國立中山大學 應用數學系

碩士論文

題目:電磁波干擾下標準場地衰減測量值之統計模型探討

研究生：鄭竣平 撰

指導教授：羅夢娜 教授

中華民國　九十二　年 六 月

A Study on the Statistical Models of Normalized Site Attenuation(NSA) Measurements for Electromagnetic Interference(EMI).

by

C. P. Cheng

Advisor

M.-N. L. Huang

Department of Applied Mathematics,
National Sun Yat-sen University
Kaohsiung, Taiwan 804 R.O.C.

June 2003

中山 eThesys
學位論文

Contents

Abstract

In this work, we discuss the accuracy of measurements for electromagnetic. The two kinds of antenna we use are Dipole antenna and Broadband antenna. In general, if the antenna measurements we recorded at different frequencies do not exceed the ideal value ±4dB, we would regard this site as a normalized site, otherwise it is not a normalized site(just a measurement exceeds the range). Traditionally, all we use is Dipole antenna, but due to difficulty of operation and inaccuracy of Dipole antenna, we investigate by statistical methods if we may use the Broadband antenna to replace the traditional Dipole antenna to measure. First of all, we introduce the data and procedure in the experiments, and fit a statistical regression model to predict the measurements at different frequencies in different test setups. Then, according to the data we collected, use the change point models to modify the statistical models. Our goal is to find a suitable statistical model for the measurements. Finally, we compare the measurements of Broadband antenna with Dipole antenna in the other experimental conditions keep the same.

Keywords: Broadband antenna, Dipole antenna, Regression model, Change point model, Piecewise linear regression.

List of Tables

List of Figures

1. Introduction

EMI is electromagnetic interference. The source of EMI includes a switch circle, static electricity, lightning, power. Basically, the EMI spread in the space all the time. The electric products for human using will send out EMI either large or small. Every electronic products have EMI's problems, and all products have to be tested and inspected on a EMI's unit. The electronic products are allowed to enter the market after passing the EMI's test. It will cause waste of time and money if the designer ignores the problems of EMI. The electromagnetic interference for products will appear and cause to the failure of products. So in this work we are interested in studying the EMI's problems.

There are many different test setups in this experiment such as two different kinds of antenna: Dipole antenna and Broadband antenna; the polarization of antenna: horizontal and vertical; and the different transmit antenna height and receive antenna height. At every test setup, besides the 27 data points from frequency 30 MHz to 1000 MHz recorded, we also take records on Vdir, Afrx, Aftx, and Vsite values at each data point. The six test setups for this experiment are:

1 Broadband, hor, h1=1M, h2=1-4M

2 Broadband, hor, h1=2M, h2-1-4M

3 Broadband, ver, h1=1M,h2=1-4M

4 Broadband, ver, h1=1.5M, h2=1-4M

5 Dipole, hor, h1=2M, h2=1-4M

6 Dipole, ver, h1=2.75M, h2=2.75-4M

And at each test setup we repeat the experiment three times during three days(1^{st}, 2^{nd}, 3^{rd}). In the data we collected, the *dev* values are the measurements after subtracting the ideal values. All our focus is on the difference between measurements($meas$ =Vdir-Vsite-Afrx-Aftx) and ideal values. In Section 2 we begin analyzing the experimental data: we first compare the difference between the repeating measurements, and fit the statistical models. In Section 3, we use the change point models to modify the statistical models we fitted in Section 2, discuss the pure error for model estimation, and try to find a better fit model. In Section 4, we contrast the measuring difference between the two kinds of antenna, including the comparison of Dipole and Broadband antenna and the comparison of different polarization and different transmit antenna height in Broadband antenna. The final section is the conclusion.

2. Fitting models

After introducing the data obtained in this experiment, we analyze the data using the statistical model method. Firstly, we use the statistical models to compare and test the replication data, then fit the polynomial regression models according to the result of tests.

We have repeated the data at each test setup three times. Therefore, before fitting the models, we should consider whether there are differences of these measurements or not. In these models, the dependent variables are the frequencies which take log_{10}, denoted below as $log_{10}(freq)$, and do the following analysis.

a. We fit one model respectively to the repeated measurements, and contrast directly these three models at each test setup.

b. Then the test methods are as following: in each test setup, we fit another model with combining the repeat data and add two indicator variables to this model in order to distinguish three repeat measurements. We examine whether the effects of two indicator variables are dominated or not in this model.

According to the analysis, the three models at each test setup are not obviously different(see **Tables I-1∼I-6**). We formally test the three repeated data by adding two indicator variables into the combined models:

(2.1) $Y_t = \alpha_0 + \alpha_1 X_1 + \alpha_2 X_2 + \alpha_3 X_3 + \varepsilon_t$

and $Y_t = meas, X_1 = log_{10}(freq), (X_2, X_3) = (0, 0)$, represents the 1^{st} measurement, and $(X_2, X_3) = (1, 0)$ represents the 2^{nd} measurement, and $(X_2, X_3) = (0, 1)$ represents the 3^{rd} measurement, ε_t i.i.d is normal distribution; In this model, the null hypothesis is :

$$H_0 : \alpha_2 = 0 \quad \text{v.s} \quad H_1 : \alpha_2 \neq 0 \qquad \text{and} \qquad H_0 : \alpha_3 = 0 \quad \text{v.s} \quad H_1 : \alpha_3 \neq 0$$

at significant level α, we reject H_0 iff $| t^* | > t(1 - \frac{\alpha}{2}, n - 2)$. As the result, all the R^2 value of each model are close 1(**Table I-7**) and all the conclusions receive H_0 at six different test setups. So we obtain that these three repeated measurements at different test setups do not have measuring differences according to our statistical analysis.

Before fitting models, firstly we draw some scatter plots for measurements to see the relations of the data and the forms of models(**Figure 2.1,2.2**). We can get some information from scatter plots: the two kinds of antenna have different tendencies on measuring and an apparent variability appears when the antenna's polarization is vertical and low frequencies. By the scatter plots for measurements versus frequencies and ideal versus frequencies, they tend to square curves, so we take frequencies to $log_{10}(freq)$ in order to remove the square tendency because of the unequal distance in original frequency interval.

As the result, we fit linear and square polynomial regression models at each test setup without the repeated data(**Table I-8,I-9**). What we interest in is the difference between measurements and ideal values at different frequencies in each test setup. No matter what linear or square models we fitted, R^2 for models are high. But these models are still not perfect because the residual plots appear some obvious tendencies. So we consider to use another models to modify the residual.

3. The change point models

If the simple regression models do not correspond to our research data, we have two basic choices:

a. Find another better models.

b. Transform the original data to fit models.

We hope that we can find a much better model to predict the measurements, but we cannot get a better result after transforming original data(because of the residuals). And we cannot explain its physical meaning easily after transforming. According to the principles of EMI, low frequencies have a obvious difference than high frequencies when measuring. Our thought is to find a change point to divide data into two groups(low frequencies and high frequencies), and use a change point model to fit data which we got. Among many previous papers have studied this subject. Generally speaking, what we interest in change point models are these two points(*Krishnaiah* and *Miao* (1988)): (1)to determine whether the change points exist in models or not (2)to estimate the numbers of change points and positions of these change points in models. Many papers had investigated the estimation of change points, such as *Quandt* (1958, 1960), *Quandt* and *Ramsey* (1978), and *Robison* (1964), they all discussed the problems of switch regression models. *Page* (1954, 1955, 1957) firstly used the cumulative sum(CUSUM) method to estimate change points. *Hudson* (1966) and other authors also put forward the maximum likelihood and least square methods to estimate and test change points. Then, *Hinkley* (1971) proved the approximate distribution of the intersections of two models. In addition, *Chenrnoff* and *Zacks* (1964), *Smith* (1975), *Carter* and *Blight* (1981) used the bayesian estimator to estimate. And *Freeman* (1984, 1986)discussed the goodness of fit for two-phase regression and an unknown change point model. However, in the change point models, they become more complicated as long as the numbers of change point are more than one. So we suppose that there is only one change point in our models, because of the feature of data(low frequencies and high frequencies). And we suppose that the errors of model ε_t are i.i.d $N(0, \sigma^2)$(σ is unknown).In this section, we discuss the methods to fit change point models at six different test setups. Here is the procedures:

a. Find the change point

b. Test the change point we found

c. Test the change points are jump points or not

d. Examine the change point models

3.1 Estimation of change points

When estimating change points, we use two methods to deal with; one is to use the theory values of NSA and E_D^{MAX} to estimate the position of change points, another is to use the piecewise regression model to estimate. The two methods are as following:
a. Use the formula of NSA's theory value:
The formula of NSA's theory values were introduced by *Akira*(1990, 1992).
$NSA_{TH} = -20log(f_M) + 48.9 - E_D^{MAX}$.
f_M= frequency in MHz.
E_D^{MAX}= maximum received field.
The formula of E_D^{MAX}(vertical and horizontal) is

$$E_{DH}^{MAX} = \frac{\sqrt{49.2}(d_2^2 + d_1^2 \mid \rho_H \mid^2 + 2d_1 d_2 \mid \rho_H \mid cos[\Phi_H - \frac{2\pi}{\lambda}(d_2 - d_1)]^{1/2})}{d_1 d_2}$$

$$E_{DV}^{MAX} = \frac{\sqrt{49.2}R^2(d_2^6 + d_1^6 \mid \rho_V \mid^2 + 2d_1^3 d_2^3 \mid \rho_V \mid cos[\Phi_V - \frac{2\pi}{\lambda}(d_2 - d_1)]^{1/2})}{d_1^3 d_2^3}$$

$$E_D^{MAX}(ideal) = 48.9 - 20log(f_M) - ideal$$

$E_D^{MAX}(meas) = 48.9 - 20log(f_M) - meas$

draw $E_D^{MAX}(ideal)$ versus $log_{10}(freq)$(**Figure 3.1**), and directly determine the initial position of change point according to the figures. After deciding the position, we can get two linear models. We take the point which make the residual sum of these two models being the smallest. Here is the least square estimation model:
(3.1)

$$\begin{aligned} y_t &= \begin{cases} \alpha_{10} + \alpha_{11}t + \varepsilon_t & , \quad 0 < t \leq t_m \\ \alpha_{20} + \alpha_{21}t + \varepsilon_t & , \quad t > t_m \end{cases} \\ &= \begin{cases} \alpha_1' X(t) + \varepsilon_t & , \quad 0 < t \leq t_m \\ \alpha_2' X(t) + \varepsilon_t & , \quad t > t_m \end{cases} \end{aligned}$$

$Q(\alpha, t) = \Sigma_{k=1}^{m}(y_k - \widehat{\alpha}_1' X(t))^2 + \sum_{k=m+1}^{n}(y_k - \widehat{\alpha}_2' X(t))^2$

$Q(\widehat{\alpha}, \widehat{t}) = minQ(\alpha, t)$

$y_t = measurements$, $t = log_{10}(freq)$, we take a change point(t_M) firstly, then we can get the parametric estimations of these two models. We put t_M into $Q(\alpha, t)$. Find $\hat{t_M}$ which makes the $Q(\alpha, t)$ be the smallest by repeating the procedure, after that $\hat{t_M}$ is the estimation of change point. The estimations of change point at six different test setups are individually 180MHz, 90MHz, 250MHz, 300MHz, 90MHz, 140MHz.
b. Piecewise Regression:
model is as follows.
(3.2) $y_t = \beta_0 + \beta_1 t + \beta_2(t - t_m)I_{[t>t_m]} + \varepsilon_t$
$y_t = measurements$, $t = log_{10}(freq)$, $t_m = log_{10}(changepoint)$, $I_{[t>t_m]}$ is equal to 0, if $t \leq t_m$; $I_{[t>t_m]}$ is equal to 1, if $t > t_m$. Find the $\hat{t_M}$ which makes the SSE of models be the smallest. The SSE's figures are shown as in **Figure 3.2**, and the estimations of change point at six different test setups are individually 160MHz, 80MHz, 400MHz, 250MHz, 80MHz, 60MHz.

3.2 Test the change points

We use two methods to estimate change points in Section 3.1. After estimating the positions of change point, we test respectively the change points we found:

a. Let the model is:

(3.3)

$$\begin{aligned} y_t &= \begin{cases} \alpha_{10} + \alpha_{11} t + \varepsilon_t & , \quad 0 < t \le t_m \\ \alpha_{20} + \alpha_{21} t + \varepsilon_t & , \quad t > t_m \end{cases} \\ &= \alpha_{10} + \alpha_{11} t + (\alpha_{20} + \alpha_{21} t) \cdot I_{[t>t_m]} + \varepsilon_t \end{aligned}$$

here the null hypothesis is $H_0 : \alpha_2' = (\alpha_{20}, \alpha_{21}) = 0$, versus $H_1 : \alpha_2' = (\alpha_{20}, \alpha_{21}) \neq 0$ and the test statistic is :

$$F^* = \frac{SSE(R) - SSE(F)}{df_R - df_F} / \frac{SSE(F)}{df_F}$$

at the significant level $\alpha = 0.05$, reject H_0 iff $\mid F^* \mid > F_{2,77}(0.05)$. The result is as **Table 1**, and the conclusion rejects H_0. The points we found are obviously change points in the models.

Then we test whether the change points are jump points or not:

Here the null hypothesis $H_0 : \alpha_{20} = 0$ versus $H_1 : \alpha_{20} \neq 0$, and use the t-test, reject H_0 iff $\mid t^* \mid > t(0.975, 79)$. The result is as **Table 1**, the critical value = t(0.975,79)=2.2849, and we obtain that the conclusion is H_1. And it means that the change point we found are jump points. In other words, slope changes in the change point models, and "jump" on the position of change point.

b.The method we use is that find the $\hat{t}_M$ makes the model have the smallest SSE. So we use the same models to test the change point by using statistical test:

consider the model as in (3.2). Here the null hypothesis $H_0 : \beta_2 = 0$ versus $H_1 : \beta_2 \neq 0$ and we reject H_0 iff $\mid F^* \mid > F_{2,77}(0.05)$. The result is as **Table 2**, and the conclusion is reject H_0. The points we found are also obviously change points in the models.

Then we test whether the change points are jump points or not:

Add $\beta_3 I_{[t>t_m]}$ into the model (3.2) because we want to check whether the model is continuous on the position of change point or not, consider the model:

(3.4) $$y_t = \beta_0 + \beta_1 t + \beta_2 (t - t_m) I_{[t>t_m]} + \beta_3 I_[t > t_m] + \varepsilon_t$$

the null hypothesis $H_0 : \beta_3 = 0$ versus $H_1 : \beta_3 \neq 0$, and the critical value is t(0.975,79)=2.2849. The result is as **Table 2**, and the conclusion is H_0. It means that the slope change without jump points in models for method 2.

Table 1 Test statistic and p-value of model (3.3)

	F^*	p-value	t^*	p-value	c.p
BD(hor,h1=1M,h2=1-4M)	188.25	0.0000	-15.017	0.0000	180MHz
BD(hor,h1=2M,h2=1-4M)	195.476	0.0000	-19.434	0.0000	90MHz
BD(ver,h1=1M,h2=1-4M)	27.914	0.0000	6.866	0.0000	250MHz
BD(ver,h1=1.5M,h2=1-4M)	41.735	0.0000	2.744	0.0075	300MHz
DP(hor,h1=2M,h2=1-4M)	167.337	0.0000	-18.028	0.0000	90MHz
DP(ver,h1=2.75M,h2=2.75-4M)	12.083	0.0000	3.154	0.0023	140MHz

Table 2 Test statistic and p-value of model (3.2) and (3.4)

	F^*	p-value	t^*	p-value	c.p
BD(hor,h1=1M,h2=1-4M)	190.66	0.0000	0.255	0.7994	160MHz
BD(hor,h1=2M,h2=1-4M)	172.668	0.0000	-0.098	0.9222	80Mhz
BD(ver,h1=1M,h2=1-4M)	26.062	0.0000	-1.778	0.0793	400MHz
BD(ver,h1=1.5M,h2=1-4M)	38.697	0.0000	-1.248	0.2157	250MHz
DP(hor,h1=2M,h2=1-4M)	159.85	0.0000	0.692	0.4909	80MHz
DP(ver,h1=2.75M,h2=2.75-4M)	10.872	0.0000	0.078	0.9380	60MHz

3.3 Goodness of fit test

As the results, we used two methods to find change points, then test change points we found. So we can obtain that the change point models found by method 1 consist of two straight lines with different slopes. The other change point models found by method 2 are continuous broken lines with different slopes. Hence, we discuss the goodness of fit of these two change point models separately.

a. Model found by method 1:

By **Figure 3.3**, the goodness of fit is pretty good. We put emphasis on the continuous change point models.

b. Model found by method 2:

As the result, we can confirm the model is as in model (3.2), and the model information and estimations of change points are shown as **Table 3.1**:

Table 3.1 Coefficients of model (3.2) and estimations of change point

	β_0	β_1	β_2	R^2	c.p	MSE
BD(hor,h1=1M,h2=1-4M)	83.921	-36.44	14.27	0.996	160MHz	0.669
BD(hor,h1=2M,h2=1-4M)	79.128	-36.692	15.959	0.996	80MHz	0.458
BD(ver,h1=1M,h2=1-4M)	37.27	-15.214	-14.435	0.986	400MHz	1.899
BD(ver,h1=1.5M,h2=1-4M)	34.928	-13.792	-12.039	0.976	250MHz	1.774
DP(hor,h1=2M,h2=1-4M)	83.469	-39.261	17.538	0.996	80MHz	0.596
DP(ver,h1=2.75M,h2=2.75-4M)	34.041	-11.375	-11.237	0.979	60MHz	1.999

Table 3.2 The front and rear model information of model (3.2)

	c.p	R_1^2	R_2^2	MSE_1	MSE_2
BD(hor,h1=1M,h2=1-4M)	160MHz	0.988	0.990	0.9088	0.3845
BD(hor,h1=2M,h2=1-4M)	80MHz	0.972	0.993	0.7880	0.3296
BD(ver,h1=1M,h2=1-4M)	400MHz	0.908	0.934	2.1965	0.5184
BD(ver,h1=1.5M,h2=1-4M)	250MHz	0.865	0.960	2.1613	0.7616
DP(hor,h1=2M,h2=1-4M)	80MHz	0.990	0.987	0.3213	0.7215
DP(ver,h1=2.75M,h2=2.75-4M)	60MHz	0.311	0.981	4.6193	1.3795

The residual plots of model are shown as in **Figure 3.4**. Although the R^2 of models are quite high, we have to check the residuals:

(1)The normal distribution of residual
Normal P-P plots are shown as in **Figure 3.5**. We can obtain that residuals didn't vary from the normal hypothesis too far.

(2)The correlations of residual(Durbin-Watson test)
Durbin-Watson test is used to check whether the correlation of residual is AR(1) model or not(**Figure 3.6**). The critical value of Durbin-Watson test is difficult to obtain, but the smaller DW value lead to the auto correlation relation. The result of test is shown as **Table 4**.

By Table 4, the DW value of 2^{nd} test setup is 0.75, it seems to have some auto correlation tendencies in 2^{nd} test setup. We cannot make sure it has AR(1) model by **Figure 3.6**. And we try to fit a AR(1) model at 2^{nd} test setup, but it does not improve the auto correlation tendencies. Notice that the Durbin-Watson test is not robust when the correlation is not AR(1) model(for example: AR(2)). So we need more data to confirm its model at 2^{nd} test setup.

(3)The variability of residual

We use the Levene test(*Levene*(1960)) to check the variability of residuals. The result is shown as **Table 4**, and the critical value= t(0.025,79)=2.285. The p-values of 2^{nd} and 6^{th} test setups are significant. We consider using the weighted least square to modify, and the coefficients are shown as **Table 4**. But the WLS models don't improve too much, so we also need more data to check models at 2^{nd} and 6^{th} test setups.

Table 4 DW-statistic and coefficients of WLS

	DW	t_L^*	p-value	β_0'	β_1'	β_2'
BD(hor,h1=1M,h2=1-4M)	1.382	1.686	0.09574	83.735	-36.34	14.102
BD(hor,h1=2M,h2=1-4M)	0.75	2.298	0.0242	78.783	-36.457	15.62
BD(ver,h1=1M,h2=1-4M)	1.105	2.147	0.03486	37.937	-15.538	-13.444
BD(ver,h1=1.5M,h2=1-4M)	1.239	1.944	0.05546	35.47	-14.061	-11.567
DP(hor,h1=2M,h2=1-4M)	1.21	-1.758	0.08262	84.02	-39.591	17.963
DP(ver,h1=2.75M,h2=2.75-4M)	1.306	2.473	0.01555	31.602	-9.928	-13.088

Based on our analysis, the change point models found by piecewise regression are better than polynomial regression models, especially in the residual diagnosis. We also want to know that the difference in the values fitted by change point models, original measurements, and ideal values.(see **Figure 3.8**)

In this section, we use two kinds of change point models: one is discontinuous model and the other is continuous model. Even if the change point models are more suitable than polynomial regression models, we have to be careful when using them. There are still some trends we cannot find in residual diagnosing, especially when the polarization is vertical.

3.4 The estimation of error in models

In this section, we focus on the continuous broken line change point models. After analyzing the measurements fitted by change point models, we want to know that the difference between these fitted values and ideal values. The formula deduces estimation of predict values and bias of ideal values, it is shown as below:

$\sum_i (y_i - m_i)^2$
$= \sum_i (y_i - Ey_i + Ey_i - m_i)^2 = \sum_i (y_i - Ey_i)^2 + \sum_i (Ey_i - m_i)^2$

$E[\sum_i (y_i - m_i)^2]$
$= E[\sum_i (y_i - Ey_i)^2] + E[\sum_i (Ey_i - m_i)^2]$
$= \sum_i [E(y_i - Ey_i)^2] + \sum_i [E(Ey_i - m_i)^2]$
$= \sum_i (Var(y_i)) + \sum_i [E(Ey_i - m_i)^2]$ ——————— (1)
and

$\sum_i (\widehat{y}_i - m_i)^2$
$= \sum_i (\widehat{y}_i - Ey_i + Ey_i - m_i)^2 = \sum_i (\widehat{y}_i - Ey_i)^2 + \sum_i (Ey_i - m_i)^2$

$E[\sum_i (\widehat{y}_i - m_i)^2]$
$= \sum_i [E(\widehat{y}_i - Ey_i)^2] + \sum_i [E(Ey_i - m_i)^2]$
$= \sum_i (Var(\widehat{y}_i)) + \sum_i [E(Ey_i - m_i)^2]$
$= p\widehat{\sigma}^2 + \sum_i [E(Ey_i - m_i)^2]$($p$ is the number of coefficients)
then

$\sum_i [E(Ey_i - m_i)^2] = E[\sum_i (\widehat{y}_i - m_i)^2] - p\widehat{\sigma}^2$
and put it into (1), so we can get that

$E[\sum_i (y_i - m_i)^2] = n\widehat{\sigma}^2 + E[\sum_i (\widehat{y}_i - m_i)^2] - p\widehat{\sigma}^2$
$= (n-p)\widehat{\sigma}^2 + E[\sum_i (\widehat{y}_i - m_i)^2]$ (here $\widehat{\sigma}^2 = \frac{1}{n-p} \sum_i (y_i - \widehat{y}_i)^2$)

By the results, we can use the models we fitted to estimate the difference between measurements and ideal values.

3.5 Comparison with other models

In previous sections, because we know the difference between low frequencies and high frequencies innately, we consider the change point model methods, and divide the original data into two groups to analyze. Therefore, it is reasonable for us to use change point models to analyze. When analyzing the original data, we find not only the scatter plots but also residual plots show the form of sawtooth on the certain test setup. Therefore, we think that it may be the periodic relation of data. In this section, we add the Fourier polynomial into model to find a better model.

1. Fourier model

$$y_t = \theta_0 + \Sigma_{r=1}^{k}[\theta_r sin(2\pi r \cdot \frac{t}{log_{10}(1000)}) + \varphi_r cos(2\pi r \cdot \frac{t}{log_{10}(1000)})] \tag{3.5}$$

2. Linear + Fourier model

$$y_t = \alpha_0 + \alpha_1 t + \Sigma_{r=1}^{k}[\theta_r sin(2\pi r \cdot \frac{t}{log_{10}(1000)}) + \varphi_r cos(2\pi r \cdot \frac{t}{log_{10}(1000)})] \tag{3.6}$$

3. Change point + Fourier model

$$y_t = \alpha_0 + \alpha_1 t + \alpha_2 (t - t_m) \cdot I_{t>t_m} + \Sigma_{r=1}^{k}[\theta_r sin(2\pi r \cdot \frac{t}{log_{10}(1000)}) + \varphi_r cos(2\pi r \cdot \frac{t}{log_{10}(1000)})] \tag{3.7}$$

4. Square change point model

According to Section 3.4, when the polarization is horizontal, we are satisfied with the linear change point models. However, the residuals of model reveal some tendencies when the polarization is vertical. So we add the square polynomial into change point models when the polarization is vertical:

$$Y(t) = \beta_0 + \beta_1 t + \beta_2 (t - t_m) I_{[t>t_m]} + \beta_3 (t - t_m)^2 I_{[t>t_m]} + \varepsilon_t \tag{3.8}$$

and the coefficients of model is as below:

Table 5 Coefficients of square change point model (3.8)

	β_0	β_1	β_2	β_3	c.p
BD(ver,h1=1M,h2=1-4M)	36.498	-14.799	-29.388	41.257	400MHz
BD(ver,h1=1.5M,h2=1-4M)	34.142	-13.355	-20.546	15.102	250MHz
DP(ver,h1=2.75M,h2=2.75-4M)	31.767	-9.915	-14.774	1.74	80MHz

For the data of vertical polarization, the square change point models are still better than the linear change point models.

4. Comparison

In the preceding research, we fitted the statistical model in order to predict the measurements at different frequencies in different test setups. Moreover, what interest us is the difference between these two kinds of antenna in measurement. So, in this section, we discuss the measuring difference between these two kinds of antenna under the same test setup, and the comparison of Broadband antenna's test setups.

4.1 Comparison between Broadband antenna and Dipole antenna

We choose the 2^{nd} and 5^{th} test setups to analyze. These two test setups are Broadband antenna and Dipole antenna which are on (hor,h1=1M,h2=1-4M). **Figure 4.1** are scatter plots of measurements and *dev* values for these two data. It seems that there is no difference in the plots of measurement, but in the plots of *dev*, *dev* values of Broadband antenna seems to be bigger than Dipole antenna almost all frequencies. It means that the measurements of Broadband antenna are bigger than Dipole antenna, especially at high frequencies, and the *dev* values of Dipole antenna disperse more disorderly than Broadband antenna does. It seems that the values measured by Dipole antenna are more different from ideal values. And we consider this model:

$$y_t = \beta_0 + \beta_1 t + \beta_2 (t - t_m)_+ + \beta_3 \cdot I + \beta_4 t \cdot I + \varepsilon_t \tag{4.1}$$

here $y_t = measurements$, $t = log_{10}(freq)$, I is a indicator, and I is equal to 1, iff antenna is Broadband; I is equal to 0, iff antenna is Dipole. And t_M in this model is the change point of Broadband and Dipole antenna(80 MHz). The model's information are shown as **Table 6**:

Table 6 Coefficients of model (4.1)

	coefficient	t^*	p-value
β_0	82.357	82.303	0.000
β_1	-38.625	-69.885	0.000
β_2	16.749	25.812	0.000
β_3	-2.117	-3.671	0.000
β_4	1.298	5.068	0.000

According to Table 6, we can obtain that the effect of antenna is obvious(p-value= 0.0000). It means that these two kinds of antenna have obvious difference in measurement. And this will prove the hypothesis we considered before. We use Levene test(*Levene*(1960)) to check the variability of errors because of we put two different kinds of antenna data into the same model:

$$| t_L^* | = 1.3649 < t(0.975, 160) = 2.263,$$

so the result is that the variability of error for these two kinds of antenna are not different.

4.2 Comparison of Broadband antenna

In this experiment, we want to know whether difference exist in the polarization or not. So we choose the 1^{st} and 3^{rd} test setups to compare. From the scatter plots shown as **Figure 4.1**, we can obtain some information from them. It seems that there are the obvious differences between these two data. We consider this model:

(4.2) $$y_t = \beta_0 + \beta_1 t + \beta_2 I + \beta_3 t \cdot I + \varepsilon_t$$

here $y_t = measurements$, $t = log_{10}(freq)$, I is a indicator, and I is equal to 1, iff the polarization is horizontal; I is equal to 0, iff the polarization is vertical. The model information are shown as **Table 7**:

Table 7 Coefficients of model (4.2)

	coefficient	t^*	p-value
β_0	42.896	40.650	0.000
β_1	-18.142	-38.727	0.000
β_2	26.469	17.736	0.000
β_3	-10.506	-15.858	0.000

As to Table 7, the difference in measurements will come out when the polarization of Broadband antenna changes. Moreover, we also want to know that the difference in measurements on Broadband antenna when the height of transmit antenna changes. So we choose the data of 1^{st} and 2^{nd} test setups for horizontal polarization, and the data of 3^{rd} and 4^{th} test setups for vertical polarization. For the scatter plots shown as in **Figure 4.2**. It seems to have measuring difference on different height of transmit antenna for horizontal polarization. But the difference in measurement are not so obvious when the polarization is vertical. We consider the same model as below, and the model information are shown as **Table 8** and **Table 9**:

Table 8 Coefficients of model(horizontal)

	coefficient	t^*	p-value
β_0	56.653	56.419	0.000
β_1	-23.848	-53.498	0.000
β_2	12.711	8.951	0.000
β_3	-4.800	-7.614	0.000

Table 9 Coefficients of model(vertical)

	coefficient	t^*	p-value
β_0	43.420	42.429	0.000
β_1	-18.259	-40.191	0.000
β_2	-0.524	-0.362	0.718
β_3	0.117	0.182	0.856

As what we think, the effect of coefficients in model prove that there are obvious difference in measurements when the height of transmit antenna changes for horizontal polarization. But for vertical polarization, there have no difference(p-value= 0.718, 0.856).

As the result, there is difference between the traditional Dipole and Broadband antenna at high frequencies when measuring. The sum of square *dev* values are shown as **Table 10**. If we regard ideal values as the standard of antenna measurements, it is shown that Broadband antenna is more exact than Dipole antenna(the $\Sigma(dev)^2$ of 2^{nd} and 5^{th} test setups are 72.95 and 97.67). And when the Broadband antenna are measuring, there have difference of different polarizations. It is shown that the horizontal polarization is more exact than the vertical polarization when the Broadband antenna is measuring. We can get some ideals from previous scatter plots. In this section, we offer more powerful evidence to support this statement. When the polarization is vertical, the difference of height of transmit antenna could probably cause different results. However, the variability of measurement would be extreme when the polarization is vertical.

Table 10 The value of $\Sigma(dev)^2$ and $\Sigma(predict - ideal)^2$

	c.p	$\Sigma(dev)^2$	$\Sigma(predict - ideal)^2$
BD(hor,h1=1M,h2=1-4M)	160MHz	109.82	56.19
BD(hor,h1=2M,h2=1-4M)	80MHz	72.95	36.6
BD(ver,h1=1M,h2=1-4M)	400MHz	233.08	88.47
BD(ver,h1=1.5M,h2=1-4M)	250MHz	236.58	100.95
DP(hor,h1=2M,h2=1-4M)	80MHz	97.67	46.32
DP(ver,h1=2.75M,h2=2.75-4M)	60MHz	175.48	39.72

5. Conclusion

In electromagnetism, the theoretical values of NSA were set up by a ideal Dipole antenna with infinite plane, and superconductor ground. These situations do not exist in real site. For the most part, reasons for measuring errors are defect of sites, inaccurate character of antenna, and the errors of instruments in experiment.

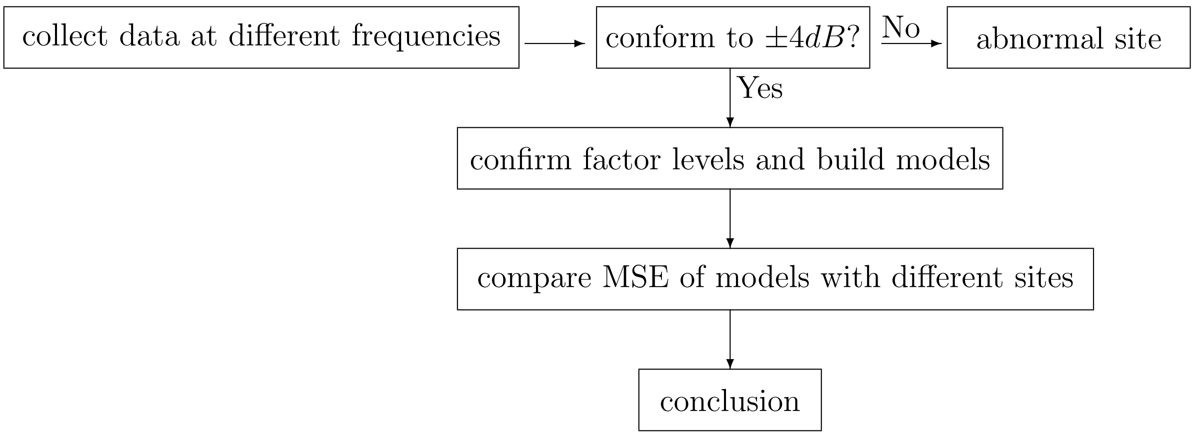
In the study we have mentioned before, we try to fit some different statistical models for NSA's measurements. We hope that we can find the fittest model to analyze the accuracy of antenna measurements, and contrast the measuring difference between different antenna. Moreover, after we find the fitted model, we do the following for model validation before using it(*Ronald* (1977)).

a. Check the coefficients of model and compare predict values with the results of theory or experience.

b. Collect new data to confirm the prediction from the fitted models.

c. If we cannot get new data, then divide the original data into two parts, use one part data to fit models and estimate coefficients of model, and use the other to confirm the models.

In our research, we cannot obtain a full factor level model because of the lack of data(only six test setups). If we can collect more data, that will be benefit to the model fitting and confirm. Moreover, the *dev* values for both sites conforming within $\pm 4dB$, we put forward a testing procedure of comparing two normal sites at the same test setup by the model we fitted:



Based on results from this work, it would be of interest to collect more data and information about EMI to build a more suitable model under all test setups in the future. And we can use it to predict and compare the measurements of antenna at different frequencies in different test setups. Although the ideal value is calculated under a very perfect condition, this condition does not exist in our daily life. The measurements of EMI will be different in different environment and physical conditions. It is not reasonable to use the unrealistic ideal value to be the standard value when measuring. In the future, we will try to establish standard operation procedures to validate the uncertainties on the accuracy of antenna measurements.

References

Akira.S. (1990). Fomulation of Normalized Site Attenuation in Terms of Antenna Impedances. *IEEE Transactions on Electromagnetic Compatibility* **32**, no.4, 257-263.

Akira.S. (1992). Correction Factors for Normalized Site Attenuation. *IEEE Transactions on Electromagnetic Compatibility* **34**, no.4, 461-470.

Cater, R.L. and Blight, B.J.N. (1981). A bayesian change-point problem with an application to the prediction and detection of ovulation in women. *Biometrics* **37**, 743-751.

Chernoff, H. and Zacks, S. (1964). Estimation the current mean of a normal distribution which is subjected to changes in time. *Annals Math. Statist* **35**, 999-1018.

Chung, B.L. (1995). Estimating the number of change points in a sequence of independent normal random variables. *Statistics and Probability Letters* **25**, 241-248.

Freeman, J.M. (1984). Two-phase regression and goodness of fit. *Communication in Statistics, Theorey and Methods* **13**, 1321-1334.

Freeman, J.M. (1986). An unknown change point and goodness of fit. *The Statistician* **35**, 335-344.

Hinkley, D.V. (1969). Inference about the intersection in two-phase regression. *Biometrika* **56**, 495-504.

Hinkley, D.V. and Hinkley, E.A. (1970). Inference about the change point in a sequence of random variables. *Biometrika* **57**, 1-17.

Hinkley, D.V. (1971). Inference about the change point from cumulative sum tests. *Biometrika* **58**, 509-523.

Hinkley, D.V. (1972). Time ordered classification. *Biometrika* **59**, 509-523.

Hudson, D.J. (1966). Fitting segmented curves whose join points have to estimated. *J. Amer. Statist. Assoc.* **61**, 1097-1129.

Krishnaiah, P.R. and Miao, B.Q. (1988). Review About Estimation of Change Points. *Handbook of Statistics* **7**, 374-402.

Levene, H.(1960). Robust Tests for Equality of Variances. *Contributions to Probability and Statistics*, ed. I. Olkin. Palo Alto, Calif.: Stanford University Press, 278-292.

Neter, J., Kutner, K.H., Nachtshem, C.J. and Wasserman, W. (1999). *Applied Linear Regression Models*, third edition, McGraw-Hill Int, N.Y.

Page, E.S. (1954). Continuous inspection schemes. *Biometrika* **41**, 100-114.

Page, E.S. (1955). A test for a change in a parameter occurring at an unknown point. *Biometrika* **42**, 523-527.

Page, E.S. (1957). On problems in which a change in parameter occurs at an unknown point. *Biometrika* **44**, 248-252.

Quandt, R.E. (1958). The estimation of the parameter of a linear regression system obeying two separate regimes. *J. Amer. Statist. Assoc.* **53**, 873-880.

Quandt, R.E. (1960). Tests of the hypotheses that a linear regression system obeys two separate regimes. *J. Amer. Statist. Assoc.* **55**, 324.

Quandt, R.E. and Ramsey, B. (1978). Estimating mixtures of normal distributions and switching regression. *J. Amer. Statist. Assoc.* **73**, 730-738.

Robison, D.E. (1964). Estimating for the points of intersection of two polynomial regressions. *J. Amer. Statist. Assoc.* **59**, 214-224.

Ronald, D.S. (1977). Validation of Regression Models: Methods and Examples. *Technometrics* **19**, 415-428.

Smith, A.F.M. (1975). A Bayesian approach to inference about a change-point in a sequence of random variables. *Biometrika* **62**, 407-416.

Table I

Table I-1 The model comparison of repeated measurements for Broadband antenna on (hor,h1=1M,h2=1-4M)

	α_0	α_1	R^2	MSE
1st	70.295	-28.816	0.976	4.442
2nd	68.296	-28.296	0.979	3.688
3rd	69.504	-28.798	0.979	3.744

Table I-2 The model comparison of repeated measurements for Broadband antenna on (hor,h1=2M,h2=1-4M)

	α_0	α_1	R^2	MSE
1st	56.703	-23.831	0.980	2.517
2nd	56.382	-23.770	0.979	2.585
3rd	56.876	-23.945	0.979	2.699

Table I-3 The model comparison of repeated measurements for Broadband antenna on (ver,h1=1M,h2=1-4M)

	α_0	α_1	R^2	MSE
1st	43.608	-18.419	0.949	3.935
2nd	42.056	-17.790	0.956	3.157
3rd	43.024	-18.218	0.962	2.781

Table I-4 The model comparison of repeated measurements for Broadband antenna on (ver,h1=1.5M,h2=1-4M)

	α_0	α_1	R^2	MSE
1st	43.619	-18.244	0.945	4.125
2nd	42.821	-18.086	0.953	3.457
3rd	43.820	-18.447	0.956	3.393

Table I-5 The model comparison of repeated measurements for Dipole antenna on (hor,h1=1M,h2=1-4M)

	α_0	α_1	R^2	MSE
1st	58.967	-25.217	0.976	3.296
2nd	58.565	-25.109	0.977	3.199
3rd	58.780	-25.112	0.978	3.026

Table I-6 The model comparison of repeated measurements for Dipole antenna on (ver,h1=2.75M,h2=2.75-4M)

	α_0	α_1	R^2	MSE
1st	48.845	-20.403	0.972	2.541
2nd	52.467	-22.156	0.973	2.939
3rd	51.757	-21.720	0.980	2.115

Table I-7 Test of model's coefficients effect

	α_0	α_1	α_2	α_3	p_2	p_3
BD(hor,h1=1M,h2=1-4M)	69.924	-28.648	-0.926	-0.752	0.088	0.164
BD(hor,h1=2M,h2=1-4M)	56.741	-23.848	-0.185	-0.078	0.670	0.858
BD(ver,h1=1M,h2=1-4M)	42.997	-18.142	-0.163	-0.141	0.740	0.774
BD(ver,h1=1.5M,h2=1-4M)	43.652	-18.259	-0.448	-0.248	0.386	0.631
DP(hor,h1=2M,h2=1-4M)	56.810	-25.146	-0.163	-0.044	0.0735	0.926
DP(ver,h1=2.75M,h2=2.75-4M)	51.105	-21.426	-0.248	-0.0037	0.572	0.993

Table I-8 Linear regression model

	α_0	α_1	R^2	MSE
BD(hor,h1=1M,h2=1-4M)	69.365	-28.648	0.977	3.933
BD(hor,h1=2M,h2=1-4M)	56.653	-23.848	0.979	2.476
BD(ver,h1=1M,h2=1-4M)	43.896	-18.142	0.955	3.144
BD(ver,h1=1.5M,h2=1-4M)	43.420	-18.259	0.951	3.512
DP(hor,h1=2M,h2=1-4M)	58.771	-25.146	0.977	3.033
DP(ver,h1=2.75M,h2=2.75-4M)	51.023	-21.426	0.974	2.531

Table I-9 Square regression model

	β_0	β_1	β_2	R^2	MSE
BD(hor,h1=1M,h2=1-4M)	113.037	-68.976	8.942	0.995	0.843
BD(hor,h1=2M,h2=1-4M)	87.113	-51.975	6.236	0.992	0.980
BD(ver,h1=1M,h2=1-4M)	16.203	6.507	-5.465	0.972	2.011
BD(ver,h1=1.5M,h2=1-4M)	11.752	10.984	-6.484	0.974	1.906
DP(hor,h1=2M,h2=1-4M)	91.539	-55.405	6.709	0.990	1.304
DP(ver,h1=2.75M,h2=2.75-4M)	40.630	-11.829	-2.128	0.976	2.385

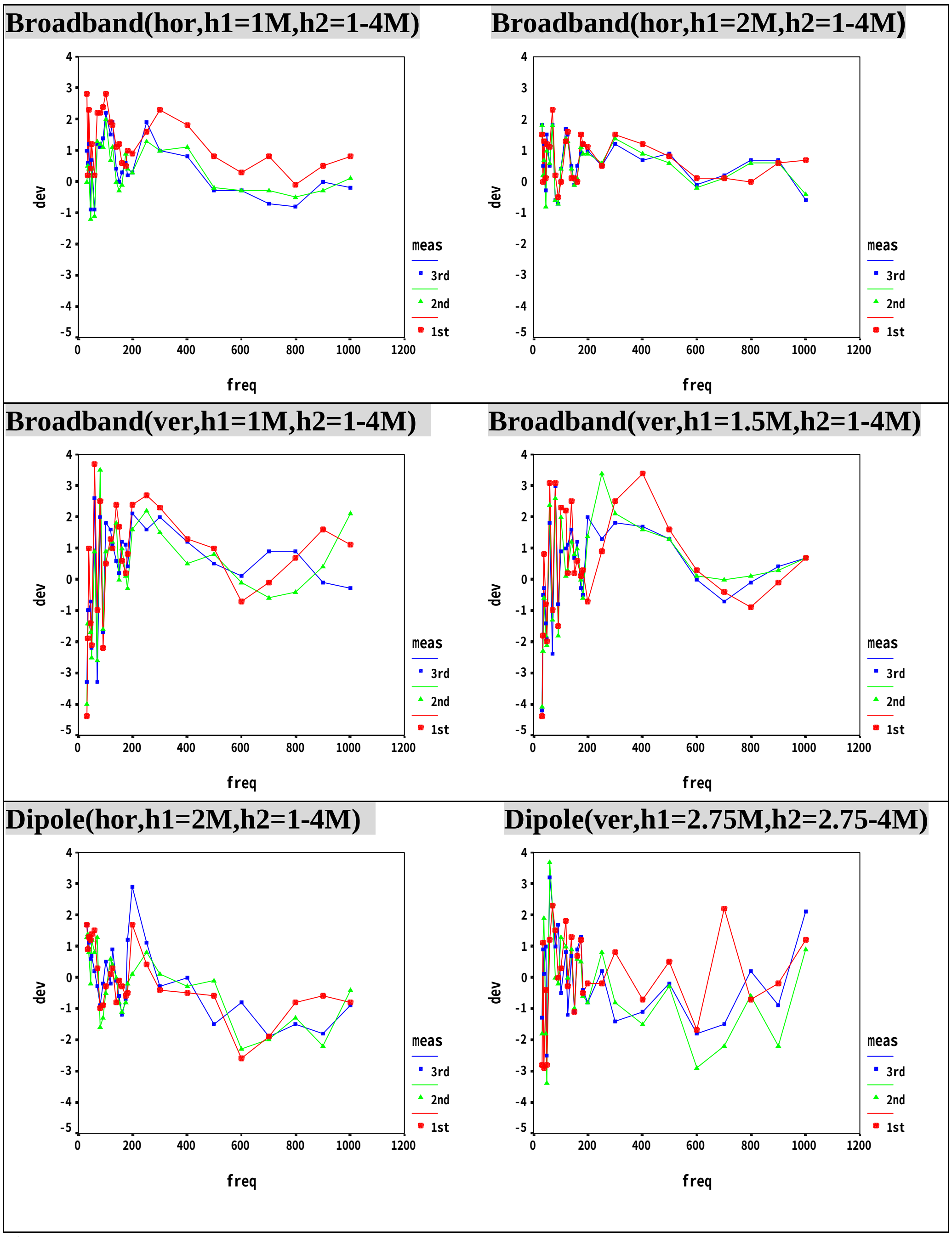


Figure 2.1 *Dev* values at different test setups

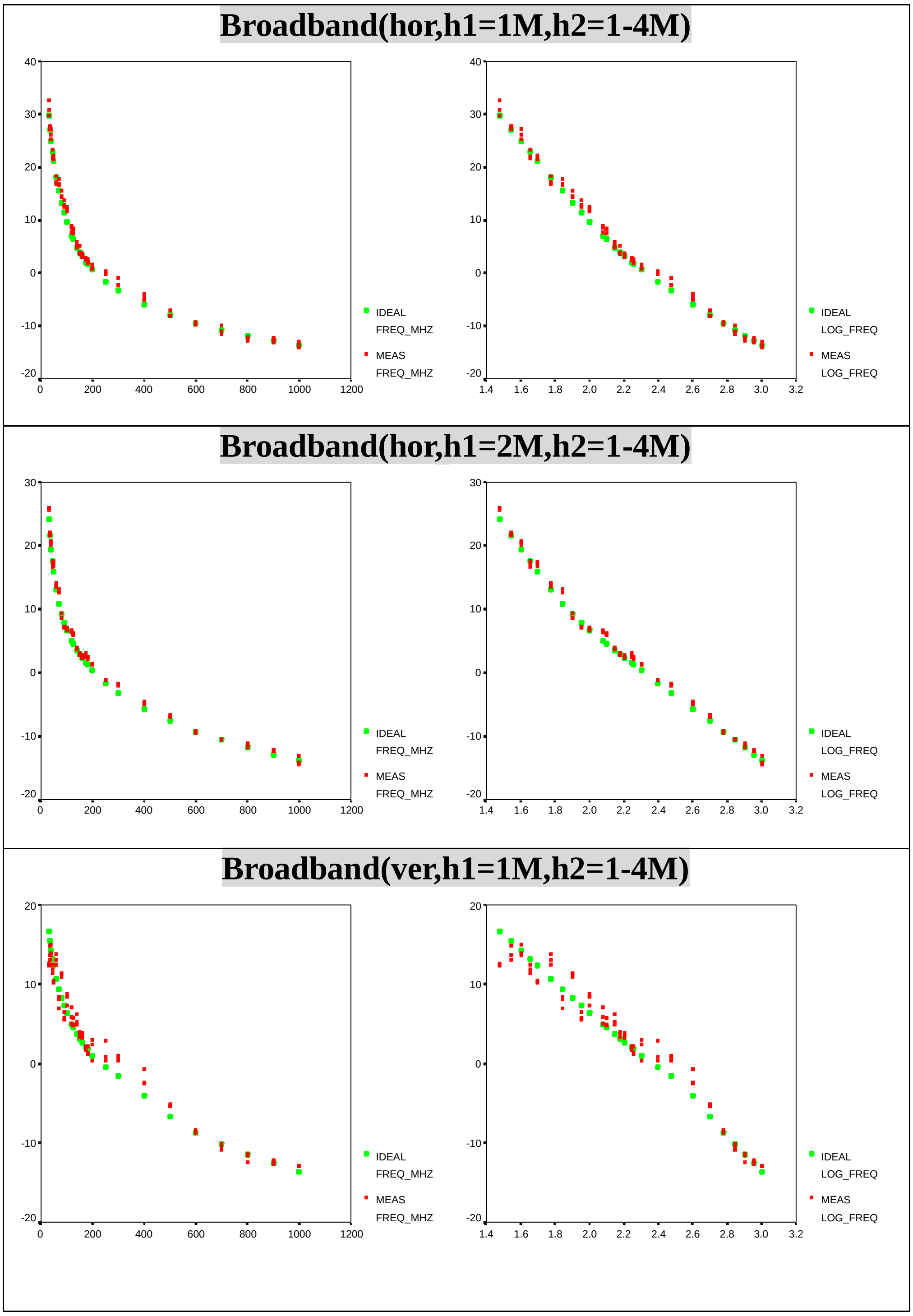


Figure 2.2(1) Measurements versus $\log_{10}(\text{freq})$

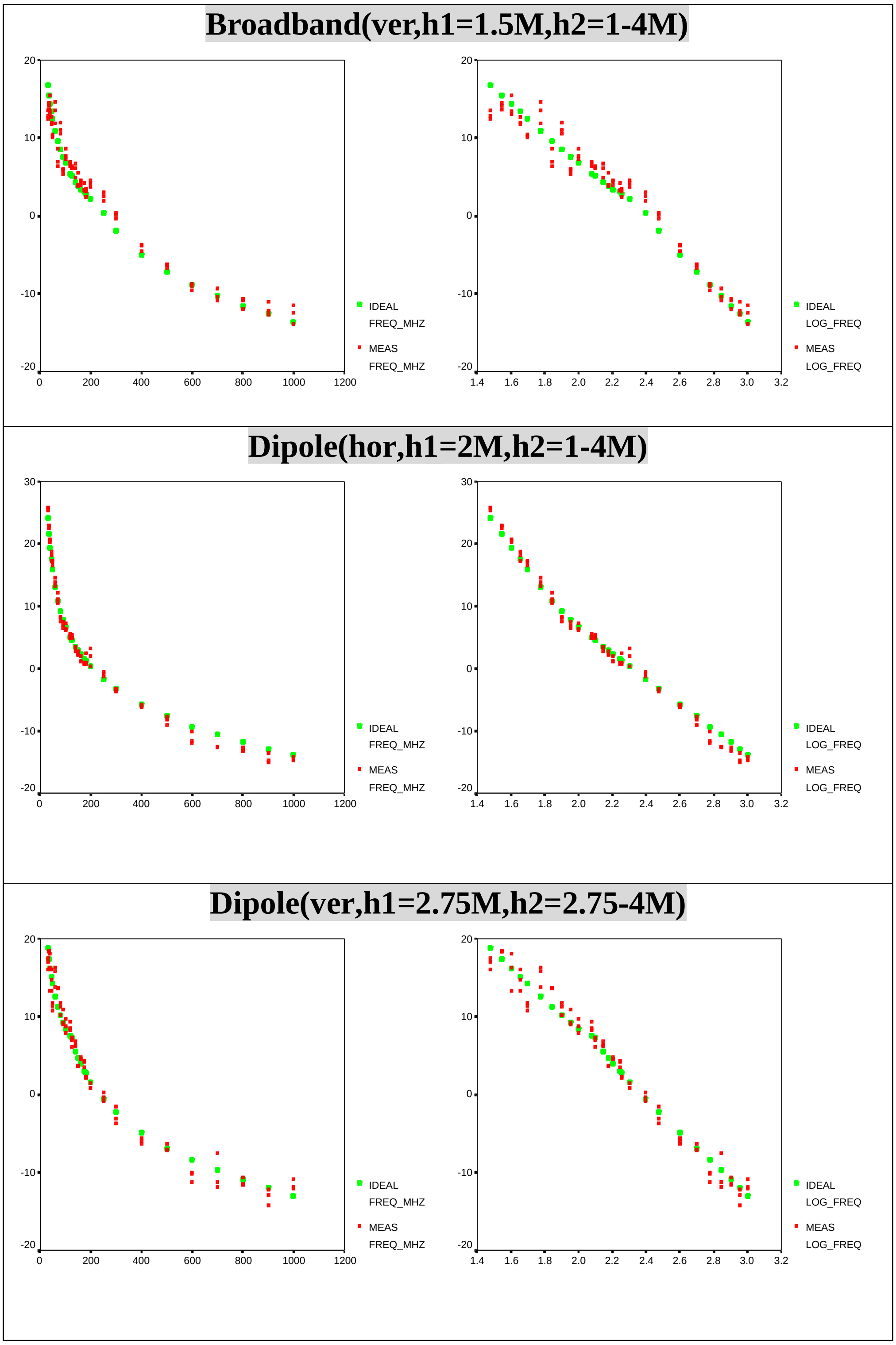


Figure 2.2(2) Measurements versus $\log_{10}(\text{freq})$

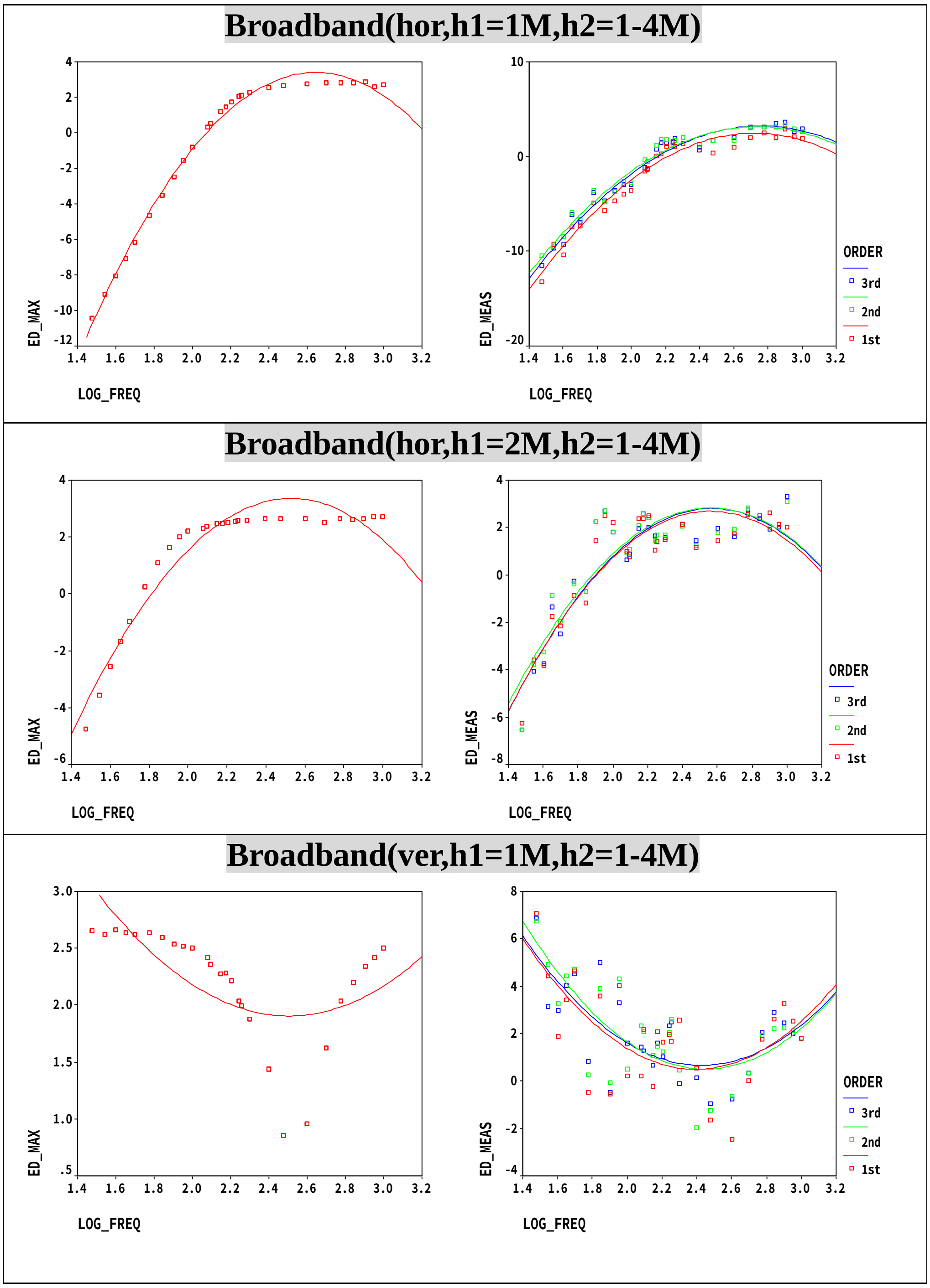


Figure 3.1(1) Graphs of $E_D^{MAX}(ideal)$ and $E_D^{MAX}(meas)$

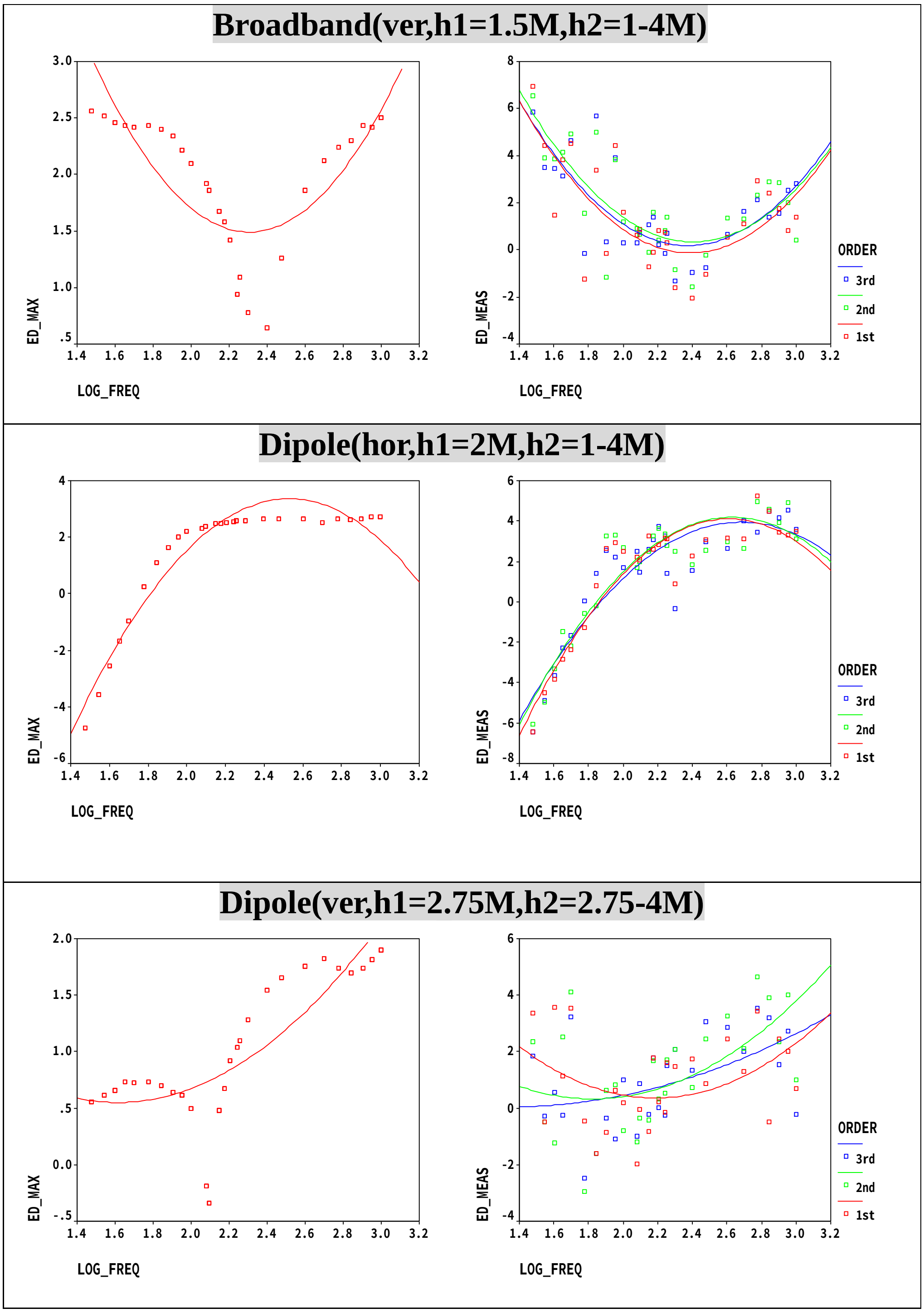


Figure 3.1(2) Graphs of $E_D^{MAX}(ideal)$ and $E_D^{MAX}(meas)$

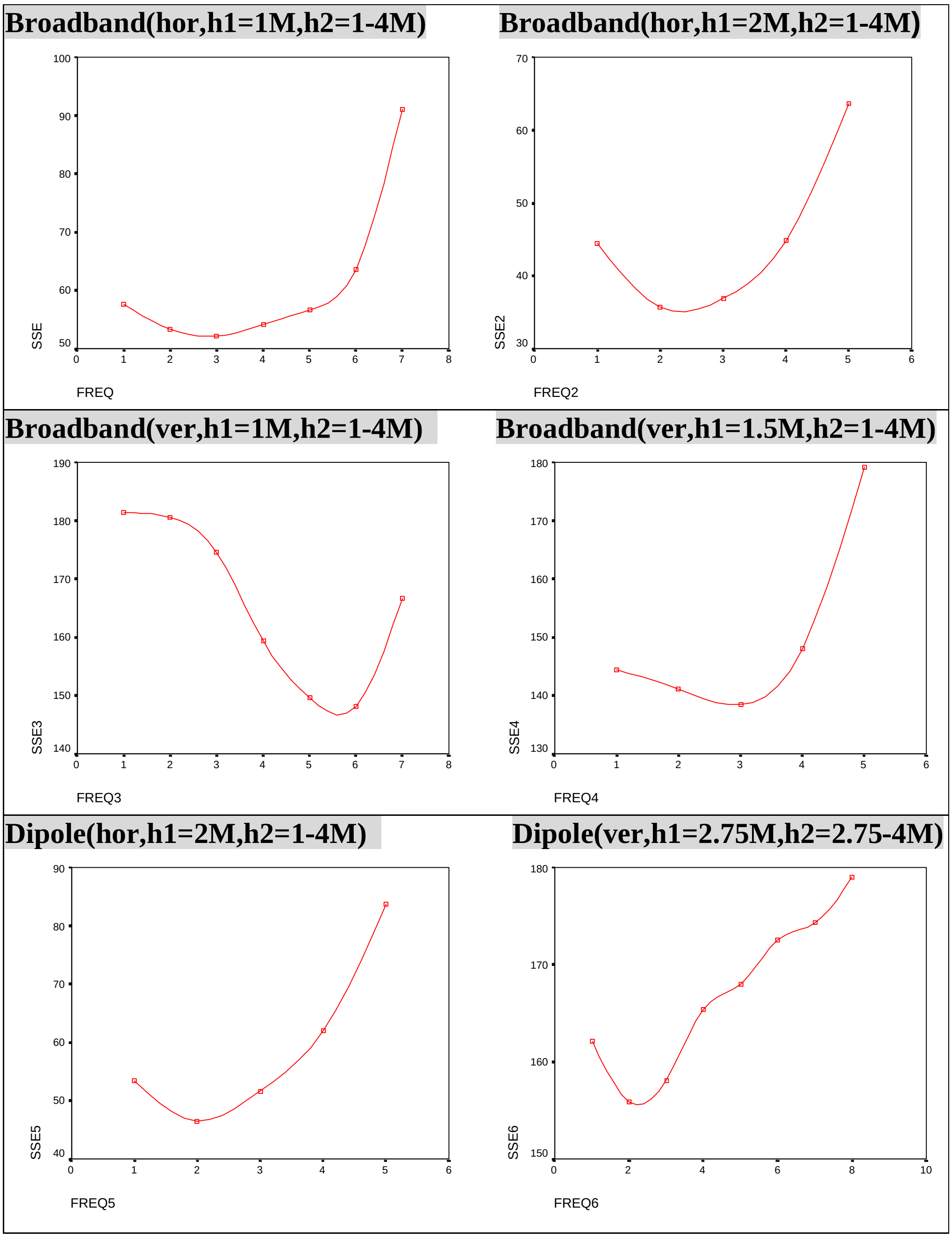


Figure 3.2 SSE of piecewise regression model

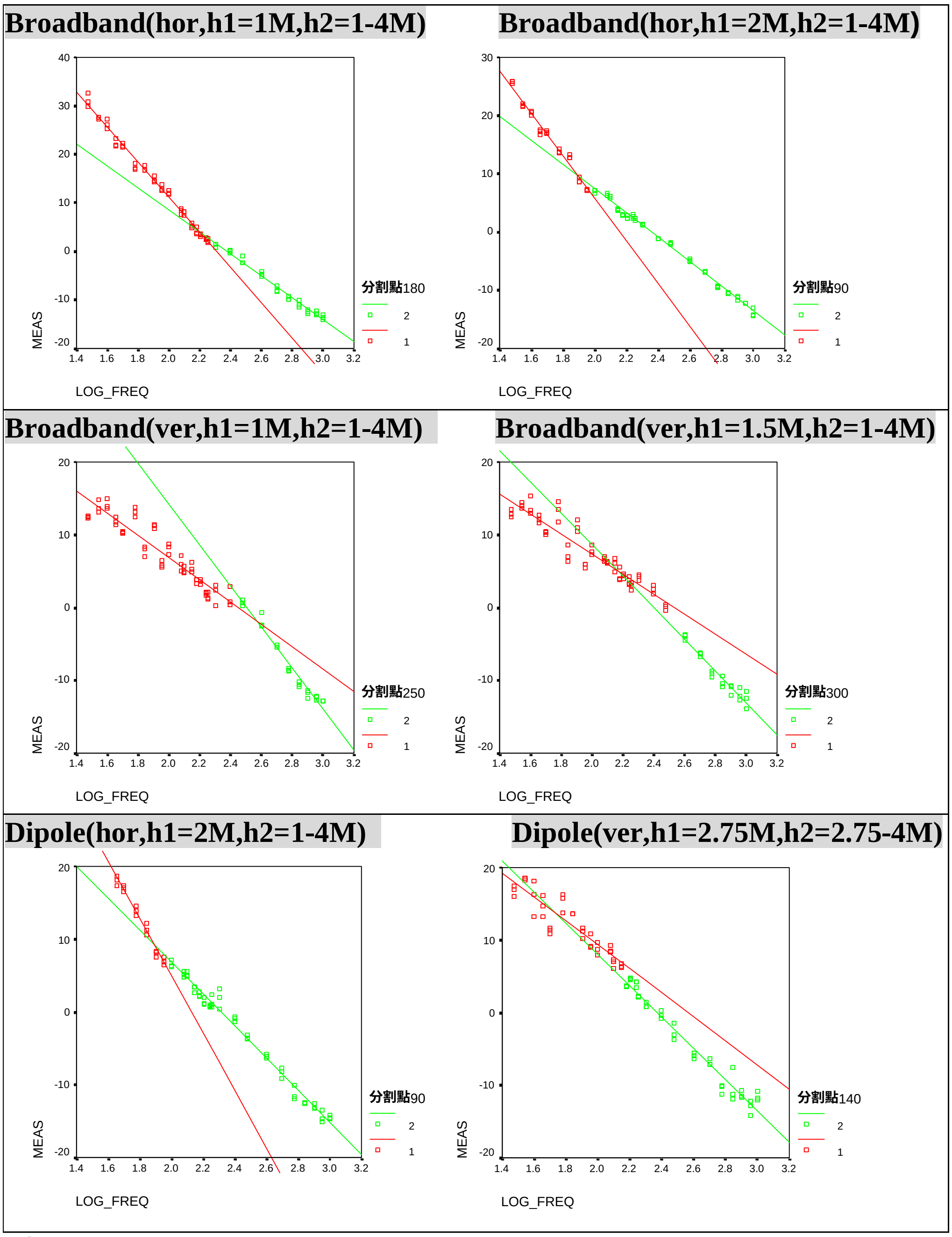


Figure 3.3 Fitted values at different test setups

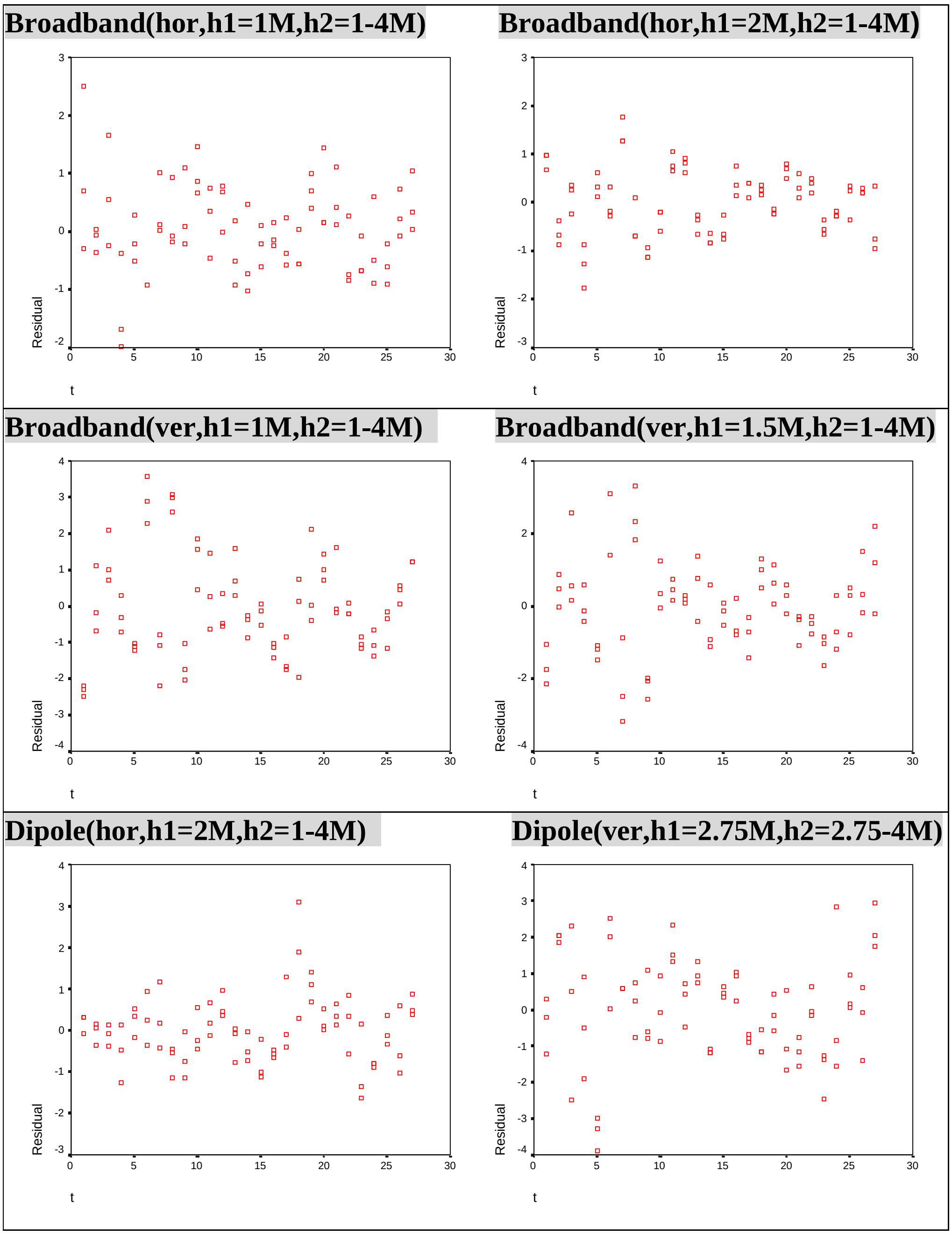


Figure 3.4 Residual plots of change point models

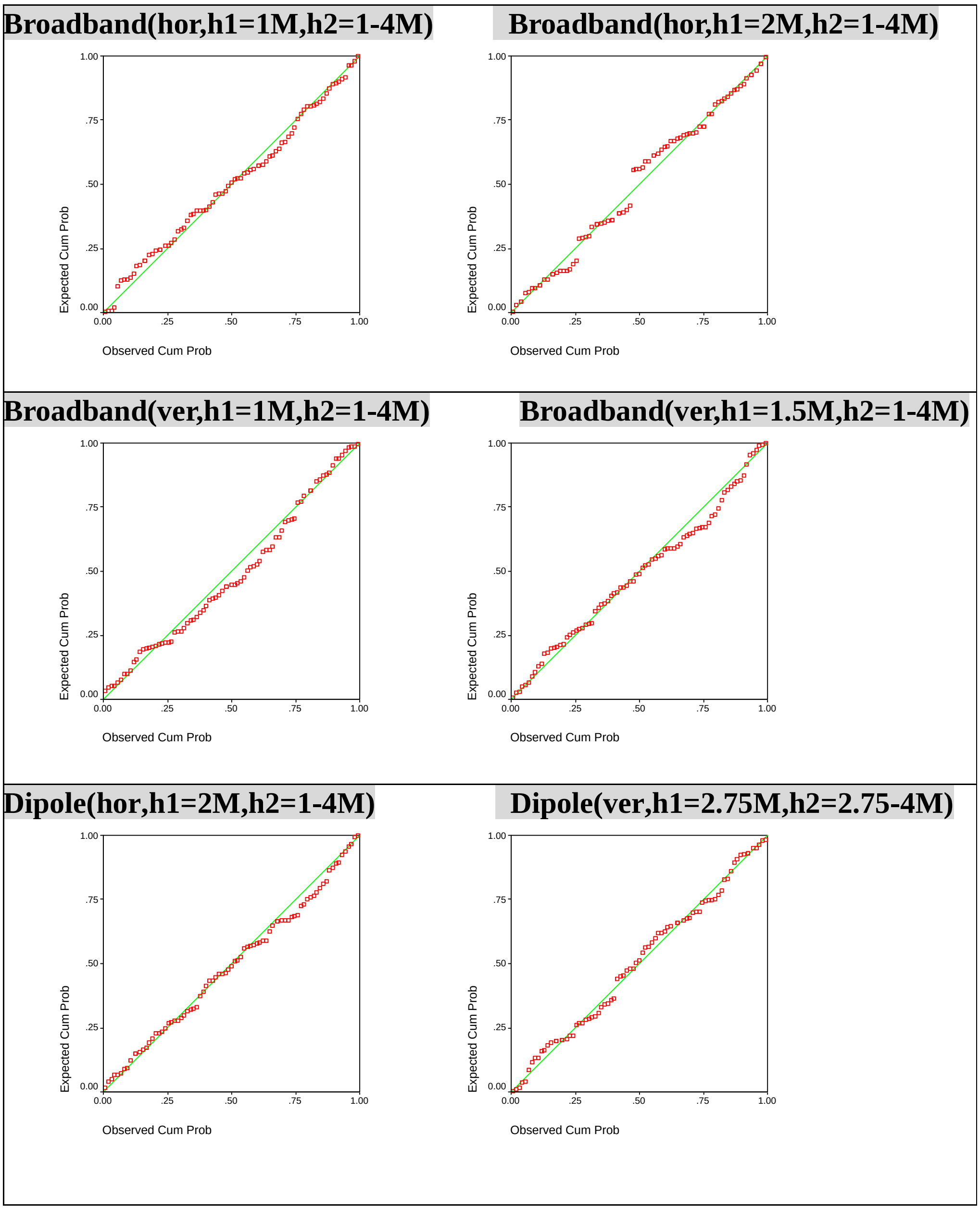


Figure 3.5 Normal P-P plots

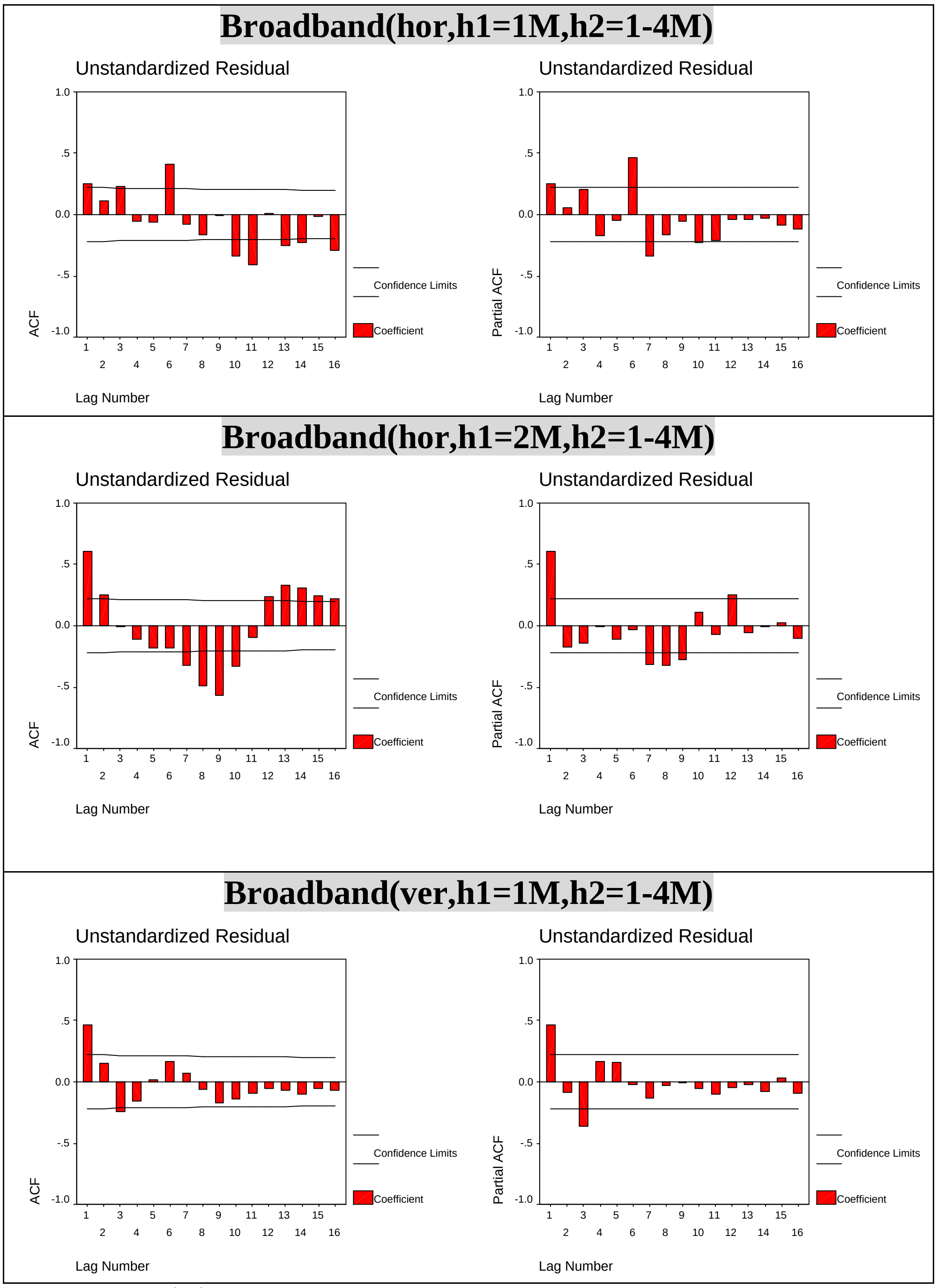


Figure 3.6(1) ACF and PACF plots of piecewise regression model

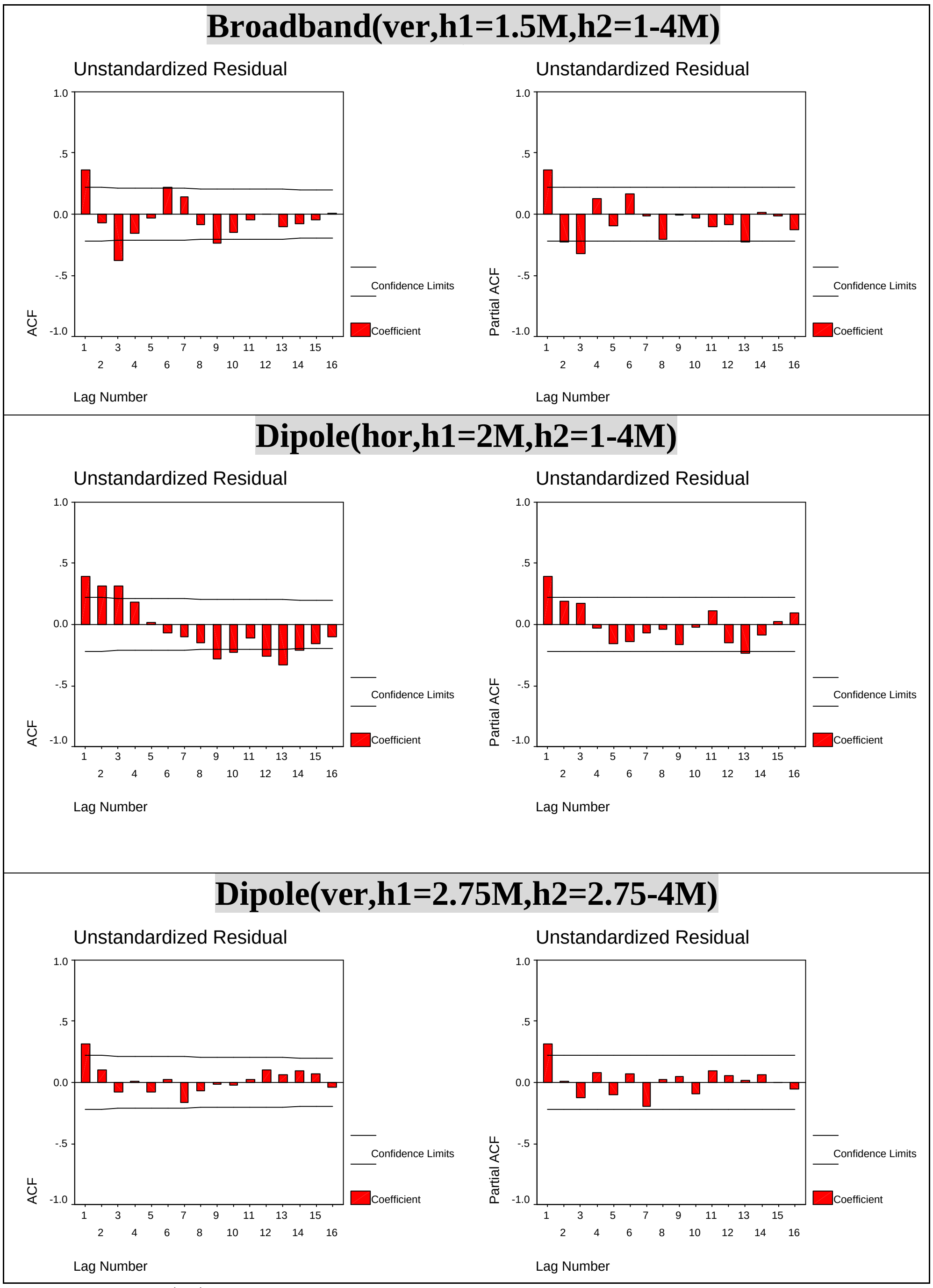


Figure 3.6(2) ACF and PACF plots of piecewise regression model

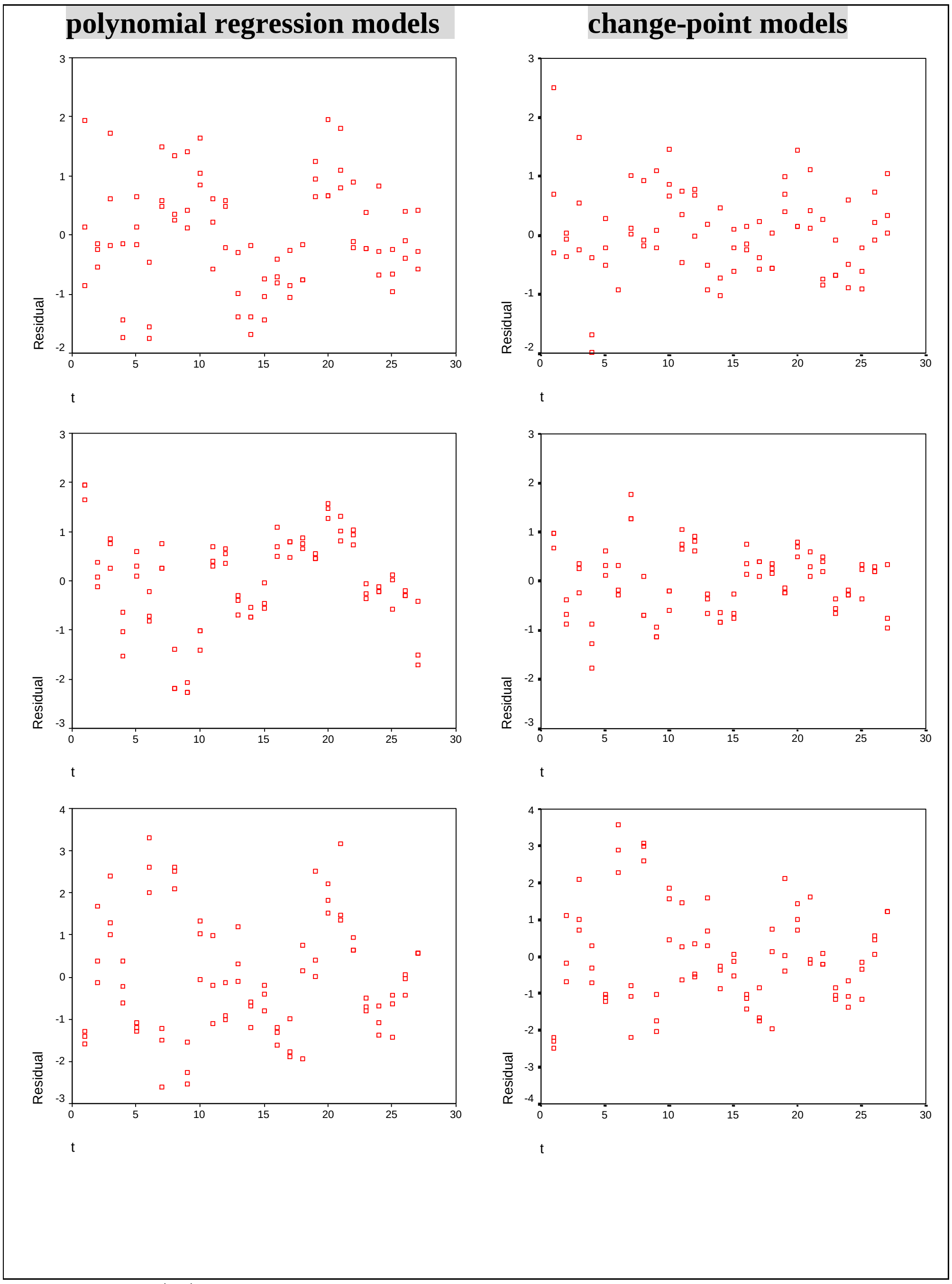


Figure 3.7(1) **Residual plots of polynomial regression models and piecewise regression models**

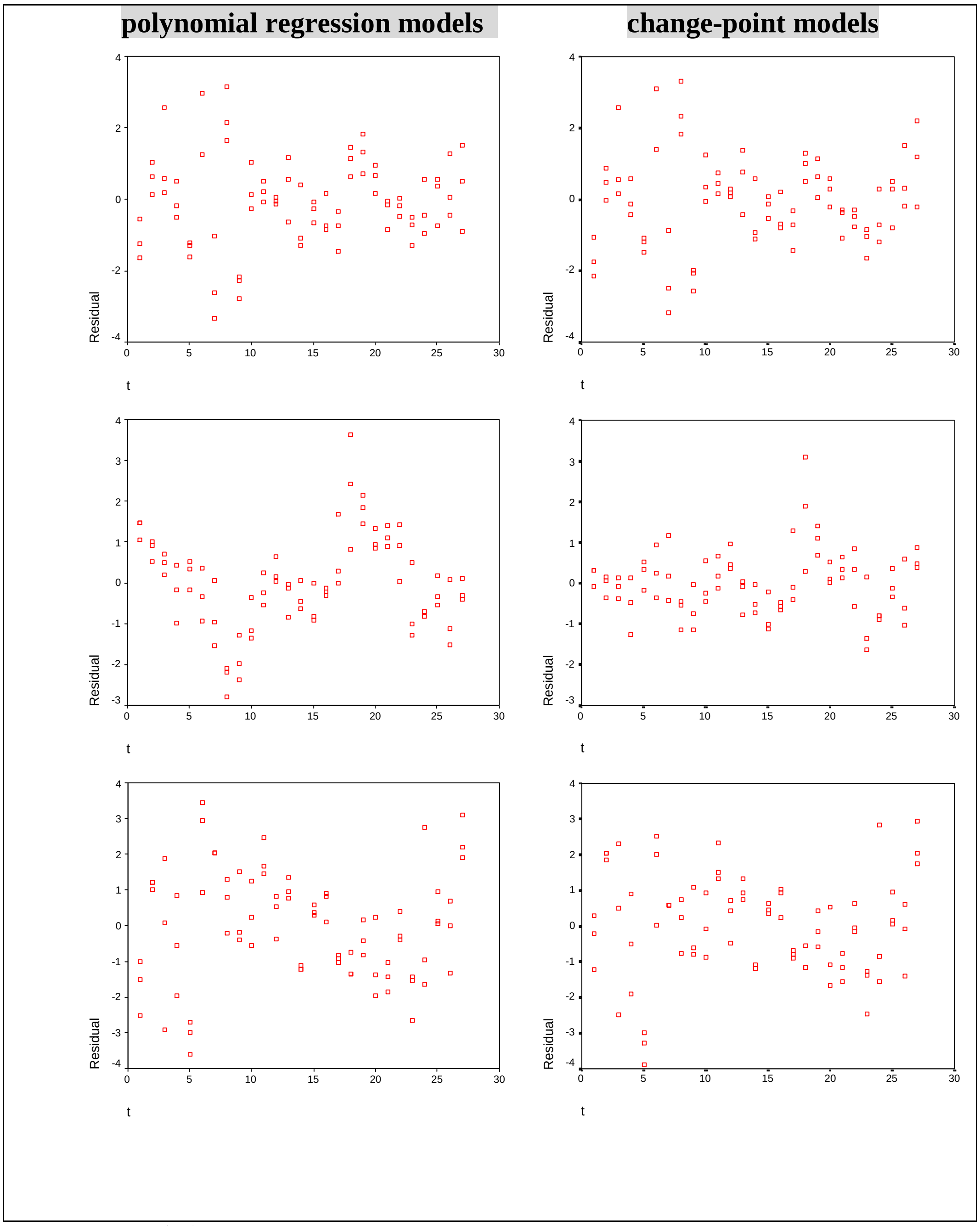


Figure 3.7(2) Residual plots of polynomial regression models and piecewise regression models

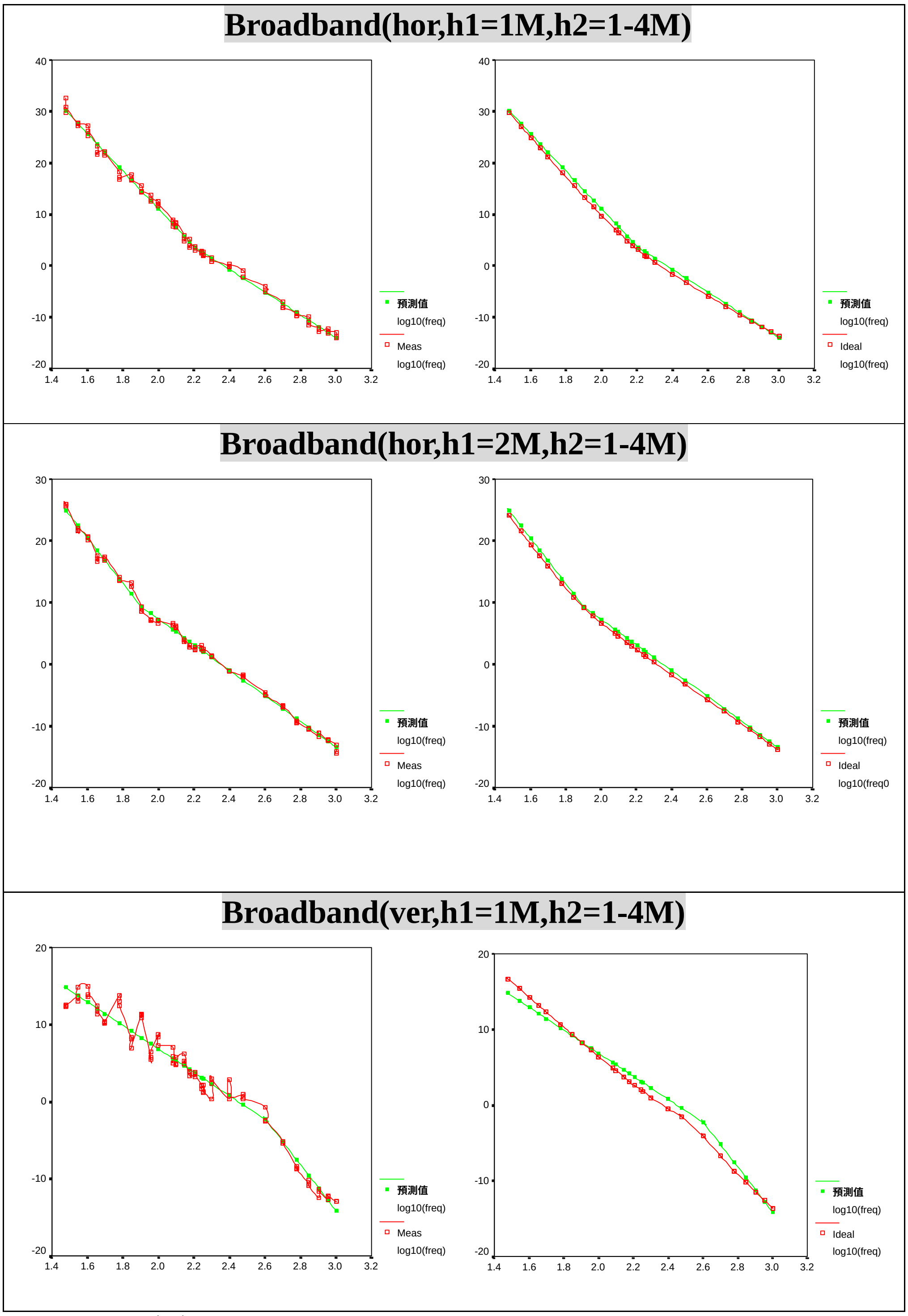


Figure 3.8(1) Fitted values of piecewise regression model, measurements, and ideal values

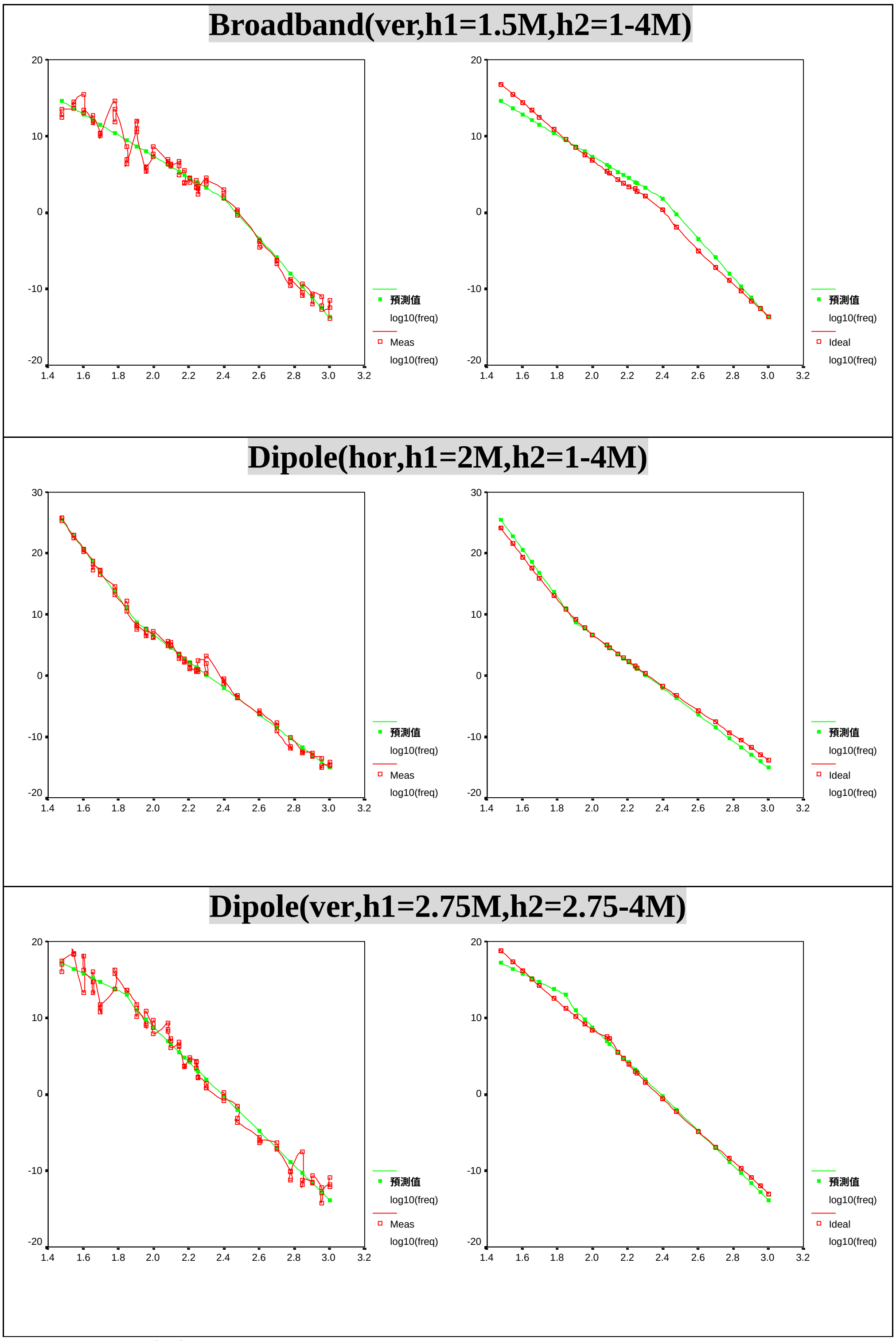


Figure 3.8(2) Fitted values of piecewise regression model, measurements, and ideal values

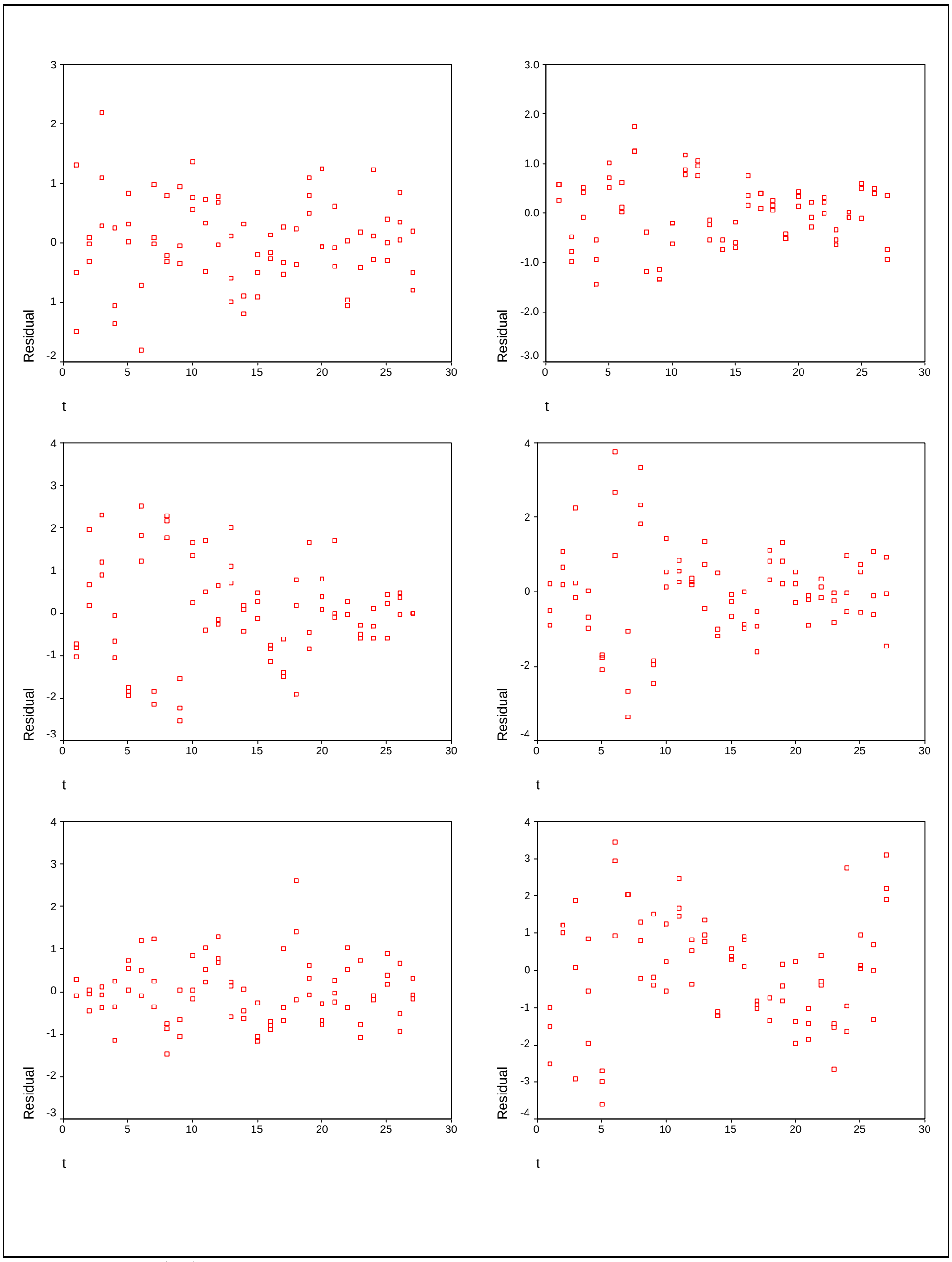


Figure 3.9(1) Residuals of Fourier Model(k=3)

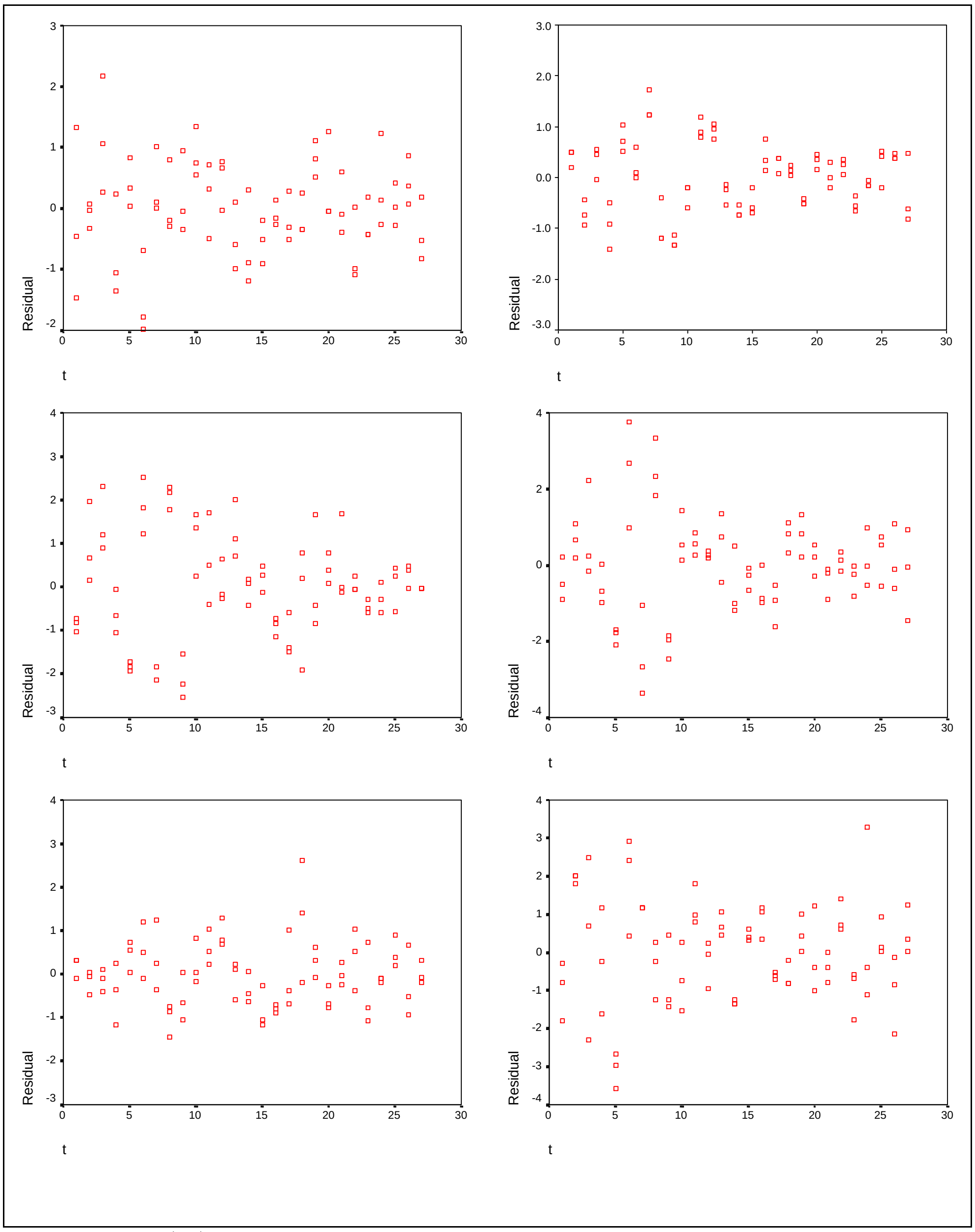


Figure 3.9(2) **Residuals of Linear + Fourier Model(*k=3*)**

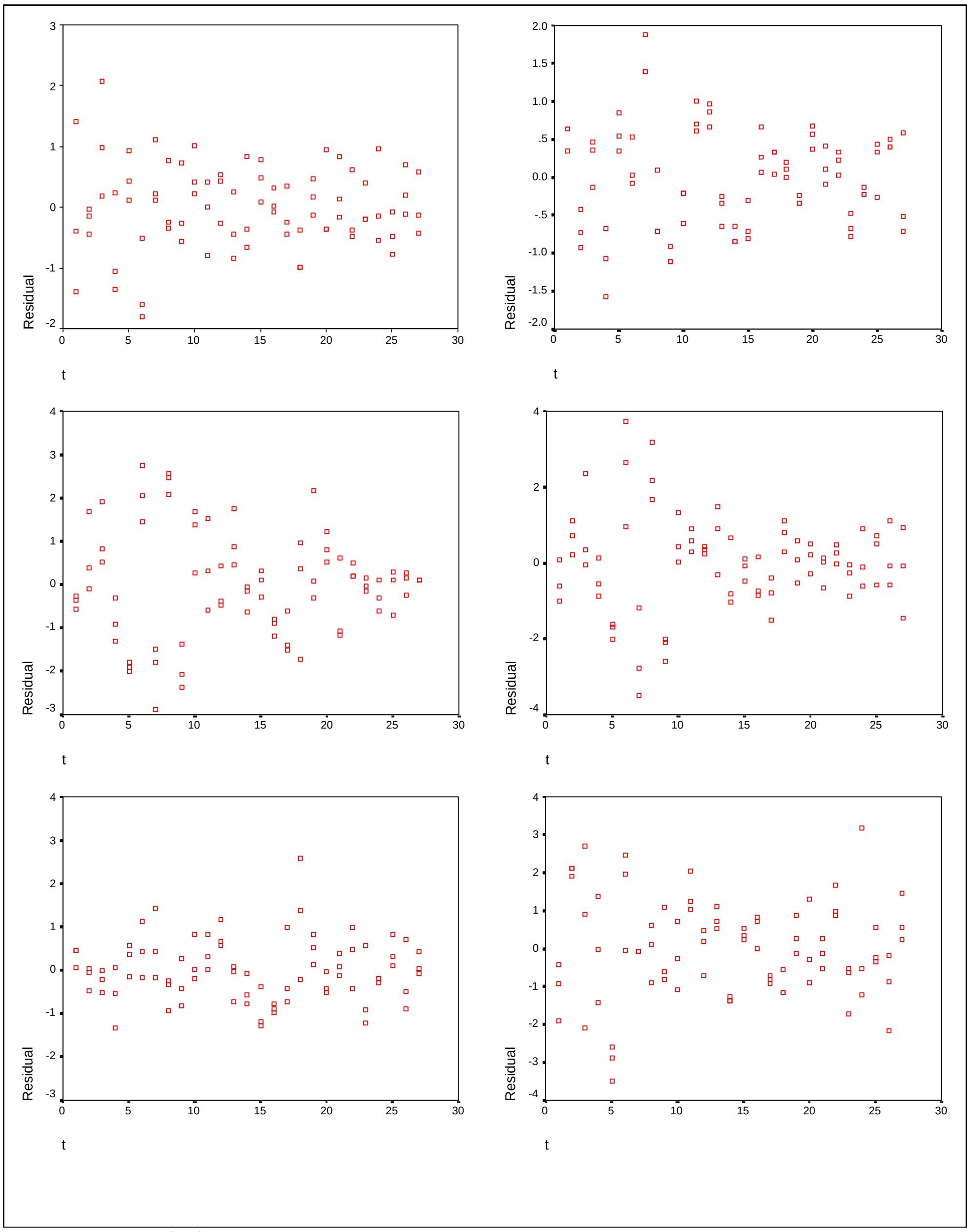
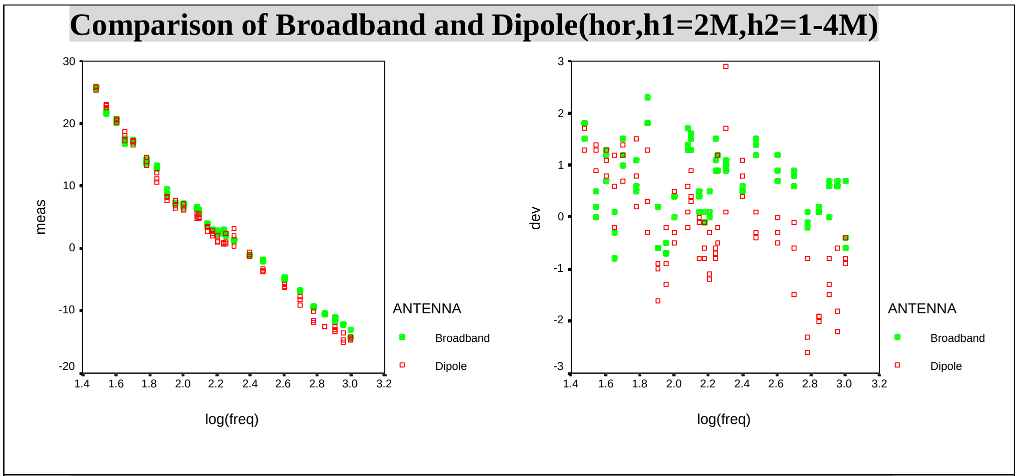


Figure 3.9(3) Residuals of Change point + Fourier Model(*k=2*)



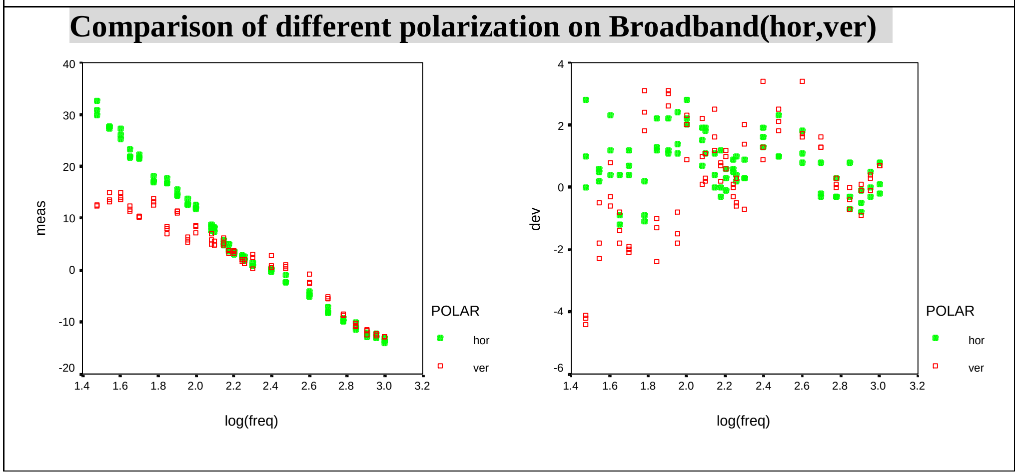


Figure 4.1 Measurements and ideal values at the same test setup

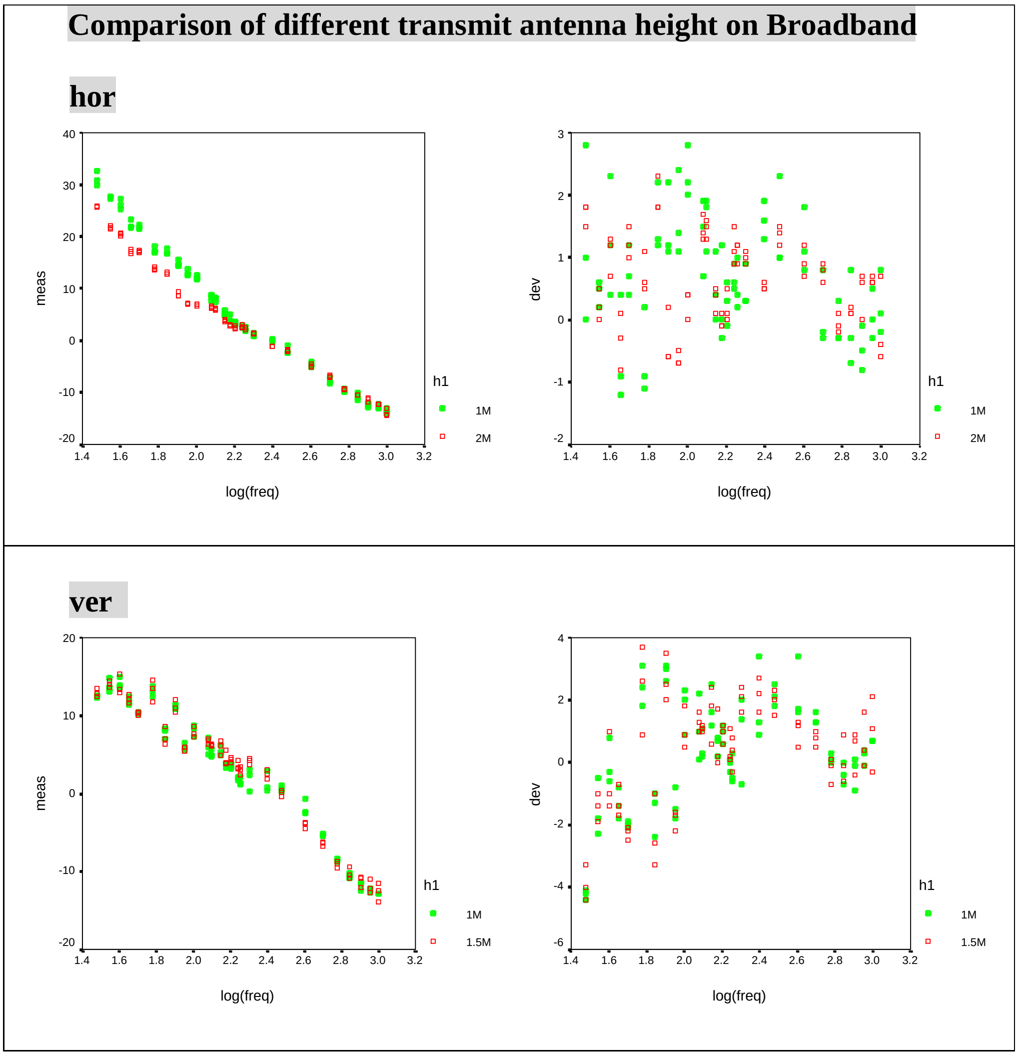


Figure 4.2 Measurements and ideal values at the same test setup