DAR model Overleaf

**Intro Abstract**

In this article we propose an extension of the first order Double Autoregressive Model (DAR(1)) from Ling 2004, which we refer to as the DARMA(1,1)-X model, and apply it to perform 1-step forecasting of hourly Nord Pool day-ahead prices. By explicitly modelling first order autocorrelation in the standardised residuals, we allow ARCH-type processes, augmented with exogenous variables, to enter the conditional mean. This enables us to capture the stylised fact that high price volatility is associated with high mean prices, which stems from the convexity of the production. As this article focuses only on forecasting and not on inference, we can allow relatively loose bounds on the parameter space. For both Gaussian and heavy-tailed Student-t errors, we derive and simulate the necessary and sufficient conditions for the existence of the fractional moment (strict stationarity), absolute first order moment (sufficient to forecast), and second order moment (week stationarity). Following the literature and best practices (Weron 2021), we benchmark the performance against at least one state of the art model falling within the two main paradigms in timeseries forecasting: Econometric and ML/Hybrid.

**Structure:**

The Danish Energy Market

Stylised Facts and Price Characteristics

Time Series Modelling

* The DARMA(p,1)-X Model
  + Extending the DAR(p) Model
  + Exogenous Variables
  + Stationarity
  + Estimation (SLSQP)
* Benchmark Models
  + Statistical/Econometric: ARIMA
  + Deep Learning: Nothing
  + Hybrid: LEAR

Data Processing

* Endogenous Variables
  + Deseasonalisation
* Exogenous Variables
  + Standardisation
  + Dimensionality Reduction Using PCA

Forecasting:

Forecasting Scheme:

* + Training, Validation and Test Data
  + Recalibration: Rolling, expanding, ect

Forecast Evaluation:

* + Accuracy Measures
  + Statistical Testing

Empirical Results:

* Data
* Validation Data:
  + DARMA-X: Standardised Errors
  + LEAR Hyperparameters
* Forecasting Results

Discussion:

* Higher lag order
* Non-linear dimensionality reduction
* Long-term seasonality

PCA

A close-up of a text

Description automatically generated

**Maximum Likelihood Estimation**

**Gaussian**

Ignoring all constants, i.e. , the Gaussian log-likelihood function for the DARMA model is defined as,

**Student-t estimation**

We can drop the constant term since it is independt of the parameters and data,

**The DAR Model**

**Function**

with

**Number of parameters**

**The DARMA-X Model**

**Function**

It is argues by (XXX) that due to

**DARMA(p,1)-X**

Using that ,

Where we have that is defined as,

Then,

For this reduces to,

Defining and , the DARMA(p,1)-X model can then be written as,

With,

**Forecast**

**Number of parameters**

DARMA(1,1)-X

Where denotes the cardinality , i.e. the length of .

If and ,

Max lag for is: 2

**Extending the DAR Model**

The property of “volatility-induced stationarity” allows us to estimate the DAR() model in levels and thus retain more information about the timeseries compared to ARIMA-type models which rely on first-differing to handle unitroots.

From the stylised facts, we have seen that modelling conditional mean and volatility together is an important step towards capturing the dynamics of electricity prices. One specification is the Double AR() – also called the DAR() – model, it combines an autoregressive process with conditional heteroskedasticity, and has been studied extensively in, amongst other, (XXX), (XXX) and (XXX). The DAR() model is defined as,

, where , , and . From derivations in Appendix XXX and Jiang et al. (2019), the strict stationarity and moment conditions for the DAR() with Gaussian innovations, , are tabualted in Table XXX

|  |  |  |
| --- | --- | --- |
| **Existence of Moments for** | | |
|  | DAR(1) | DARMA(1,1) |
| **Condition** | **Constraint** | |
| Strict Stationarity: |  | Working on it |
|  |  |  |
| Weak Stationarity: |  |  |

PDF of scaled student-t distribution:

Heavy-tailed:

Gaussian:

Since is stricter than , then garuantees . Therefore, is a sufficient (but not necessary) condition to perform 1-step forecasts.

As we see in Fig XXX the autoregressive coefficeont is allowed to be , which contrasts starkly with the classicial AR() model that requires for weak stationarity. ARIMA-type models deal with this problem by performing estimation on the differenced series. However, this approch removes valuable information about the stochastic process, and is therefore suboptimal. As argued in (Ling 2004), the reason why the DAR() is stationary in the precense of unit-roots (or even explosive roots), is that – for – the conditional variance can control the log-likelihood function, score function and information matrix in such a way that they are bounded.

Ling 2014 shows that this holds for many different, and heavy-tailed, distributions of .

**DAR(P) Moments**

With we consider the drift function ,

Some useful rules:

* and are independent for all
* Multiplicativeness:
* Subadditivity 1:
* Subadditivity 2:

For ,

This gives,

First part:

Second part:

**DAR Moments**

**First Order Moment**

Some usefull rules:

* and are independent for all
* Multiplicativeness:
* Subadditivity 1:
* Subadditivity 2:

Let be the drift function that bounds :

Since we have restricted and , then,

Using Subadditivity 2 and Multiplicativeness we get that . Therefore,

With this can be written as,

, where the constant is defined as . To mimic a first order autoregressive process, this can be reformulated as,

, with such that by definition. Therefore, it is clear that if the drift criterion holds and the process is bounded such that .

**Try 2:**

SRE of the DAR:

We try again,

Defining and

**DAR(p) SRE**

The DAR(),

Then,

Let . Then has the following SRE representation,

Where is the -order identity matrix and is a matrix of zero. This can be written as,

, where is a matrix and is a vector.

**DARMA(,1) SRE**

The DARMA() model

Combined,

Using the same arguments as above,

Let . Then has the following SRE representation,

Where is the -order identity matrix and is a matrix of zero. This can be written as,

**Stationarity conditions**

Lemma A.3: With a real matrix, and any matrix norm , it holds that:

1. if and only if

Let denote spectral radius of matrix . By Lemma A.3 in Section 1, we have that,

Let denote the top Lyapunov exponent. Then using the above, we can rewrite as,

Or,

Therefore, it implies that

for

From Theorem 4 in (Tweedie 1988) we have that for any positive integer then if the following conditions holds:

1. The eigenvalues of have modulus less than unity, or equivalently

, where “” denotes the Kronecker product.

**DARMA Moments**

**Second Order Moment**

If , the Euclidian norm corresponds to the drift criterion that bounds the aboslute value on vector form,

Then the drift criterion that bounds the second order moments is,

With we consider the drift function ,

Examining first first term:

All the cross-terms cancel out since ,

We have that therefore and are independent,

We will now derive ,

Using the distributive property,

Then the taking the expectation yields,

Since and for from the assumption of independence. If then sum converges to a infinite geometric series,

Inserting back into term 1 yields,

Examining second term:

All the cross-terms cancel out since ,

Using the arguments as before reveals that the solution to is independent of the lag since and , yielding the same solution for any . Therefore,

Inserting back into term 2 yields,

Putting the two terms together we get,

Defining constants and as and ,

Using that this can be rewritten as,

, where . Further, to minic an AR(1) process we have,

With and by defintion. Assumming , we can reduce to the following,

Therefore, if the drift criterion holds and the second order moment is bounded.

**First Order Moment (Manhattan Norm)**

With we consider the drift function ,

Some usefull rules:

* and are independent for all
* Multiplicativeness:
* Subadditivity 1:
* Subadditivity 2:

Examnining the first term:

Where we have ussed Subadditivity. Using that is strictly positive and in the information set, then

To determine we first need to examine the distribution of . Since is assumed to be i.i.d., then .

By assumption, which holds for over all , thus . Likewise, for all . Therefore, using standard variance-rules it follows that,

Since for all then,

Therefore the distribution of becomes,

Let be any random Gaussian variable such that . Then by definition,

Therefore, it follows that,

Inserting back into the first term yields,

Using Subadditivity 2 and Multiplicativeness we get that . Therefore,

Examnining the second term:

Using the exact same arguments as before, increasing the lag by one, we get,

Putting the two terms together yields,

Defining constants and as and ,

Using that this can be rewritten as,

, where . Further, to minic an AR(1) process this can be represented as,

With and by defintion. Assuming , we can reduce to the following,

Therefore, if the drift criterion holds and the first order moment is bounded.

**DARMA-X Moments**

.

For we have that and for we have that

Assuming .

**Second Order Moment**

With we consider the drift function

Examining first term:

All the cross-terms cancel out since ,

Examining second term:

Same answer, except variables are lagged by one more,

Putting the two terms together we get,

Defining constants , and as , and ,

Using that can be rewritten as,

, where . We can further rewrite the expression to mimic an AR(1) process,

With and by construction. Assumming , we can reduce to the following,

Therefore, if , and is stationary, the drift criterion holds and the second order moment is bounded.