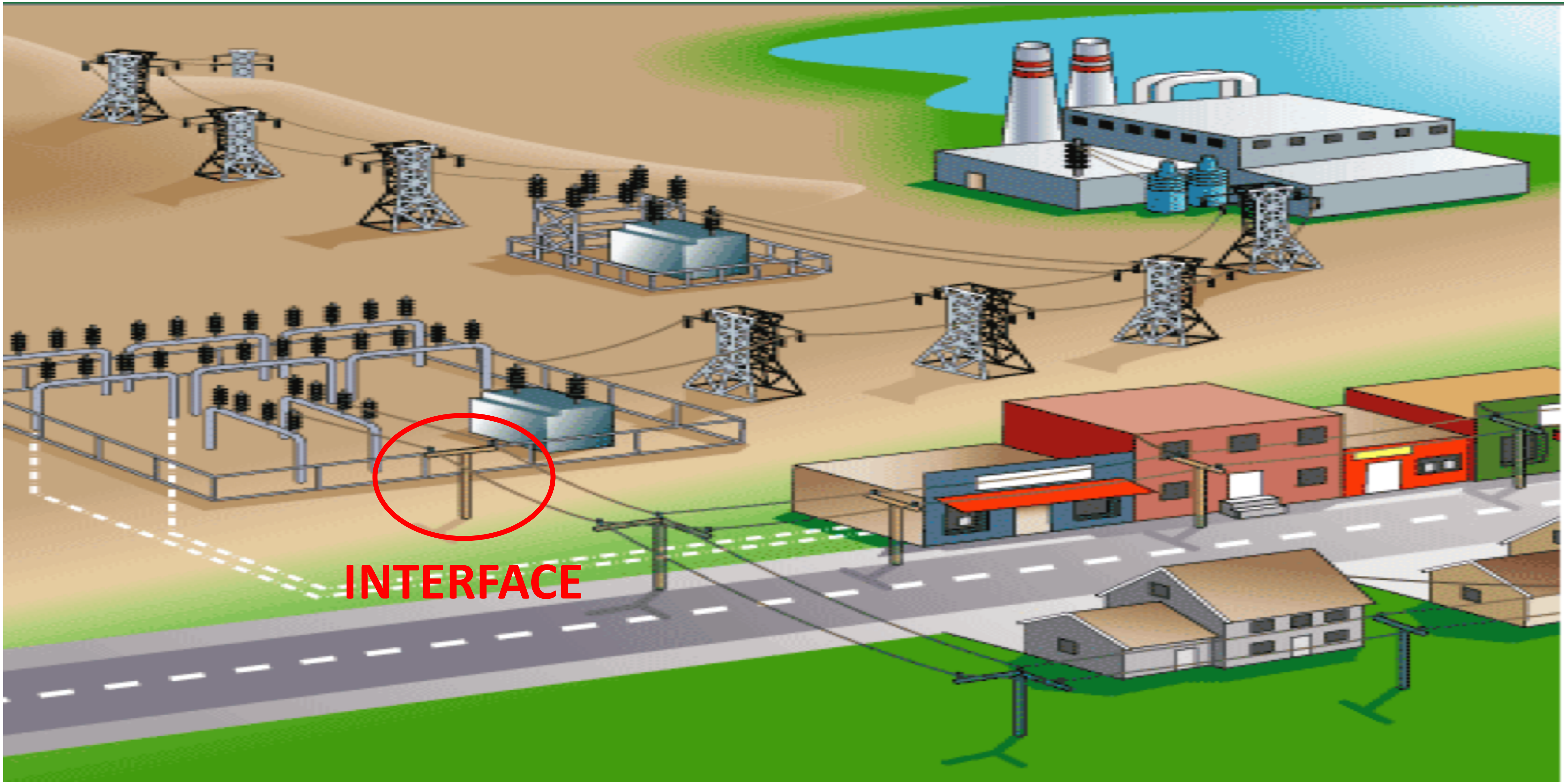
Modeling of Power Grid Loads based on High Rate Grid Data

Andrew Miller
Alexander Dimitrov
Washington State University

Outline

- **Introduction:**
  - What is Load Modeling?
  - What is Phasor Measurement Unit (PMU) data?
- **Methods:**
  - Nondimensionalization of PMU data
  - Discrete Time model
  - Least Squares
  - Linear model Assumptions check
  - Model ranking via AICc
- **Results:**
  - Linear model assumptions analysis
  - Ranking via AICc
  - Test data applied to load model
- **Discussion**

Introduction-Power System

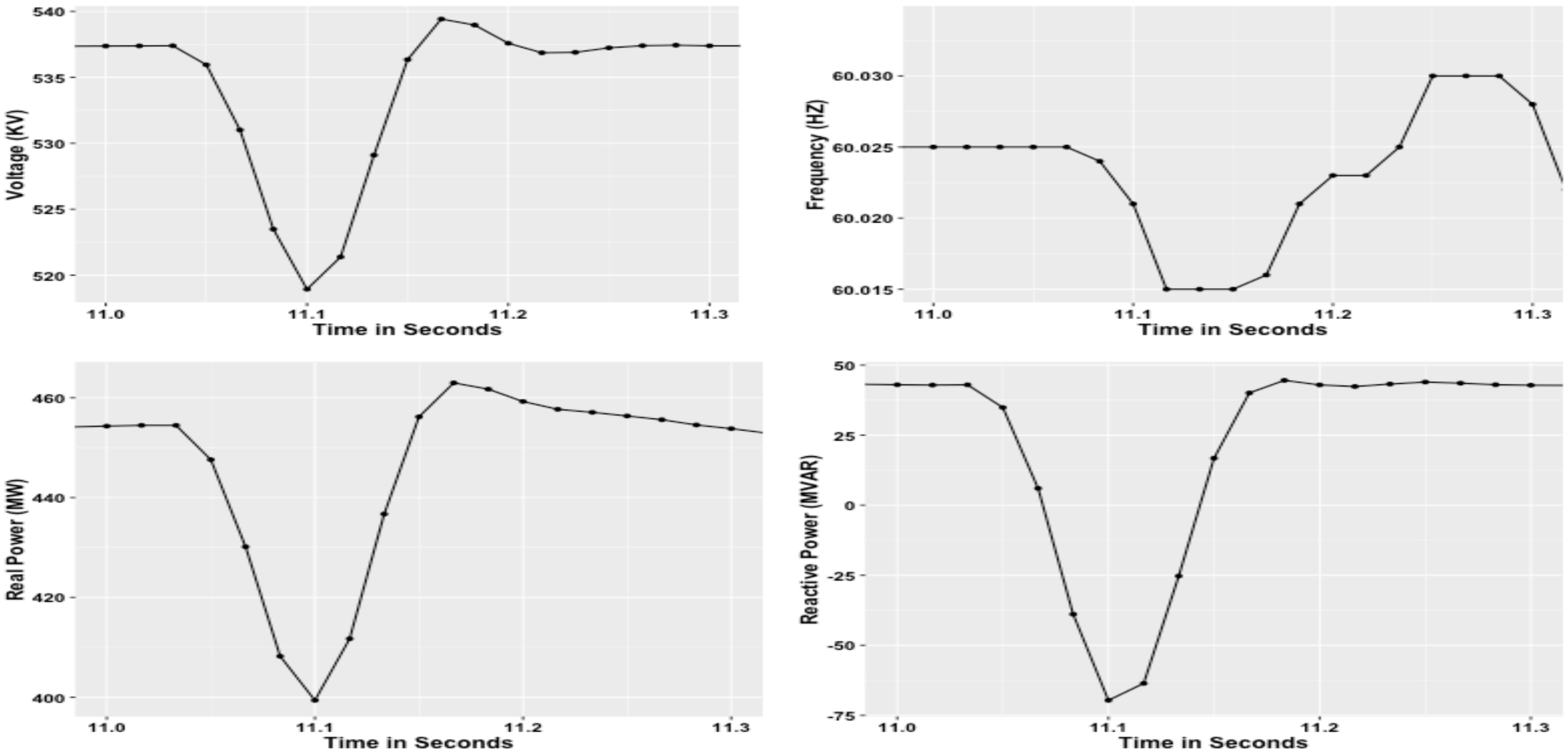


Introduction-Motivation

- **Motivation**
  - Load refers to the total power demand at a particular substation
  - Dynamic responses of loads critical for system stability and controls

- **Objective**
  - Predict power system response to perturbations of voltage/frequency through mathematical models

Introduction-Motivation (Single Line to Ground)



Introduction-Phasor Measurement Unit (PMU)

- Sensors that are at different locations along power system
- Measures phasors of voltages and currents
- GPS synchronized
- 60 samples/second
- ~10 years old

Introduction- Measurement-based Load Modeling

- Mathematical relationship between **response**:
  - Real power (MW)
  - Reactive power (MVAR)

  and **predictors**:
  - Voltage (KV)
  - Frequency (Hz)

- PMU data to create load model (estimate parameters)

- Useful for online monitoring

Methods-Nondimensionalization

- Eliminates physical units (MW, KV, MVAR)
- Intrinsic properties (natural units) remain
- Intrinsic system's dynamics observed
- Given by

$$Z_{n,k} = \frac{Y_{n,k} - \widehat{\mu_k}}{\widehat{\sigma_k}},$$

where

$Z_{n,k}$ K variables and n observations per variable;
$Y_{n,k}$ observed responses;
$\widehat{\mu_k}$ mean for each variable k;
$\widehat{\sigma_k}$ standard deviation for each variable k.

Methods-Discrete time load model

- Voltage and frequency dependent model:

$$\begin{cases} \widehat{P_t} = a_1 P_{t-1} + \cdots + a_n P_{t-n} + b_0 V_t + \cdots + b_n V_{t-n} + c_0 F_t + \cdots + c_n F_{t-n} \\ \widehat{Q_t} = d_1 Q_{t-1} + \cdots + d_n Q_{t-n} + b_0 V_t + \cdots + b_n V_{t-n} + c_0 F_t + \cdots + c_n F_{t-n} \end{cases},$$

where

$P_t,\ Q_t$	Measured real and reactive power at time t, respectively;
$V_t,\ F_t$	Measured voltage and frequency at time t, respectively;
$t-n$	n number of time lags;
$a_1, \ldots, a_n$	real power parameters;
$b_0, \ldots, b_n$	voltage parameters;
$c_0, \ldots, c_n$	frequency parameters.

- Voltage dependent load model will also be considered

Methods-Least Squares

- Assumes that data is well represented by

$$Y_i = \sum_{j=1}^{n} \theta_j X_{i,j} + \varepsilon_i,$$

where

n	number of parameters;
Y_i	i^{th} response from data;
$X_{i,j}$	input data matrix;
θ_j	vector containing j parameters;
ε_i	i^{th} random error (i.i.d with $\varepsilon \sim N(0, \sigma^2)$)

Methods-Least Squares

- Error function to be minimized is defined in terms of variance

$$E(\theta) = \frac{1}{N}\sum_{i=1}^{N}(Y_i - \sum_{j=1}^{n}\theta_j X_{i,j})^2,$$

where

θ	vector of parameters (e.g. $\theta = [a_1, \dots, a_n, b_0, \dots, b_n]$);
N	number of observations.

- $E(\theta)$ can be minimized by using optimization techniques from calculus

Methods- Assumptions Check $\boldsymbol{\varepsilon \sim N(0, \sigma^2)}$

- **Assumptions:**

  1) Normality

  2) Independence

  3) Homoscedasticity (constant variance)

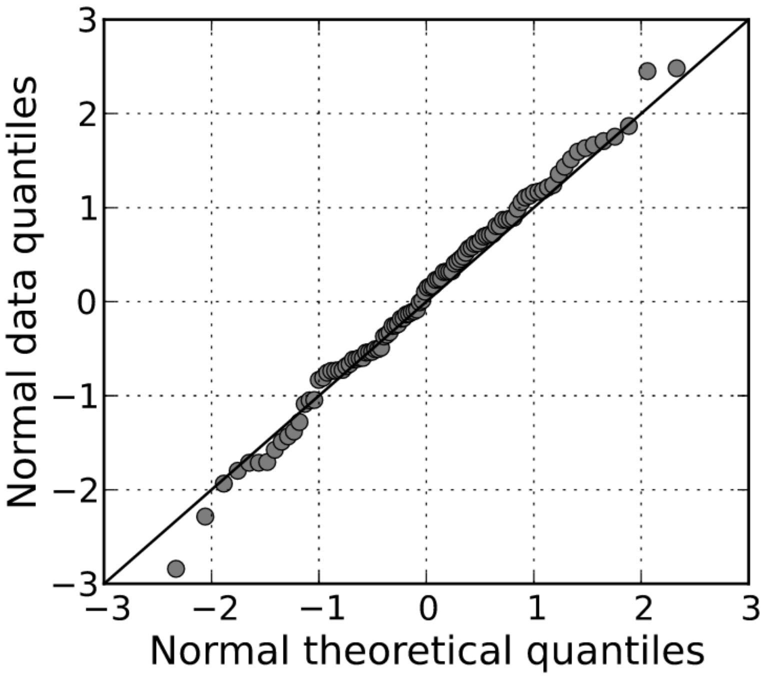
Methods-Assumption of Normality

- Anderson-Darling test as a **quantitative** assessment
  - Compares $f(x)$ to $\widehat{f(x)}$
  - $\begin{cases} H_0\text{: Errors are normally distributed} \\ H_a: Errors\ are\ not\ normally\ distributed, \end{cases}$

  where we reject if P-value < 0.05

Methods-Assumption of Normality

- QQ-Plots as a **qualitative** assessment
  - Theoretical vs. Actual quantiles
  - If data normally distributed, then



$$y_n = \hat{\sigma} z_n + \hat{\mu},$$

where

y_n	observed response from data;
$\hat{\sigma}$	sample standard deviation;
$\hat{\mu}$	sample mean;
z_n	lower quantiles corresponding to y_n.

Methods-Assumption of Independence

- Durbin-Watson test as a **quantitative** assessment

- $\begin{cases} H_0\text{: Errors are serially uncorrelated} \\ H_a: Errors\ are\ serially\ correlated \quad (\varepsilon_t = \rho\varepsilon_{t-1} + \alpha_t), \end{cases}$

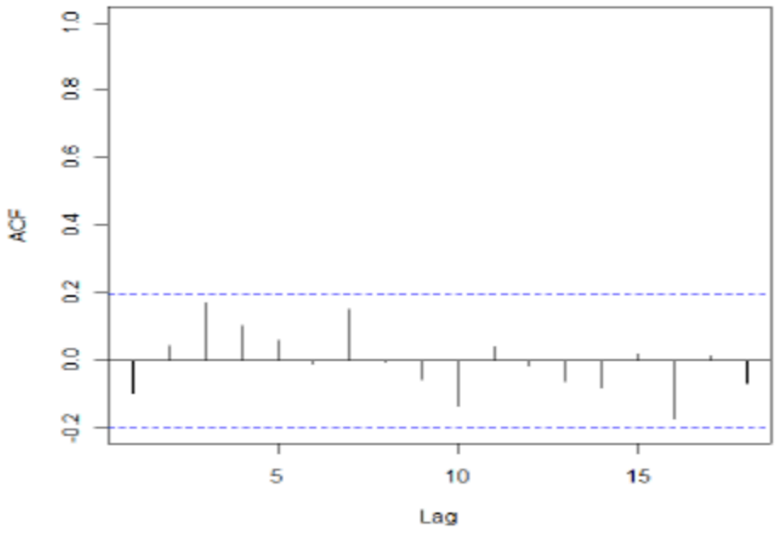
where

α	reject if bootstrapped P-value < $\alpha = 0.05$;
ρ	autocorrelation parameter and $\lvert\rho\rvert < 1$;
α_t	is i.i.d. $NID(0, \sigma_\alpha^2)$ random variable.

Methods-Assumption of Independence

- Autocorrelation Function (ACF)-Plots as a **qualitative** assessment
  - Relationship between ε_t and ε_{t-h}, for lag $h = 1,..,n$
  - r_t are autocorrelation coefficients given by

$$r_t = \frac{\sum_{i=1}^{t-h}(\varepsilon_i - \bar{\varepsilon})(\varepsilon_{i+h} - \bar{\varepsilon})}{\sum_{i=1}^{t}(\varepsilon_i - \bar{\varepsilon})^2}$$



where

$\bar{\varepsilon}$ error mean;

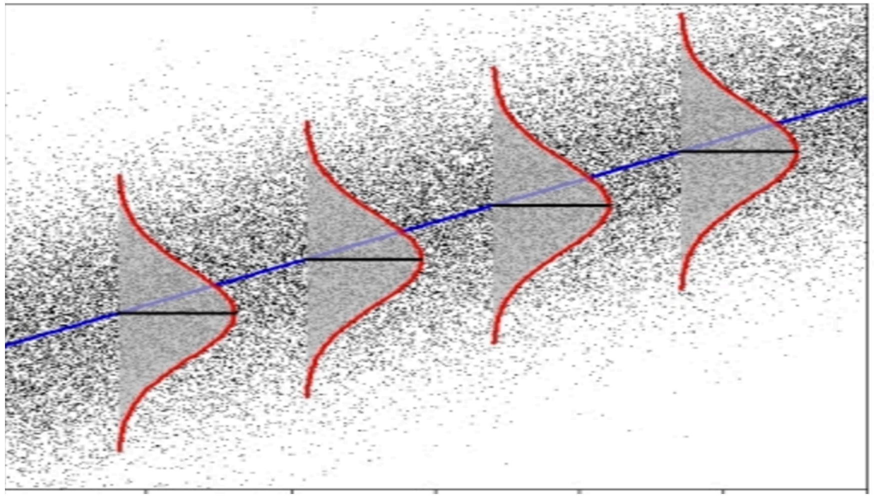
$t - h$ time series of errors having h lags.

Methods-Assumption of Homoscedasticity

- Breusch-Pagan test as a **quantitative** assessment

  - $\begin{cases} H_0\text{: Errors have constant variance} \\ H_a\text{:} Errors\ can\ be\ written\ as\ \ \varepsilon_t^2 = \alpha_0 + \alpha_1 V_{t,1} + \cdots + \alpha_p F_{t,p}, \end{cases}$

  where
  α reject if P-value < $\alpha = 0.05$;
  $V_{t,1}, F_{t,p}$ Voltage and frequency.



- Fitted vs. Errors as a **qualitative** assessment

  - Change of errors' spread around mean with fitted values
  - Points $x_i > 4\sigma$ from the mean indicate severe violation

Methods-Akaike's Information Criterion (AIC)

- Assesses discrepancy between f and g using Kullback-Leibler
- Given by

$$AIC = -2 \ln \mathcal{L}(\hat{\theta} \,|data)) + 2K,$$

where

K number of parameters;

$\hat{\theta}$ vector of parameters to be optimized.

- AICc given by

$$AICc = AIC + \frac{2K(K+1)}{n-K-1}.$$

- AICc converges to AIC as $n \to \infty$

Methods-Akaike's Information Criterion (AIC)

- Candidate equations ranked using

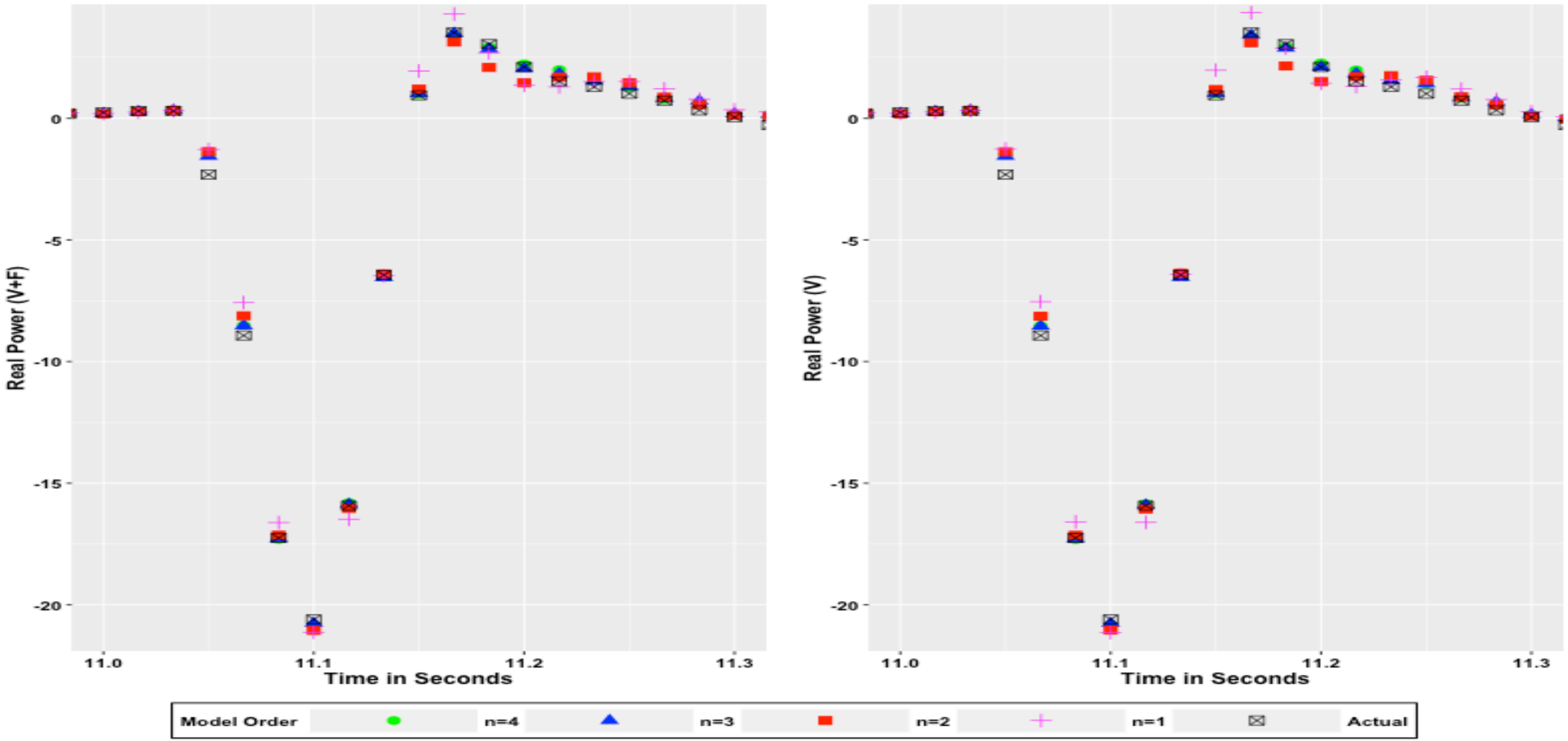
$$\begin{cases} \Delta_i = AIC_i - AIC_{min} \\ \mathcal{L}(g_i|data) = e^{-0.5\Delta_i} \\ Evidence\ ratio = \dfrac{1}{e^{-0.5\Delta_i}}, \end{cases}$$

where

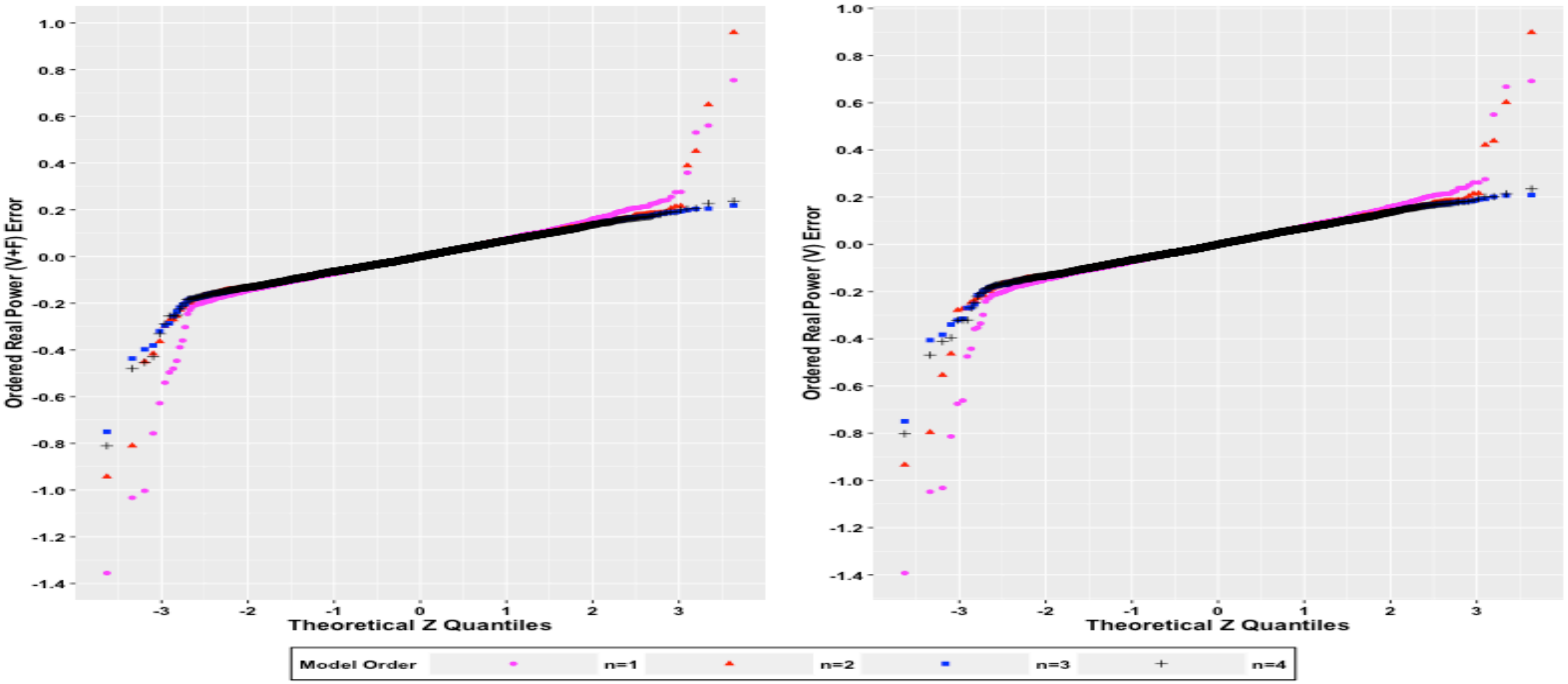
g_i	i candidate equations g;
AIC_{min}	best model;
$\mathcal{L}(g_i\|data)$	relative likelihood of each model given the data.

- Results interpreted as "evidence is x times stronger for best model"

Results-Real Power



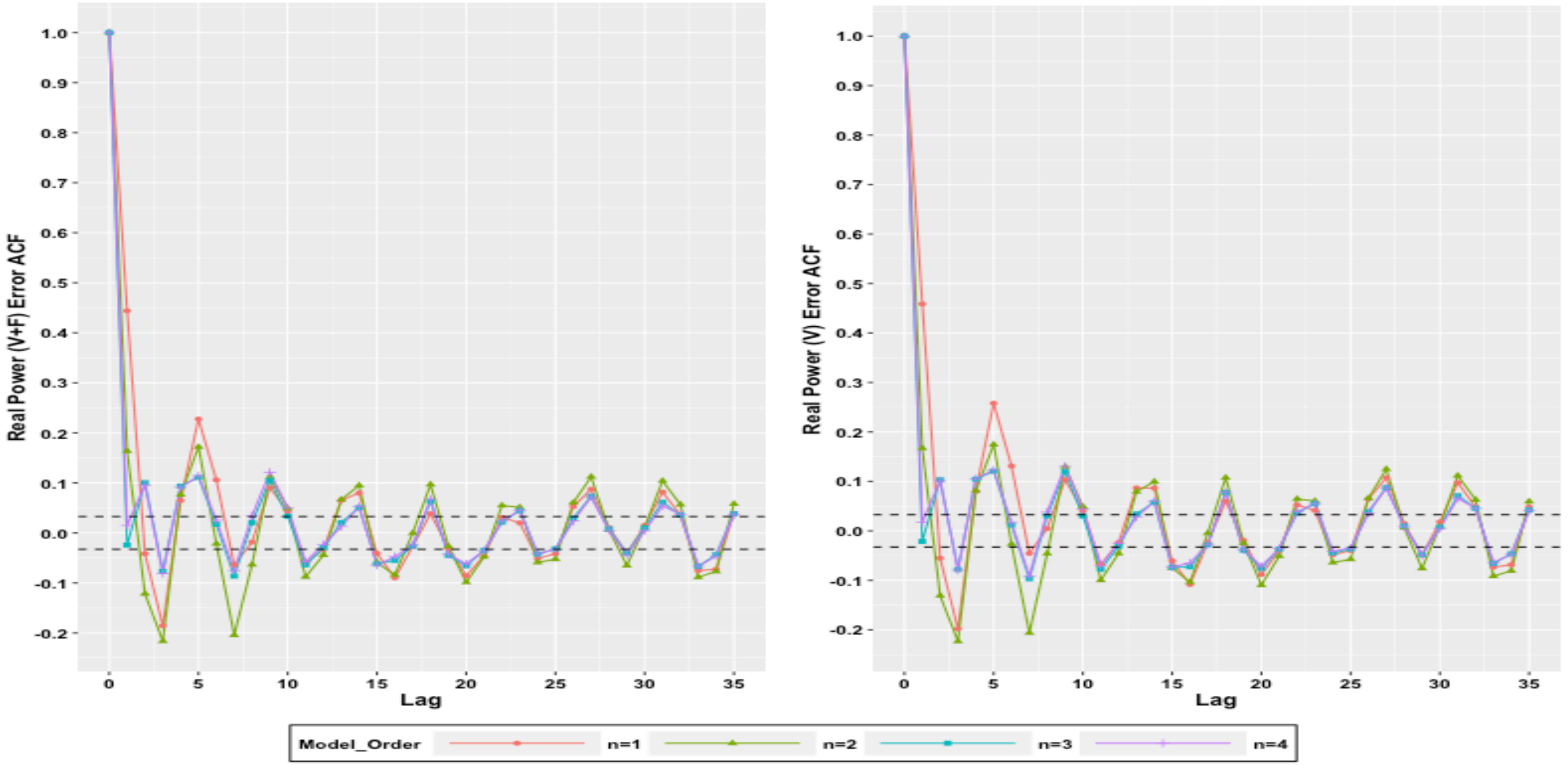
Results-QQ Plots for Real Power



Results- Anderson-Darling for Real Power

Model Order	Independent Variables	P_t Test Statistic (A)	P_t P-value
n = 1	V+F	21.808	2.2e-16
n = 1	V	21.728	2.2e-16
n = 2	V+F	8.5404	2.2e-16
n = 2	V	7.5586	2.2e-16
n = 3	V+F	1.3351	0.001842
n = 3	V	1.1237	0.006101
n = 4	V+F	1.8968	7.711e-05
n = 4	V	1.6646	0.0002858

Results-Autocorrelation Function for Real Power



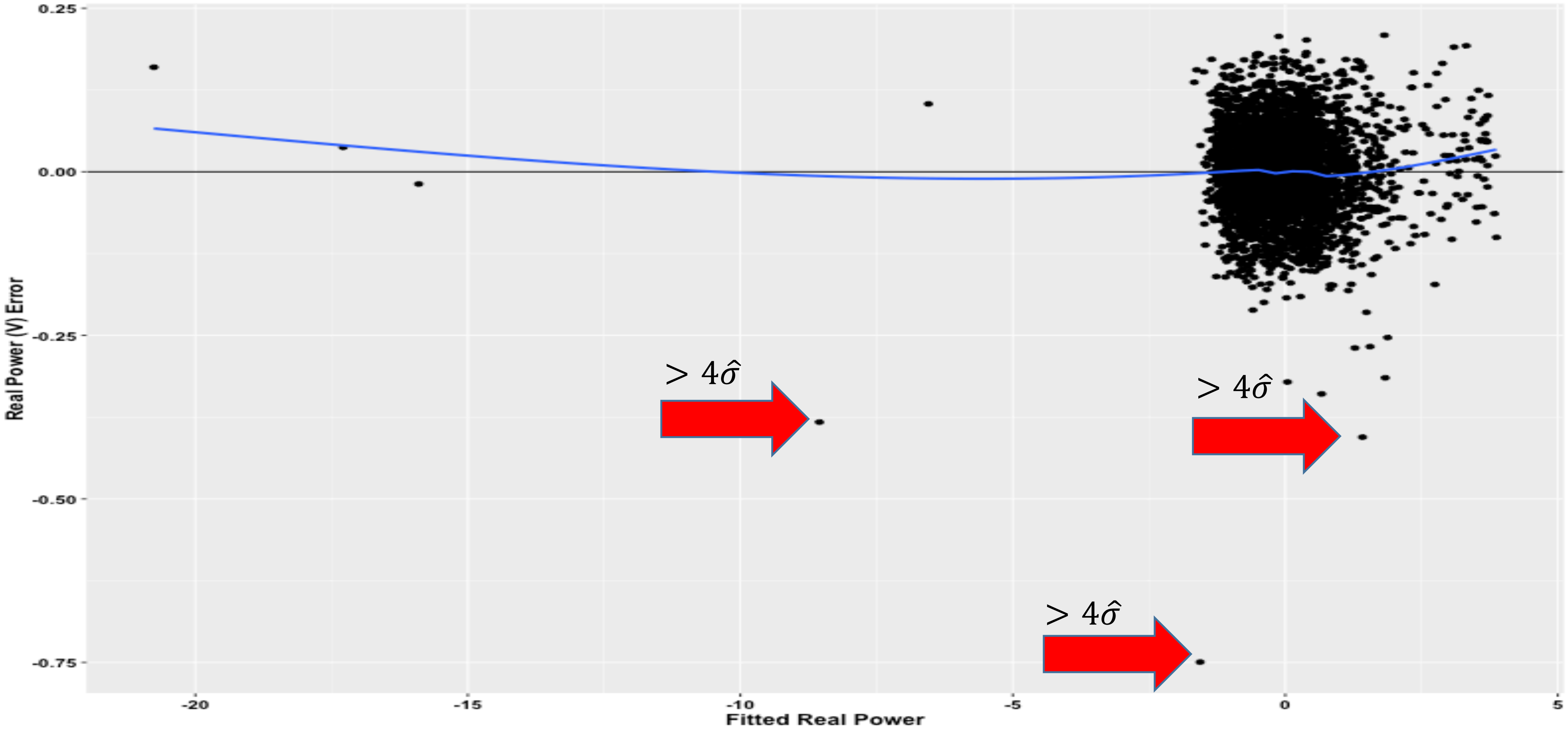
Results-Durbin Watson for Real Power

- # of uncorrelated lags/15

Model Order	Independent Variables	# of uncorrelated lags	Best
n = 1	V+F	1	
n = 1	V	2	
n = 2	V+F	1	
n = 2	V	1	
n = 3	V+F	5	
n = 3	V	5	select as best since simplest model.
n = 4	V+F	5	
n = 4	V	4	

Results-Fitted values vs. Residuals for Real Power

- Third order



Results- Breusch Pagan for Real Power

Model Order	Independent Variables	Chi-square	P-value
n = 1	V+F	1424.587	9.543696e-312
n = 1	V	1824.349	0
n = 2	V+F	175.0876	5.728838e-40
n = 2	V	152.4671	5.008834e-35
n = 3	V+F	9.110667	0.002541231
n = 3	V	5.606244	0.01789659
n = 4	V+F	28.13365	1.132202e-07
n = 4	V	14.74218	0.0001232578

Results- Conclusions for Real Power

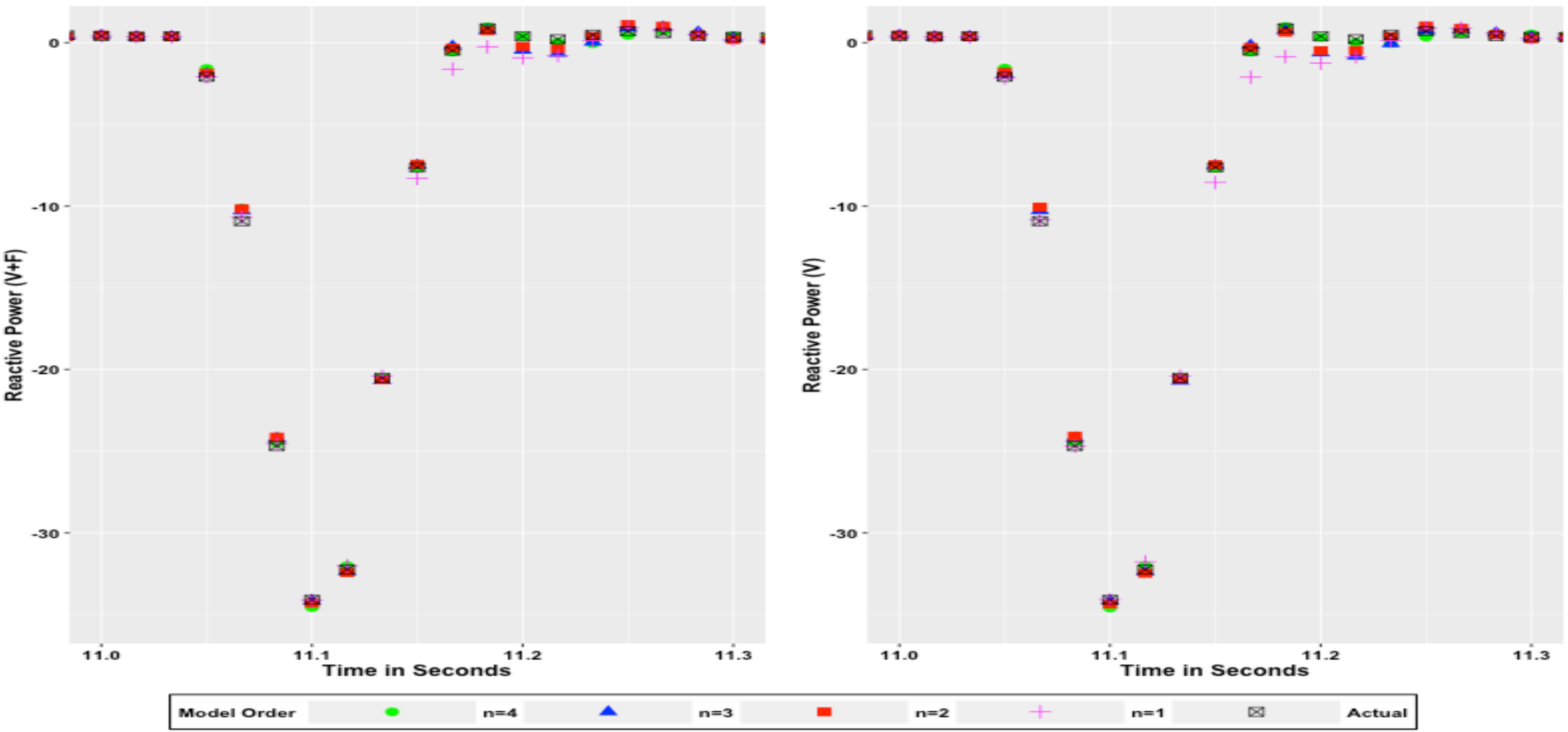
- **Normality:**
  a) QQ-Plot: 3rd and 4th order models are about the same for both V and V+F
  b) Anderson-Darling test: 3rd order voltage model is best

- **Independence:**
  a) Durbin-Watson test: no benefit gained from going beyond 3rd order voltage model
  b) ACF-Plot: agrees with the results of the DW test

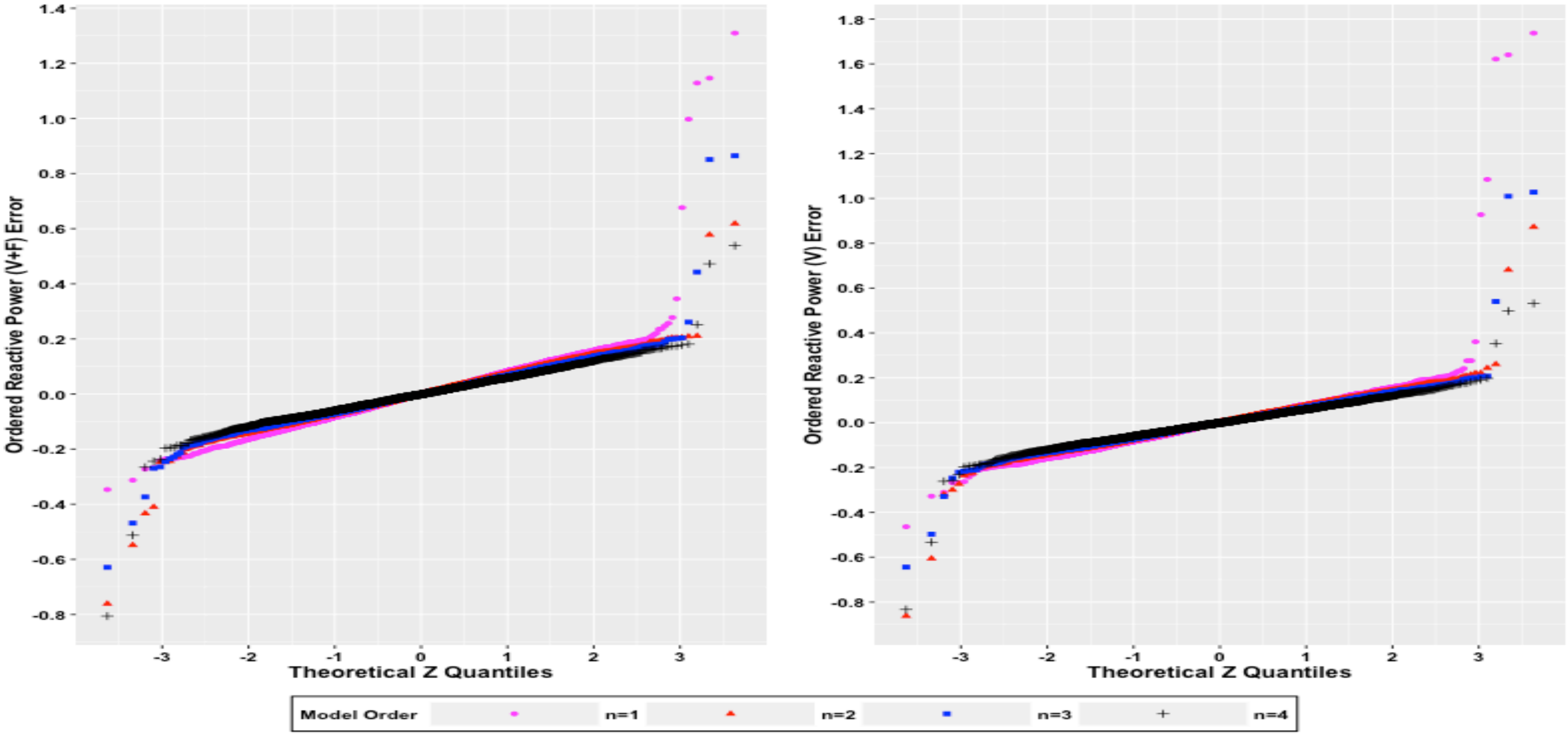
- **Homoscedasticity:**
  a) Breusch-Pagan test: no benefit gained from going beyond 3rd order voltage model
  b) Fitted vs. Residuals: 3rd order voltage model confirmed BP results

- **Conclusion for P_t:** 3rd order voltage model gives optimal results and is simplest

Results-Reactive



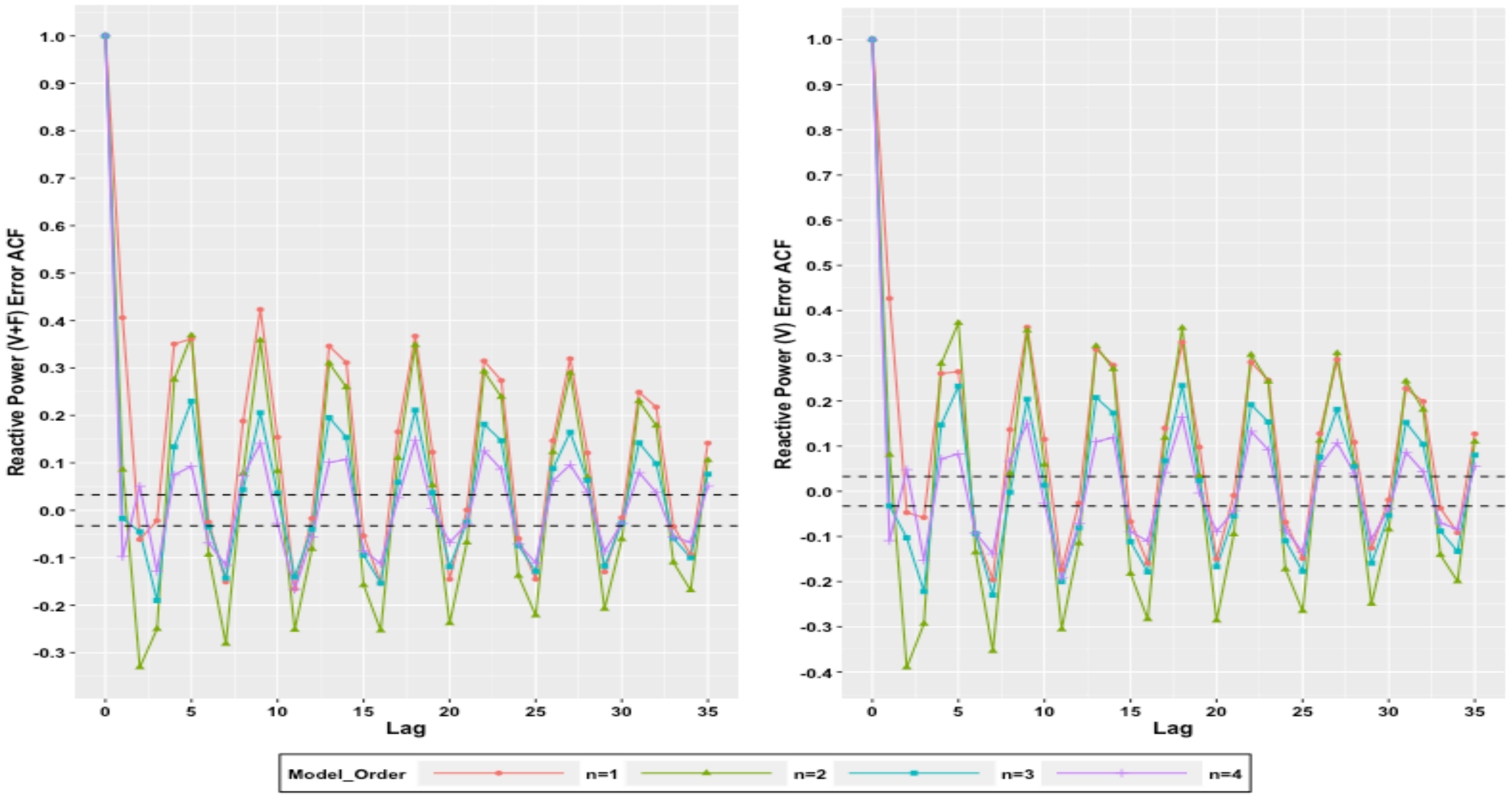
Results-QQ Plots for Reactive Power



Results-Anderson Darling for Reactive Power

Model Order	Independent Variables	Q_t Test Statistic (A)	Q_t P-value
n = 1	V+F	8.8153	2.2e-16
n = 1	V	26.36	2.2e-16
n = 2	V+F	1.4725	0.0008467
n = 2	V	2.305	7.736e-06
n = 3	V+F	3.8779	1.166e-09
n = 3	V	5.0148	2.133e-12
n = 4	V+F	3.2677	3.503e-08
n = 4	V	3.6999	3.141e-09

Results-Autocorrelation for Reactive Power



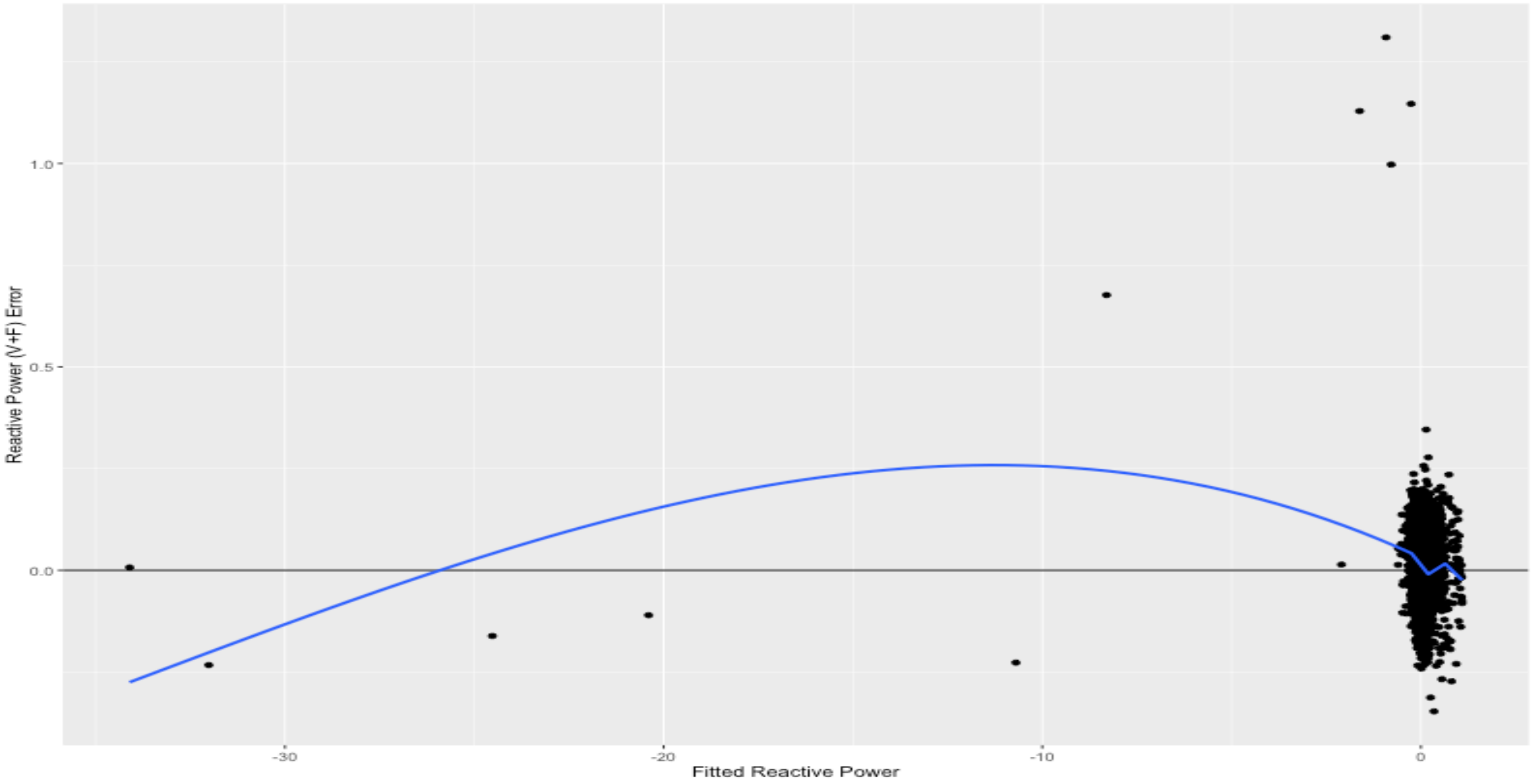
Results-Durbin Watson for Reactive Power

- # of uncorrelated lags/15

Model Order	Independent Variables	# of uncorrelated lags	Best
n = 1	V+F	3	
n = 1	V	1	
n = 2	V+F	0	
n = 2	V	0	
n = 3	V+F	1	
n = 3	V	3	Select as best since simplest model.
n = 4	V+F	1	
n = 4	V	1	

Results-Fitted values vs. Residuals for Reactive Power

- Third order



Results- Breusch Pagan for Reactive Power

Model Order	Independent Variables	Chi-square	P-value
n = 1	V+F	208.4134	3.048111e-47
n = 1	V	1008.033	3.22238e-221
n = 2	V+F	988.6812	5.182923e-217
n = 2	V	1551.624	0
n = 3	V+F	622.6902	1.943818e-137
n = 3	V	766.2074	1.200529e-168
n = 4	V+F	2548.82	0
n = 4	V	2903.575	0

Results- Conclusions for Reactive Power

- **Normality:**
  a) QQ-Plot: 4^{th} order models are about the same for both V and V+F
  b) Anderson-Darling test: states that 2^{nd} order V+F is best

- **Independence:**
  a) Durbin-Watson test: no benefit gained from going beyond 3^{rd} order V model
  b) ACF-Plot: agrees with the results of the DW test

- **Homoscedasticity:**
  a) Breusch-Pagan test: no benefit gained from going beyond 1^{st} order V+F model
  b) Fitted vs. Residuals: 1^{st} order V+F model confirmed BP results

- **Conclusion for Q_t:** All models considered are performing poorly. 4^{th} order V model selected based on figures from acf and QQ-plot

Next Method

AICc based Ranking

Results-AICc based Ranking for Real Power

Model Order	Number of parameters K	Independent Variables	AICc	Δ_i	$\mathcal{L}(g_i\|data)$	Evidence Ratio
n = 4	K = 14	V+F	-19169.22	0.00	1.0	1
n = 3	K = 11	V+F	-19124.46	44.76021	1.907406e-10	5.242723e+09
n = 4	K = 9	V	-19103.05	66.17212	4.274713e-15	2.339338e+14
n = 3	K = 7	V	-19076.08	93.14021	5.954715e-21	1.679342e+20
n = 2	K = 8	V+F	-18640.93	528.29351	1.916556e-115	5.217692e+114
n = 2	K = 5	V	-18619.30	549.91742	3.863226e-120	2.588510e+119
n = 1	K = 5	V+F	-17555.56	1613.66177	0.00	∞
n = 1	K = 3	V	-17465.17	1704.04749	0.00	∞

Results-AICc based Ranking for Reactive Power

Model Order	Number of Parameters K	Independent Variables	AICc	Δ_i	$\mathcal{L}(g_i \mid data)$	Evidence Ratio
n = 4	K = 14	V+F	-19908.81	0.0000	1	1
n = 4	K = 9	V	-19813.68	95.13445	2.196933e-21	4.551801e+20
n = 3	K = 11	V+F	-18912.08	996.73146	3.651856e-217	2.738333e+216
n = 3	K = 7	V	-18697.87	1210.94449	1.113643e-263	8.979536e+262
n = 2	K = 8	V+F	-18550.95	1357.86493	1.391158e-295	7.188254e+294
n = 2	K = 5	V	-18341.09	1567.72144	0	∞
n = 1	K = 5	V+F	-17199.56	2709.25215	0	∞
n = 1	K = 3	V	-16720.93	3187.88793	0	∞

Results-AICc Ranking

- Correctly identifies best models
- Evidence ratio values hard to reconcile with results
- **For $\boldsymbol{P_t}$:**
  a) AICc states that the most complicated model is the best
  b) No benefit beyond 3rd order V model
  c) Conclude that AICc differences between top 4 models not significant
  d) Choose simplest model (3rd order voltage model)

- **For $\boldsymbol{Q_t}$:**
  a) AICc states that the most complicated model is the best
  b) No benefit beyond 4th order V model
  c) Conclude that AICc differences between top 2 models not significant
  d) Choose simplest model (4th order voltage model)

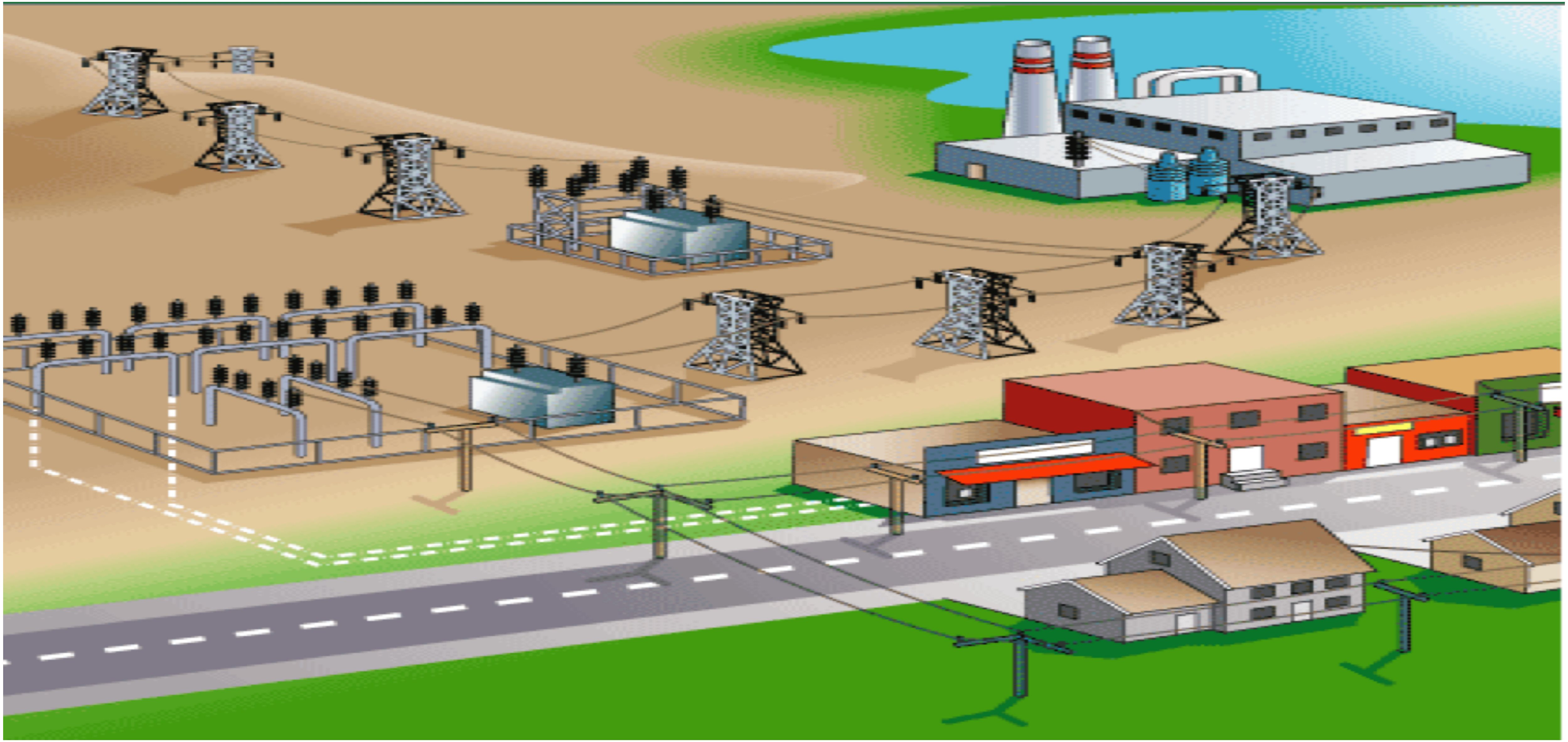
Results-Final Load Model

- Final equations selected for P_t and Q_t

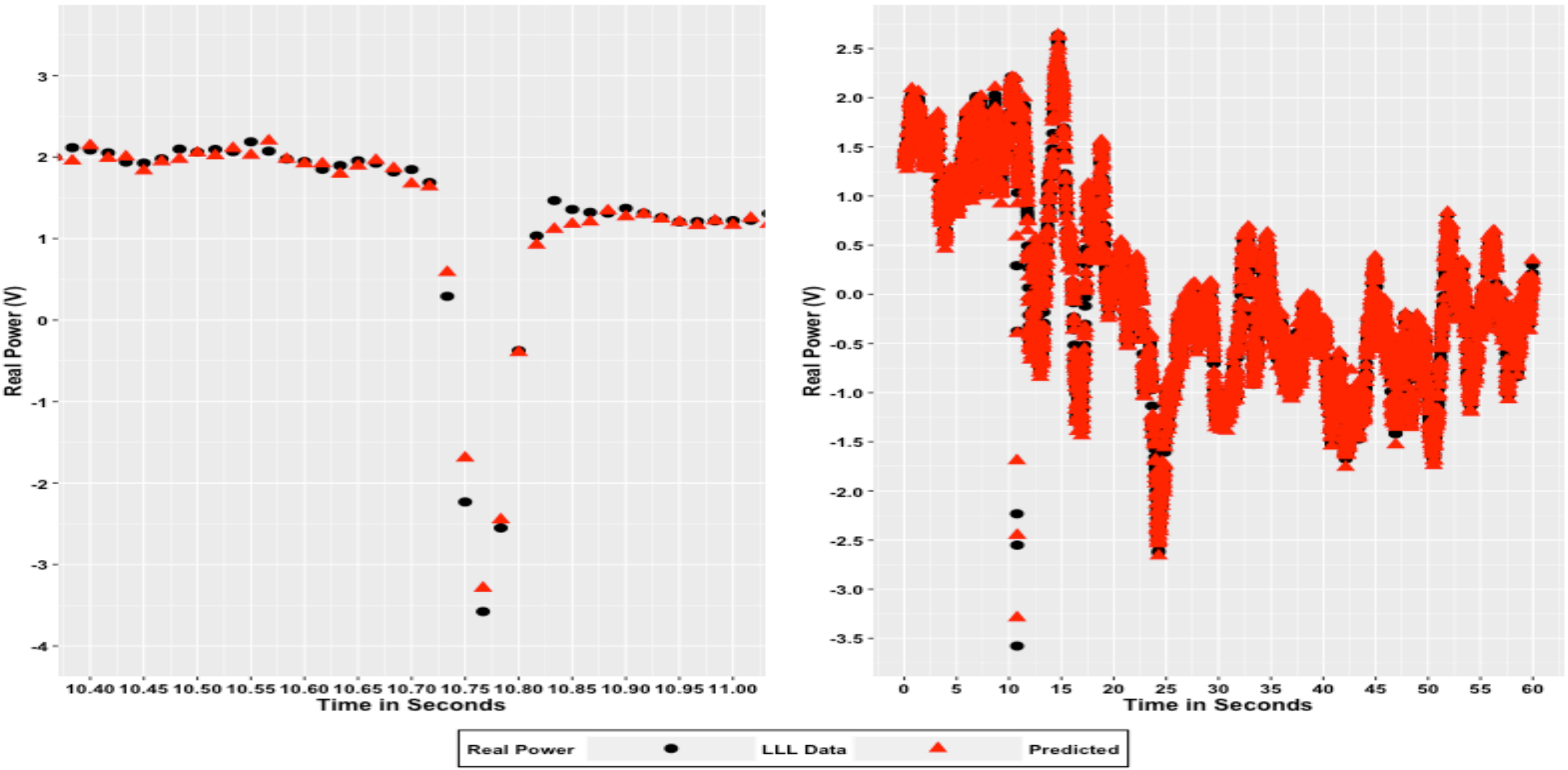
$$\begin{cases} \widehat{P_t} = 1.49P_{t-1} - 0.67P_{t-2} + 0.18P_{t-3} + \\ \qquad 0.66V_t - 1.13V_{t-1} + 0.65V_{t-2} - 0.198V_{t-3} \\ \widehat{Q_t} = 1.44Q_{t-1} - 1.07Q_{t-2} + 0.73Q_{t-3} - 0.11Q_{t-4} + \\ \qquad 0.71V_t - 0.65V_{t-1} + 0.12V_{t-2} + 0.03V_{t-3} - 0.197V_{t-4} \end{cases}$$

- Set of equations defines load model

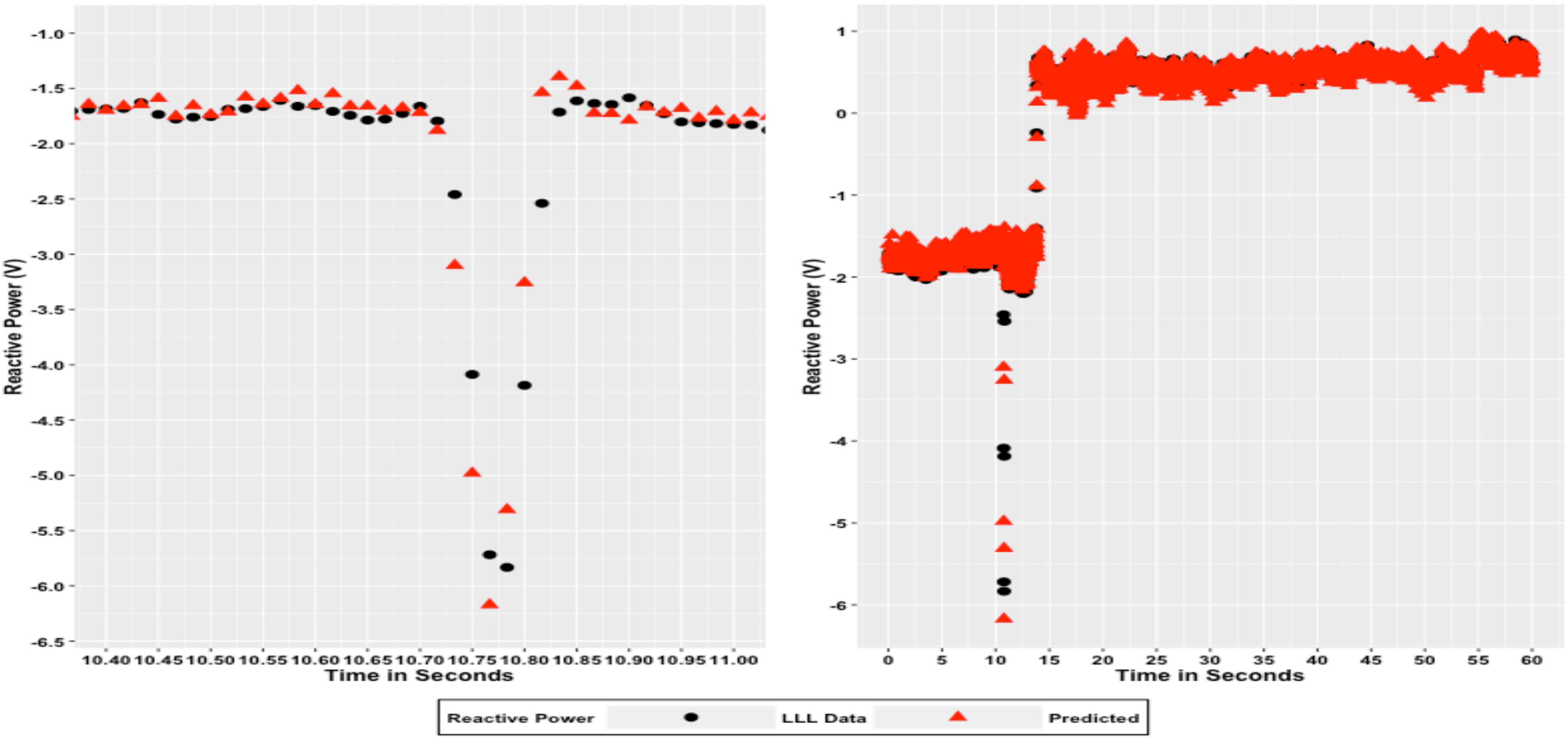
Results-Line to Line to Line (LLL) Fault



Results-Real Power Tested on LLL Fault Data



Results-Reactive Power Tested on LLL Fault Data



Results-Validation of Real Power

Model Order n	# of Parameters k	Model	RMSE Train	RMSE Validation	Factor
N = 2	8	V+F	0.0747	0.0776	1.0
N = 1	5	V+F	0.0869	0.0868	1.0
N = 3	7	V	0.07035	0.0834	1.2
N = 3	11	V+F	0.0699	0.0845	1.2
N = 4	14	V+F	0.0694	0.0842	1.2
N = 1	3	V	0.0015	0.0853	57
N = 2	5	V	0.0013	0.0777	60
N= 4	9	V	0.0012	0.0834	70

Results-Validation of Reactive Power

Model Order n	# of Parameters k	Model	RMSE Train	RMSE Validation	Factor
N = 3	7	V	1.7593	0.0742	24 up
N = 4	14	V+F	0.0625	0.0788	1.3
N = 2	8	V+F	0.0757	0.1744	2.3
N = 1	5	V+F	0.0914	0.3373	3.7
N = 3	11	V+F	0.0719	1.6931	24
N = 1	3	V	0.0016	1.6666	1023
N = 2	5	V	0.0013	1.7437	1342
N= 4	9	V	0.0011	1.7843	1687

Discussion

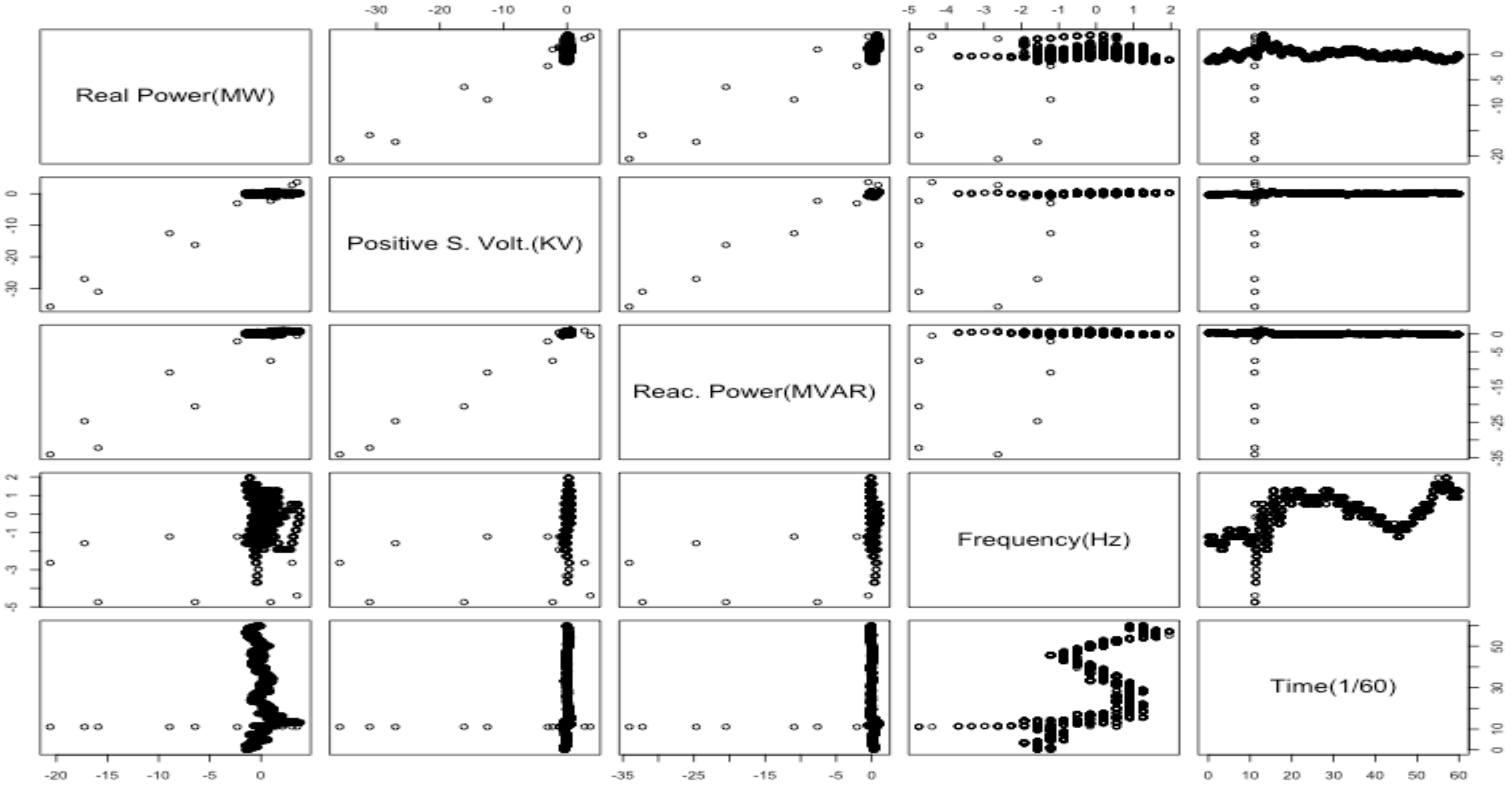
- Linear models capture real power acceptably
- No improvement beyond 3rd order real power linear voltage model
- Reactive power not linear
- No improvement beyond 4th order reactive power linear voltage model
- Congruency between statistical methods

Future Work

- Different model types should be explored
- Different parameter estimators
- Model errors
- Better test data

Results

- Pairs plot of SLG data



AIC

III.V Method 2: Akaike Information Criterion for Model Selection

In method 2 the second order Akaike Information Criterion (AICc) is used to rank a set of linear regression models. AIC assesses the discrepancy between two continuous models based on the Kullback-Leibler (K-L) distance given by

$$I(f,g) = \int f(x) \cdot \log(f(x)) \cdot dx - \int f(x) \cdot \log(g(x|\theta)) \cdot dx, \tag{20}$$

where

$I(f,g)$ is defined as the "information loss;"
$f(x)$ is the unknown reality with conceptually infinite parameters;
$g(x|\theta)$ is the approximating model being compared to $f(x)$ [20].

Now, equation 20 is nothing more than

$$I(f,g) = \mathbf{E}_f[\log(f(x))] - \mathbf{E}_f[\log(g(x|\theta))], \tag{21}$$

AIC

$\mathbf{E}_f$ is the expectation with respect to the truth.

Realizing that the first term on the right-hand side of equation 21 is a constant (an unknown AIC number), equation 21 becomes

$$I(f,g) - C = -\mathbf{E}_f[\log(g(x|\theta))] \tag{22}$$

From equation 22, the problem becomes one of minimizing the right-hand side, but this term is not computable as it stands. It has been shown that when data is introduced to the derivation, the expectation of the K-L information can be estimated by

$$\mathbf{E}_y\mathbf{E}_x[\log(g(x|\hat{\theta}(y)))] = \log(\mathcal{L}(\hat{\theta}|\text{data})) - K, \tag{23}$$

where

θ is a vector of parameters;

K is the number of parameters [21].

For historical reasons, 23 is multiplied by -2. In practice, it is recommended that AICc be used due to the fact that it has a correction term for small sample sizes. AICc is given by