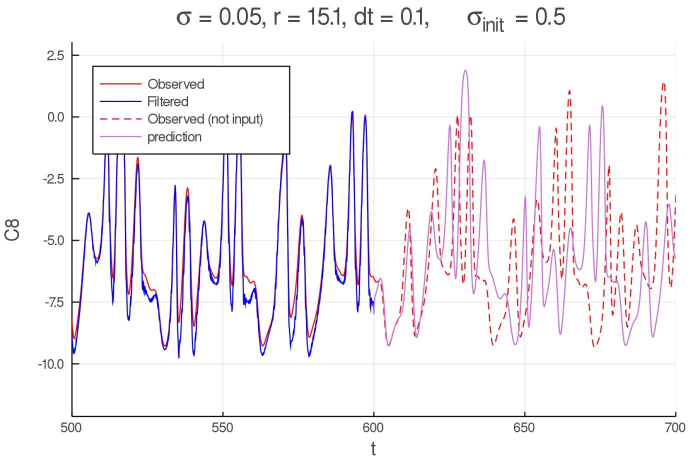
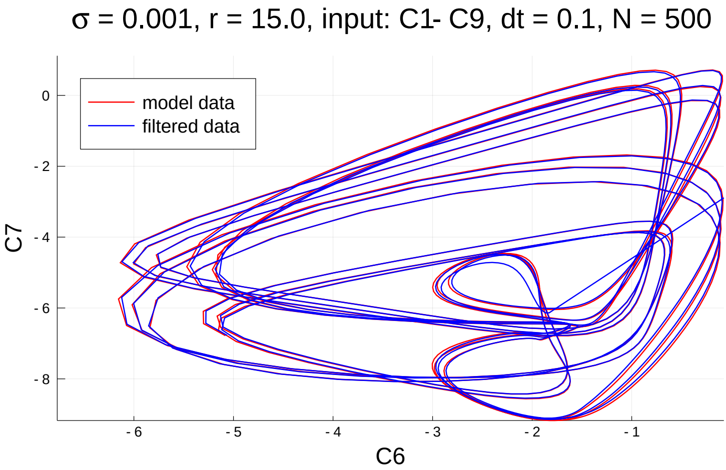


Figure: Attempt to forecast the time series (Input C1-C6)

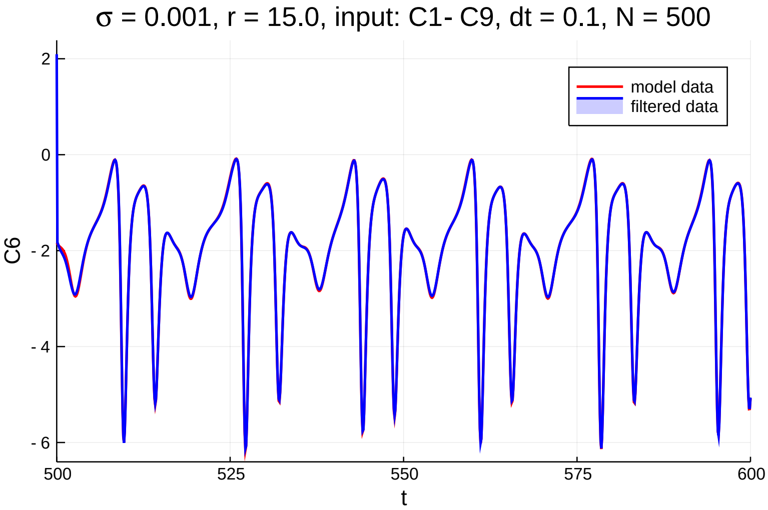
The ensemble Kalman filter

Lorenz9 - periodic - all variables



The ensemble Kalman filter

Lorenz9 - periodic - all variables



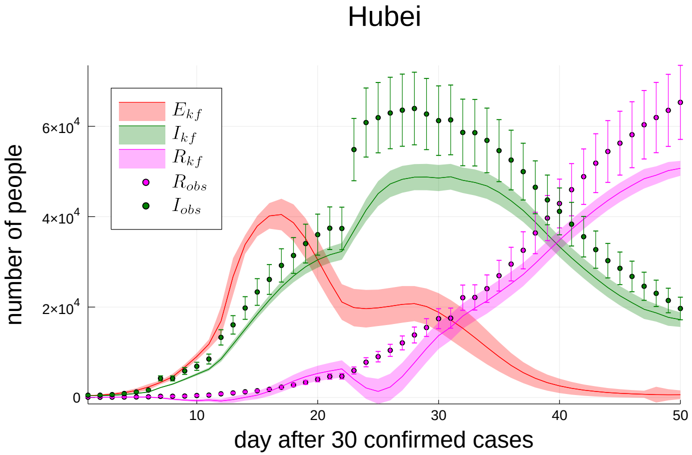


Figure: Caption

The Particle Filter

Parameter estimation of Real World Data

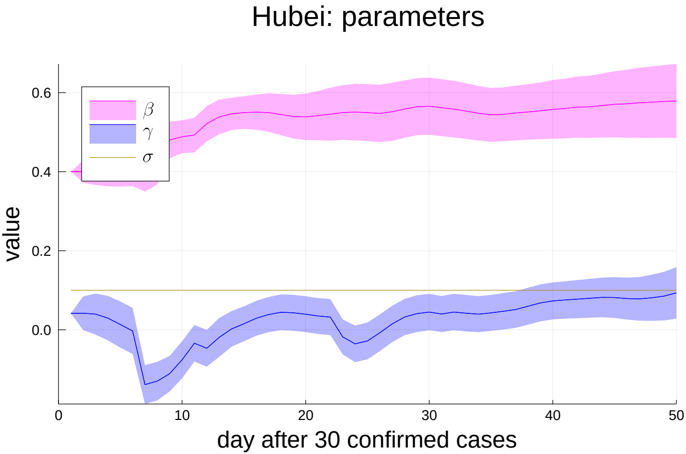


Figure: Caption

The Particle Filter

Parameter estimation of Real World Data

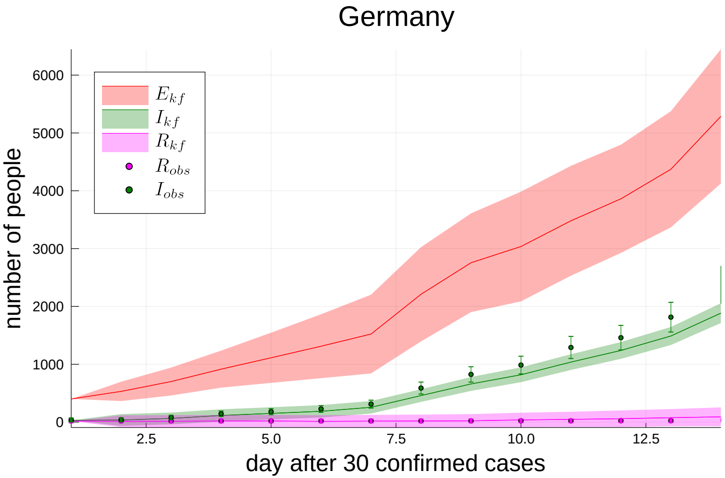


Figure: Plot of observed vs. Kalman-estimated data.

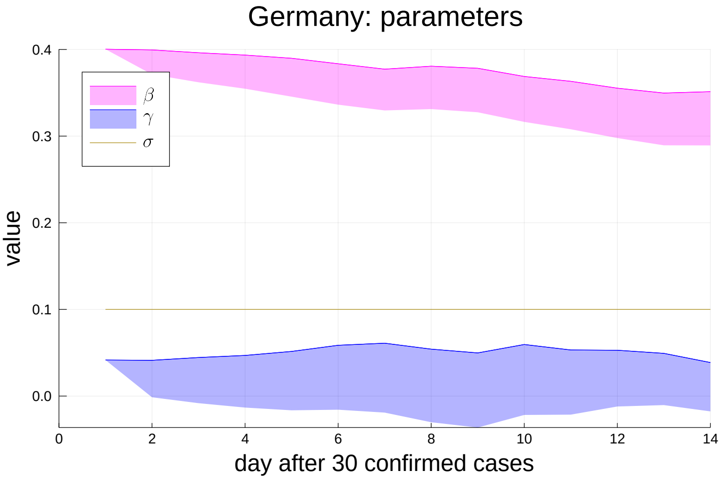
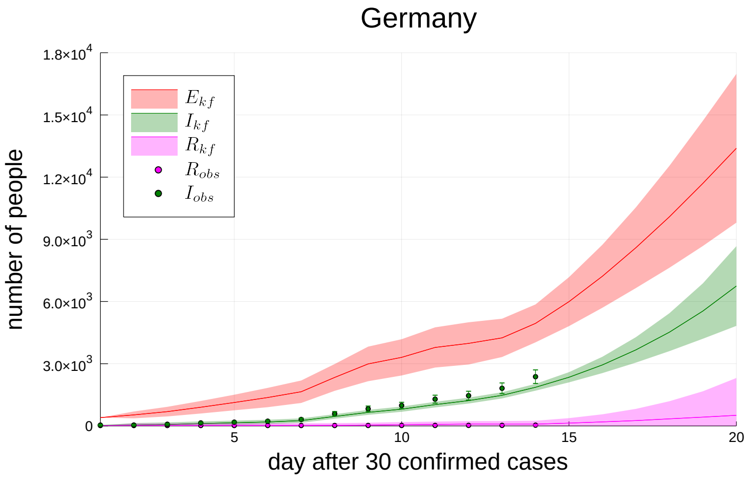
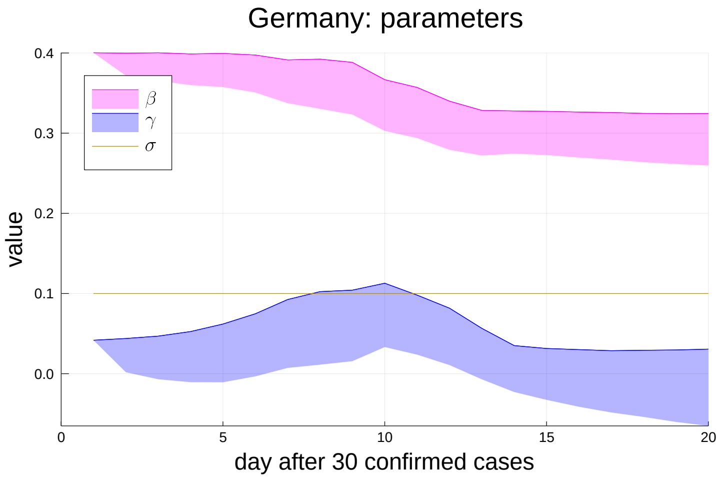


Figure: Plot of the Kalman-filtered parameters

Forecast


Germany
number of people
1.8×10^4
1.5×10^4
1.2×10^4
9.0×10^3
6.0×10^3
3.0×10^3
0
5
10
15
20
day after 30 confirmed cases
E_{kf}
I_{kf}
R_{kf}
R_{obs}
I_{obs}

Forecast


Germany: parameters
value
day after 30 confirmed cases
0.4
0.3
0.2
0.1
0.0
0
5
10
15
20
β
γ
σ

The Particle Filter

Results - the best

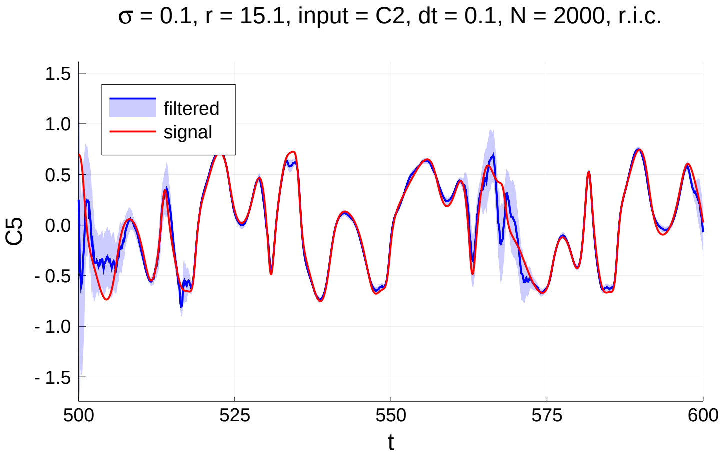


Figure: Particle Filter with unknown initial condition and large noise added to ODE in chaotic regime. First rough adaptive resampling used.

Summary

Filters with exact initial condition and low noise

	CKF	ExKF	EnKF	ParF
$r = 15.0$, $\#\text{obs} = 9$	$\times$	✓	✓	✓
$r = 15.0$, $\#\text{obs} = 1$	$\times$	$\times$	[✓]	✓
$r = 15.1$, $\#\text{obs} = 9$	$\times$	✓	✓	✓
$r = 15.1$, $\#\text{obs} = 1$	$\times$	$\times$	[✓]	✓

Table: Comparison of the implemented filters on grounds of performance with the Lorenz9-model. Here, the filters were given the exact initial condition of the model.

Summary

Filters with exact initial condition and larger noise in SDE

	CKF	ExKF	EnKF	ParF
$r = 15.0$, $\#\text{obs} = 9$	$\times$	$\checkmark$	$\checkmark$	$\checkmark$
$r = 15.0$, $\#\text{obs} = 1$	$\times$	$\times$	$[\checkmark]$	$\checkmark$
$r = 15.1$, $\#\text{obs} = 9$	$\times$	$\checkmark$	$\checkmark$	$\checkmark$
$r = 15.1$, $\#\text{obs} = 1$	$\times$	$\times$	$[\checkmark]$	$\checkmark$

Table: Comparison of the implemented filters on grounds of performance with the Lorenz9-model. Here, the filters were given the exact initial condition of the model, but a larger noise was added.

The Particle Filter

Results - the worst

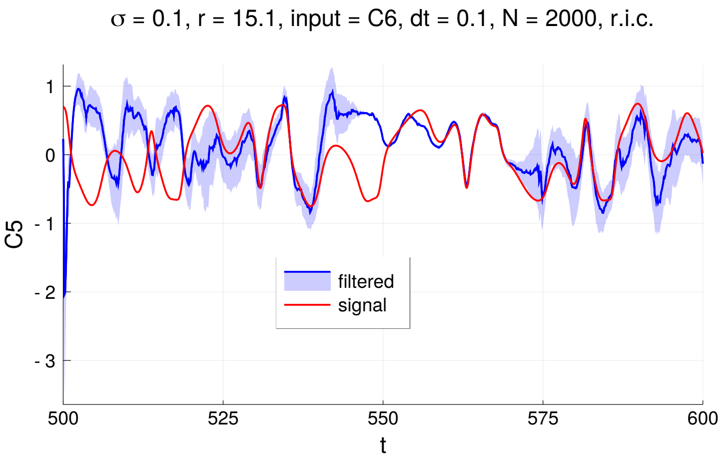


Figure: Particle Filter with unknown initial condition and large noise added to ODE in chaotic regime. Same rough adaptive resampling used as before.

Summary

Reasons?

$$\dot{C}_1 = -\sigma b_1 C_1 - C_2 C_4 + b_4 C_4^2 + b_3 C_3 C_5 - \sigma b_2 C_7$$
$$\dot{C}_2 = -\sigma C_2 + C_1 C_4 - C_2 C_5 + C_4 C_5 - \sigma C_9/2$$
$$\dot{C}_3 = -\sigma b_1 C_3 + C_2 C_4 - b_4 C_2^2 - b_3 C_1 C_5 + \sigma b_2 C_8$$
$$\dot{C}_4 = -\sigma C_4 - C_2 C_3 - C_2 C_5 + C_4 C_5 + \sigma C_9/2$$
$$\dot{C}_5 = -\sigma b_5 C_5 + C_2^2/2 - C_4^2/2$$
$$\dot{C}_6 = -b_6 C_6 + C_2 C_9 - C_4 C_9$$
$$\dot{C}_7 = -b_1 C_7 - rC_1 + 2C_5 C_8 - C_4 C_9$$
$$\dot{C}_8 = -b_1 C_8 + rC_3 - 2C_5 C_7 + C_2 C_9$$
$$\dot{C}_9 = -C_9 - rC_2 + rC_4 - 2C_2 C_6 + 2C_4 C_6 + C_4 C_7 - C_2 C_8$$

Summary

What's variable 2 doing over there?

$$
\begin{aligned}
\dot{C}_1 &= -\sigma b_1 C_1 - \mathbf{C}_2 C_4 + b_4 C_4^2 + b_3 C_3 C_5 - \sigma b_2 C_7 \\
\dot{C}_2 &= -\sigma \mathbf{C}_2 + C_1 C_4 - \mathbf{C}_2 C_5 + C_4 C_5 - \sigma C_9/2 \\
\dot{C}_3 &= -\sigma b_1 C_3 + \mathbf{C}_2 C_4 - b_4 {\mathbf{C}_2}^2 - b_3 C_1 C_5 + \sigma b_2 C_8 \\
\dot{C}_4 &= -\sigma C_4 - \mathbf{C}_2 C_3 - \mathbf{C}_2 C_5 + C_4 C_5 + \sigma C_9/2 \\
\dot{C}_5 &= -\sigma b_5 C_5 + {\mathbf{C}_2}^2/2 - C_4^2/2 \\
\dot{C}_6 &= -b_6 C_6 + \mathbf{C}_2 C_9 - C_4 C_9 \\
\dot{C}_7 &= -b_1 C_7 - rC_1 + 2 C_5 C_8 - C_4 C_9 \\
\dot{C}_8 &= -b_1 C_8 + rC_3 - 2 C_5 C_7 + \mathbf{C}_2 C_9 \\
\dot{C}_9 &= -C_9 - r\mathbf{C}_2 + rC_4 - 2\mathbf{C}_2 C_6 + 2 C_4 C_6 + C_4 C_7 - \mathbf{C}_2 C_8
\end{aligned}
$$

Summary

Where's the 6th variable?

$$\dot{C}_1 = -\sigma b_1 C_1 - C_2 C_4 + b_4 C_4^2 + b_3 C_3 C_5 - \sigma b_2 C_7$$
$$\dot{C}_2 = -\sigma C_2 + C_1 C_4 - C_2 C_5 + C_4 C_5 - \sigma C_9/2$$
$$\dot{C}_3 = -\sigma b_1 C_3 + C_2 C_4 - b_4 C_2^2 - b_3 C_1 C_5 + \sigma b_2 C_8$$
$$\dot{C}_4 = -\sigma C_4 - C_2 C_3 - C_2 C_5 + C_4 C_5 + \sigma C_9/2$$
$$\dot{C}_5 = -\sigma b_5 C_5 + C_2^2/2 - C_4^2/2$$
$$\dot{C}_6 = -b_6 \mathbf{C_6} + C_2 C_9 - C_4 C_9$$
$$\dot{C}_7 = -b_1 C_7 - rC_1 + 2 C_5 C_8 - C_4 C_9$$
$$\dot{C}_8 = -b_1 C_8 + rC_3 - 2 C_5 C_7 + C_2 C_9$$
$$\dot{C}_9 = -C_9 - rC_2 + rC_4 - 2 C_2 \mathbf{C_6} + 2 C_4 \mathbf{C_6} + C_4 C_7 - C_2 C_8$$