# PhD Progress

Marco Cucchi

July 12, 2019

## Contents

# 1 Introduction

This document is a log-book for all the work done during my PhD project. All code used con be found on github repository `https://github.com/marco-cucchi/L96gev`.

# 2 EVT and the Lorenz-96 model

Aim of this work is to find EVT parameters for observables of the Lorenz-96 (L96) model, and compare them with the bounds provided in [1].

## 2.1 Lorenz-96 model simulations

As a first step, a number of independent simulations of the L96 model are performed. The L96 model is defined as follows. For $i = 1, ..., N$:

$$\frac{dx_i}{dt} = (x_{i+1} - x_{i-2})x_{i-1} - x_i + F \tag{1}$$

where it is assumed that $x_{-1} = x_{N-1}$, $x_0 = x_N$ and $x_{N+1} = x_1$. Here $x_i$ is the state of the system on the $i$-th coordinate, and $F$ is the forcing constant. For this set of simulations, values of $F$ and $N$ have been set to $F = 8$ and $N = 32$.
Integration has been conducted using 4-*th* order Runge-Kutta scheme, with integration step $dt = 10^{-2}$. Initial conditions for each simulation have been set equal to

$$x_i^0 = 8 + \epsilon, \quad \epsilon \sim U([-0.05, +0.05]) \tag{2}$$

Different levels of spatial aggregation, defined as $A = 32, 16, 8, 4, 2, 1$, have been considered:

- For $A = 32$ no aggregation is performed, and each value $x_i$ is treated independently;
- For $A = 1$ all original $N$ $x_i$ values are spatially averaged into one single value $\overline{x}$ for each time-step;
- More in general, for $A = K$ the $N$ spatial coordinates indicated by the index $i$ are divided into $K$ non-overlapping clusters $c_j$ fixed in time, and corresponding $x_i$ values belonging to the same cluster are averaged at each time-step.

As observable, the local energy of the system for different levels af aggregation

$$E_j = \frac{1}{2}x_j^2, \quad x_j = \begin{cases} x_i, & A = 32 \\ \overline{x} = \frac{1}{N}\sum_{i=1}^{N} x_i & A = 1 \\ \frac{1}{\#c_j}\sum_{i \in c_{j,K}} x_i & A = K \end{cases} \tag{3}$$

is considered.

In order to extract information on the statistics of extremes, very long simulations have to be performed. In order to find a good compromise between this requirement and the limited amount of disk space, a similar procedure to the one adopted in [2] has been followed: instead of keeping all values of each simulation, only block-maxima are retained, with block size $\Delta t = 0.5$. It is important to highlight that block-maxima are computed *after* aggregation (spatial average).

Following this procedure, for each simulation (initial condition) 6 different files are obtained, each corresponding to one particular aggregation level $A$: each of these files, then, contain $A$ time-series of block-maxima, one for each of the $A$ clusters.

Script: c003e11.

## 2.2 Statistics of Extreme Events

Parameters defining GEV distribution are estimated using three different approaches:

- Direct fit using *block-maxima* approach;
- Direct fit using *POT* approach (still not described hear);
- Method of *moments* described in [1].

Finally, estimates derived with these approaches are compared among them and with bounds related to attractor's dimensions described in [2].

### 2.2.1 Block-maxima approach

In this approach, each time-series for each different cluster of each simulation is fitted against GEV family of distribution separately. More specifically, for each cluster time-series belonging to a different simulation the following procedure is carried out:

1. Percentiles' orders $p$ of interest are fixed (e.g. 0.99, 0.995, ...), and corresponding percentiles (thresholds) $T_p$ are computed; [1]
2. Time-series is divided in $n$ blocks, where $n = length(\text{time-series})(1 - p)$;
3. Compute maxima for each block;
4. Fit GEVD family to the block-maxima series.

The fit is performed with the R function `gevFit` from the package `fExtremes`, using MLE approach. As a result estimations of shape parameter $\xi$, location parameter $\mu$ and scale parameter $\sigma$ are returned, with respective uncertainties as computed via MLE.
Location parameter $\mu$ is actually assigned the value $T_p$; the *absolute maximum* from each time-series is also kept; *modified scale parameter* $\sigma^*$ is computed as

$$\sigma^* = \sigma - \xi T_p \tag{4}$$

Error on $\sigma^*$ is estimated via propagation of error.
As explained in [3], in order to find a valid threshold value $T_0$ for excess to follow generalized Pareto distribution (and, consequently, GEV distribution), it is a good practice to plot $\xi$ and $\sigma^*$ against $T_p$ and look for the value where both start to be approximately constant: that value is $T_0$.
Once parameters have been estimated for all clusters in a simulation, a single estimation of each parameter is saved as the average among all estimates.[2] Furthermore, the following parameters are estimated for each simulation:

- *scale parameter* $\sigma$ is computed with inverse of equation 4, and relative error is computed via propagation of errors; [3]
- *upper end-point* is computed as[4]

  $$u\hat{e}p = \hat{\mu} - \frac{\hat{\sigma}}{\hat{\xi}} \tag{5}$$

  and the relative error is computed via propagation of errors.

Finally, for each different aggregation, ensemble averages of *shape* and *modified scale* among all simulations are computed.

[1] This could be something to think upon; in this way I have (slightly?) different percentiles for different time-series in the same simulation and for different simulations. Is this right? The underlying assumption in this procedure should is that all time-series belonging to all simulations should come from the same distribution. So shouldn't the percentiles be the same for all of them?

[2] The *absolute maximum* is also averaged, and this could be an error. The average of the *location parameter* $\mu$ is also a little disturbing, but this could be solved following reasoning in footnote **??**. Error computation should be checked.

[3] This sounds very stupid, since $\sigma$ was originally estimated (but not saved) via MLE fit to GEVD.

[4] Find reference

### 2.2.2 Method of Moments

Following theory described in [1], we want to estimate *shape* and *scale* parameters using the following equations (Par 8.2.6 in [1]):

$$\xi_A^T = \frac{1}{2}\left(1 - \frac{(\langle \tilde{A}_1^T \rangle)^2}{\langle \tilde{A}_0^T \rangle \langle \tilde{A}_2^T \rangle - (\langle \tilde{A}_1^T \rangle)^2}\right) \tag{6}$$

$$\sigma_A^T = \frac{1}{2}\frac{\langle \tilde{A}_1^T \rangle \langle \tilde{A}_2^T \rangle}{\langle \tilde{A}_2^T \rangle \langle \tilde{A}_0^T \rangle - \langle \tilde{A}_1^T \rangle^2} \tag{7}$$

where $A(x)$ is an observable of the system, $T$ is a threshold value and

$$\langle \tilde{A}_n^T \rangle = \int \mu(dx)\Theta(A(x) - T)(A(x) - T)^n, \tag{8}$$

being $\Theta$ the Heaviside distribution. This results are exact in the limit for $T \rightarrow A_{max}$.
In order to perform this computation, the following procedure has been adopted. First, for each cluster time-series belonging to a different simulation:

1. Percentiles' orders $p$ of interest are fixed, and corresponding percentile (thresholds) $T_p$ are computed (footnote 1);
2. $\langle \tilde{A}_n^{T_p} \rangle$ for $n = 0, 1, 2$ are computed, using temporal average in place of ensemble average (assuming ergodicity).[5]

Once moments have been estimated for all clusters in a simulation, a single estimation of each moment is saved as the average among all estimates, and relative standard deviations are computed.
Using these estimates, *shape* parameter is computed through equation 6 and estimation of uncertainty is computed via propagation of error. Finally, for each different aggregation, ensemble averages among all simulations are computed.

### 2.2.3 Bounds to the *shape* parameter from the attractor's dimensions

We want to verify relation (8.2.15) in [1], which states that

$$(d_s + d_u + d_n)/2 \leq \delta \leq d_s + (d_u + d_n)/2, \tag{9}$$

where

- $d_u$ is equal to the number of positive Lyapunov exponents of the system [4];
- $d_n$ is equal to the number of zero Lyapunov exponents of the system, and in particular it is 1 for Axiom A systems [6];
- $d_s = n + \sum_{k=1}^{n} \lambda_k / |\lambda_{n+1}| - d_u - d_n$ [2], with $\lambda_k$ denoting the Lyapunov exponents of the system, in a descending order, and $n$ is such that $\sum_{k=1}^{n} \lambda_k$ is positive and $\sum_{k=1}^{n+1} \lambda_k$ is negative;
- $\xi = -1/\delta$;
- $\sigma = (A_{max} - T)/\delta$, with $A_{max}$ and $T$ denoting the maximum observed value of the observable[7] and the threshold value.

Lyapunov exponents have been computed using Benettin algorithm with QR decomposition. Bounds have been computed and averaged over 50 iterations (simulations).
Script:

- Lyapunov exponents computation: 136afd4
- Average bounds computation:

[5] No standard deviation has been computed at this stage!
[6] We are taking this for true in our system
[7] or the *upper end point*?

## 2.3 Statistics of Extreme Events: Corrections and Results

In this section results of the analyses reported in Sec. 2.2 are described, after issues highlighted in the footnotes 1,2,3.
Script:

- quantiles computation: 6880d20

### 2.3.1 Block-maxima approach

The following corrections have been applied:

- Percentiles are computed once, concatenating the first 80 simulations of the first clusters for each aggregation;
- Shape, scale and location parameters from fit procedures are saved for each cluster in each simulation. Averages and computation of derived parameters come after;

Results are shown in Fig. 1 and 2.
Script:

- fit: 240d6d0
- parameters derivation and plots: 97925ef

### 2.3.2 Method of Moments

Results are shown in Fig. 3 and 4.
Script:

- moments computation: 240d6d0
- parameters derivation and plots: 97925ef

# 3 LRT and the Lorenz-96 model

We aim to apply Ruelle response theory [5][6] to predict the response of different observables to the action of both constant and time-dependent forcings to our Lorenz-96 model. Among these observables, we will focus our efforts on ones describing statistics of extreme events.

## 3.1 Background

Given a *nonautonomous dissipative dynamical system* in the form

$$\dot{x} = F(x) + \epsilon g(x) f(t) \tag{10}$$

and a scalar observable $\Psi(x)$, Ruelle's response theory [5] asserts that its mean $\langle\Psi\rangle = \int \mu_t(dx)\Psi(x)$ can be decomposed as

$$\langle\Psi\rangle(t) = \sum_{j=1}^{\infty} \epsilon^j \langle\Psi\rangle^{(j)} + \langle\Psi\rangle_0 \tag{11}$$

where the $\langle\Psi\rangle$ can be expressed as multiple convolution integrals involving the pertinent Green's functions [7]. In the linear (first-order) approximation, we can thus express the response to the forcing $f(t)$ as

$$\Delta\langle\Psi\rangle(t) = \langle\Psi\rangle^{(1)}(t) = G_{\Psi}^{(1)}(t) * f(t) = \int_{-\infty}^{\infty} d\tau G_{\Psi}^{(1)}(\tau) f(t-\tau) \tag{12}$$

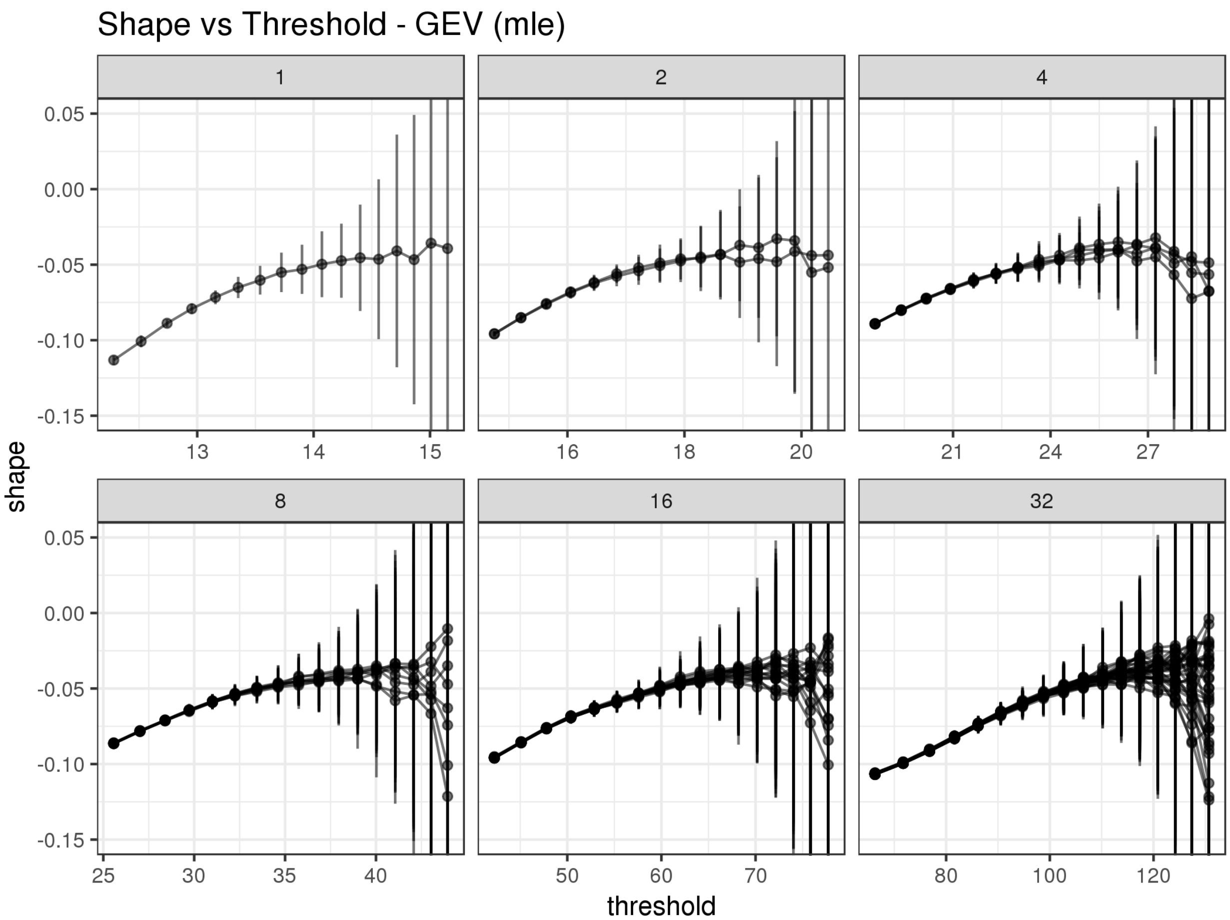



Figure 1: Ensemble average of *shape* parameter over 92 simulations. Each cluster is treated separately.

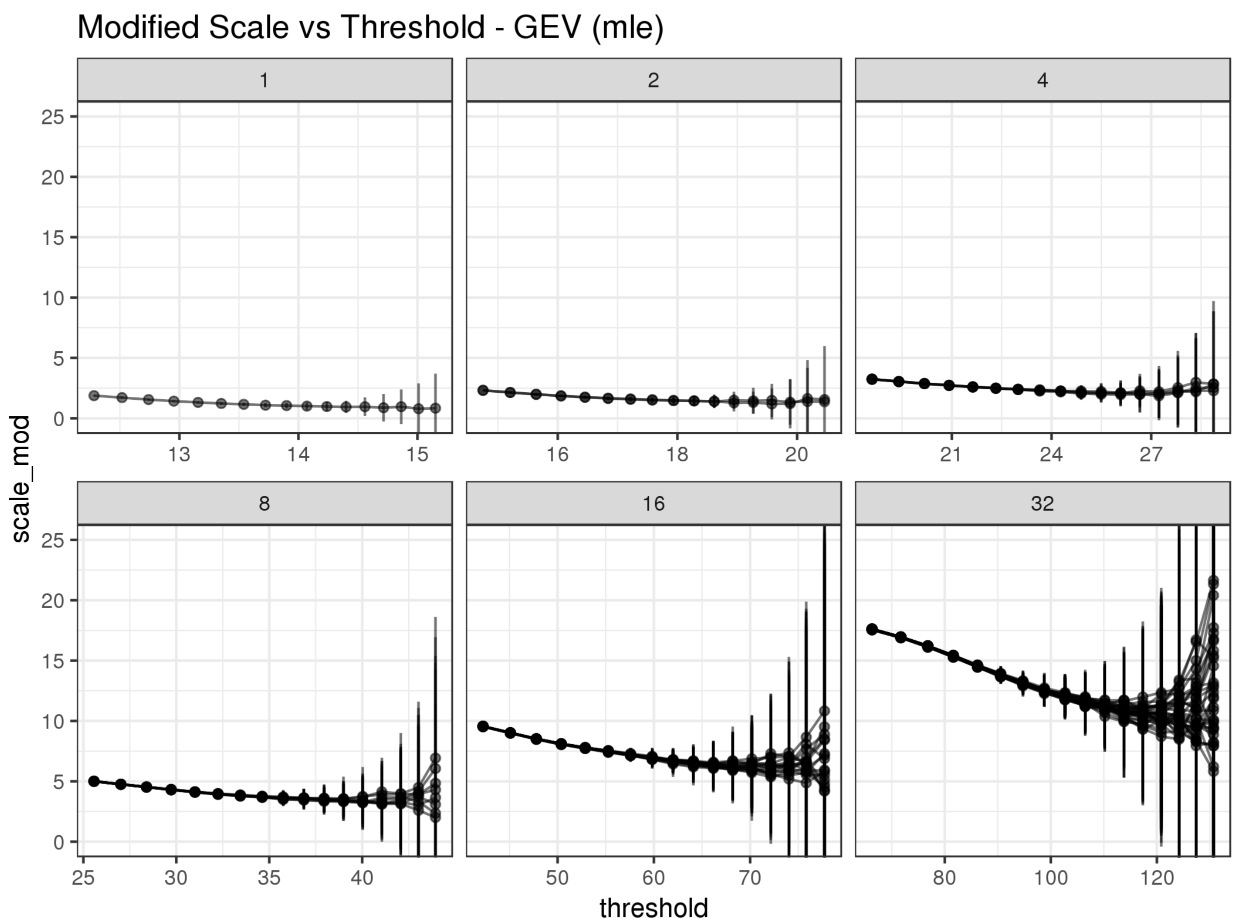



Figure 2: Ensemble average of *modified scale* parameter over 92 simulations. Each cluster is treated separately.

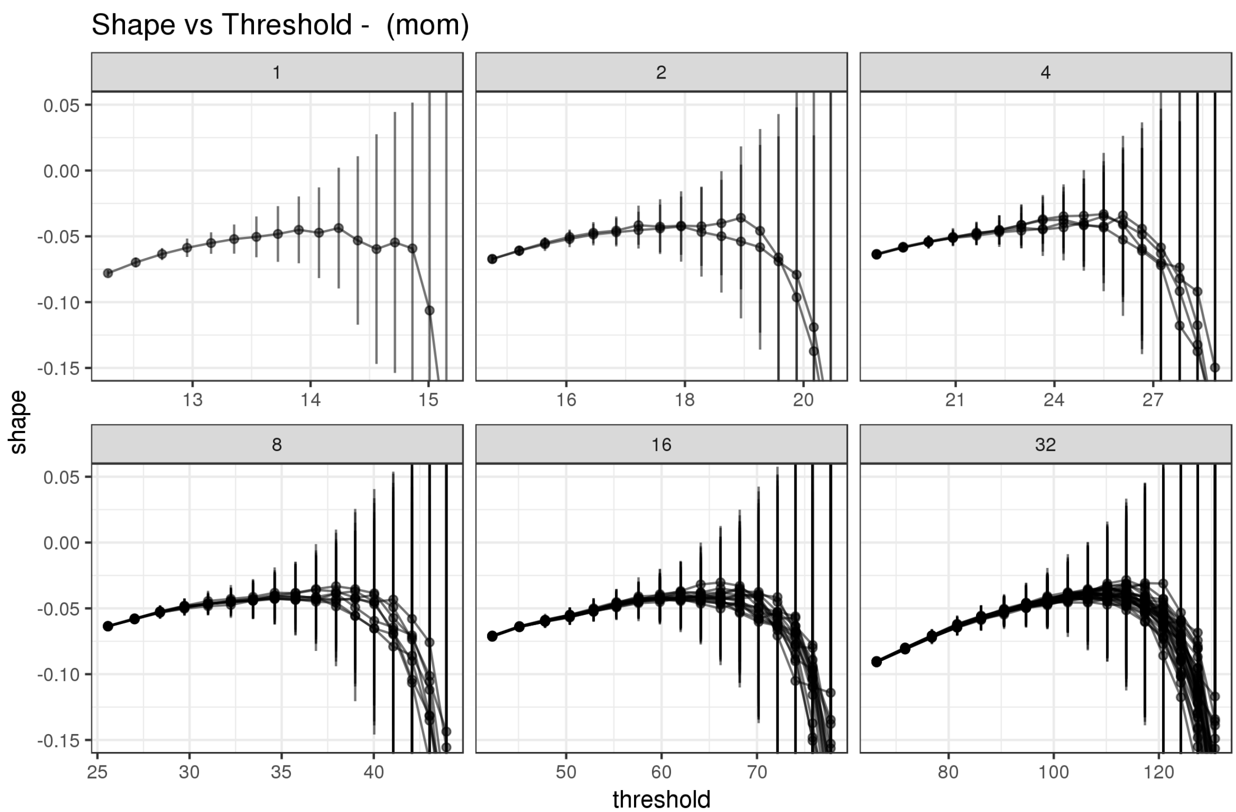



Figure 3: Ensemble average of *shape* parameter over 92 simulations. Each cluster is treated separately.

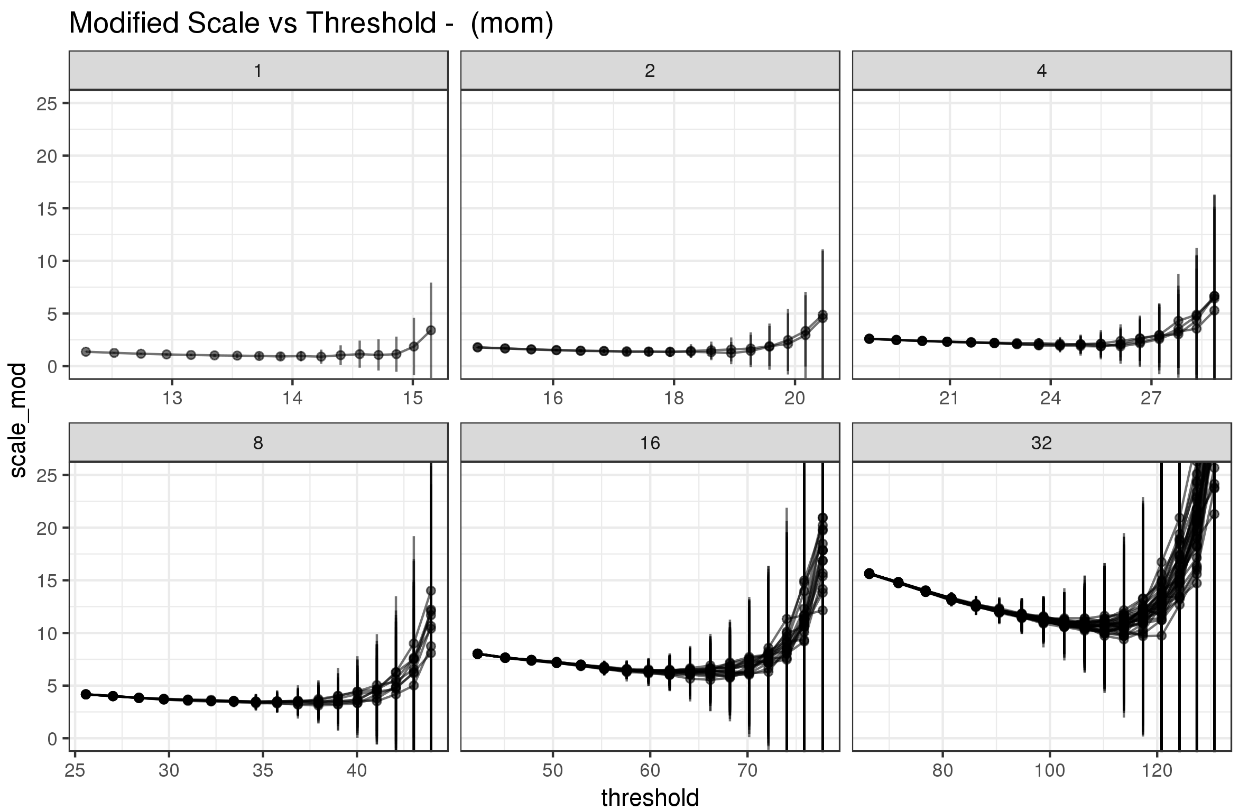



Figure 4: Ensemble average of *modified scale* parameter over 92 simulations. Each cluster is treated separately.

where the Green's function has been established by Ruelle to take the form of

$$G_{\Psi}^{(1)}(t) = \int dx \Psi(x) \left( \exp\left[ tL_f \right] \left[ L_g \overline{\mu} \right] \right)(x) \tag{13}$$

where $\overline{\mu}(dx)$ is the natural invariant measure/probability distribution of the autonomous system ($f = 0$), and operators are defined as $L_f \mu = -\operatorname{div}(f\mu)$ and $L_g \mu = -\operatorname{div}(g\mu)$, in the notation of [8].

Following [6], in a discrete-time scenario sample values of the response with the sampling $\Psi[n] = \Psi(t = (n+\nu)T)$ at any phase $\nu \in [0,1]$ obey:

$$\langle \hat{\Psi} \rangle^{(1)}[n] = \sum_{k=-\infty}^{\infty} h_{\Psi}[k] f[n-k] = h_{\Psi}[n] * f[n] \tag{14}$$

where the discrete-time (DT) impulse response or DT Green's function $h_{\Psi}[n]$ is the response $\langle \hat{\Psi}_{\perp} \rangle$ to a Kronecker delta function forcing: $f[n] = \delta[n] = 1$ if $n = 0$ and 0 otherwise. Unlike the Dirac delta [8], the Kronecker delta can be realised for numerical purposes. It is equivalent to applying a step forcing and taking the difference:

$$h_{\Psi}[n] = \langle \hat{\Psi}_{\Gamma} \rangle [n] - \langle \hat{\Psi}_{\Gamma} \rangle [n-1] \tag{15}$$

In the case of a *finite* time-series, $f[l]$ and $h_{\Psi}[l]$, $l = 0, \dots, L-1$, the response $h_{\Psi} * f[l], l = 0, \dots, L-1$ can be computed as:

$$(h * f_N)[n = 0, \dots, N-1] = \sum_{k=0}^{N-1} h[k] f_N[n-k] = \mathrm{DFT}^{-1}\{\mathrm{DFT}\{h\}\mathrm{DFT}\{f\}\}, \tag{16}$$

where DFT is the discrete Fourier transform.

## 3.2 The Experiment

We want to make predictions on the average response of observables of the Lorenz-96 system to different forcings.

We use as identification forcing [9] a *step forcing of unit intensity*, activated at time $t_0 = 0$. An ensemble of 10000 simulations of length $T = 100$ is generated [10] with this forcing and for the system at rest ($F = 8$), and responses for the different observables are averaged among ensemble members.

The same procedure is repeated using as identification forcing a *step forcing of negative unit intensity*. The semi-difference of the two average responses is then used to compute the DT Green's function $h_{\Psi}[n]$ as described in (15) [11]

Results are compared with the ones obtained using only response to the (positive) unit step forcing to compute response function.

For each simulation, only one node is considered.

Two observables are taken into considerations:

- Energy
- Above threshold (between thresholds) occurrence of energy values

### 3.2.1 Energy

Energy is computed at single locations, as

$$e_n = \frac{1}{2} x_n^2 \tag{17}$$

[8] Never mentioned before.

[9] Never mentioned before. See Par. 2.2 in [6].

[10] Describe more how simulations are generated

[11] This procedure less noisy (and more accurate?) results. **Ask Valerio for reference**.

### 3.2.2 Above threshold occurrence of energy values

Starting from energy values $e_n$, this observable $o_{n,e_t}$ is computed as

$$o_{n,t} = H[e_n - e_t], \tag{18}$$

where $H$ is the Heaviside step function

$$H[x] = \begin{cases} 0, & x < 0 \\ 1, & x \geq 0 \end{cases} \tag{19}$$

and $e_t$ is some energy value.

Once averaged over ensemble members, this observable yields the average frequency of energy values above a certain threshold. Using as thresholds $e_t$ high quantiles of the energy distribution for the system at rest, if LRT works we could use it to predict changes in the frequency of extreme (above high thresholds) events.

More generally, we can look at **between-thresholds occurrence**, computed as

$$o_{n,t_1,t_2} = H[e_n - e_{t_1}] - H[e_n - e_{t_2}], \tag{20}$$

with $e_{t_1} < e_{t_2}$. Again, averaging this observable over ensemble members yields the average frequency of energy values between the two thresholds $e_{t_1}$ and $e_{t_2}$. Using LRT, **we could thus theoretically predict changes in the shape of energy values probability distribution**.

## 3.3 Results

In this section, comparisons between actual and predicted (via LRT) responses are shown for **above** and **between-thresholds occurrences** observables. As thresholds, energy percentiles (computed on the system at rest) have been taken (see Figures' titles). The following forcings have been considered:

- **Step forcing**
$$F[t] = \begin{cases} 8, & t \leq 0 \\ 11, & t > 0 \end{cases} \tag{21}$$

- **Linearly increasing forcing**
$$F[t] = \begin{cases} F_0 + Ct, & t \leq t_f \\ F_0 + Ct_f, & t > t_f \end{cases} \tag{22}$$

  with two sets of parameters $(F_0, C, t_f)$, namely $(8, 0.03, 100)$ and $(8, 0.3, 10)$.

Results are shown in Figures 5, 6 and 7.

### 3.3.1 Comments

Why predicted responses in linear forcing cases are so much less noisy than in the step forcing case?

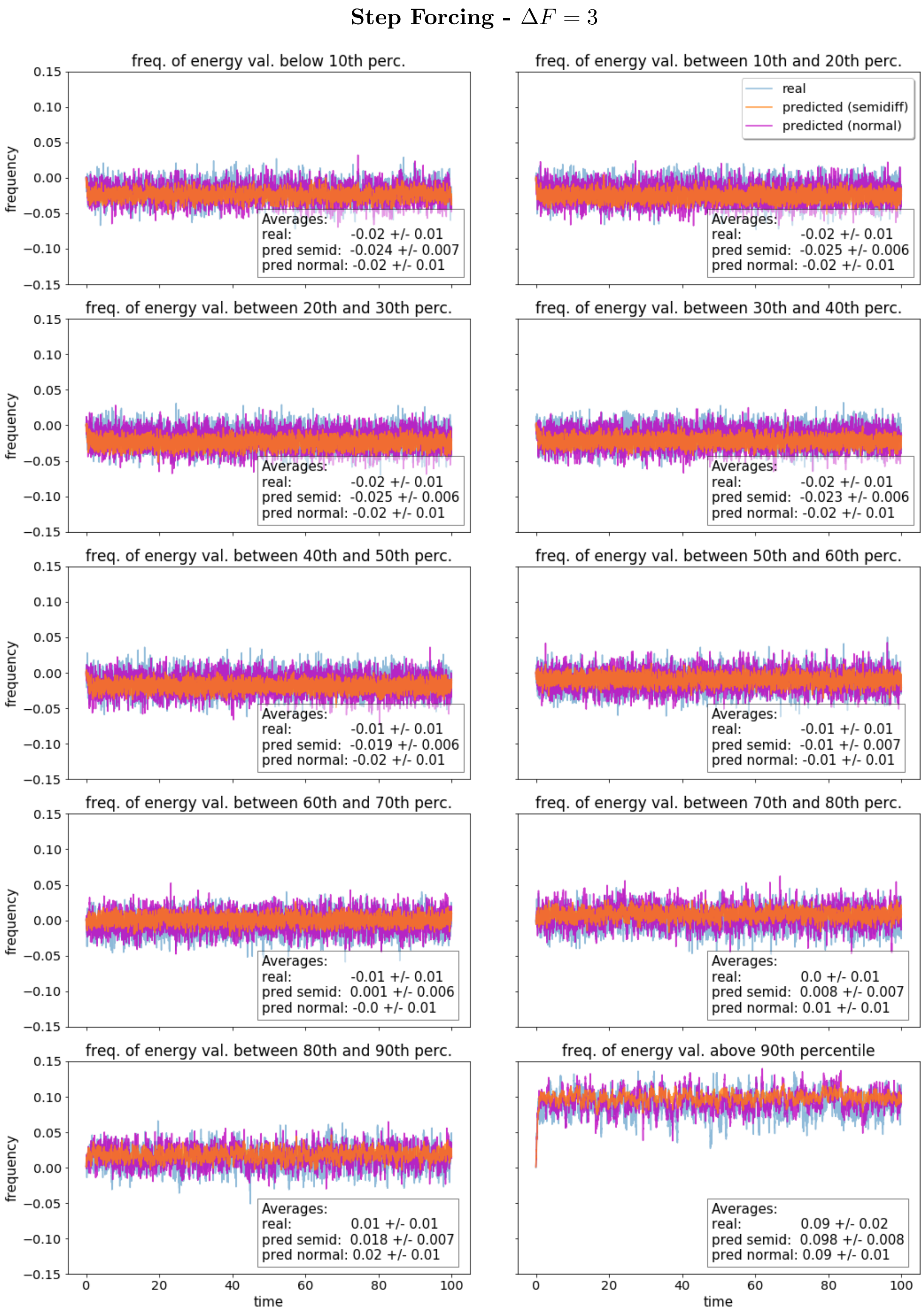



Figure 5: Real vs. predicted average responses of **occurrence observables** for **step forcing**, activated at time $t = 0$ with intensity $\Delta F = 3$.

**Linear Forcing -** $(F_0 = 8, C = 0.03, t_f = 100)$

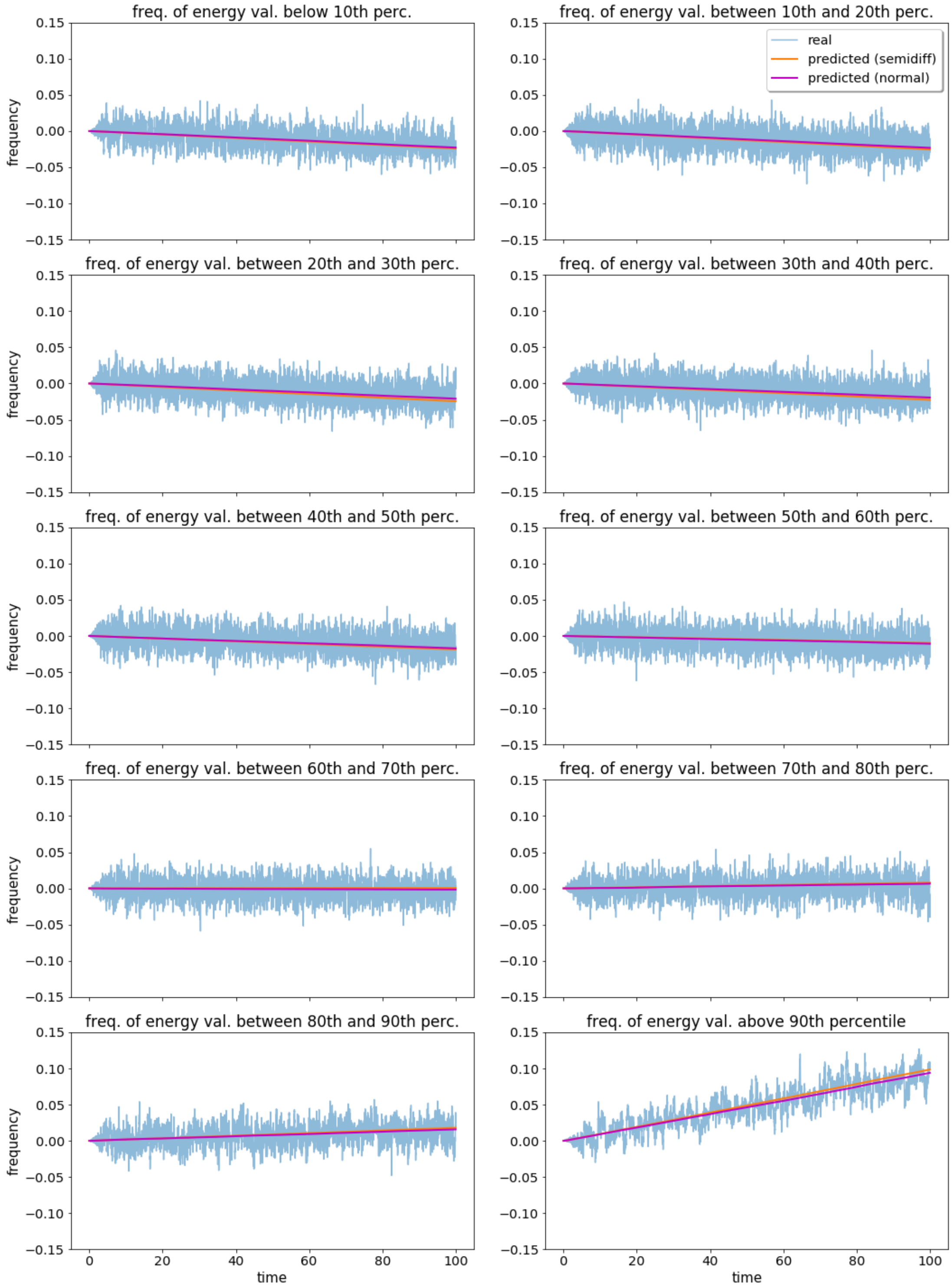



Figure 6: Real vs. predicted average responses of **occurrence observables** for **linear forcing**, activated at time $t_i = 0$, deactivated at time $t_f = 100$, with linear coefficient $C = 0.03$.

**Linear Forcing -** $(F_0 = 8, C = 0.3, t_f = 10)$

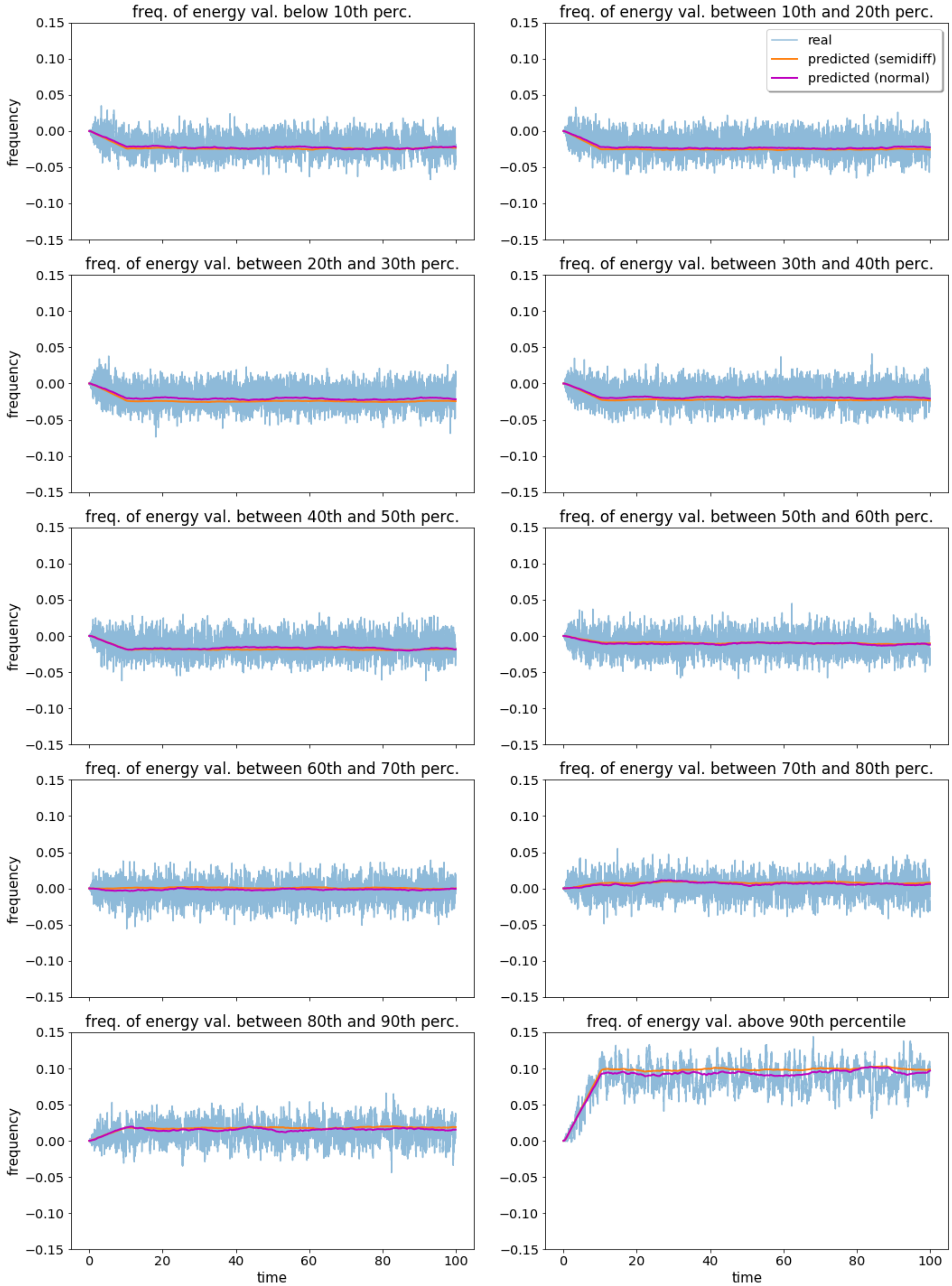



Figure 7: Real vs. predicted average responses of **occurrence observables** for **linear forcing**, activated at time $t_i = 0$, deactivated at time $t_f = 10$, with linear coefficient $C = 0.3$.

### 3.4 A closer look at higher percentiles (and smaller bins)

In Figure 8 results are shown for the same forcings explored in 3.3, but for exceedances above a higher percentile ($99^{th}$, Figs. figs. 8b, 8d and 8f) and occurrences inside smaller bin ($98^{th}$-$99^{th}$, Figs. figs. 8a, 8c and 8e).

From the plots *a discrepancy is evident between observations and predictions*, which was not so obvious for larger percentiles' intervals. Furthermore, *the discrepancy is bigger* (especially for exceedances of $99^{th}$ percentile) *when using semi-difference of responses for computing response function.*

The discrepancy could be due to the fact that *the applied forcings are too strong for the linear approximation to be valid.* In order to explore this, Fig. 9 shows predictions for exceedances of the $99^{th}$ percentile for the following forcings:

- Linear Forcing - $F_0 = 8, C = 0.03, t_f = 100$
- Linear Forcing - $F_0 = 8, C = 0.02, t_f = 100$
- Linear Forcing - $F_0 = 8, C = 0.01, t_f = 100$
- Linear Forcing - $F_0 = 8, C = 0.3, t_f = 10$
- Linear Forcing - $F_0 = 8, C = 0.2, t_f = 10$
- Linear Forcing - $F_0 = 8, C = 0.1, t_f = 10$

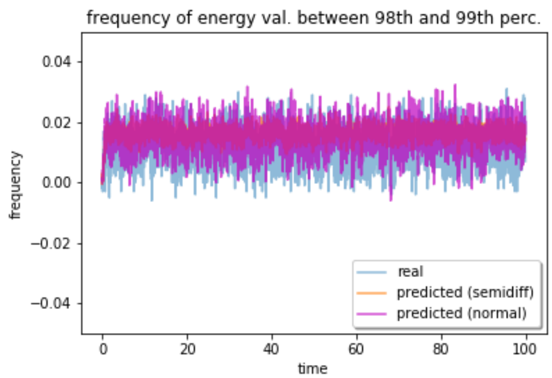
From Fig. 9, it is clear how **things get better lowering the forcing intensity**, and again it is confirmed the **higher accuracy of predictions made using only the unit step forcing to derive response function**.

As a consequence, for the moment, **predictions will be made against forcings having up to $\Delta F = 2$ as maximum intensity, and deriving response functions using only unit step forcing**.

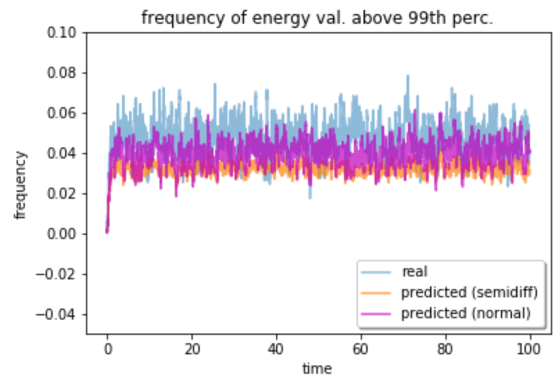
### 3.5 Prediction of changes in Energy values distribution

In this section predictions of changes in the Energy values distribution are shown for the following forcings:

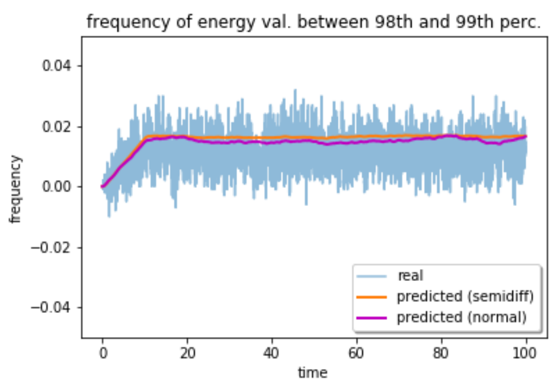
- Linear Forcing - $F_0 = 8, C = 0.01, t_f = 100$
- Linear Forcing - $F_0 = 8, C = 0.02, t_f = 100$



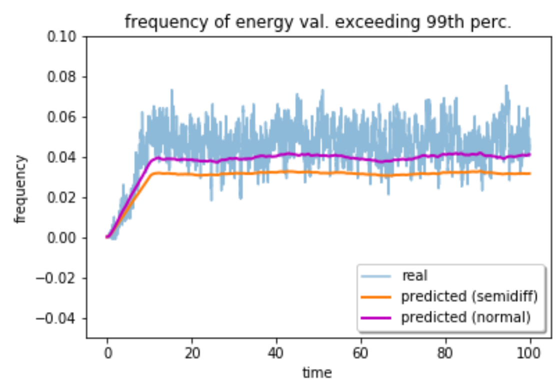

(a) Step forcing, $\Delta F = 3$



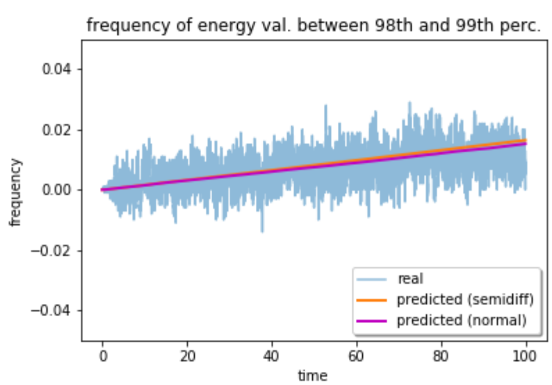

(b) Step forcing, $\Delta F = 3$



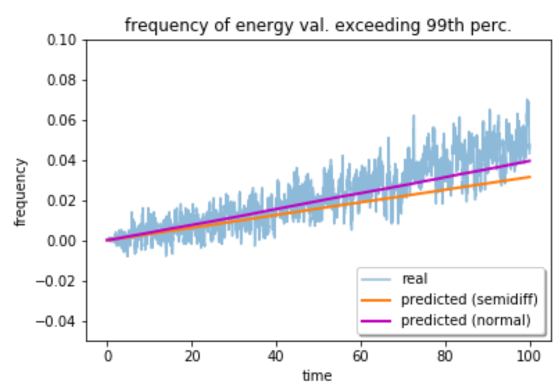

(c) Linear Forcing - $(F_0 = 8, C = 0.3, t_f = 10)$




(d) Linear Forcing - $(F_0 = 8, C = 0.3, t_f = 10)$

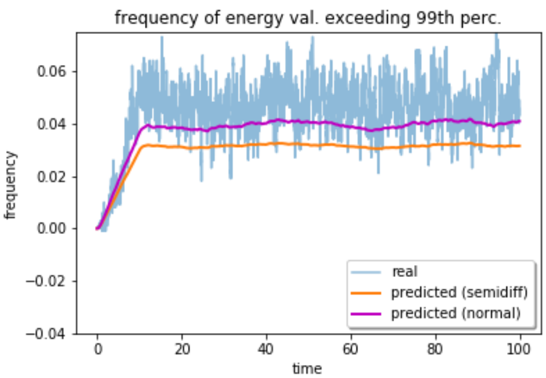



(e) Linear Forcing - $(F_0 = 8, C = 0.03, t_f = 100)$

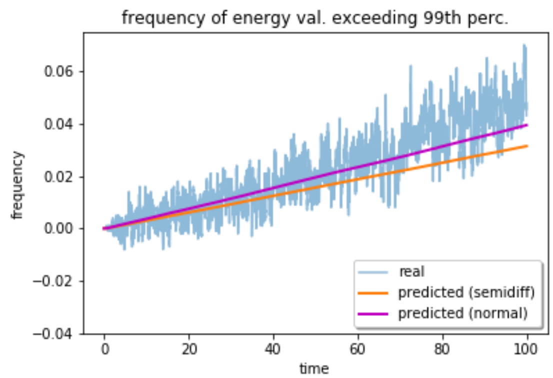



(f) Linear Forcing - $(F_0 = 8, C = 0.03, t_f = 100)$

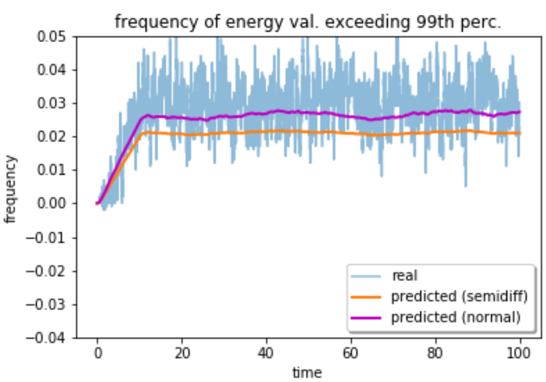
Figure 8: predictions for exceedance of $99^{th}$ percentile and occurrence between $98^{th}$ and $99^{th}$ percentiles



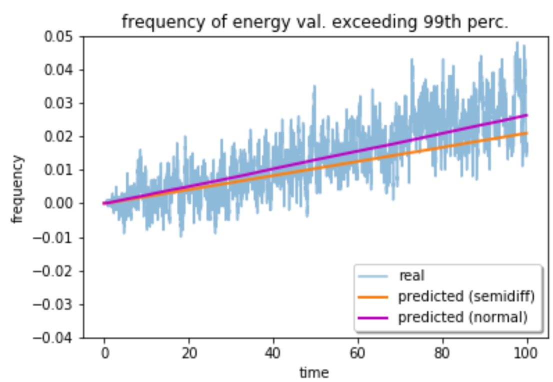

(a) Linear Forcing - ($F_0 = 8, C = 0.3, t_f = 10$)



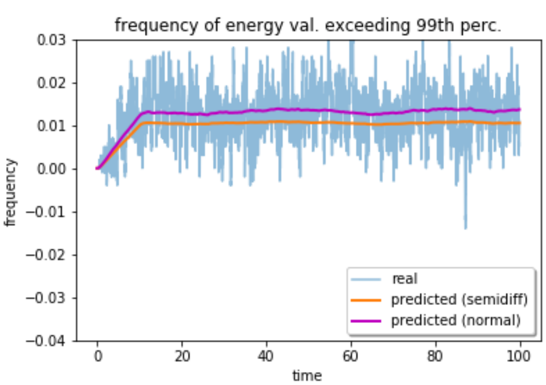

(b) Linear Forcing - ($F_0 = 8, C = 0.03, t_f = 100$)



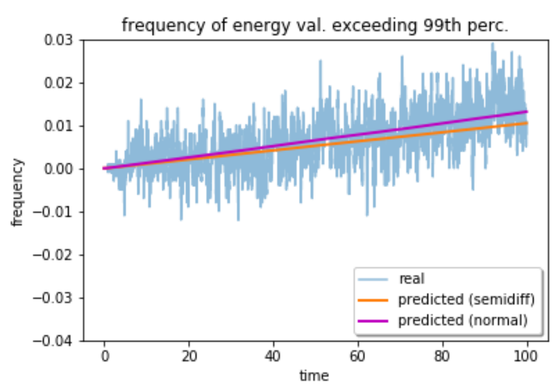

(c) Linear Forcing - ($F_0 = 8, C = 0.2, t_f = 10$)




(d) Linear Forcing - ($F_0 = 8, C = 0.02, t_f = 10$)




(e) Linear Forcing - ($F_0 = 8, C = 0.1, t_f = 100$)




(f) Linear Forcing - ($F_0 = 8, C = 0.01, t_f = 100$)

Figure 9: predictions for exceedance of $99^{th}$ percentile and occurrence between $98^{th}$ and $99^{th}$ percentiles

## References


[1] V. Lucarini, *Extremes and Recurrence in Dynamical Systems*. Wiley, 2016.

[2] M. Galfi, T. Bòdai, and V. Lucarini, "Convergence of extreme value statistics in a two-layer quasi-geostrophic atmospheric model," *Complexity*, 2017.

[3] S. Coles, *An Introduction to Statistical Modeling of Extreme Values*. Springer, 2001.

[4] E. Ott, *Chaos in Dynamical Systems*. Cambridge University Press, 2002.

[5] D. Ruelle, "A review of linear response theory for general differentiable dynamical systems," *Nonlinearity*, 2009.

[6] T. Bòdai, V. Lucarini, and F. Lunkeit, "Linear response theory applied to geoengineering," *???*, 2018.

[7] V. Lucarini, F. Ragone, and F. Lunkeit, "Predicting climate change using response theory: Global averages and spatial patterns," *Journal of Statistical Physics*, 2017.

[8] R. V. Abramov and A. J. Majda, "New approximations and tests of linear fluctuation-response for chaotic nonlinear forced-dissipative dynamical systems," *Journal of Nonlinear Science*, 2008.