

Article

# Measuring the Distance between Sets of ARMA Models

Umberto Triacca

Department of Computer Engineering, Computer Science and Mathematics, University of L'Aquila, Via Vetoio Coppito, L'Aquila I-67010, Italy; umberto.triacca@univaq.it; Tel.: +39-06-4456100

Academic Editor: Kerry Patterson

Received: 5 May 2016; Accepted: 8 July 2016; Published: 15 July 2016

**Abstract:** A distance between pairs of sets of autoregressive moving average (ARMA) processes is proposed. Its main properties are discussed. The paper also shows how the proposed distance finds application in time series analysis. In particular it can be used to evaluate the distance between portfolios of ARMA models or the distance between vector autoregressive (VAR) models.

**Keywords:** ARMA models; distance; time series; VAR models

**JEL:** C1; C32

---

## 1. Introduction

Measuring the similarity and dissimilarity between models is crucial in many fields of time series analysis. We limit ourselves to mentioning some applications: clustering time series data (see [1,2]), data mining problems (see [3]), time series classification (see [4]), selecting between direct and indirect model-based seasonal adjustment (see [5]), the analysis of Granger causality (see [6]), comparing autocorrelation structures of multiple time series (see [7]).

A number of measures of dissimilarity between univariate linear models have been suggested in the literature. Among these, we find the Mahalanobis distance between autoregressive models used in [8], the autoregressive distances introduced in [9–11], the cepstral distance proposed in [12] to be of note. In this paper, we introduce a distance measure between sets of invertible autoregressive moving average (ARMA) models and we show how it finds application in time series analysis. In particular, two specific applications are described. The first application consists of measuring the distance between two single ARMA series,  $x$  and  $y$ , by means of the proposed distance between portfolios of ARMA models providing reasonably good fits of  $x$  and  $y$ . The second application uses our distance to compute distances between two vector autoregressive (VAR) models once these models are properly represented in terms of univariate ARMA models. The remainder of the paper is organized as follows: Section 2 presents the distance. Section 3 shows how this distance can be used. An empirical application is presented in Section 4. Section 5 concludes.

## 2. Proposed Distance

The intrinsic nature of a time series is usually that the observations are dependent or correlated. The autoregressive moving average (ARMA) processes are a very general class of parametric models useful for describing such correlations. For a general reference on ARMA models, see Brockwell and Davis [13].

We have that  $x = \{x_t; t \in \mathbb{Z}\}$  is an ARMA( $p, q$ ) process if, for every  $t$ ,

$$\phi(L)x_t = \theta(L)\epsilon_t,$$

where  $\phi(L)$  and  $\theta(L)$  are polynomials in the lag operator  $L$ , with no common factors, and  $\epsilon_t$  is a white noise process with constant variance  $\sigma^2$ . The ARMA process  $x$  is said to be invertible if it admits the representation:

$$\pi(L)x_t = \epsilon_t,$$

where the  $AR(\infty)$  operator is defined by

$$\pi(L) = \phi(L)\theta(L)^{-1} = 1 - \sum_{i=1}^{\infty} \pi_i L^i$$

with  $\sum_{i=1}^{\infty} |\pi_i| < \infty$ .

Let  $M$  be the class of ARMA invertible processes and let  $d(x, y)$  be a metric on  $M$ . Now, we consider the family  $\mathcal{F}_M$  of all non-empty finite subsets of  $M$ . We want to introduce a metric on  $\mathcal{F}_M$ .

If  $S$  is an element of  $\mathcal{F}_M$  and  $x$  is a point of  $M$ , the distance from  $x$  to  $S$  is defined as

$$\delta(x, S) = \min \{d(x, s); s \in S\}. \tag{1}$$

Equation (1) represents the smallest distance between  $x \in M$  and any point  $s \in S$ . Given a subset  $T$  of  $M$ , let us define the function

$$h(S, T) = \max \{\delta(s, T); s \in S\},$$

that measures the largest among all distances  $\delta(s, T)$ , with  $s \in S$ . However, this function is not symmetric since in general  $h(S, T)$  is not equal to  $h(T, S)$ . Thus, we prefer to define the distance,  $H(S, T)$ , between two finite sets of invertible ARMA processes,  $S$  and  $T$ , as their Hausdorff distance. Formally, we have:

$$H(S, T) = \max \{h(S, T), h(T, S)\}.$$

The distance between two finite sets of invertible ARMA processes, defined in this way satisfies all the desirable metric properties. In fact, we can prove the following:

**Proposition 1.** *The function  $H : \mathcal{F}_M \times \mathcal{F}_M \rightarrow \mathbb{R}$  defined as*

$$H(S, T) = \max \{h(S, T), h(T, S)\} \quad S, T \in \mathcal{F}_M, \tag{2}$$

*is a metric on  $\mathcal{F}_M$ .*

**Proof.** The function (2) is obviously non-negative and the symmetry follows straight from the definition.

In order to prove the triangle inequality, we note that if  $h(S, T) < \alpha$ , then  $\delta(s, T) < \alpha$  for all  $s \in S$ , so there exists  $s \in S, t \in T$  such that  $d(s, t) < \alpha$ . Let  $S, F, T$  be three elements of  $\mathcal{F}_M$  and put  $h(S, F) = \alpha$  and  $h(F, T) = \beta$ . For each  $\epsilon > 0$  we have that  $h(F, T) < \beta + \epsilon$  and there exists  $f \in F, t \in T$  such that  $d(f, t) < \beta + \epsilon$ . Analogously, from  $h(S, F) < \alpha + \epsilon$  there exists  $s \in S$  such that  $d(s, f) < \alpha + \epsilon$ . Since  $d$  is a metric on  $M$ ,

$$d(s, t) \leq d(s, f) + d(f, t) < \alpha + \beta + 2\epsilon.$$

It follows that

$$h(S, T) \leq h(S, F) + h(F, T).$$

and

$$h(T, S) \leq h(T, F) + h(F, S).$$

Thus

$$H(S, F) + H(F, C) \geq H(S, T).$$

Finally, we show that  $H(S, T) = 0$  if and only if  $S = T$ . It is clear that if  $S = T$ , then  $H(S, T) = 0$ . Now, we assume that  $H(S, T) = 0$ . This implies that  $h(S, T) = 0$ . It follows that  $\delta(s, T) = 0 \forall s \in S$  and hence for every  $s \in S$  there exists a  $t \in T$  such that  $\delta(s, t) = 0$ . Since  $\delta(s, t)$  is a metric on  $M$ , this implies that  $s = t$ . Thus  $S \subset T$ . In an analogous way, we can show that  $T \subset S$ . Thus we can conclude that  $S = T$ . The non-negative real-valued function (2) is a metric on  $\mathcal{F}_M$ .  $\square$

In order to make operative the notion of distance between finite sets of invertible ARMA models introduced above, it is necessary to specify a particular distance  $d$  between invertible ARMA models. If  $x, y \in M$ , following Piccolo [9], we can consider the Euclidean distance between the corresponding  $\pi$ -weights sequence,  $\{\pi_j\}$ , that is

$$d_P(x, y) = \left[ \sum_{i=1}^{\infty} (\pi_{xi} - \pi_{yi})^2 \right]^{\frac{1}{2}} .$$

where  $\{\pi_{xi}\}$  and  $\{\pi_{yi}\}$  denote the sequences of AR weights of  $x$  and  $y$ , respectively.

The measure  $d_P(x, y)$  satisfies the following properties:

- i. Non-negativity:  $d_P(x, y) \geq 0 \forall x, y \in M$ ;
- ii. Symmetry:  $d_P(x, y) = d_P(y, x) \forall x, y \in M$ ;
- iii. Subadditivity:  $d_P(x, y) \leq d_P(x, z) + d_P(z, y) \forall x, y, z \in M$ .

We note that the distance between two ARMA processes, measured by  $d_P(x, y)$ , is allowed to be zero even if they are generated by different white noise processes,  $\epsilon_x$  and  $\epsilon_y$ .<sup>1</sup> This implies that  $d_P$  is a pseudometric on  $M$ . It follows that also  $H$ , with  $d = d_P$ , becomes a pseudometric on  $\mathcal{F}_M$ .

### 3. Applications

This section discusses two applications of the distance function defined in (2) in the field of time series analysis.

#### 3.1. Distance between Portfolios of ARMA Models

The first application consists of measuring the distance between two univariate ARMA series,  $x$  and  $y$ , by means of the proposed distance between portfolios of ARMA models providing reasonably good fits of  $x$  and  $y$ . The notion of portfolio of ARMA models has been introduced by Poskitt and Tremayne [14]. When modeling time series data an analyst typically searches over a range of models. Then a single model is usually selected as a satisfactory representation of the true, but unknown, underlying data generating mechanism. However, given that a wrong model may be selected or that a “best” model may not exist anyway (see Chatfield [15], p. 80), a better strategy could be to allow the possibility that there may be more than one model which may be regarded as a reasonable representation of the data generating mechanism.

Consider a process  $x = \{x_t; t \in \mathbb{Z}\}$  that admits an invertible ARMA representation

$$x_t - \phi_1 x_{t-1} - \dots - \phi_p x_{t-p} = u_t + \theta_1 \epsilon_{t-1} + \dots + \theta_q \epsilon_{t-q},$$

$$\epsilon_t \sim WN(0, \sigma^2).$$

Following Box and Jenkins [16], given a realization  $x_1, \dots, x_T$ , a common model building strategy is to initially select plausible values of  $p$  and  $q$  based on statistics (the Akaike information criterion (AIC), the Bayesian information criterion (BIC), etc.) calculated from the data.

---

<sup>1</sup> It is important to note that this is a good property of the distance measure  $d_P(x, y)$ . The white noise variance is simply a scale parameter and it is not relevant to measure distance between ARMA models.

Consider, for example, the BIC criterion

$$\text{BIC}(p, q) = \log \hat{\sigma}^2 + \frac{(p+q)\log T}{T},$$

where  $\hat{\sigma}^2$  is the maximum likelihood estimate of  $\sigma^2$ . Let  $p_1$  and  $q_1$  be the orders of the ARMA selected by using the BIC criterion, that is

$$\text{BIC}(p_1, q_1) = \min_{p \in \bar{P}, q \in \bar{Q}} \text{BIC}(p, q),$$

where  $\bar{P} = \{0, 1, \dots, P\}$  and  $\bar{Q} = \{0, 1, \dots, Q\}$ . We denote with  $M_j$  an ARMA( $p_j, q_j$ ) model (with  $p_j \in \bar{P}$  and  $q_j \in \bar{Q}$ ) for the process  $x$  and, following Poskitt and Tremayne [14], we consider the quantity

$$R(M_j) = \exp \left[ -\frac{1}{2} T \{ \text{BIC}(p_1, q_1) - \text{BIC}(p_j, q_j) \} \right].$$

We note that this quantity is related to the  $Y$  parameter introduced and studied in [17,18].

When  $1 < R(M_j) < \sqrt{10}$  the ARMA model  $M_j$  is said to be a “close competitor” to the criterion-minimizing ARMA( $p_1, q_1$ ) model. The set of closely competing models

$$P(x) = \{ M_j | 1 < R(M_j) < \sqrt{10} \}$$

is termed a model portfolio for the process  $x$ .

The concept of model portfolio suggests not only that the model minimizing the criterion should be selected but also that any additional specifications closely competing with ARMA( $p_1, q_1$ ) should not be discarded.

Using the Hausdorff distance  $H$ , we can evaluate the distance between portfolios of ARMA models. Let  $x$  and  $y$  be two invertible ARMA processes and let  $P(x)$  and  $P(y)$  be the model portfolios of the process  $x$  and  $y$ , respectively. Since  $P(x), P(y) \in \mathcal{F}_M$ , the distance between the portfolios  $P(x)$  and  $P(y)$  is given by  $H(P(x), P(y))$ .

### 3.2. Distance between VAR Processes

The vector autoregressive (VAR) process is a generalization of the univariate autoregressive process. The VAR processes has been popularized by Sims [19]. In this subsection, we will show how the proposed distance can be used to evaluate the distance between VAR processes.

Recall that a  $k$ -dimensional process  $y = \{ y_t = (y_{1t}, \dots, y_{kt})' ; t \in \mathbb{Z} \}$  is a VAR( $p$ ) process if it can be represented as

$$y_t = A_1 y_{t-1} + \dots + A_p y_{t-p} + \epsilon_t$$

where  $A_1, \dots, A_p$  are  $k \times k$  matrices of coefficients and  $\epsilon_t = (\epsilon_{1t}, \dots, \epsilon_{kt})'$  is a  $k$ -dimensional white noise, with non-singular covariance matrix  $\Sigma_\epsilon$ . Using the lag operator,  $L$ , we can rewrite the VAR model in the following way

$$A(L)y_t = \epsilon_t \tag{3}$$

where  $A(L) = I - A_1 L - \dots - A_p L^p$ .

We have

$$A(L)^{adj} A(L) = \det[A(L)] I$$

where  $A(L)^{adj}$  and  $\det[A(L)]$  are, respectively, the adjoint matrix and the determinant of the matrix  $A(L)$ .

Following Zellner and Palm [20], premultiplying both sides of (3) by  $A(L)^{adj}$ , we obtain the “final equations”:

$$\det [A(L)] y_t = A(L)^{adj} \epsilon_t \tag{4}$$

For the  $i$ th element of  $y_t$ , expression (4) becomes

$$\det [A(L)] y_{it} = A(L)_i^{adj} \epsilon_t \tag{5}$$

where  $A(L)_i^{adj}$  denotes the  $i$ th row of the adjoint matrix  $A(L)^{adj}$ .

Since the right-hand side of (5) is the sum of  $k$  finite moving averages, it can also be represented as a finite moving average  $\theta_i(L)u_{it}$ , where  $u_{it}$  is a white noise process, such that

$$\theta_i(L)u_{it} = A(L)_i^{adj} \epsilon_t \tag{6}$$

The coefficients of the polynomial  $\theta_i(L)$  are found by equating the autocovariances in the two representations. In this way we obtain a non-linear system of  $q^* + 1$  equations in  $q^* + 1$  unknowns  $\theta_{i1}, \theta_{i2}, \dots, \theta_{iq^*}$  and  $\sigma_{u_i}^2 = \text{var}(u_{it})$ . It is important to note that we consider the invertible solution of this system.

Considering (5) and (6), the univariate models implied by (3) are given by

$$\det [A(L)] y_{it} = \theta_i(L)u_{it} \quad i = 1, \dots, k.$$

Thus, we can conclude that

$$y_{it} \sim \text{ARMA}(p^*, q^*) \quad i = 1, \dots, k,$$

where it is well known that  $p^* \leq kp$  and  $q^* \leq (k - 1)p$ . We denote with  $A_y$  this set of univariate ARMA processes. Given two VAR processes,  $y$  and  $x$ , we can obtain their final forms,  $A_y$  and  $A_x$ . Since  $A_y, A_x \in \mathcal{F}_M$ , we can calculate

$$H(A_y, A_x) = \max \left\{ \max_{y_i \in A_y} \min_{x_j \in A_x} d_P(y_i, x_j), \max_{x_i \in A_x} \min_{y_j \in A_y} d_P(y_j, x_i) \right\}$$

and consider it as distance between the VAR processes  $y$  and  $x$ .

### An Illustrative Example

Here, we present an example of the proposed distance. We consider two VAR(1) processes,  $y$  and  $x$ , defined, respectively, by the following equations

$$\begin{bmatrix} 1 - 0.5L & 0.66L \\ 0.5L & 1 + 0.3L \end{bmatrix} \begin{bmatrix} y_{1t} \\ y_{2t} \end{bmatrix} = \begin{bmatrix} a_{1t} \\ a_{2t} \end{bmatrix} \tag{7}$$

with

$$E \left( \begin{bmatrix} a_{1t} \\ a_{2t} \end{bmatrix} \begin{bmatrix} a_{1t} & a_{2t} \end{bmatrix} \right) = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

and

$$\begin{bmatrix} 1 - 0.5L & 0 \\ 0.5L & 1 + 0.3L \end{bmatrix} \begin{bmatrix} x_{1t} \\ x_{2t} \end{bmatrix} = \begin{bmatrix} b_{1t} \\ b_{2t} \end{bmatrix} \tag{8}$$

with

$$E \left( \begin{bmatrix} b_{1t} \\ b_{2t} \end{bmatrix} \begin{bmatrix} b_{1t} & b_{2t} \end{bmatrix} \right) = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

where the bivariate white noise processes are such that

$$E \left( \begin{bmatrix} a_{1t} \\ a_{2t} \end{bmatrix} \begin{bmatrix} b_{1t-h} & b_{2t-h} \end{bmatrix} \right) = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} \text{ for } h = 0, \pm 1, \dots \quad (9)$$

The univariate models implied by (7) are given by:

$$\begin{aligned} (1 - 0.2L - 0.48L^2) y_{1t} &= \epsilon_{1t} + 0.204\epsilon_{1t-1} \text{ with } \sigma_{\epsilon_1}^2 = 1.469, \\ (1 - 0.2L - 0.48L^2) y_{2t} &= \epsilon_{2t} - 0.382\epsilon_{2t-1} \text{ with } \sigma_{\epsilon_2}^2 = 1.309. \end{aligned}$$

The univariate models implied by (8) are:

$$\begin{aligned} (1 - 0.5L) x_{1t} &= e_{1t} \text{ with } \sigma_{e_1}^2 = 1, \\ (1 - 0.2L - 0.15L^2) x_{2t} &= e_{2t} - 0.382e_{2t-1} \text{ with } \sigma_{e_2}^2 = 1.309. \end{aligned}$$

We have that  $d_p(y_1, x_1) = 0.42$ ,  $d_p(y_1, x_2) = 0.67$ ,  $d_p(y_2, x_1) = 0.81$  and  $d_p(y_2, x_2) = 0.36$ . Thus we obtain  $H(A_y, A_x) = 0.42$ .

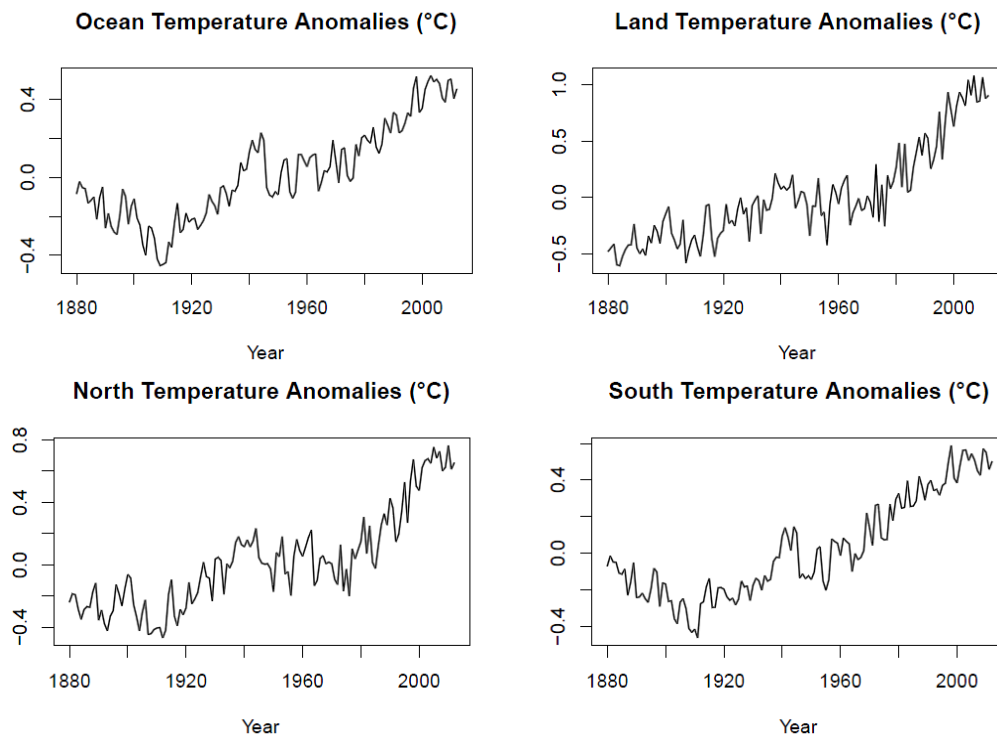
#### 4. An Empirical Application

To show the applicability of the derived results we consider the following time series:

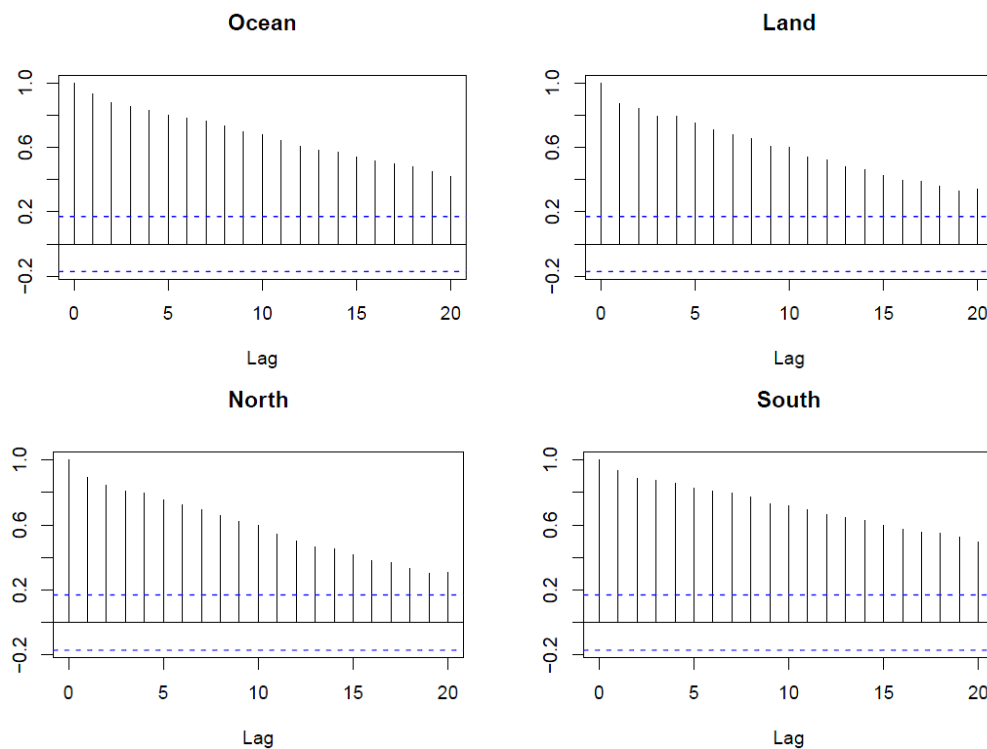
- the northern hemisphere annual temperature anomalies (N);
- the southern hemisphere annual temperature anomalies (S);
- the annual global land temperature anomalies (L);
- the annual global ocean temperature anomalies (O).

These global temperature anomaly data, with respect to the 20th century average, come from the Global Historical Climatology Network-Monthly (GHCN-M) data set and International Comprehensive Ocean-Atmosphere Data Set (ICOADS). The data span is from 1980 to 2012. These time series, shown in Figure 1, are available at [21]. Various subsets of this data set have been used in a number of studies (see, for example, [22,23]). The Auto Correlation Functions (ACFs) of the series and the Cross Correlation Functions (CCFs) are presented in Figures 2 and 3, respectively. The ACF plots indicate that the series are nonstationary, since the ACFs decay very slowly.

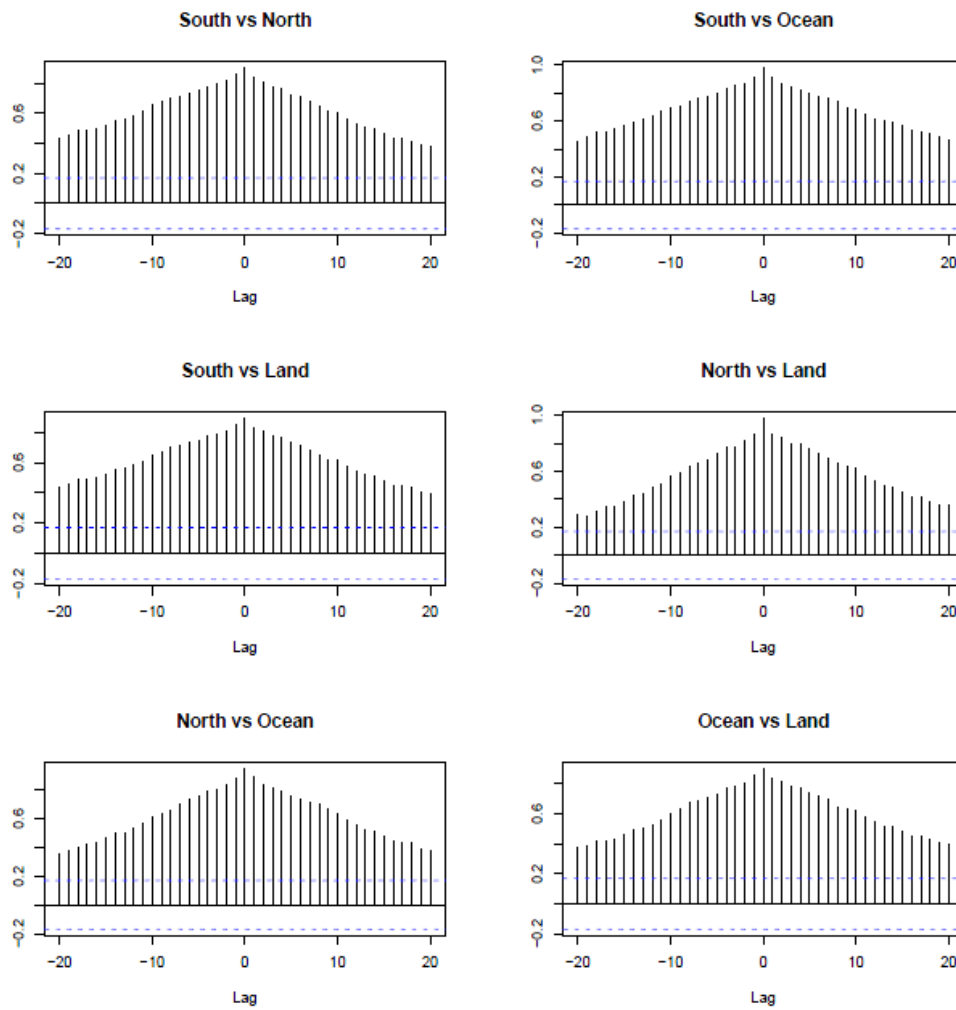
We will build portfolios of ARMA models for these four series, and hence, by estimating the models for each portfolio, an estimate of the Hausdorff distance  $H$  among our time series will be obtained. Preliminarily, we have used the augmented Dickey Fuller (ADF) test in order to establish whether the considered series are stationary or non-stationary. We concluded that all four temperature time series are integrated of order one. These results are consistent with those obtained by Stern and Kaufmann ([24]), Liu and Rodriguez ([25]) and Mills ([26]). Now, in order to build a model portfolio for the considered time series, we identify ARMA models for the first differences of our time series. Bayesian Information Criterion (BIC) has been used to find the appropriate orders  $p$  and  $q$  of the AR and MA polynomials respectively.



**Figure 1.** Global Ocean temperature anomalies, Land temperature anomalies, Northern and Southern Hemisphere temperature anomalies (clockwise from top left).



**Figure 2.** Sample Auto Correlation Functions (ACFs) of the series: Global Ocean temperature anomalies, Land temperature anomalies, Northern and Southern Hemisphere temperature anomalies (clockwise from top left). The lag axes are in terms of years.



**Figure 3.** Sample Cross Correlation Functions (CCFs). The lag axes are in terms of years.

The BIC values has been calculated for the  $ARMA(p, q)$  models with  $p, q \leq 3$ . Table 1 shows the orders  $(p, q)$  of ARMA models (for each time series) for which the value of the BIC, has been found to be minimum. The residual plots, reported in Figure 4, indicate no presence of serial correlation on the error term. The obtained portfolios of ARMA models for the four differenced series are:

- $P(O) = \{AR(2), MA(2)\}$ ,
- $P(S) = \{AR(2), MA(2), AR(3)\}$ ,
- $P(N) = \{MA(1), AR(3)\}$ ,
- $P(L) = \{MA(1)\}$ .

**Table 1.** Selection of the orders  $(p, q)$  of autoregressive moving average (ARIMA) $(p, 1, q)$  models.

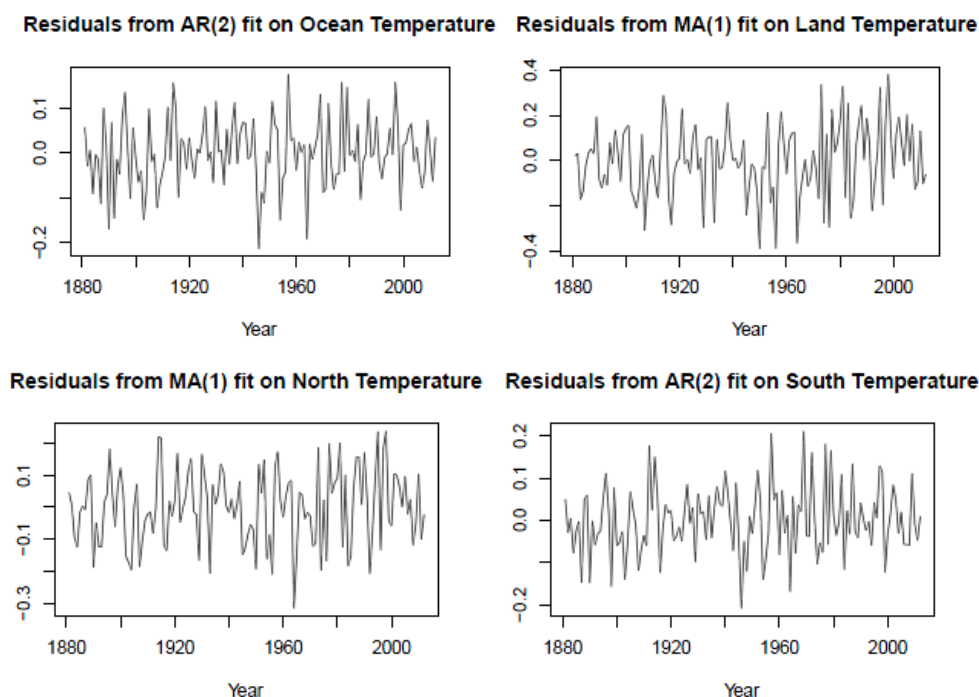
Series	O	L	N	S
Order	(2, 0)	(0, 1)	(0, 1)	(2, 0)

By estimating the models for each portfolio, an estimate of the Hausdorff distance  $H$  is obtained. Table 2 summarizes the results.



**Table 2.** Estimates of the Hausdorff distance  $H$ .

Differenced Series	O	S	N	L
O	0.000	0.199	0.908	1.128
S	0.199	0.000	0.911	1.153
N	0.908	0.911	0.000	0.344
L	1.128	1.153	0.344	0.000

**Figure 4.** The residuals of the chosen models.

According to our notion of distance, the dynamic structure of the ocean temperature (O) and the dynamic structure of the southern temperature (S) seem to be very similar. On the contrary, both the series, O and S, seem to be very distant by the land temperature (L) and by the northern hemisphere temperature (N). Instead, a smaller distance separates the series L and N. The big thermal inertia of oceans, together with the uneven distribution of the land and sea (the Southern Hemisphere contains 80.9 percent water and 19.1 percent land, while the Northern Hemisphere is 60.7 percent water and 39.3 percent land) represents a possible explanation of our results. This conclusion is attractive since it links a statistical result to a physical mechanism.

## 5. Conclusions

This manuscript focuses on measuring the distance between two finite sets of invertible ARMA processes using the Hausdorff distance. In order to make this distance operative, a dissimilarity measure between two single ARMA processes (the AR metric) is considered. Of course other dissimilarity measures, e.g., the Mahalanobis distance between autoregressive models or the cepstral distance, could be used. We have chosen the AR metric since it is simple to compute and it is robust with respect to the presence of anomalous behavior in the data. Further, it is implemented for both stationary and non-stationary time series.

In order to show the usefulness of the proposed distance, two specific applications have been described. The first application consists of measuring the distance between two single ARMA series, let's say  $x$  and  $y$ , by means of the Hausdorff distance between portfolios of ARMA models providing reasonably good fits of  $x$  and  $y$ . The second application uses the Hausdorff distance to compute distances between two VAR processes once these models are properly represented in terms of univariate ARMA processes. An empirical application is also presented.

**Acknowledgments:** We thank three anonymous reviewers and Diane Bergman for their constructive comments, which helped us to improve the manuscript.

**Conflicts of Interest:** The author declares no conflict of interest.

## References

- Maharaj, E.A. Clusters of time series. *J. Classif.* **2000**, *17*, 297–314.
- Liao, T. Clustering time series data—A survey. *Pattern Recognit.* **2005**, *38*, 1857–1874.
- Agrawal, R.; Faloutsos, C.; Swami, A. Efficient similarity search in sequence databases. In Proceedings of the 4th International Conference on Data Organization and Algorithms FODO '93, Chicago, IL, USA, 13–15 October 1993; Springer Verlag: New York, NY, USA, 1993; Number 730 in Lecture Notes in Computer Science; pp. 69–84.
- Caiado, J.; Crato, N.; Peña, D. A periodogram-based metric for time series classification. *Comput. Stat. Data Anal.* **2006**, *50*, 2668–2684.
- Otranto, E.; Triacca, U. Measures to evaluate the discrepancy between direct and indirect model-based seasonal adjustment. *J. Off. Stat.* **2002**, *18*, 551–530.
- Di Iorio, F.; Triacca, U. Testing for Granger non-causality using the autoregressive metric. *Econ. Model.* **2013**, *33*, 120–125.
- Jin, L. Comparing autocorrelation structures of multiple time series via the maximum distance between two groups of time series. *J. Stat. Comput. Simul.* **2015**, *85*, 3535–3548.
- Thomson, P.J.; de Souza, P. Speech recognition using LPC distance measures. In *Handbook of Statistics 5 Time Series in the Time Domain*; Hannan, E.J., Krishnaiah, P.R., Rao, M.M., Eds.; North Holland: Amsterdam, The Netherlands, 1985; pp. 389–412.
- Piccolo, D. A distance measure for classifying ARIMA models. *J. Time Ser. Anal.* **1990**, *11*, 153–164.
- Corduas, M.; Piccolo, D. Time series clustering and classification by the autoregressive metric. *Comput. Stat. Data Anal.* **2008**, *52*, 1860–1872.
- Maharaj, E.A. A significance test for classifying ARMA models. *J. Stat. Comput. Simul.* **1996**, *54*, 305–331.
- Martin, R.J. A metric for ARMA processes. *IEEE Trans. Signal Proces.* **2000**, *48*, 1164–1170.
- Brockwell, P.; Davies, R. *Time Series: Theory and Methods*, 2nd ed.; Springer-Verlag: New York, NY, USA, 1991.
- Poskitt, D.S.; Tremayne, A.R. Determining a portfolio of linear time series models. *Biometrika* **1987**, *74*, 125–137.
- Chatfield, C. *Time-Series Forecasting*; Chapman & Hall: London, UK, 2000.
- Box, G.E.P.; Jenkins, G.M. *Time Series Analysis: Forecasting and Control*; Holden-Day: San Francisco, CA, USA, 1970.
- Faranda, D.; Dubrulle, B.; Daviaud, F.; Pons, F.M.E. Probing turbulence intermittency via autoregressive moving-average models. *Phys. Rev. E* **2014**, *90*, 061001.
- Faranda, D.; Pons, F.M.E.; Giachino, E.; Vaianti, S.; Dubrulle, B. Early warnings indicators of financial crises via auto regressive moving average models. *Commun. Nonlinear Sci. Numer. Simul.* **2015**, *29*, 233–239.
- Sims, A. Macroeconomics and reality. *Econometrica* **1980**, *48*, 1–48.
- Zellner, A.; Palm, F. Time series analysis and simultaneous equation econometric models. *J. Econom.* **1974**, *2*, 17–54.
- Climate at a Glance. NOAA. Available online: <http://www.ncdc.noaa.gov/cag/time-series/global> (accessed on 10 June 2016).
- Trenberth, K.E. Has there been a hiatus? *Science* **2015**, *6249*, 691–692.
- Triacca, U.; Pasini, A.; Attanasio, A. Measuring persistence in time series of temperature anomalies. *Theor. Appl. Climatol.* **2014**, *118*, 491–495.
- Stern, D.I.; Kaufmann, R.K. Econometric analysis of global climate change. *Environ. Model. Softw.* **1999**, *14*, 597–605.

25. Liu, H.; Rodríguez, G. Human activities and global warming: A cointegration analysis. *Environ. Model. Softw.* **2005**, *20*, 761–773.
26. Mills, T.C. Breaks and unit roots in global and hemispheric temperatures: An updated analysis. *Clim. Chang.* **2013**, *118*, 745–755.



© 2016 by the author; licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC-BY) license (<http://creativecommons.org/licenses/by/4.0/>).