engineering proceedings

MDPI

Proceeding Paper

Outliers Impact on Parameter Estimation of Gaussian and Non-Gaussian State Space Models: A Simulation Study †

Fernanda Catarina Pereira [1,*], Arminda Manuela Gonçalves [2,‡] and Marco Costa [3,‡]

[1] Centre of Mathematics, University of Minho, 4710-057 Braga, Portugal
[2] Department of Mathematics and Centre of Mathematics, University of Minho, 4710-057 Braga, Portugal; mneves@math.uminho.pt
[3] Centre for Research and Development in Mathematics and Applications, Águeda School of Technology and Management, University of Aveiro, 3810-193 Aveiro, Portugal; marco@ua.pt
* Correspondence: id9976@alunos.uminho.pt
† Presented at the 8th International Conference on Time Series and Forecasting, Gran Canaria, Spain, 27–30 June 2022.
‡ These authors contributed equally to this work.

Abstract: State space models are powerful and quite flexible tools that allow systems that vary significantly over time due to their formulation to be dealt with, because the models' parameters vary over time. Assuming a known distribution of errors, in particular the Gaussian distribution, parameter estimation is usually performed by maximum likelihood. However, in time series data, it is common to have discrepant values that can impact statistical data analysis. This paper presents a simulation study with several scenarios to find out in which situations outliers can affect the maximum likelihood estimators. The results obtained were evaluated in terms of the difference between the maximum likelihood estimate and the true value of the parameter and the rate of valid estimates. It was found that both for Gaussian and exponential errors, outliers had more impact in two situations: when the sample size is small and the autoregressive parameter is close to 1, and when the sample size is large and the autoregressive parameter is close to 0.25.

Keywords: state space models; parameter estimation; outliers; simulation study

Citation: Pereira, F.C.; Gonçalves, A.M.; Costa, M. Outliers Impact on Parameter Estimation of Gaussian and Non-Gaussian State Space Models: A Simulation Study. *Eng. Proc.* **2022**, *18*, 31. https://doi.org/10.3390/engproc2022018031

Academic Editors: Ignacio Rojas, Hector Pomares, Olga Valenzuela, Fernando Rojas and Luis Javier Herrera

Published: 22 June 2022

1. Introduction

There are several books in the literature that describe state space models in detail [1–5]. A major advantage of these models is the possibility of explicitly integrating the unobservable components of a time series by relating to each other stochastically.

State space models have in their structure a latent process, the state, which is not observed. The Kalman filter is typically used to estimate it, as it is a recursive algorithm that, at each time, computes the optimal estimator in the sense that it has the minimum mean squared error of the state when the model is fully specified, and one-step-ahead predictions by updating and improving the predictions of the state vector in real time when new observations become available. The Kalman filter was originally developed by control engineering in the 1960s in one of Kalman's papers [6] describing a recursive solution to the linear filter problem for discrete time. Today, this algorithm is applied in various areas of study.

Usually, to estimate the unknown parameters of the model, the maximum likelihood method is used by assuming normality of the errors; however, this assumption cannot always be guaranteed. Non-parametric estimation methods can be a strong contribution when it comes to the initial values of iterative methods used to optimize the likelihood function, which often do not verify the convergence of the algorithms due to the initial choice of these parameters. For example, ref. [7] propose estimators based on the generalized method of moments, the distribution-free estimators, where these estimators do not depend on the distribution of errors.

Nevertheless, even if the assumption of normality of errors is not verified, the Kalman filter still returns optimal predictions within the class of all linear estimators. However, the optimal properties of Kalman filter predictors can only be ensured when all state space models' parameters are known. When the unknown parameter vector is replaced by its estimate, the mean squared error of the estimators is underestimated.

The analysis and modeling of dynamic systems through state space models has been quite useful given its flexibility. In its formulation, the state process is assumed to be a Markov process, allowing optimal predictions of the states and, consequently, observations based only on the optimal estimator of the current state to be obtained.

Despite these advantages, any prediction model is dependent on the quality of the data. Particularly, in many cases, meteorological time series are subject to higher uncertainties, and Kalman filter solutions can be biased [8].

In particular, outliers are an important issue in time series modeling. Time series data are typically dependent on each other and the presence of outliers can impact parameter estimates, forecasting and also inference results [9]. In the presence of incomplete data and outliers in the observed data, ref. [10] developed a modified robust Kalman filter. Ref. [11] showed that linear Gaussian state space models are suitable for estimating the unknown parameters and can consequently affect the state predictions, especially when the measurement error was much larger than the stochasticity of the process. Ref. [12] proposed a non-parametric estimation method based on statistical data depth functions to obtain robust estimates of the mean and the covariance matrix of the asset returns, which is more robust in the presence of outliers, and also does not require parametric assumptions.

This work arose from the project "TO CHAIR—The Optimal Challenges in Irrigation", in which short-term forecast models, with the state space representation, were developed to model the time series of maximum air temperature. For this project, we analyzed data provided by the University of Trás-os-Montes and Alto Douro, corresponding to the maximum air temperature observed in a farm, located in the district of Bragança, between 20 February and 11 October 2019, and data from the website weatherstack.com, corresponding to the forecasts with a time horizon of 1 to 6 days of the same meteorological variable for the same location. The main goal focused on improving the accuracy of the forecasts for the farm. However, there were some modeling problems, particularly regarding the convergence of the numerical method, which arose in the presence of outliers.

Therefore, to evaluate and compare the quality of the estimates of the unknown parameters of the linear invariant state space model in the presence of outliers, this paper presents four simulation studies: the first is based on the linear Gaussian state space model; the second is based on the linear Gaussian state space model with contaminated observations; the third is based on the linear non-Gaussian state space model with exponential errors; and the last one is based on the linear non-Gaussian state space model with exponential errors and contaminated observations. For each of the four studies, several scenarios were tested, in which 2000 samples with valid estimates of size n $(n = 50, 200, 500)$ were simulated. The results obtained were evaluated in terms of the difference between the maximum likelihood estimate and the true value of the parameter and the rate of valid estimates.

2. Simulation Design

In general, the linear univariate state space model is given as follows:

$$Y_t = \beta_t W_t + e_t, \;\; \text{observation equation} \tag{1}$$

$$\beta_t = \mu + \phi(\beta_{t-1} - \mu) + \varepsilon_t, \;\; \text{state equation} \tag{2}$$

where $t = 1, \ldots, n$ is the discrete time and

- Y_t is the observed data;
- W_t is a factor, assumed to be known, that relates the observation Y_t to the state β_t at time t;

- $\{\beta_t\}_{t=1,\ldots,n} \sim AR(1)$, $-1 < \phi < 1$, $E(\beta_t) = \mu$, and $var(\beta_t) = \dfrac{\sigma_\varepsilon^2}{1-\phi^2}$;
- $E(e_t) = 0$, $E(e_t e_s) = 0$, $\forall t \neq s$, and $var(e_t) = \sigma_e^2$;
- $E(\varepsilon_t) = 0$, $E(\varepsilon_t \varepsilon_s) = 0$, $\forall t \neq s$, and $var(\varepsilon_t) = \sigma_\varepsilon^2$;
- $E(e_t \varepsilon_s) = 0, \forall t, s$.

This paper aims to investigate under what conditions the presence of outliers affects the estimation of parameters and states in the state space model. Thus, we simulate time series of size n ($n = 50, 200, 500$) using the model defined by Equations (1) and (2). For simplicity's sake, we consider for all simulation studies $W_t = 1$, $\forall t$, and $\mu = 0$, that is

$$Y_t = \beta_t + e_t, \tag{3}$$

$$\beta_t = \phi\beta_{t-1} + \varepsilon_t, \quad t = 1, \ldots, n. \tag{4}$$

To create the contamination scenario, we study real time series concerning maximum air temperature. We used data from two different sources: the first corresponds to daily records of maximum air temperature between 20 February and 11 October 2019 (234 observations) through a portable weather station installed on a farm located in the Bragança district in northeastern Portugal; the second database corresponds to forecasts from the weatherstack.com website. These forecasts have a time horizon of up to 6 days; this means that, for a certain time t, we have forecasts given at times $t-6, t-5, \ldots, t-1$.

So, first we took the difference between the recorded/observed maximum temperature and the website's forecasts, say, $\Lambda_{t,(h)}$, where t is the time, in days, and h is the time horizon of the forecasts, $h = 1, \ldots, 6$ days. Next, we calculated the percentage of outliers of $\Lambda_{t,(h)}$, whose percentage was on average 5%. Regarding the variable $\Lambda_{t,(h)}$, outliers were removed and replaced by linear interpolation, say, $\Lambda^*_{t,(h)}$, in order to remove the contamination present in the data, and its mean was subtracted, $\Lambda^*_{t,(h)} - mean(\Lambda^*_{t,(h)})$, so that it had zero mean. Then, for each time horizon h ($h = 1, \ldots, 6$), the model with a state space representation presented by Equations (3) and (4) was fitted to the data $\Lambda^*_{t,(h)} - mean(\Lambda^*_{t,(h)})$.

In order to establish a relationship between the estimates of parameters ϕ, σ_ε^2 and σ_e^2, that were obtained from the "non-contaminated" data, and the magnitude of the outliers of $\Lambda_{t,(h)}$, the linear regression model was fitted, whose relationship is given by

$$k = 1.8874 + 3.5161\sqrt{\frac{\sigma_\varepsilon^2}{1-\phi^2} + \sigma_e^2} \tag{5}$$

where $k = |\text{outliers of } \Lambda_{t,(h)} - \text{mean of } \Lambda_{t,(h)} \text{without outliers}|$, is the magnitude of the outliers, and $\dfrac{\sigma_\varepsilon^2}{1-\phi^2} + \sigma_e^2$ is the total variance of Y_t. In total, $\Lambda_{t,(h)}$ ($h = 1, \ldots, 6$) shows 59 outliers.

In this work, four simulation scenarios were tested:

1. The first is based on the linear Gaussian state space model given by

$$Y_t = \beta_t + e_t, \quad e_t \sim \mathcal{N}(0, \sigma_e^2)$$
$$\beta_t = \phi\beta_{t-1} + \varepsilon_t, \quad \varepsilon_t \sim \mathcal{N}(0, \sigma_\varepsilon^2), \quad t = 1, \ldots, n$$

2. The second is based on the linear Gaussian state space model with contaminated observations.
   To contaminate the model, the deterministic factor k, given in (5), is added in this way

$$Y_t = \beta_t + e_t + I_t k, \;\; e_t \sim \mathcal{N}(0, \sigma_e^2)$$
$$\beta_t = \phi\beta_{t-1} + \varepsilon_t, \;\; \varepsilon_t \sim \mathcal{N}(0, \sigma_\varepsilon^2),$$

where $I_t \sim \mathcal{B}(1, 0.05)$.

3. The third is based on the linear non-Gaussian state space model with exponential errors defined by

$$Y_t = \beta_t + e_t, \;\; e_t \sim \text{Exp}(\lambda_e) - \frac{1}{\lambda_e}$$
$$\beta_t = \phi\beta_{t-1} + \varepsilon_t, \;\; \varepsilon_t \sim \text{Exp}(\lambda_\varepsilon) - \frac{1}{\lambda_\varepsilon}, \;\; t = 1, \ldots, n$$

4. The last one is based on the linear non-Gaussian state space model with exponential errors and contaminated observations. Similar to scenario 2, we have

$$Y_t = \beta_t + e_t + I_t k, \;\; e_t \sim \text{Exp}(\lambda_e) - \frac{1}{\lambda_e}$$
$$\beta_t = \phi\beta_{t-1} + \varepsilon_t, \;\; \varepsilon_t \sim \text{Exp}(\lambda_\varepsilon) - \frac{1}{\lambda_\varepsilon}, \;\; t = 1, \ldots, n$$

where $I_t \sim \mathcal{B}(1, 0.05)$, and k given in (5).

For each of the four scenarios, sample sizes of $n = 50, 200, 500$ were simulated. In this study, a range of values were simulated for ϕ (0.25, 0.75), and σ_ε^2 and σ_e^2 (0.10, 1.00, 5.00, 0.10, 2.00, 0.05). For each parameter combination, 2000 replicates with valid estimates were considered, i.e., estimates within the parameter space: $-1 < \phi < 1$, $\sigma_\varepsilon > 0$, and $\sigma_e > 0$. In all simulations, we take the initial state $\beta_0 = 0$ in the Kalman filter.

To evaluate the quality of the parameter estimates, we considered the Root Mean Square Error (RMSE),

$$\text{RMSE}(\Theta) = \sqrt{\frac{1}{2000}\sum_{i=1}^{2000}\left(\Theta_i - \widehat{\Theta}_i\right)^2}$$

the Mean Absolute Error (MAE),

$$\text{MAE}(\Theta) = \frac{1}{2000}\sum_{i=1}^{2000}\left|\Theta_i - \widehat{\Theta}_i\right|$$

the Mean Absolute Percentage Error (MAPE),

$$\text{MAPE}(\Theta) = \frac{1}{2000}\sum_{i=1}^{2000}\left|\frac{\Theta_i - \widehat{\Theta}_i}{\Theta_i}\right| \times 100$$

$\Theta = (\phi, \sigma_\varepsilon^2, \sigma_e^2)$ and the convergence rate. The convergence rate provides information about the percentage of valid estimates among all simulations (simulations with valid and non-valid estimates). The convergence rate is given by the number of valid simulated estimates (in this case, 2000) divided by the number of total simulations.

To estimate the unknown parameters of the state space model (3) and (4) $\Theta = (\phi, \sigma_\varepsilon^2, \sigma_e^2)$ of each simulation, the maximum likelihood method was used by assuming the normality of the disturbances for all four scenarios. Log-likelihood maximization was performed by the Newton–Raphson numerical method. In this study, the R package "astsa" was used [3,13,14].

3. Results

In this section, the simulation results are presented. Tables 1–3 present the results of the simulations in terms of the RMSE, MAE, MAPE (%) and the convergence rate (%) for sample sizes n = 50, n = 200 and n = 500, respectively, considering both non-contaminated (NC) and contaminated Gaussian errors. Tables 4–6 show the simulation results considering contaminated and non-contaminated exponential errors.

Table 1. RMSE, MAE, MAPE and convergence rate of Θ with 2000 simulations of sample sizes $n = 50$, considering Gaussian errors (NC = Non-Contaminated; C = Contaminated).

Parameters				**RMSE**			**MAE**			**MAPE (%)**			**Convergence Rate**
ϕ	σ^2_ε	σ^2_e		ϕ	σ^2_ε	σ^2_e	ϕ	σ^2_ε	σ^2_e	ϕ	σ^2_ε	σ^2_e	**(%)**
0.25	0.10	0.05	NC	0.2542	0.0563	0.0502	0.1894	0.0484	0.0452	75.7423	48.3383	90.4927	$2000/2679 \simeq 75\%$
			C	0.3823	0.3325	0.3095	0.2983	0.2375	0.2052	119.3057	237.5420	410.4236	$2000/3059 \simeq 65\%$
	1.00	0.10	NC	0.2598	0.4638	0.3699	0.1934	0.3634	0.2632	77.3528	36.3408	263.1570	$2000/2466 \simeq 81\%$
			C	0.3514	1.3520	1.2319	0.2717	1.0957	0.7552	108.6993	109.5651	755.1521	$2000/2295 \simeq 87\%$
	5.00	2.00	NC	0.2880	2.7078	2.2864	0.2184	2.2837	1.9787	87.3520	45.6746	98.9341	$2000/2515 \simeq 80\%$
			C	0.3820	6.3580	5.6467	0.2994	5.1446	4.1499	119.7610	102.8927	207.4966	$2000/2223 \simeq 90\%$
	0.10	1.00	NC	0.3320	0.6580	0.6630	0.2634	0.4794	0.5380	105.3703	479.3548	53.7974	$2000/3520 \simeq 57\%$
			C	0.4851	1.5345	1.1760	0.3916	1.0672	0.9876	156.6558	1067.1860	98.7612	$2000/2533 \simeq 79\%$
	2.00	5.00	NC	0.3097	3.3616	3.3738	0.2404	2.6656	2.7813	96.1717	133.2785	55.6251	$2000/3137 \simeq 64\%$
			C	0.4735	6.8657	5.6124	0.3706	4.8976	4.7397	148.2420	244.8814	94.7940	$2000/2265 \simeq 88\%$
	0.05	0.10	NC	0.2672	0.0750	0.0727	0.2077	0.0621	0.0620	83.0659	124.1018	62.0471	$2000/3015 \simeq 66\%$
			C	0.4473	0.3198	0.3676	0.3585	0.2216	0.2602	143.4173	443.1005	260.2065	$2000/3033 \simeq 66\%$
0.75	0.10	0.05	NC	0.1595	0.0503	0.0367	0.1228	0.0413	0.0309	16.3687	41.2797	61.7597	$2000/2265 \simeq 88\%$
			C	0.4356	0.3444	0.5165	0.3019	0.1916	0.3533	40.2592	191.5928	706.5637	$2000/3843 \simeq 52\%$
	1.00	0.10	NC	0.1190	0.3430	0.1885	0.0917	0.2728	0.1408	12.2261	27.2783	140.7727	$2000/2374 \simeq 84\%$
			C	0.3056	1.3890	2.3812	0.2071	0.9184	1.7949	27.6161	91.8402	1794.8840	$2000/2552 \simeq 78\%$
	5.00	2.00	NC	0.1364	2.2249	1.5857	0.1062	1.8382	1.3111	14.1653	36.7633	65.5527	$2000/2220 \simeq 90\%$
			C	0.2899	6.8220	10.8105	0.1897	4.5106	8.5797	25.2983	90.2117	428.9829	$2000/2192 \simeq 91\%$
	0.10	1.00	NC	0.3152	0.4849	0.4972	0.2410	0.3009	0.3666	32.1341	300.8559	36.6612	$2000/2695 \simeq 74\%$
			C	0.5611	1.4225	1.4693	0.3981	0.8159	1.2037	53.0740	815.9315	120.3714	$2000/2751 \simeq 73\%$
	2.00	5.00	NC	0.2362	2.6149	2.4755	0.1784	1.8878	1.9114	23.7931	94.3914	38.2272	$2000/2228 \simeq 90\%$
			C	0.4479	6.7006	7.5337	0.3085	4.2198	6.1402	41.1287	210.9902	122.8036	$2000/2302 \simeq 87\%$
	0.05	0.10	NC	0.2296	0.0582	0.0526	0.1743	0.0429	0.0414	23.2456	85.7212	41.3812	$2000/2223 \simeq 90\%$
			C	0.5148	0.3089	0.4349	0.3731	0.1787	0.3175	49.7412	357.4894	317.4564	$2000/3234 \simeq 62\%$

As expected, contamination had an impact on the performance of the maximum likelihood estimators.

First, it is seen that for small sample sizes and non-contaminated errors, the convergence rate tends to decrease. For example, for $n = 500$ in the case of non-contaminated Gaussian errors, the convergence rate was over 72%, while for $n = 50$, it was over 57%. For contaminated Gaussian and exponential errors, the convergence rate decreased compared to non-contaminated errors.

Overall, an improvement in the rate of valid estimates (convergence rate) is noticeable when $\phi = 0.75$ compared to $\phi = 0.25$ in the case of non-contaminated Gaussian and exponential errors. In the case of contaminated Gaussian and exponential errors, this behavior only occurred when $n = 500$.

Table 2. RMSE, MAE, MAPE and convergence rate of Θ with 2000 simulations of sample sizes n = 200, considering Gaussian errors (NC = Non-Contaminated; C = Contaminated).

Parameters				**RMSE**			**MAE**			**MAPE (%)**			**Convergence Rate (%)**
ϕ	σ_ε^2	σ_e^2		ϕ	σ_ε^2	σ_e^2	ϕ	σ_ε^2	σ_e^2	ϕ	σ_ε^2	σ_e^2	
0.25	0.10	0.05	NC	0.2125	0.0533	0.0486	0.1552	0.0481	0.0445	62.0843	48.0637	89.0227	$^{2000}/_{2142} \simeq 93\%$
			C	0.4414	0.2937	0.3973	0.3511	0.2170	0.3170	140.4550	216.9786	634.0003	$^{2000}/_{3415} \simeq 59\%$
	1.00	0.10	NC	0.1827	0.3872	0.3342	0.1263	0.2890	0.2466	50.5172	28.8980	246.6108	$^{2000}/_{2158} \simeq 93\%$
			C	0.3302	1.2494	1.3655	0.2531	1.1183	0.8983	101.2448	111.8291	898.3335	$^{2000}/_{2339} \simeq 86\%$
	5.00	2.00	NC	0.2257	2.4857	2.2298	0.1647	2.1584	1.9656	65.8709	43.1676	98.2816	$^{2000}/_{2114} \simeq 95\%$
			C	0.3210	6.0353	5.7585	0.2525	5.2843	4.1940	101.0001	105.6860	209.6988	$^{2000}/_{2171} \simeq 92\%$
	0.10	1.00	NC	0.3294	0.5910	0.5952	0.2693	0.4203	0.4418	107.7206	420.3230	44.1772	$^{2000}/_{3064} \simeq 65\%$
			C	0.5079	1.1490	1.2565	0.4159	0.7094	1.1202	166.3406	709.3999	112.0214	$^{2000}/_{2934} \simeq 68\%$
	2.00	5.00	NC	0.2942	3.1051	3.0346	0.2352	2.4888	2.4204	94.0959	124.4377	48.4077	$^{2000}/_{2432} \simeq 82\%$
			C	0.4888	5.2005	5.9769	0.3932	3.6031	5.3909	157.2640	180.1539	107.8174	$^{2000}/_{2355} \simeq 85\%$
	0.05	0.10	NC	0.2512	0.0690	0.0672	0.1992	0.0574	0.0555	79.6690	114.7299	55.4981	$^{2000}/_{2353} \simeq 85\%$
			C	0.4992	0.2898	0.3841	0.4026	0.1951	0.3182	161.0268	390.1951	318.2143	$^{2000}/_{3550} \simeq 56\%$
0.75	0.10	0.05	NC	0.0791	0.0306	0.0223	0.0613	0.0243	0.0177	8.1741	24.2574	35.4005	$^{2000}/_{2020} \simeq 99\%$
			C	0.3249	0.1774	0.6307	0.1998	0.1042	0.5622	26.6395	104.1552	1124.3160	$^{2000}/_{5655} \simeq 35\%$
	1.00	0.10	NC	0.0557	0.1838	0.1057	0.0442	0.1468	0.0859	5.8966	14.6823	85.8500	$^{2000}/_{2175} \simeq 92\%$
			C	0.2726	0.7185	2.5998	0.1459	0.5113	2.3158	19.4529	51.1345	2315.7740	$^{2000}/_{3564} \simeq 56\%$
	5.00	2.00	NC	0.0763	1.4259	0.9946	0.0596	1.1414	0.7971	7.9484	22.8271	39.8563	$^{2000}/_{2022} \simeq 99\%$
			C	0.1241	3.4324	10.9591	0.0944	2.3950	10.0756	12.5843	47.8992	503.7812	$^{2000}/_{2054} \simeq 97\%$
	0.10	1.00	NC	0.2457	0.3409	0.3348	0.1779	0.1836	0.2116	23.7169	183.5833	21.1571	$^{2000}/_{2139} \simeq 94\%$
			C	0.4690	0.8662	1.5181	0.3137	0.4036	1.3994	41.8257	403.6151	139.9393	$^{2000}/_{2673} \simeq 75\%$
	2.00	5.00	NC	0.1293	1.4609	1.3363	0.0943	1.0084	0.9691	12.5672	50.4175	19.3819	$^{2000}/_{2012} \simeq 99\%$
			C	0.3233	3.2270	7.9895	0.1826	1.8840	7.3360	24.3493	94.1989	146.7208	$^{2000}/_{2320} \simeq 86\%$
	0.05	0.10	NC	0.1246	0.0326	0.0291	0.0927	0.0232	0.0216	12.3547	46.3956	21.6304	$^{2000}/_{2025} \simeq 99\%$
			C	0.3633	0.2014	0.4895	0.2410	0.1021	0.4373	32.1288	204.1943	437.3301	$^{2000}/_{4233} \simeq 47\%$

When the errors are not contaminated, the RMSE, MAE and MAPE tend to decrease with increasing sample size. However, this premise is not true when the errors are contaminated. In fact, it was found that for both Gaussian and exponential errors, outliers had more impact in two situations: when $\phi = 0.75$ and $n = 50$ (Tables 1 and 4); and when $\phi = 0.25$ and $n = 500$ (Tables 3 and 6). This impact is reflected in the RMSE, MAE and MAPE, which produced very high values.

Furthermore, there are many cases where, for example, the RMSE of the estimators of the contaminated errors are 3 times higher than the RMSE of the non-contaminated errors. For example, in the case of the Gaussian errors with n = 500, $\phi = 0.25$, $\sigma_\varepsilon^2 = 0.10$ and $\sigma_e^2 = 0.05$, the RMSE of ϕ, σ_ε^2 and σ_e^2 of the contaminated Gaussian errors were about 3, 6 and 11 times higher, respectively, compared to the non-contaminated Gaussian errors (Table 3).

On the other hand, comparing both the Gaussian and exponential error cases, we find that there are no significant differences in the convergence rate, as well as in the efficiency of the autoregressive ϕ estimator. However, the RMSE, MAE and MAPE of the variance estimators, σ_ε^2 and σ_e^2, are in general higher in the case of exponential errors.

Table 3. RMSE, MAE, MAPE and convergence rate of Θ with 2000 simulations of sample sizes $n = 500$, considering Gaussian errors (NC = non-contaminated; C = contaminated).

Parameters				**RMSE**			**MAE**			**MAPE (%)**			**Convergence Rate**
ϕ	σ_ε^2	σ_e^2		ϕ	σ_ε^2	σ_e^2	ϕ	σ_ε^2	σ_e^2	ϕ	σ_ε^2	σ_e^2	**(%)**
0.25	0.10	0.05	NC	0.1670	0.0487	0.0451	0.1246	0.0440	0.0410	49.8262	43.9675	81.9983	$2000/2090 \simeq 96\%$
			C	0.5099	0.2843	0.4946	0.4266	0.2027	0.4261	170.6503	202.6937	852.1144	$2000/3936 \simeq 51\%$
	1.00	0.10	NC	0.1322	0.3212	0.2834	0.0926	0.2411	0.2142	37.0469	24.1096	214.2324	$2000/2073 \simeq 96\%$
			C	0.3757	1.0511	1.6760	0.2954	0.9210	1.3679	118.1795	92.1025	1367.9340	$2000/2309 \simeq 87\%$
	5.00	2.00	NC	0.1665	2.1973	2.0204	0.1219	1.9401	1.8018	48.7737	38.8022	90.0924	$2000/2015 \simeq 99\%$
			C	0.3571	5.2313	7.1197	0.2745	4.4927	6.0116	109.7836	89.8549	300.5794	$2000/2154 \simeq 93\%$
	0.10	1.00	NC	0.3186	0.5157	0.5157	0.2655	0.3497	0.3555	106.1993	349.7172	35.5477	$2000/2793 \simeq 72\%$
			C	0.5660	1.0072	1.2870	0.4735	0.5918	1.1856	189.3834	591.7733	118.5649	$2000/2666 \simeq 75\%$
	2.00	5.00	NC	0.2596	2.7002	2.6426	0.2080	2.1282	2.0529	83.2062	106.4111	41.0575	$2000/2102 \simeq 95\%$
			C	0.4327	4.9215	5.7057	0.3449	3.4861	5.2227	137.9698	174.3029	104.4536	$2000/2347 \simeq 85\%$
	0.05	0.10	NC	0.2375	0.0645	0.0628	0.1900	0.0528	0.0510	75.9833	105.5881	50.9948	$2000/2082 \simeq 96\%$
			C	0.5787	0.2691	0.4908	0.5039	0.1635	0.4430	201.5559	326.9245	443.0177	$2000/4751 \simeq 42\%$
0.75	0.10	0.05	NC	0.0477	0.0195	0.0142	0.0373	0.0154	0.0114	4.9771	15.3516	22.7729	$2000/2003 \simeq 100\%$
			C	0.1696	0.0817	0.6618	0.1106	0.0549	0.6455	14.7532	54.8753	1291.0500	$2000/2501 \simeq 80\%$
	1.00	0.10	NC	0.0395	0.1343	0.0782	0.0318	0.1081	0.0647	4.2341	10.8147	64.6665	$2000/2090 \simeq 96\%$
			C	0.0732	0.3663	2.5760	0.0587	0.2834	2.5003	7.8213	28.3405	2500.3230	$2000/2815 \simeq 71\%$
	5.00	2.00	NC	0.0474	0.9427	0.6600	0.0371	0.7485	0.5228	4.9477	14.9700	26.1379	$2000/2005 \simeq 100\%$
			C	0.0744	1.9273	10.4652	0.0578	1.4293	10.1291	7.7126	28.5863	506.4527	$2000/2020 \simeq 99\%$
	0.10	1.00	NC	0.1732	0.2068	0.2027	0.1219	0.1011	0.1174	16.2546	101.1061	11.7441	$2000/2001 \simeq 100\%$
			C	0.2439	0.5109	1.5706	0.2001	0.2151	1.5087	26.6834	215.1086	150.8725	$2000/2390 \simeq 84\%$
	2.00	5.00	NC	0.0723	0.7514	0.7300	0.0554	0.5623	0.5663	7.3812	28.1170	11.3252	$2000/2000 \simeq 100\%$
			C	0.1162	1.6095	7.9601	0.0903	1.0526	7.6541	12.0342	52.6321	153.0817	$2000/2014 \simeq 99\%$
	0.05	0.10	NC	0.0689	0.0175	0.0159	0.0528	0.0131	0.0122	7.0341	26.2519	12.2090	$2000/2002 \simeq 100\%$
			C	0.1914	0.1201	0.5832	0.1534	0.0547	0.5635	20.4470	109.4875	563.4770	$2000/2293 \simeq 87\%$

Table 4. RMSE, MAE, MAPE and convergence rate of Θ with 2000 simulations of sample sizes $n = 50$, considering exponential errors (NC = non-contaminated; C = contaminated).

Parameters				**RMSE**			**MAE**			**MAPE (%)**			**Convergence rate**
ϕ	σ_ε^2	σ_e^2		ϕ	σ_ε^2	σ_e^2	ϕ	σ_ε^2	σ_e^2	ϕ	σ_ε^2	σ_e^2	**(%)**
0.25	0.10	0.05	NC	0.2403	0.0621	0.0520	0.1799	0.0504	0.0457	71.9672	50.3763	91.4927	$2000/2635 \simeq 76\%$
			C	0.3995	0.3399	0.3274	0.3134	0.2428	0.2110	125.3731	242.7797	422.0013	$2000/3048 \simeq 66\%$
	1.00	0.10	NC	0.2591	0.5315	0.3983	0.1934	0.4320	0.2635	77.3567	43.2020	263.5235	$2000/2442 \simeq 82\%$
			C	0.3516	1.3525	1.3138	0.2744	1.0954	0.7735	109.7515	109.5363	773.5130	$2000/2313 \simeq 86\%$
	5.00	2.00	NC	0.2810	3.0601	2.4320	0.2140	2.4909	1.9947	85.6126	49.8187	99.7367	$2000/2489 \simeq 80\%$
			C	0.3788	6.3086	5.6419	0.2971	5.0720	4.0209	118.8333	101.4390	201.0444	$2000/2221 \simeq 90\%$
	0.10	1.00	NC	0.3352	0.7070	0.7121	0.2656	0.4958	0.6145	106.2327	495.7865	61.4506	$2000/3845 \simeq 52\%$
			C	0.4979	1.4952	1.2090	0.4045	1.0352	1.0202	161.7847	1035.2270	102.0164	$2000/2500 \simeq 80\%$
	2.00	5.00	NC	0.3036	3.5020	3.6159	0.2359	2.7018	3.0956	94.3541	135.0912	61.9127	$2000/3273 \simeq 61\%$
			C	0.4756	7.4254	5.7965	0.3764	5.2303	4.8901	150.5748	261.5164	97.8024	$2000/2281 \simeq 88\%$
	0.05	0.10	NC	0.2712	0.0818	0.0775	0.2089	0.0644	0.0676	83.5522	128.8785	67.6088	$2000/3045 \simeq 66\%$
			C	0.4505	0.3498	0.3421	0.3573	0.2485	0.2372	142.9146	496.9232	237.1900	$2000/3014 \simeq 66\%$

Table 4. *Cont.*

Parameters				RMSE			MAE			MAPE (%)			Convergence Rate
ϕ	σ_ε^2	σ_e^2		ϕ	σ_ε^2	σ_e^2	ϕ	σ_ε^2	σ_e^2	ϕ	σ_ε^2	σ_e^2	(%)
0.75	0.10	0.05	NC	0.1611	0.0564	0.0398	0.1246	0.0450	0.0322	16.6099	45.0059	64.3029	$^{2000}/_{2273} \simeq 88\%$
			C	0.4359	0.3223	0.5405	0.3033	0.1875	0.3750	40.4453	187.5322	750.0313	$^{2000}/_{3694} \simeq 54\%$
	1.00	0.10	NC	0.1175	0.4488	0.1929	0.0925	0.3574	0.1397	12.3275	35.7367	139.7342	$^{2000}/_{2364} \simeq 85\%$
			C	0.3433	1.3662	2.5133	0.2284	0.9272	1.8790	30.4537	92.7218	1879.0050	$^{2000}/_{2470} \simeq 81\%$
	5.00	2.00	NC	0.1448	2.7020	1.7189	0.1120	2.1216	1.3609	14.9294	42.4323	68.0429	$^{2000}/_{2181} \simeq 92\%$
			C	0.3000	6.8179	11.0149	0.1977	4.5487	8.6278	26.3555	90.9748	431.3905	$^{2000}/_{2176} \simeq 92\%$
	0.10	1.00	NC	0.3093	0.4945	0.5622	0.2368	0.3007	0.4524	31.5769	300.7228	45.2368	$^{2000}/_{2672} \simeq 75\%$
			C	0.5765	1.3806	1.5393	0.4103	0.7817	1.2490	54.7124	781.7267	124.9006	$^{2000}/_{2792} \simeq 72\%$
	2.00	5.00	NC	0.2394	2.8641	2.9144	0.1801	1.9810	2.3408	24.0172	99.0487	46.8162	$^{2000}/_{2221} \simeq 90\%$
			C	0.4688	6.9340	7.5328	0.3241	4.3006	6.0393	43.2112	215.0307	120.7854	$^{2000}/_{2277} \simeq 88\%$
	0.05	0.10	NC	0.2345	0.0614	0.0594	0.1767	0.0437	0.0484	23.5599	87.3926	48.3987	$^{2000}/_{2246} \simeq 89\%$
			C	0.5399	0.3405	0.4259	0.4039	0.2083	0.3024	53.8548	416.5115	302.3952	$^{2000}/_{3314} \simeq 60\%$

Table 5. RMSE, MAE, MAPE and convergence rate of Θ with 2000 simulations of sample sizes $n = 200$, considering exponential errors (NC = non-contaminated; C = contaminated).

Parameters				RMSE			MAE			MAPE (%)			Convergence Rate
ϕ	σ_ε^2	σ_e^2		ϕ	σ_ε^2	σ_e^2	ϕ	σ_ε^2	σ_e^2	ϕ	σ_ε^2	σ_e^2	(%)
0.25	0.10	0.05	NC	0.2195	0.0567	0.0502	0.1605	0.0501	0.0458	64.1779	50.0747	91.6168	$^{2000}/_{2185} \simeq 92\%$
			C	0.4489	0.2919	0.4062	0.3589	0.2159	0.3230	143.5695	215.8788	645.9832	$^{2000}/_{3303} \simeq 61\%$
	1.00	0.10	NC	0.1942	0.4247	0.3510	0.1343	0.3304	0.2550	53.7222	33.0384	255.0102	$^{2000}/_{2194} \simeq 91\%$
			C	0.3275	1.2229	1.3829	0.2520	1.0916	0.8967	100.8199	109.1563	896.7205	$^{2000}/_{2299} \simeq 87\%$
	5.00	2.00	NC	0.2352	2.6233	2.2958	0.1709	2.2450	2.0030	68.3418	44.8991	100.1478	$^{2000}/_{2120} \simeq 94\%$
			C	0.3157	6.0485	5.8589	0.2491	5.2427	4.2303	99.6284	104.8537	211.5132	$^{2000}/_{2170} \simeq 92\%$
	0.10	1.00	NC	0.3263	0.6068	0.6071	0.2698	0.4264	0.4734	107.9090	426.4313	47.3421	$^{2000}/_{3235} \simeq 62\%$
			C	0.5063	1.1797	1.2615	0.4136	0.7430	1.1099	165.4252	743.0232	110.9928	$^{2000}/_{2793} \simeq 72\%$
	2.00	5.00	NC	0.2871	3.0873	3.0756	0.2286	2.4294	2.4630	91.4476	121.4710	49.2608	$^{2000}/_{2409} \simeq 83\%$
			C	0.4800	5.3433	5.8893	0.3878	3.7179	5.2539	155.1106	185.8963	105.0775	$^{2000}/_{2349} \simeq 85\%$
	0.05	0.10	NC	0.2547	0.0706	0.0689	0.2040	0.0582	0.0576	81.5998	116.3832	57.6243	$^{2000}/_{2381} \simeq 84\%$
			C	0.4861	0.2824	0.3672	0.3959	0.1923	0.3042	158.3634	384.6917	304.2286	$^{2000}/_{3600} \simeq 56\%$
0.75	0.10	0.05	NC	0.0796	0.0350	0.0233	0.0617	0.0274	0.0186	8.2315	27.4339	37.1357	$^{2000}/_{2037} \simeq 98\%$
			C	0.3272	0.2014	0.6101	0.1953	0.1114	0.5404	26.0350	111.3510	1080.8030	$^{2000}/_{5646} \simeq 35\%$
	1.00	0.10	NC	0.0597	0.2508	0.1085	0.0474	0.1994	0.0873	6.3241	19.9427	87.2755	$^{2000}/_{2177} \simeq 92\%$
			C	0.2891	0.7251	2.6068	0.1467	0.5148	2.3368	19.5576	51.4762	2336.7630	$^{2000}/_{3530} \simeq 57\%$
	5.00	2.00	NC	0.0746	1.6329	1.0466	0.0594	1.3146	0.8439	7.9157	26.2910	42.1946	$^{2000}/_{2025} \simeq 99\%$
			C	0.1272	3.6999	10.6605	0.0940	2.5298	9.7850	12.5288	50.5952	489.2506	$^{2000}/_{2058} \simeq 97\%$
	0.10	1.00	NC	0.2397	0.3397	0.3613	0.1728	0.1807	0.2566	23.0433	180.7138	25.6634	$^{2000}/_{2212} \simeq 90\%$
			C	0.4477	0.8176	1.5542	0.3042	0.3849	1.4218	40.5591	384.9222	142.1828	$^{2000}/_{2667} \simeq 75\%$
	2.00	5.00	NC	0.1296	1.5155	1.5429	0.0951	1.0538	1.1744	12.6814	52.6910	23.4870	$^{2000}/_{2015} \simeq 99\%$
			C	0.3411	3.2286	8.1396	0.1906	1.8699	7.4604	25.4199	93.4933	149.2072	$^{2000}/_{2354} \simeq 85\%$
	0.05	0.10	NC	0.1199	0.0326	0.0327	0.0888	0.0235	0.0253	11.8369	46.9942	25.2911	$^{2000}/_{2029} \simeq 99\%$
			C	0.4057	0.1971	0.4953	0.2636	0.0980	0.4410	35.1415	196.0500	440.9552	$^{2000}/_{4253} \simeq 47\%$

Table 6. RMSE, MAE, MAPE and convergence rate of Θ with 2000 simulations of sample sizes $n = 500$, considering exponential errors (NC = non-contaminated; C = contaminated).

Parameters				RMSE			MAE			MAPE (%)			Convergence Rate
ϕ	σ_ε^2	σ_e^2		ϕ	σ_ε^2	σ_e^2	ϕ	σ_ε^2	σ_e^2	ϕ	σ_ε^2	σ_e^2	(%)
0.25	0.10	0.05	NC	0.1774	0.0502	0.0459	0.1295	0.0449	0.0415	51.8133	44.9298	83.0451	$2000/2101 \simeq 95\%$
			C	0.5105	0.2796	0.4923	0.4312	0.2000	0.4234	172.4715	200.0235	846.7976	$2000/3882 \simeq 52\%$
	1.00	0.10	NC	0.1360	0.3393	0.2890	0.0925	0.2583	0.2157	37.0141	25.8315	215.7496	$2000/2068 \simeq 97\%$
			C	0.3751	1.0426	1.6910	0.2938	0.9117	1.3735	117.5078	91.1686	1373.5180	$2000/2331 \simeq 86\%$
	5.00	2.00	NC	0.1704	2.2503	2.0470	0.1234	1.9625	1.8020	49.3534	39.2493	90.0986	$2000/2017 \simeq 99\%$
			C	0.3479	5.3039	7.0687	0.2679	4.5285	5.9227	107.1780	90.5707	296.1330	$2000/2129 \simeq 94\%$
	0.10	1.00	NC	0.3131	0.5300	0.5411	0.2597	0.3703	0.3968	103.8980	370.3198	39.6789	$2000/2816 \simeq 71\%$
			C	0.5597	1.0212	1.2769	0.4684	0.6074	1.1618	187.3698	607.3962	116.1793	$2000/2652 \simeq 75\%$
	2.00	5.00	NC	0.2601	2.7288	2.7233	0.2073	2.1620	2.1463	82.9108	108.0979	42.9251	$2000/2126 \simeq 94\%$
			C	0.4306	4.9082	5.7249	0.3452	3.5071	5.1919	138.0972	175.3540	103.8379	$2000/2317 \simeq 86\%$
	0.05	0.10	NC	0.2376	0.0645	0.0627	0.1901	0.0524	0.0510	76.0432	104.7235	50.9923	$2000/2074 \simeq 96\%$
			C	0.5858	0.2454	0.4984	0.5118	0.1446	0.4560	204.7275	289.2989	456.0194	$2000/4715 \simeq 42\%$
0.75	0.10	0.05	NC	0.0475	0.0231	0.0155	0.0372	0.0181	0.0123	4.9720	18.1002	24.5343	$2000/2004 \simeq 100\%$
			C	0.1591	0.0842	0.6609	0.1062	0.0545	0.6454	14.1550	54.5197	1290.8090	$2000/2474 \simeq 81\%$
	1.00	0.10	NC	0.0384	0.1704	0.0773	0.0307	0.1373	0.0643	4.0941	13.7260	64.3178	$2000/2080 \simeq 96\%$
			C	0.0721	0.3674	2.5627	0.0581	0.2882	2.4862	7.7527	28.8233	2486.2120	$2000/2794 \simeq 72\%$
	5.00	2.00	NC	0.0466	1.0946	0.6805	0.0370	0.8577	0.5377	4.9267	17.1536	26.8825	$2000/2003 \simeq 100\%$
			C	0.0724	1.9116	10.6844	0.0572	1.4456	10.3496	7.6297	28.9121	517.4798	$2000/2039 \simeq 98\%$
	0.10	1.00	NC	0.1725	0.1945	0.2149	0.1211	0.0988	0.1473	16.1486	98.8105	14.7318	$2000/2010 \simeq 100\%$
			C	0.2492	0.5142	1.5627	0.1972	0.2080	1.4974	26.2946	208.0308	149.7398	$2000/2384 \simeq 84\%$
	2.00	5.00	NC	0.0760	0.8252	0.9283	0.0576	0.6068	0.7308	7.6842	30.3421	14.6165	$2000/2001 \simeq 100\%$
			C	0.1204	1.7123	7.9869	0.0934	1.0745	7.6659	12.4592	53.7261	153.3177	$2000/2013 \simeq 99\%$
	0.05	0.10	NC	0.0684	0.0185	0.0190	0.0524	0.0138	0.0149	6.9855	27.6962	14.9040	$2000/2001 \simeq 100\%$
			C	0.1908	0.1162	0.5854	0.1535	0.0531	0.5665	20.4642	106.2381	566.4536	$2000/2334 \simeq 86\%$

4. Discussion

In this work, outliers were found to impact the performance of the Maximum Likelihood estimators. In particular, it was found through the simulation study that outliers have a very significant impact in both cases: when the sample size is small and the autoregressive parameter is close to 1, and when the sample size is large and the autoregressive parameter is close to 0.25. This impact was reflected in the RMSE, MAE and MAPE values which, in many cases, were higher compared to the case of non-contaminated errors.

Moreover, we notice that the rate of valid estimates (convergence rate) is higher for large sample sizes, and is more evident for non-contaminated Gaussian and exponential errors. On the other hand, it is also important to have large sample sizes to avoid problems related to parameter estimation [11]. In general, the convergence rate is lower when Gaussian and exponential errors are contaminated.

Therefore, our next step is to develop methods to detect outliers in time series and/or to establish other estimation methods that are more robust, in the sense that they do not assume a distribution of the data and are less sensitive to outliers.

In this work, the outliers were generated from a regression model that established a linear relationship between the magnitude of the outliers and the total variance of the model with the state space representation of maximum air temperature real data. The rate of outliers from the real data was 5%; thus, this was the percentage used in this work.

In the literature, we did not find a unanimous approach for doing this. For example, ref. [15] contaminated the error of the zero-mean Gaussian equation of state by replacing the standard deviation of the observation error with a 10-times-higher standard deviation with a probability of 10% (symmetric outliers). They also considered the case of asymmetric outliers, where the zero mean of the observation error was replaced with a value 10 times higher than the standard deviation with a probability of 10%. Ref. [16] followed the same line as [15], but in this case they call symmetric outliers "zero-mean" and asymmetric outliers "non-zero", considering the probability of contamination to be 5%. Ref. [9] contaminated both the observation and state equation errors, considering the magnitude of the outliers equal to 2.5 the standard deviation from the diagonal elements of the observation and state covariance matrices, respectively.

Author Contributions: F.C.P., A.M.G. and M.C. contributed to this work. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by FEDER/COMPETE/NORTE 2020/POCI/FCT funds through grants UID/EEA/-00147/20 13/UID/IEEA/00147/006933-SYSTEC project and To CHAIR - POCI-01-0145-FEDER-028247. A. Manuela Gonçalves was partially financed by Portuguese Funds through FCT (Fundação para a Ciência e a Tecnologia) within the Projects UIDB/00013/2020 and UIDP/00013/2020 of CMAT-UM. Marco Costa was partially supported by The Center for Research and Development in Mathematics and Applications (CIDMA) through the Portuguese Foundation for Science and Technology (FCT - Fundação para a Ciência e a Tecnologia), references UIDB/04106/2020 and UIDP/04106/2020. F. Catarina Pereira was financed by national funds through FCT (Fundação para a Ciência e a Tecnologia) through the individual PhD research grant UI/BD/150967/2021 of CMAT-UM.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Hamilton, J.D. *Time Series Analysis*; Princeton University Press: Princeton, NJ, USA, 1994.
2. Harvey, A.C. *Forecasting, Structural Time Series Models and the Kalman Filter*; Cambridge University Press: Cambridge, UK, 2009.
3. Shumway, R.H.; Stoffer, D.S. *Time Series Analysis and its Applications: With R Examples*; Springer: New York, NY, USA, 2017.
4. Petris, G.; Petrone, S.; Campagnoli, P. *Dynamic Linear Models with R*; Springer: Berlin/Heidelberg, Germany, 2009.
5. Durbin, J.; Koopman, S. *Time Series Analysis by State Space Methods*; Oxford University Press: Oxford, UK, 2001.
6. Kalman, R. A New Approach to Linear Filtering and Prediction Problems. *ASME J. Basic Eng.* **1960**, *82*, 35–45. [CrossRef]
7. Costa, M.; Alpuim, T. Parameter estimation of state space models for univariate observations. *J. Stat. Plan. Inference* **2010**, *140*, 1889–1902. [CrossRef]
8. Costa, M.; Monteiro, M. Bias-correction of kalman filter estimators associated to a linear state space model with estimated parameters. *J. Stat. Plan. Inference* **2016**, *176*, 22–32. [CrossRef]
9. You, D.; Hunter, M.; Chen, M.; Chow, S.M. A diagnostic procedure for detecting outliers in linear state-space models. *Multivar. Behav. Res.* **2020**, *55*, 231–255. [CrossRef] [PubMed]
10. Cipra, T.; Romera, R. Kalman filter with outliers and missing observations. *Test* **1997**, *6*, 379–395. [CrossRef]
11. Auger-Méthé, M.; Field, C.; Albertsen, C.M.; Derocher, A.E.; Lewis, M.A.; Jonsen, I.D.; Flemming, J.M. State-space models' dirty little secrets: Even simple linear Gaussian models can have estimation problems. *Sci. Rep.* **2016**, *6*, 26677. [CrossRef] [PubMed]
12. Pandolfo, G.; Iorio, C.; Siciliano, R.; D'Ambrosio, A. Robust mean-variance portfolio through the weighted L^p depth function. In *Annals of Operations Research*; Springer: Berlin/Heidelberg, Germany, 2020; Volume 292, pp. 519–531.
13. Shumway, R.H.; Stoffer, D.S. *Time Series: A Data Analysis Approach Using R*; CRC Press: Boca Raton, FL, USA, 2019.
14. R Core Team. *R: A Language and Environment for Statistical Computing*; R Foundation for Statistical Computing: Vienna, Austria, 2021.
15. Crevits, R.; Croux, C. Robust estimation of linear state space models. *Commun. Stat.-Simul. Comput.* **2019**, *48*, 1694–1705. [CrossRef]
16. Ali, K.; Tahir, M. Maximum likelihood-based robust state estimation over a horizon length during measurement outliers. *Trans. Inst. Meas. Control* **2021**, *43*, 510–518. [CrossRef]