# Predicting Meteorological Values on a Spatial Grid

**Felipe Hernández**
felipeh@andrew.cmu.edu

**Ben Humberston**
bhumbers@cs.cmu.edu

## Abstract

TODO

## 1 Introduction

### 1.1 Motivation

Weather forecasts are broadly used; from personal activity planning (What clothes should I wear today? Will it be a good time to plan outdoor events?); to large-scale economical decisionmaking (What activities should be prioritized on a crop next week? What precautions should the air-traffic controllers enforce on a given day?); to emergency preparation and response (What alternate routes should become available due to snow?, Where should the emergency vehicles be sent during a flood event?). The availability and accuracy of forecasts thus have a profound impact on human activities at many levels, both in measurable and unmeasurable aspects. However, predicting weather is a difficult research problem. Most often, physically-based models with global scale are used to forecast future conditions. In this project, we will instead take a machine learning approach focused on a local scale and attempt to predict atmospheric variables at a specific geographic location. The forecasts will be based on prior atmospheric states in the neighborhood of the selected location. In particular, we will attempt to predict variables such as pressure, precipitation, and temperature based on the previous values of these variables based on regular historical snapshots produced from satellite observations by NASA.

### 1.2 Related Work

TODO

## 2 Method

We will be using a vecor autoregressive (VAR) model in order to model and predict the evolution of our meteorological system over time. We will adopt the formulation given in [1]. The model describes the system values at time $t$ given some number of prior system states $p$, which is the lag order of the model:

$$X_t = \boldsymbol{c} + \boldsymbol{\Pi}_1 X_{t-1} + ... + \boldsymbol{\Pi}_p X_{t-p} + \epsilon_t, t = 1, ..., T \tag{1}$$

TODO: Describe each term in the above

Note that each variable is offset to zero-mean over all time steps and normalized to have unit variance.

TODO: Describe how to solve the above

TODO: Equation noting how $\boldsymbol{\Pi}$ is concatenation of each individual $\boldsymbol{\Pi}_i$

### 2.1 Lasso

TODO: Describe how basic lasso works and why we would use it (sparsity, feature selection). Be sure to note original [2] source for LASSO.

TODO: Equation giving our problem formulated for lasso regularization

### 2.2 Fused Lasso

The fused lasso is a modified version of lasso and was introduced by [3]. It is useful for time series data because it encodes the relationship between successive parameters

TODO: Equation for fused lasso in our context

In particular, this may be useful if we examine the VAR model behavior over longer time scales than hour-by-hour weather. There is, of course, seasonal variations in meteorological variables, and it is possible that the system behavior is also dynamic based on the time of year.

TODO: Discuss different optimization framework needed for fused lasso (see source work)

### 2.3 Evaluation

The success of our model depends on its ability to correctly forecast (predict) the new system state at time $t$ based on the system state at times $t-p$ to $t-1$. Since there is some inherent noise $\epsilon$ in the variables of the system, the primary goal is to minimize the residual of predicted system state $\hat{X}(t)$ relative to the actual historical observation $X(t)$ to within some reasonable range of this noise term. The size of our data set allows us to easily train on a subset of the full data and test on a completely independent set (eg: train on data from the year 2004, test on data from the year 2007).

## 3 Results

Currently, we have implemented both a simple linear regression model as well as a standard VAR($p$) model which predicts new system values given prior system states. We illustrate an example of the output from the VAR model in Figure 1.

So far, both models have shown poor performance, where the errors in the model output are very large relative to the expected noise in each observed variable.

TODO: Image showing heat-map evolution of temperature, pressure values, with 1 row per timestep and 1 column per variable. Indicate how grid of values is remapped to a single section of a row.

There are several possible reasons for this which we intend to investigate. First, the VAR model assumes a *stationary* process, and our current preprocessing of the data before fitting the model may not guarantee this assumption. Additionally, there may be exogenous variables which affect the predictions which we are not yet including in the model (eg: wind speeds, elevation, etc.). It may be necessary to include these terms by augmenting our model to be as follows:

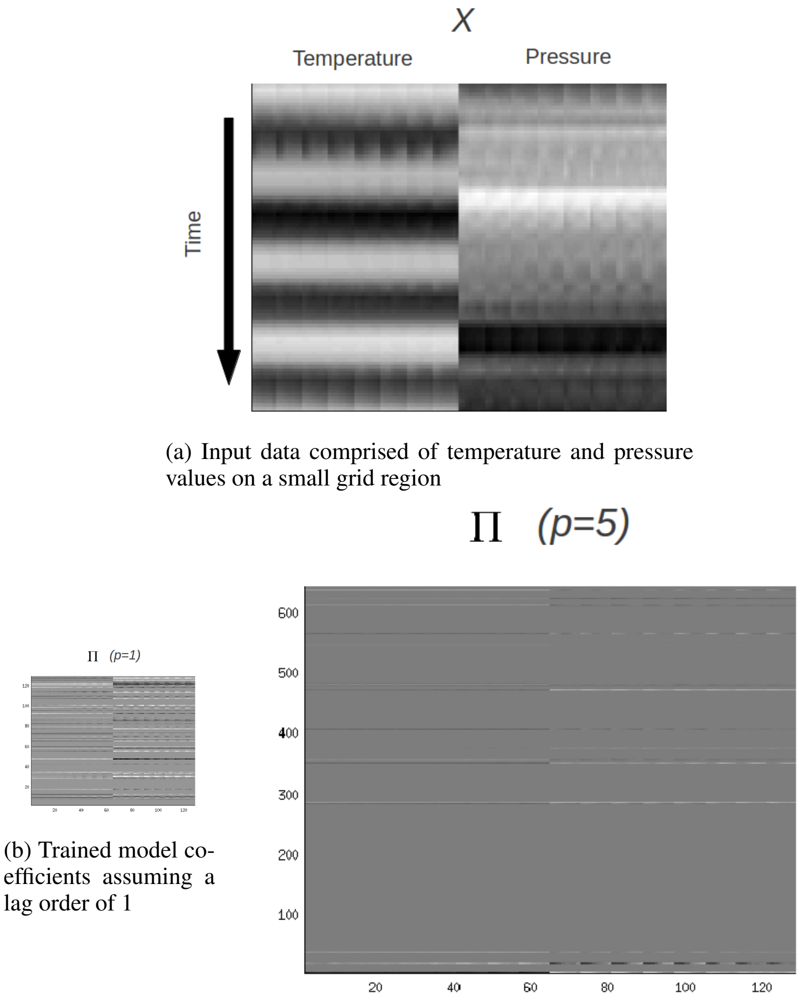
$$TODO \tag{2}$$

## References

[1] Vector autoregressive models for multivariate time series. In *Modeling Financial Time Series with S-PLUS*, pages 385–429. Springer New York, 2006.

[2] Robert Tibshirani. Regression shrinkage and selection via the lasso. *Journal of the Royal Statistical Society. Series B (Methodological)*, pages 267–288, 1996.

[3] Robert Tibshirani, Michael Saunders, Saharon Rosset, Ji Zhu, and Keith Knight. Sparsity and smoothness via the fused lasso. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 67(1):91–108, 2005.



(a) Input data comprised of temperature and pressure values on a small grid region

(b) Trained model coefficients assuming a lag order of 1

(c) Trained model coefficients assuming a lag order of 5

Figure 1: An illustrative example of the input data $X$ and trained output model $\mathbf{\Pi}$ obtained using different lag parameters. This particular data is comprised of hourly samples over a 4-day window. Each row corresponds to all the observed values at a particular timestep, where the 2D gridded values for each variable type (in this case, surface temperature and pressure) are reshaped in into a row vector. Note that the large, monochrome gray regions in each $\mathbf{\Pi}$ indicates significant sparsity of the model, which was not expected before using lasso regularization.