SPATIO-TEMPORAL COVARIANCE MODELING WITH SOME ARMA TEMPORAL MARGINS

Samuel Seth Demel

Juan Du

Follow this and additional works at: http://newprairiepress.org/agstatconference

Part of the Agriculture Commons, and the Applied Statistics Commons

This work is licensed under a Creative Commons Attribution-Noncommercial-No Derivative Works 4.0 License.

Recommended Citation

SPATIO-TEMPORAL COVARIANCE MODELING WITH SOME ARMA TEMPORAL MARGINS

Samuel Seth Demel
The Department of Statistics, Kansas State University, Manhattan KS 66506-0802, USA.
E-mail: ssdemel@k-state.edu

Juan Du
The Department of Statistics, Kansas State University, Manhattan KS 66506-0802, USA.
E-mail: dujuan@k-state.edu

A valid covariance structure is needed to model spatio-temporal data in various disciplines, such as environmental science, climatology and agriculture. In this work we propose a collection of spatio-temporal functions whose discrete temporal margins are some autoregressive and moving average (ARMA) models, obtain a necessary and sufficient condition for them to be covariance functions. An asymmetric version of this model is also provided to account for space-time irreversibility property in practice. Finally, a spatio-temporal model with AR(2) discrete margin is fitted to wind data from Ireland for estimation and prediction, which are compared with some general existing parametric models in terms of likelihood and mean squared prediction error.

Keywords: Autoregressive and moving average process; Fourier transform; Spatio-temporal covariance function; Stationary.
1 Introduction

In the ever growing world of data analysis one area, spatio-temporal statistics, has allowed us to model our world as a whole. The world we live in is four dimensional, three dimensions making up space and the fourth dimension time. Often times appealing to computational convenience, we study space and time separately not taking into account their possible interaction. However many environmentally processes are dependent on the space-time interaction, examples can be found in climatology, agriculture, and other environmental studies. See Cressie and Huang (1999), de Luna and Genton (2005), Wikle and Royle (2005), Xu et al. (2005), Le and Zidek (2006), Pearce et al. (2006), among others. Combining these two processes together creates the spatio-temporal process, which is a random field given by,

\[
\{ Z(\mathbf{s}; t), \mathbf{s} \in \mathbb{R}^d, t \in \mathcal{T} \},
\]

where \( \mathcal{T} = \mathbb{R} \) or \( \mathbb{Z} \). The spatial marginal of the spatio-temporal process is given by, \( \{ Z(\mathbf{s}; 0), \mathbf{s} \in \mathbb{R}^d \} \), and the temporal marginal of the process is given by, \( \{ Z(0; t), t \in \mathcal{T} \} \).

There have been many recent developments in space-time modeling, mainly dealing with the construction of the space-time covariance functions

\[
C(\mathbf{s}_1, \mathbf{s}_2; t_1, t_2) = \text{Cov}(Z(\mathbf{s}_1; t_1), Z(\mathbf{s}_2; t_2)), \quad (\mathbf{s}_1; t_1), \quad (\mathbf{s}_2; t_2) \in \mathbb{R}^d \times \mathcal{T},
\]

with \( \mathcal{T} = \mathbb{R} \), to mention a few, Haslett and Raftery (1989), Gneiting (2002), Ma (2003, 2005), Stein (2005a) and Gneiting, Genton, and Guttrop (2007). Within all of these references the authors build valid space-time covariance structures under the framework of continuous time and mention a need to contend with difficulties about the fact that time data are usually measured at discrete time points and normally viewed as a component of time series. However, the attempts in the spatio-temporal modeling with discrete time series margin, i.e. \( \mathcal{T} = \mathbb{Z} \), are limited. Most of these works were either based on spectral representation or stochastic equations (see e.g., Storvik et al. (2002), Stein (2005b)). In these cases the model choice is very difficult to justify in practice. The goal of this paper is to construct a covariance function to model the dependence structure of continuous space discrete time data using an intuitive approach which will take advantage of existing time series and spatial statistics tools to identify an interpretable model as well as an easier model to apply in practice. When trying to model the covariance, we tend to make some simplifications and often justifiable assumptions, such as stationarity, separability, and symmetry.

Stationarity of a space time process occurs when the process is both stationary in time and in space simultaneously, meaning the covariance \( C(\mathbf{s}_1, \mathbf{s}_2; t_1, t_2) \) depends only on the lag \( \mathbf{s}_1 - \mathbf{s}_2 = \mathbf{s} \) over space and lag \( t_1 - t_2 = t \) over time so that (1) can be denoted \( C(\mathbf{s}; t) \). All of the models presented herein will assume stationarity. In the history of this field the initial idea was to develop a valid covariance functions for both space and time independently of each other then multiply them together. This created what is known as the separable model. The authors mentioned above explain that the separable model, although simple and computationally attractive, does not really capture the interaction between space and time. This limits the modeling on the actual behavior of predictors. Some further discussion of the limitations of separable models in kriging can be found in literature e.g. (Stein, 2005) and references therein. Symmetry versus asymmetry is a more recent problem that tries to take into account underlying space time effects such as prevailing winds, ocean currents, or
thermal waves. Symmetric models often do not have the structure to include these effects, therefore we have to introduce the asymmetric model that will be discussed in Section 3. The proposed model is given in next section and its related proof is postponed to the appendix. In Section 4, the proposed model is fitted to Irish wind data and the results are presented by comparing with those using the model developed by Gneiting (2002), and Gneiting et al. (2007) in both symmetric and asymmetric cases.

2 A Spatio-temporal Model with Discrete temporal margins

In practice space time data are often collected at monitored discrete time lags and it is natural to assume the underlying process $Z(s,t)$ inhabits $\mathbb{R}^d \times \mathbb{Z}$. To model this type of data, we usually start by breaking the problem into two parts, the time series part and the spatial part. Typically we start at a fixed location and analyze the time series usually with an ARMA-type model then we analyze spatial field at a fixed time. In many research fields, experimenters use this approach in looking at time on one hand and space on the other. There is huge need for a model that enables the researcher to use the gathered information of the spatial marginal analysis and the temporal marginal analysis, then apply it to the spatio-temporal analysis. This is the underlying motivation to use our time series knowledge to aid us in building space-time covariance functions. In all cases we must show nonnegative definiteness, specially, under the assumption of stationary, it means for any $a_1, \ldots, a_k \in \mathbb{R}$ and any $k$ locations and time points, the following inequality holds

$$
\sum_{i=1}^{k} \sum_{j=1}^{k} a_i a_j C(s_i - s_j; t_i - t_j) \geq 0.
$$

The following theorem will provide a sufficient and necessary condition for the proposed function to satisfy (2), and therefore become a valid covariance spatio-temporal function. The proof is based on the well-known Bochner’s Theorem (Rudin (1962)) and is given in the Appendix.

**Theorem 1.** Assume that $\alpha_1, \alpha_2, \beta_1, \text{ and } \beta_2$ are constants with $0 < \alpha_1 < \alpha_2$ and $-1 < \beta_1 < \beta_2 < 1$. A necessary and sufficient condition for the function

$$
C(s,t) = \theta \exp(-\alpha_1 ||s||)\beta_1^{||t||} + (1 - \theta) \exp(-\alpha_2 ||s||)\beta_2^{||t||}, \quad s \in \mathbb{R}^d, \ t \in \mathbb{Z},
$$

(3) to be a stationary covariance function on $\mathbb{R}^d \times \mathbb{Z}$ is that the constant $\theta$ satisfies

$$
\left\{ 1 - \frac{\alpha_2^d (1 - \beta_1)(1 + \beta_2)}{\alpha_1^d (1 + \beta_1)(1 - \beta_2)} \right\}^{-1} \leq \theta \leq \left\{ 1 - \frac{\alpha_1 (1 - \beta_2)(1 + \beta_1)}{\alpha_2 (1 + \beta_2)(1 - \beta_1)} \right\}^{-1}.
$$

(4)
Remark: $\theta$ in (4) can be negative and it is not simply a convex combination of two separable models, but rather a more general model that enables the modeling of the spatio-temporal interaction.

Note here that when the spatial lag is set to zero the temporal margin of (3) is

$$C(0; t) = \theta \beta_1 |t| + (1 - \theta) \beta_2 |t|, \quad t \in \mathbb{Z},$$

which includes families of correlation functions of stationary AR(1), AR(2), and ARMA (2, 1) time series. Roughly speaking, $\alpha_i$'s can be looked at as the scaling parameter for the spatial covariance exponential model. The $\beta_i$'s are the corresponding coefficients when fitting an autoregressive time series and $\theta$ plays as a balancing parameter based on strength of both space and time interaction. If $\beta_1 = \beta_2$ then the temporal margin has an AR(1) structure and (3) becomes a separable model. In the case when the time lag is zero the spatial marginal is given by,

$$C(s; 0) = \theta \exp(-\alpha_1 \|s\|) + (1 - \theta) \exp(-\alpha_2 \|s\|), \quad s \in \mathbb{R}^d,$$

which is a mixture of two spatial exponential models and it can be extended to general Matérn model. In the special case where $\alpha_1 = \alpha_2$ the spatial margin is reduced to a single exponential model. When applying this model we can use time series techniques to fit independent time series for each location developing ARMA order and starting values for $\beta_1$, $\beta_2$, and $\theta$, so the final parameter estimation can be achieved by maximum likelihood estimation or weighted least square estimation. While for spatial aspect we can use spatial statistics procedures to find starting values for $\alpha_1$ and $\alpha_2$. The advantage here is that we can employ well-established time series techniques, such as ACF, PACF to determine the model patterns and orders, since the temporal margin is treated as time series. This is worth mentioning due to increasing demand on the statistical technique for the model selection and justification given all these different theoretical models developed in continuous space and time, e.g. Gneiting (2002), Ma (2003). Another advantage of the proposed model is that the simple structure of the model gives intuitive meaning for each component, which eases the cumbersome task of determining the appropriateness of the model. The ARMA process can be easily interpreted and techniques for estimating parameters are well understood. This leads us to use these techniques to find starting values to fit the over space-time covariance function presented. All these are not readily shared in general continuous case. For example, Gneiting (2002) states that his model relies on two functions which must be complete monotone or that the derivatives are monotone. A researcher has to both determine these functions, verify their conditions, and obtain that the resulting covariance function is positive definite. There is no clear defined method proposed by Gneiting for determining these functions, which can cause some difficulties in actual application. Although this is a challenge to the researcher, Gneiting’s model is very versatile in fitting spatio-temporal data. Our proposed model presented can serve as an attempt in seeking of more straightforward approach to study spatial-temporal data where at each location the temporal process can be modeled with a ARMA type covariance structure. Moreover, some of the models for discrete time can not be expanded to continuous process which covers the discrete one as constrained version on discrete domain, which is so-called embedding problem. For instance, the autocorrelation
function (acf) of AR(1) process is \( \beta^h \) with \(-1 < \beta < 0\) and \( h \in \mathbb{Z} \), but \( \beta^t, t \in \mathbb{R} \) is not a valid acf for continuous case. More examples will be studied in subsequent work by authors.

3 Asymmetric Covariance Functions

As mentioned earlier the need of asymmetric covariance model comes into play when taking into account the underlying space-time effects caused by natural occurring forces. To construct stationary functions that are not fully symmetric, which means \( C(s, t) = C(s, -t) = C(-s, t) = C(-s, -t) \) is not necessarily true. Gneiting, Genton, and Guttrop (2007) used the general idea of a Lagrangian reference frame (May and Julien (1998)). This Lagrangian framework can be thought of as modeling the center of an air, water, or thermal mass. The resulting spatio-temporal random field has stationary covariance, \( C(s; t) = E C_S(s - V t), s \in \mathbb{R}^d, t \in \mathbb{R}, \) where \( C_S \) is a valid spatial covariance function and \( V \in \mathbb{R}^d \) is a random velocity vector. This random velocity vector has various choices depending on the physicality of the data and can be justified when included in the model. When analyzing the Irish wind data, Gneiting, Genton, and Guttrop (2007), noted that “Ireland has a prevailing westerly wind and that the simplest case is when \( V = v \) is constant and represents the mean or prevailing wind.” Applying this knowledge they form a special case of the the Lagrangian which has the form,

\[
C_{LGR}(s, t) = \left( 1 - \frac{1}{2v} |s_{\text{long}} - vt| \right)_+, \tag{5}
\]

where the spatial separation vector \( s = (s_{\text{long}}, s_{\text{lat}})' \) has longitudinal lag (east-west) component \( s_{\text{long}} \) and latitudinal lag (north-south) component \( s_{\text{lat}} \), and \( v \in \mathbb{R} \) is a longitudinal velocity. Recall that Gneiting’s Model assumes a continuous frame for time while the proposed model (3) has a discrete time domain. Based on the fact that a correlation function on \( \mathbb{R} \) restricted to a discrete domain subset of itself, say \( \mathbb{Z} \), will still be a valid correlation function on that discrete domain. Therefore to apply the Lagrangian reference frame and enrich our model to accommodate asymmetry property, we only have to change the time domain from \( \mathbb{R} \) to \( \mathbb{Z} \), and make the convex combination of the built model and the asymmetric model (5) to create

\[
C(s, t) = (1 - \lambda) \left\{ \theta \exp(-\alpha_1 ||s||) \beta_1^{||t||} + (1 - \theta) \exp(-\alpha_2 ||s||) \beta_2^{||t||} \right\} \\
+ \lambda \left( 1 - \frac{1}{2v} |s_{\text{long}} - vt| \right)_+, \quad 0 < \lambda < 1 \tag{6}
\]

Now we are ready to apply and compare the models above using the Irish wind dataset.

4 Irish Wind Data Analysis

To compare the proposed model (3) with existing models presented by Gneiting (2002), Gneiting et al. (2007). We use the Irish wind data set first analyzed by Haslett and Raftery (1989). This is a spatio-temporal data set that measures average daily wind speeds taken
from twelve synoptic meteorological weather stations in Ireland from year 1961 to 1978. As in Haslett and Raftery (1989), Gneiting (2002), Gneiting et al. (2007) and Stein (2005a, 2005b), the following steps were taken to clean the data. A square root transformation is taken to stabilize the variance over both stations and time periods so that the margins of space and time to be approximately normally distributed. This allows us to use Gaussian theory for both the temporal process and spatial process. The location of Rosslare was removed for stationarity reason, as recommended by Haslett and Raftery (1989). The seasonal trend was fit and removed using annual harmonic regression. To extract the spatial trend the station specific means where removed. The final resulting dataset becomes time series of velocities for eleven meteorological stations, to which we will apply the proposed models (3) and (6).

For evaluating the ability to predict, the data is split into a training data set from years 1961 to 1970 and a test data set from years 1971 to 1978. These steps were also taken by Haslett and Raftery (1989), Gneiting (2002), Gneiting et al. (2007) and Stein (2005a, 2005b).

4.1 Symmetric Covariance Models

To start the analysis we first compare the symmetric cases of the proposed models. To achieve a good fit of the data Gneiting (2002) noted that the space-time correlation approaches zero after a lag of three days, we will utilize this in fitting our model. The symmetric spatio-temporal covariance functions that we will compare are Gneiting’s separable and non-separable models given by,

\[ C_{G.SEP}(s; t) = \{ (1 - \nu) \exp(-c||s||) + \nu \delta_{s=0} \} \cdot \{ (1 + a|t|^{2\alpha})^{-1} \}, \quad s \in \mathbb{R}^d, \ t \in \mathbb{R}, \]  

(7)

and

\[ C_{G.NSEP}(s; t) = \frac{1 - \nu}{1 + a|t|^{2\alpha}} \left\{ \exp\left( -\frac{c||s||}{1 + a|t|^{2\alpha}}^{\beta/2} \right) + \frac{\nu}{1 - \nu} \delta_{s=0} \right\}, \quad s \in \mathbb{R}^d, \ t \in \mathbb{R}, \]  

(8)

and our proposed model denoted by \( C_{SY M}(s; t) \) and given by,

\[
C_{SY M}(s; t) = (1 - \nu) \left\{ \theta \exp(-\alpha_1||s||)\beta_1^{[t]} + (1 - \theta) \exp(-\alpha_2||s||)\beta_2^{[t]} \right\} \\
+ \nu \left\{ \theta \beta_1^{[t]} + (1 - \theta) \beta_2^{[t]} \right\} \delta_{s=0}, \quad s \in \mathbb{R}^d, \ t \in \mathbb{Z},
\]

(9)

Here \( \nu \) is a spatial nugget effect coefficient that can also account for measurement error and \( \delta \) is an indicator function that equals one only when spatial lag is zero. Using the same techniques mentioned in Section 2 we first look at space and time independently of each other. From Figure 1 we can see that it is reasonable that the temporal margin for each station could possibly have an AR(1) or AR(2) structure. Recall that by allowing \( \beta_1 = \beta_2 \) in (9) we obtain an AR(1) temporal margin structure. For comparison we fit both the AR(1) and AR(2) cases. Compared to the autoregressive structure of proposed model (3) which accounts for the discrete time points, Gneiting (2002) uses a Cauchy type function to model the temporal margin without formal justification other than using empirical time correlation. Again note the Gneiting’s methods have to chose a function to fit the temporal
process. The tools from time series allow us to make the choice of AR(1) or AR(2). Both methods use weighted least squares to fit the temporal process in order to find starting values for temporal related parameters for the covariance functions. Notice in Figure 2 that both Gneiting’s Cauchy and the presented AR(2) fit the empirical temporal correlation quite well. To fit (9) spatially we used the $\alpha_1 = \alpha_2$ simplification to obtain a single exponential. Gneiting also uses a exponential structure for fitting the spatial correlation. Looking at Figure 2 the spatial correlation plot, notice that Gneiting’s and the proposed model fit are identical.

To obtain parameter estimates we use Cressie’s weighted least squares procedure taken from Cressie (1999) and Cressie and Huang (1999), which is the same as Gneiting’s (2002) and Gneiting’s et al. (2007) approach for fitting his models. After we find the parameter estimates for the models we compare the empirical correlations with the fitted correlations for each model. The empirical correlation is calculated by finding the cross-correlation over time between all stations. The parameter estimates for Gneiting’s models (7) and (8), as well as the proposed model (9) are given in Table 4.1.

![PACFs of Irish Wind Stations](image)

**Figure 1:** Partial autocorrelation function of selected locations in the Irish wind data.

From Figure 3 we can evaluate plots of fitted correlation versus empirical correlation. For spatial correlation which is given by $C(s; 0)$ (black). Using the second set of four plots in Figure 3 we can see that the models fit the empirical correlations quite well at a time lag of 0. At a time lag of one (red) we see that all models start to under predict the empirical correlation. The second set of plots confirm this conclusion. As the time lag increases to two and three, (green and blue), we see the models start to behave better, but still have a lack of fit when compared to the empirical. To evaluate the prediction power of the models
Each column corresponds to a given model. Here the Empirical Model use the training of years (1961-1970).

Figure 2: Spatial Correlation of stations at fixed time (left) and Marginal Temporal fits at fixed location (right). Plot based of years (1961-1970)

Figure 3: Left set of four: Empirical space-time correlations versus model fitted space-time correlation. Right set of four: Fitted - Empirical versus (Fitted + Empirical)/2. Plot based of years (1961-1970)

the root-mean-square errors (RMSEs) were calculated between the fitted based on three days of lag and the test data set. When comparing these results the smaller the better. Each column corresponds to a given model. Here the Empirical Model use the training
Table 1: Irish Wind Parameter Estimates

<table>
<thead>
<tr>
<th></th>
<th>G.SEP</th>
<th>G.NSEP</th>
<th>$C_{SYM}(s; t)$ AR(1)</th>
<th>$C_{SYM}(s; t)$ AR(2)</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\nu$</td>
<td>$0.04153$</td>
<td>$0.04153$</td>
<td>$0.04267$</td>
<td>$0.04312$</td>
</tr>
<tr>
<td>$c$</td>
<td>$0.00128$</td>
<td>$0.00128$</td>
<td>$0.00126$</td>
<td>$0.00127$</td>
</tr>
<tr>
<td>$a$</td>
<td>$0.97200$</td>
<td>$0.97200$</td>
<td>$1.4318$</td>
<td>$1.07324$</td>
</tr>
<tr>
<td>$\alpha$</td>
<td>$0.83400$</td>
<td>$0.83400$</td>
<td>$0.52002$</td>
<td>$-0.16108$</td>
</tr>
<tr>
<td>$\beta$</td>
<td>$0.68060$</td>
<td>$0.68060$</td>
<td>$0.48623$</td>
<td></td>
</tr>
</tbody>
</table>

1. G.SEP is Gneiting’s separable model (7)
2. G.NSEP is Gneiting’s non-Separable model (8)
3. $C_{SYM}(s; t)$ AR(1) is model (9) taking $\alpha_1 = \alpha_2$ and $\beta_1 = \beta_2$
4. $C_{SYM}(s; t)$ AR(2) is model (9) taking $\alpha_1 = \alpha_2$

empirical correlations to predict the future correlations. Roughly this is how well the data can predicting itself. Gneiting noted in his papers that the empirical model does have the smallest RMSE overall, but these predictions are only valid at the given locations. We can not use this model to predict at unknown locations. Recall that under model (9) in the AR(1) case, the model becomes separable and comparing to Gneiting’s separable model (7) we find the results comparable. The same is true when looking at model (9) in the AR(2) case and Gneiting’s non-separable model (8). Again the presented model is an alternative to Gneiting’s that is easier to apply in that the researcher does not have to chose unknown functions to build the covariance structure.

Table 2: Irish Wind: RMSEs for one-day ahead predictions based on three previous days

<table>
<thead>
<tr>
<th>Station</th>
<th>G.SEP</th>
<th>$C_{SYM}(s; t)$ AR(1)</th>
<th>G.NSEP</th>
<th>$C_{SYM}(s; t)$ AR(2)</th>
<th>Empirical</th>
</tr>
</thead>
<tbody>
<tr>
<td>Valentia</td>
<td>0.5005178</td>
<td>0.5004045</td>
<td>0.5009972</td>
<td>0.5007395</td>
<td>0.4995971</td>
</tr>
<tr>
<td>Belmullet</td>
<td>0.4949573</td>
<td>0.4955277</td>
<td>0.4953419</td>
<td>0.4960886</td>
<td>0.4939611</td>
</tr>
<tr>
<td>Claremorris</td>
<td>0.4909316</td>
<td>0.4917950</td>
<td>0.4915745</td>
<td>0.4917838</td>
<td>0.4858922</td>
</tr>
<tr>
<td>Shannon</td>
<td>0.4675282</td>
<td>0.4672383</td>
<td>0.4683362</td>
<td>0.4673986</td>
<td>0.4540456</td>
</tr>
<tr>
<td>RochesPoint</td>
<td>0.4832071</td>
<td>0.4836868</td>
<td>0.4787071</td>
<td>0.4842802</td>
<td>0.4647548</td>
</tr>
<tr>
<td>Birr</td>
<td>0.4771278</td>
<td>0.4768426</td>
<td>0.4759683</td>
<td>0.474376</td>
<td>0.4623907</td>
</tr>
<tr>
<td>Mullingar</td>
<td>0.4274168</td>
<td>0.4275001</td>
<td>0.4242714</td>
<td>0.4274717</td>
<td>0.4143287</td>
</tr>
<tr>
<td>MalinHead</td>
<td>0.4963515</td>
<td>0.4966294</td>
<td>0.4919131</td>
<td>0.4972658</td>
<td>0.4791251</td>
</tr>
<tr>
<td>Kilkenny</td>
<td>0.4388743</td>
<td>0.4393654</td>
<td>0.4358435</td>
<td>0.4396867</td>
<td>0.4141247</td>
</tr>
<tr>
<td>Clones</td>
<td>0.4863430</td>
<td>0.4868047</td>
<td>0.4839092</td>
<td>0.4870985</td>
<td>0.4657351</td>
</tr>
<tr>
<td>Dublin</td>
<td>0.4498648</td>
<td>0.4497305</td>
<td>0.4452762</td>
<td>0.4490259</td>
<td>0.4270464</td>
</tr>
</tbody>
</table>

We also calculated the likelihood using the three days of lag and then averaging over each set of three days, which can be seen in Table 3. In all cases the log likelihoods are comparable. So for the Irish wind data the proposed models in the AR(1) and AR(2) cases are very competitive with Gneiting’s models. Gneiting’s models are very versatile, but the
Table 3: Irish Wind: Log Likelihoods

<table>
<thead>
<tr>
<th>MODELS</th>
<th>G.SEP</th>
<th>$C_{SYM}(s;t)AR(1)$</th>
<th>G.NSEP</th>
<th>$C_{SYM}(s;t)AR(2)$</th>
<th>Empirical</th>
</tr>
</thead>
</table>

our models provide nearly identical results with a procedure that is easier to apply and more intuitive when the temporal margin is close to an AR(1), AR(2), or ARMA(2, 1) structure. The ease of the application does not come from fitting. Both Gneiting’s and the presented models are fitted the same way. The ease comes from using the exploratory analysis to gain parameter estimates for the models. The presented model also utilizes basic time series techniques to justify the use of an AR(1), AR(2), or ARMA(2, 1) structure. The model proposed can use classical ARMA interpretation of the temporal related parameters, while this is not a straightforward manner in some theoretical continuous space-time models.

4.2 Asymmetric Covariance Models

Now we will take into account asymmetric cases and compare the general stationary (asymmetric) Gneiting’s model,

$$C_{G.STAT}(s; t) = \frac{(1 - \lambda)(1 - \nu)}{1 + a|t|^{2\alpha}} \left\{ \exp \left( -\frac{c||s||}{(1 + a|t|^{2\alpha})^{1/2}} \right) + \frac{\nu}{1 - \nu} \delta_{s=0} \right\} + \lambda \left( 1 - \frac{1}{2v}|s_{long} - vt| \right)$$  \hspace{1cm} (10)

versus the asymmetric case of the proposed model denoted by $C_{ASYM}(s; t)$, and given by,

$$C_{ASYM}(s; t) = (1 - \lambda) \left[ (1 - \nu) \left\{ \theta \exp(-\alpha_1||s||)\beta_1^{|t|} + (1 - \theta) \exp(-\alpha_2||s||)\beta_2^{|t|} \right\} \right. \left. + \nu \left\{ \theta \beta_1^{|t|} + (1 - \theta) \beta_2^{|t|} \right\} \delta_{s=0} \right] + \lambda \left( 1 - \frac{1}{2v}|s_{long} - vt| \right)$$  \hspace{1cm} (11)

In the asymmetric case we see the addition of two new parameters $\lambda$, the weight parameter associated with the Lagrangian reference, and $v$ is the speed in which the westerly wind systems move, measured in kilometers per day. Here we use the existing parameter estimates that we found in the symmetric case and use weighted least squares and obtain fitted estimates $\hat{\lambda} = 234$ kilometers per day and $\hat{v} = 0.0573$ as did in Gneiting et al. (2007).

From Figure 4 we can evaluate plots of fitted correlation versus empirical correlation. Spatial at a time lag of zero (black) all the models fit rather well for both Gneiting’s and our model (11). At a time lag of one (red) we see that Gneiting’s model (10) and the model (11) fits are about the same. At a lag of two Gneiting’s model is slightly better than model (11). At three days the model both asymmetrically models seem to have a similar fit. Again we
conclude that the model (11) is competitive to Gneiting’s methods. To compare prediction capability the RMSEs were calculated between the fitted based on three days of lag and the test data set. When comparing these results the smaller RMSE indicates better prediction power.

Notice that again in Table 4.2 we see similar result as the symmetric case, in that our model (11) is comparable to Gneiting’s model (10). If we compare Tables 4.1 and 4.2, notice that Gneiting’s model (10) has the lowest RMSEs, but our model (11) is very close. When applying the asymmetric case with the proposed model (11) note that the temporal margin is no longer autoregressive in its nature. This could be why our model (11) does not improve over the model (9). We also calculated the likelihood using the three days of lag and then averaging over each set of three days, which can be seen in Table 5, in which case our model (11) again is very comparable to Gneiting’s. After analyzing both the symmetric and asymmetric cases we can see that our models (9) and (11) are highly comparable to Gneiting’s models (7), (8), and (10) on all accounts. Gneiting’s models are very versatile, but the models we proposed here provide nearly identical predictions and likelihoods with a procedure that is easier to apply and more intuitive when the temporal margin is of an ARMA type time series. The advantage here is that our model takes into account the fact that the data are collected in time and are often considered as a realization of a time series. There are well developed techniques that can be borrowed to aid the choice of the models rather than trying to find an unknown function that fits the empirical time correlation.

Figure 4: Upper: Empirical versus fitted correlation Gneiting’s stationary model (10) (left) and $C_{ASYM}(s,t)$ AR(2) (right). Lower: Fitted - Empirical versus (Fitted + Empirical)/2. Plot based of years (1961-1970)
Table 4: Irish Wind: RMSEs for one-day ahead predictions based on three previous days

<table>
<thead>
<tr>
<th>Station</th>
<th>G.STAT</th>
<th>$C_{ASYM}(s; t)$ AR(2)</th>
<th>Empirical</th>
</tr>
</thead>
<tbody>
<tr>
<td>Valentia</td>
<td>0.4989789</td>
<td>0.5041726</td>
<td>0.4995971</td>
</tr>
<tr>
<td>Belmullet</td>
<td>0.4945723</td>
<td>0.4982614</td>
<td>0.4939611</td>
</tr>
<tr>
<td>Claremorris</td>
<td>0.4896983</td>
<td>0.4938673</td>
<td>0.4858922</td>
</tr>
<tr>
<td>Shannon</td>
<td>0.4657336</td>
<td>0.4694637</td>
<td>0.4540456</td>
</tr>
<tr>
<td>RochesPoint</td>
<td>0.4737622</td>
<td>0.4857265</td>
<td>0.4647548</td>
</tr>
<tr>
<td>Birr</td>
<td>0.4724635</td>
<td>0.4793497</td>
<td>0.4623907</td>
</tr>
<tr>
<td>Mullingar</td>
<td>0.4193231</td>
<td>0.4291690</td>
<td>0.4143287</td>
</tr>
<tr>
<td>MalinHead</td>
<td>0.4878435</td>
<td>0.4984323</td>
<td>0.4791251</td>
</tr>
<tr>
<td>Kilkenny</td>
<td>0.4294301</td>
<td>0.4401993</td>
<td>0.4141247</td>
</tr>
<tr>
<td>Clones</td>
<td>0.4788818</td>
<td>0.4879050</td>
<td>0.4657351</td>
</tr>
<tr>
<td>Dublin</td>
<td>0.4400079</td>
<td>0.4494564</td>
<td>0.4270464</td>
</tr>
</tbody>
</table>

1 G.STAT is Gneiting's General Stationary (Asymmetric) Model (10)
2 $C_{ASYM}(s; t)$ AR(2) ASYM is model (11) taking $\alpha_1 = \alpha_2$

Table 5: Irish Wind: Log Likelihoods Asymmetric Models

<table>
<thead>
<tr>
<th>MODELS</th>
<th>G.STAT</th>
<th>$C_{ASYM}(s; t)$ AR(2)</th>
<th>Empirical</th>
</tr>
</thead>
<tbody>
<tr>
<td>Log likelihood</td>
<td>18.267072</td>
<td>18.484894</td>
<td>19.664738</td>
</tr>
</tbody>
</table>

without much justification. In this sense we believe our proposed models is in the direction of trying to utilize convenient techniques in timer series and spatial statistics for practical continues space discrete time data modeling, which covers even larger class of models which might not have have continuous counterparts. In addition, the relative ease of interpretation of parameter estimation stems from being able to use those properties studied for time series. It would also be relevant to extend to the proposed models (9) and (11) to include more types of time series margins and the constrains for each. But the spatio-temporal models presented in this paper have successfully started to provided an alternative approach that can be used in an intuitive manner as well as being easier to apply to space-time data with some discrete temporal margins.
5 Appendix

5.1 Proof of Theorem 1:

We are going to use Bochner’s Theorem to show (3). With the Fourier transforms of $\beta_k^{[|]}$ and $e^{-\alpha_k\|s\|}$, for $k = 1, 2$

$$\sum_{t=0}^{\infty} \beta_k^{[|]} e^{it\lambda} = \frac{1 - \beta_k^2}{1 + \beta_k^2 - 2\beta_k \cos \lambda}, \quad \lambda \in [-\pi, \pi],$$

and

$$\int_{\mathbb{R}^d} e^{-is'\omega} \exp (-\alpha_k \|s\|) \, ds = c_0 \alpha_k \left(\|\omega\|^2 + \alpha_k^2\right)^{-\frac{d+1}{2}}, \quad \omega \in \mathbb{R}^d,$$

where $c_0 = 2^d \pi^\frac{d-1}{2} \Gamma\left(\frac{d+1}{2}\right)$, we obtain the Fourier transform of (3),

$$f(\omega; \lambda) = C \theta \alpha_1 \left(\|\omega\|^2 + \alpha_1^2\right)^{-\frac{d+1}{2}} \frac{1 - \beta_1^2}{1 + \beta_1^2 - 2\beta_1 \cos \lambda} + (1 - \theta) \alpha_2 \left(\|\omega\|^2 + \alpha_2^2\right)^{-\frac{d+1}{2}} \frac{1 - \beta_2^2}{1 + \beta_2^2 - 2\beta_2 \cos \lambda},$$

$$= C \left\{ \frac{\omega_1}{\omega_1^2 + \alpha_1^2} \left(\|\omega\|^2 + \alpha_1^2\right)^{-\frac{d+1}{2}} (1 - \beta_1^2) (1 + \beta_1^2 - 2\beta_1 \cos \lambda) + (1 - \theta) \frac{\omega_2}{\omega_2^2 + \alpha_2^2} \left(\|\omega\|^2 + \alpha_2^2\right)^{-\frac{d+1}{2}} (1 - \beta_2^2) (1 + \beta_2^2 - 2\beta_2 \cos \lambda) \right\} \prod_{k=1}^{2} \left(1 + \beta_k^2 - 2\beta_k \cos \lambda\right)^{-1}$$

$$= C \left\{ g_1(\omega) - 2g_2(\omega) \right\} \prod_{k=1}^{2} \left(1 + \beta_k^2 - 2\beta_k \cos \lambda\right)^{-1}, \quad \omega \in \mathbb{R}^d, \lambda \in [-\pi, \pi]$$

where $C$ is some positive constant, and

$$g_1(\omega) = \frac{\omega_1}{\omega_1^2 + \alpha_1^2} \left(\|\omega\|^2 + \alpha_1^2\right)^{-\frac{d+1}{2}} (1 - \beta_1^2) (1 + \beta_1^2 - 2\beta_1 \cos \lambda) + (1 - \theta) \frac{\omega_2}{\omega_2^2 + \alpha_2^2} \left(\|\omega\|^2 + \alpha_2^2\right)^{-\frac{d+1}{2}} (1 - \beta_2^2) (1 + \beta_2^2),$$

and

$$g_2(\omega) = \frac{\omega_1}{\omega_1^2 + \alpha_1^2} \left(\|\omega\|^2 + \alpha_1^2\right)^{-\frac{d+1}{2}} (1 - \beta_1^2) (1 + \beta_1^2 - 2\beta_1 \cos \lambda) + (1 - \theta) \frac{\omega_2}{\omega_2^2 + \alpha_2^2} \left(\|\omega\|^2 + \alpha_2^2\right)^{-\frac{d+1}{2}} (1 - \beta_2^2) (1 + \beta_2^2),$$

According to Bochner’s theorem, (3) is a stationary covariance function on $\mathbb{R}^d \times \mathbb{Z}$ if and only if its Fourier transform $f(\omega; \lambda)$ is nonnegative, or equivalently,

$$g_1(\omega) - 2g_2(\omega) \cos \lambda \geq 0, \quad \omega \in \mathbb{R}^d, \lambda \in [-\pi, \pi] \quad (12)$$

Moreover, inequality (12) holds for all $\omega \in \mathbb{R}^d$ and $\lambda \in [-\pi, \pi]$ if and only if

$$g_1(\omega) - 2g_2(\omega) \geq 0, \quad g_1(\omega) + 2g_2(\omega) \geq 0, \quad \omega \in \mathbb{R}^d. \quad (13)$$

This is true because, on one hand we obtain (13) from (12) by simply taking $\lambda = 0$ and $\pi$ in (12), and on the other hand, inequalities (13) imply

$$g_1(\omega) \geq 2|g_2(\omega)| \geq 2g_2(\omega) \cos \lambda, \quad \omega \in \mathbb{R}^d, \lambda \in [-\pi, \pi].$$
Now slightly simplifying the complete expressions of inequalities (13)

\[ \theta \alpha_1 \left( \| \omega \|^2 + \alpha_1^2 \right)^{-\frac{d+1}{2}} (1 - \beta_1^2) (1 - \beta_2^2) + (1 - \theta) \alpha_2 \left( \| \omega \|^2 + \alpha_2^2 \right)^{-\frac{d+1}{2}} (1 - \beta_1^2) (1 - \beta_2^2) \geq 0, \]

and

\[ \theta \alpha_1 \left( \| \omega \|^2 + \alpha_1^2 \right)^{-\frac{d+1}{2}} (1 - \beta_1^2) (1 + \beta_2^2) + (1 - \theta) \alpha_2 \left( \| \omega \|^2 + \alpha_2^2 \right)^{-\frac{d+1}{2}} (1 + \beta_1^2) (1 + \beta_2^2) \geq 0, \]

for all \( \omega \in \mathbb{R}^d \), yields

\[ \theta \alpha_1 \left( \| \omega \|^2 + \alpha_1^2 \right)^{-\frac{d+1}{2}} (1 + \beta_1) (1 - \beta_2) + (1 - \theta) \alpha_2 \left( \| \omega \|^2 + \alpha_2^2 \right)^{-\frac{d+1}{2}} (1 + \beta_2) (1 - \beta_1) \geq 0, \]

(14)

and

\[ \theta \alpha_1 \left( \| \omega \|^2 + \alpha_1^2 \right)^{-\frac{d+1}{2}} (1 - \beta_1) (1 + \beta_2) + (1 - \theta) \alpha_2 \left( \| \omega \|^2 + \alpha_2^2 \right)^{-\frac{d+1}{2}} (1 - \beta_2) (1 + \beta_1) \geq 0, \]

(15)

Hence, it suffices to show that inequalities (4) are necessary and sufficient for (14) and (15) to hold.

First suppose that (14) holds for any \( \omega \in \mathbb{R}^d \). Multiplying \( (\| \omega \|^2 + \alpha_1^2)^{\frac{d+1}{2}} \) on both sides of (14) and then letting \( \omega \) tend to infinity, we have

\[ \theta \alpha_1 (1 + \beta_1) (1 - \beta_2) + (1 - \theta) \alpha_2 (1 + \beta_2) (1 - \beta_1) \geq 0, \]

(16)

or

\[ \theta + (1 - \theta) \frac{\alpha_2 (1 - \beta_1) (1 + \beta_2)}{\alpha_1 (1 + \beta_1) (1 - \beta_2)} \geq 0, \]

which implies

\[ \theta \leq \left\{ 1 - \frac{\alpha_1 (1 - \beta_2) (1 + \beta_1)}{\alpha_2 (1 + \beta_2) (1 - \beta_1)} \right\}^{-1}, \]

with \( 0 < \alpha_1 < \alpha_2 \) and \( (1 + \beta_1) (1 - \beta_2) \leq (1 - \beta_1) (1 + \beta_2) \). On the other hand, substituting \( \omega = 0 \) in (15) yields

\[ \theta \alpha_1^{-d} (1 - \beta_1) (1 + \beta_2) + (1 - \theta) \alpha_2^{-d} (1 - \beta_2) (1 + \beta_1) \geq 0, \]

(17)

and thus

\[ \theta \geq \left\{ 1 - \frac{\alpha_2^{-d} (1 - \beta_1) (1 + \beta_2)}{\alpha_1^{-d} (1 + \beta_1) (1 - \beta_2)} \right\}^{-1}. \]

This concludes that inequalities (4) are necessary for (14) and (15) to hold.

Next we are going to show opposite, namely under inequalities (4), (14), and (15) hold for any \( \omega \in \mathbb{R}^d \). While this is obviously true if \( 0 \leq \theta \leq 1 \), it remains to consider the cases

\[ \left\{ 1 - \frac{\alpha_2^d (1 - \beta_1) (1 + \beta_2)}{\alpha_1^d (1 + \beta_1) (1 - \beta_2)} \right\}^{-1} \leq \theta \leq 0, \quad \text{and} \quad 1 - \theta \leq \left\{ 1 - \frac{\alpha_1 (1 - \beta_2) (1 + \beta_1)}{\alpha_2 (1 + \beta_2) (1 - \beta_1)} \right\}^{-1}. \]
Case \( \left\{ 1 - \frac{\alpha_2^d (1 - \beta_1) (1 + \beta_2)}{\alpha_1^d (1 + \beta_1) (1 - \beta_2)} \right\}^{-1} \leq \theta \leq 0: \) In this case, \( 1 - \theta \) is positive and (17) is valid. Note that \( 0 < \alpha_1 < \alpha_2 \) implies

\[
\left( \frac{\alpha_2^2}{\|\omega\|^2 + \alpha_2^2} \right)^{\frac{d+1}{2}} \geq \left( \frac{\alpha_1^2}{\|\omega\|^2 + \alpha_1^2} \right)^{\frac{d+1}{2}},
\]

then inequality (15) follows from

\[
\theta \alpha_1 (\|\omega\|^2 + \alpha_1^2)^{-\frac{d+1}{2}} (1 - \beta_1) (1 - \beta_2) + (1 - \theta) \alpha_2 (\|\omega\|^2 + \alpha_2^2)^{-\frac{d+1}{2}} (1 - \beta_2) (1 + \beta_1)
\]

\[
= \theta \alpha_1^{-d} \left( \frac{\alpha_1^2}{\|\omega\|^2 + \alpha_1^2} \right)^{\frac{d+1}{2}} (1 - \beta_1) (1 + \beta_2) (1 - \theta) \alpha_2^{-d} \left( \frac{\alpha_2^2}{\|\omega\|^2 + \alpha_2^2} \right)^{\frac{d+1}{2}} (1 - \beta_2) (1 + \beta_1)
\]

\[
\geq \left\{ \theta \alpha_1^{-d} (1 - \beta_1) (1 + \beta_2) + (1 - \theta) \alpha_2^{-d} (1 - \beta_2) (1 + \beta_1) \right\} \left( \frac{\alpha_1^2}{\|\omega\|^2 + \alpha_1^2} \right)^{\frac{d+1}{2}}
\]

\[
\geq 0.
\]

On the other hand \((1 + \beta_1) (1 - \beta_2) < (1 - \beta_1) (1 + \beta_2),\) inequality (14) follows from

\[
\theta \alpha_1 (\|\omega\|^2 + \alpha_1^2)^{-\frac{d+1}{2}} (1 - \beta_1) (1 - \beta_2) + (1 - \theta) \alpha_2 (\|\omega\|^2 + \alpha_2^2)^{-\frac{d+1}{2}} (1 - \beta_2) (1 + \beta_1)
\]

\[
\geq \left\{ \theta \alpha_1 (\|\omega\|^2 + \alpha_1^2)^{-\frac{d+1}{2}} + (1 - \theta) \alpha_2 (\|\omega\|^2 + \alpha_2^2)^{-\frac{d+1}{2}} \right\} (1 - \beta_1) (1 - \beta_2)
\]

\[
= \left\{ \theta \alpha_1^{-d} \left( \frac{\alpha_1^2}{\|\omega\|^2 + \alpha_1^2} \right)^{\frac{d+1}{2}} + (1 - \theta) \alpha_2^{-d} \left( \frac{\alpha_2^2}{\|\omega\|^2 + \alpha_2^2} \right)^{\frac{d+1}{2}} \right\} (1 - \beta_2) (1 + \beta_1)
\]

\[
\geq \left\{ \theta \alpha_1^{-d} + (1 - \theta) \alpha_2^{-d} \right\} \left( \frac{\alpha_1^2}{\|\omega\|^2 + \alpha_1^2} \right)^{\frac{d+1}{2}} (1 - \beta_2) (1 + \beta_1)
\]

\[
\geq 0,
\]

where the last inequality holds because

\[
\theta \geq \left\{ 1 - \frac{\alpha_2^d (1 - \beta_1) (1 + \beta_2)}{\alpha_1^d (1 + \beta_1) (1 - \beta_2)} \right\}^{-1} \geq \left\{ 1 - \frac{\alpha_2^d}{\alpha_1^d} \right\}^{-1}.
\]

Case \( 1 \leq \theta \leq \left\{ 1 - \frac{\alpha_1 (1 - \beta_2) (1 + \beta_1)}{\alpha_2 (1 + \beta_2) (1 - \beta_1)} \right\}^{-1} : \) In this case, \( \theta \) is positive and (16) is valid. We obtain inequality (14) from

\[
\theta \alpha_1 (\|\omega\|^2 + \alpha_1^2)^{-\frac{d+1}{2}} (1 + \beta_1) (1 - \beta_2) + (1 - \theta) \alpha_2 (\|\omega\|^2 + \alpha_2^2)^{-\frac{d+1}{2}} (1 + \beta_2) (1 - \beta_1)
\]

\[
\geq \left\{ \theta \alpha_1 (1 + \beta_1) (1 - \beta_2) + (1 - \theta) \alpha_2 (1 + \beta_2) (1 - \beta_1) \right\} (\|\omega\|^2 + \alpha_1^2)^{-\frac{d+1}{2}}
\]

\[
\geq 0,
\]
and inequality (15) from

\[ \theta \alpha_1 \left( \| \omega \|^2 + \alpha_2^2 \right)^{-\frac{d+1}{2}} (1 - \beta_1)(1 + \beta_2) + (1 - \theta) \alpha_2 \left( \| \omega \|^2 + \alpha_2^2 \right)^{-\frac{d+1}{2}} (1 - \beta_2)(1 + \beta_1) \]
\[ \geq \{ \theta \alpha_1 + (1 - \theta) \alpha_2 \} \left( \| \omega \|^2 + \alpha_2^2 \right)^{-\frac{d+1}{2}} (1 - \beta_2)(1 + \beta_1) \]
\[ \geq 0, \]

note that the last inequality follows from

\[ \theta \geq \left\{ 1 - \frac{\alpha_1 (1 - \beta_2)(1 + \beta_1)}{\alpha_2 (1 + \beta_2)(1 - \beta_1)} \right\}^{-1} \geq \left\{ 1 - \frac{\alpha_2}{\alpha_1} \right\}^{-1}. \]

This completes the proof of the theorem. ■

References


