International Journal of Academic and Applied Research (IJAAR)
ISSN: 2643-9603
Vol. 4 Issue 1, January – 2020, Pages: 14-18

Confidence Interval of Multiresponse Semiparametric Regression Model Parameters Using Truncated Spline

Lilik Hidayati[1], Nur Chamidah[2*], I Nyoman Budiantara[3]
[1]Doctoral Student of Mathematics and Natural Sciences, Airlangga University, Surabaya 60115, Indonesia
lilik.hidayati-2016@fst.unair.ac.id
[2]Departement of Mathematics, Airlangga University, Surabaya 60115, Indonesia
[*]Corresponding author: nur-c@fst.unair.ac.id
[3]Department of Statistics, Sepuluh Nopember Institute of Technology, Surabaya, Indonesia
nyomanbudiantara65@gmail.com

Abstract: *Regression model is one of the statistical methods used to determine the functional relationship between predictor and response variables. Based on the pattern of the relationship regression model can be divided into 3 namely parametric regression, nonparametric regression and semiparametric regression. In statistical inference, parameter estimation consists of point and interval estimations. In this paper, we estimate confidence interval of multiresponse semiparametric regression model parameters based on truncated spline involving inverse variance-covariance of the error weight, then it is implemented to simulation data in case of homoscedasticity. Based on the estimated parameters for both large and small samples with a significant level of 0.05 we get accuracy of 100% for using weight and get accuracy less than 100% for using no weight. It can be concluded that in the case of homoscedasticity, the estimated confidence interval of the semiparametric multiresponse model parameters by using weighted truncated spline is better than by using unweighted truncated spline.*

Keywords: Confidence interval, Multiresponse semiparametric regression, Weighted and unweighted truncated spline.

1. INTRODUCTION

Regression model is one of the statistical methods used to determine the functional relationship between predictor and response variables. If the pattern functional relationship between the predictor variable and the response variable is known, then parametric regression is used. If the pattern of functional relationships between the predictor variable and the response variable is unknown, then nonparametric regression is used. In addition to parametric regression and nonparametric regression, there is also a semiparametric regression which is a combination of two components between the parametric regression model and the nonparametric regression model [1]. This regression model is used to determine the functional relationship between the response variable and several predictors, if one of the predictor variables has a specific pattern while the other predictor variables do not have a specific pattern. Regression models can also be distinguished based on the number of response, namely a regression model consist of one response called an unirespon, while a regression model whose number of response consists of more than one response is called multiresponse regression and correlate with each other.

In the semiparametric regression model, there are several approach models to estimate the regression curve for its nonparametric components among others, kernels, local polynomial, Fourier series and splines. Among those approaches, spline has several features, namely a model that has excellent statistical and visual interpretation and can model data with changing patterns at certain sub-intervals, because spline is a type of polynomial pieces [2]. In this study, the truncated spline approach is used to estimate the curve of its nonparametric component.

Parameter estimation is one that is considered important in statistical inference, consists of point estimation and interval estimation. Confidence interval estimation is the development of point estimation, that the estimated parameter value is not focus on a point but based on a certain interval so it can minimize the error in estimating compared to the point estimate. This estimation aims to find out the predictor variable that has a significant effect on the response variable. Thus, the confidence interval is an important issue in terms of the semiparametric regression model inference. Confidence interval is an estimated range of values between two numbers, where the parameter value of a population is located within that interval.

At present many researchers have focused their research on semiparametric regression. [3-5] used the penalized spline estimator; [6-8] used the smoothing spline estimator; however all of this researchers consist of only one response.

Whereas for multiresponse semiparametric researches for instances [9] used linear local estimator; [10-11] used spline estimator; [12] used kernel estimator, [13-15] developed unipredictor multiresponse semiparametric regression model based on the penalized spline estimator, [16-17] developed biresponse and multiresponse semiparametric regression model based on the truncated spline estimator. These studies are still discuss point estimation only, but these studies have not discussed

International Journal of Academic and Applied Research (IJAAR)
ISSN: 2643-9603
Vol. 4 Issue 1, January – 2020, Pages: 14-18

confidence intervals estimation. While [18] has discussed confidence interval estimation in the semiparametric regression model by using truncated spline estimator on longitudinal data, but it is still uniresponse.

Based on the description of the development of research that has been carried out by previous researchers only limited to the estimated point and estimated confidence interval that are still uniresponse to the semiparametric regression model. Therefore, the focus of this study is to estimate the confidence intervals of the multiresponse semiparametric regression model parameters using truncated spline estimator, which is then implemented on simulation data by creating R-Code. In homoscedasticity case, study simulation is used to compare between the modeling ability which use weighted and unweighted truncated spline estimators for estimating confidence interval of parameters of the multiresponse semiparametric regression model.

2. METHOD

The method used to estimate the confidence interval of parameters in multiresponse semiparametric regression model are based on weighted and unweighted spline truncated estimator. In simulation study, we create R-code to estimate the confidence interval of parameters in multiresponse semiparametric regression model are based on weighted and unweighted spline truncated estimator.

3. ESTIMATION OF INTERVAL CONVIDENCE OF SEMIPARAMETRIC MULTI-RESPONSE REGRESSION MODEL TRUNCATD SPLINE

Before given a paired variable $(x_1^{(r)},...,x_p^{(r)},t_1^{(r)},...,t_q^{(r)},y^{(r)})$ which is approximated by a linear function for the multiresponse semiparametric regression model based on truncated spline as follows:

$$y_i^{(r)} = \beta_0^{(r)} + \sum_{j=1}^{p}\beta_j x_{ji}^{(r)} + \sum_{\ell=1}^{q}\left[\gamma_\ell^{(r)} t_{\ell i}^{(r)} + \sum_{d=1}^{S_\ell}\gamma_{\ell+d}^{(r)}(t_{\ell i}^{(r)} - K_{\ell d}^{(r)})_+\right] + \varepsilon_i^{(r)} \quad (1)$$

with $r=1,2,3,\ldots,R$

So the multiresponse truncated spline semiparametric regression model is written in another form

$$\underset{\sim}{y} = \begin{pmatrix}\underset{\sim}{x} & \underset{\sim}{t}\end{pmatrix}\begin{pmatrix}\beta\\ \gamma\end{pmatrix} + \underset{\sim}{\varepsilon} \text{ with } C = \begin{pmatrix}\underset{\sim}{x} & \underset{\sim}{t}\end{pmatrix} \quad \underset{\sim}{\theta} = \begin{pmatrix}\beta\\ \gamma\end{pmatrix}$$

Then it can be stated in matrix notation:

$$\underset{\sim}{y} = C\underset{\sim}{\theta} + \underset{\sim}{\varepsilon}, \quad \underset{\sim}{\varepsilon} \sim N(\underset{\sim}{0}, W) \quad (2)$$

where

$$\underset{\sim}{y} = \begin{bmatrix}\underset{\sim}{y}^{(1)}\\ \underset{\sim}{y}^{(2)}\\ \vdots\\ \underset{\sim}{y}^{(R)}\end{bmatrix} \quad C = \begin{bmatrix}C^{(1)} & 0 & \cdots & 0\\ 0 & C^{(2)} & 0 & 0\\ \vdots & \vdots & \ddots & \vdots\\ 0 & 0 & \cdots & C^{(R)}\end{bmatrix} \quad \underset{\sim}{\varepsilon} = \begin{bmatrix}\underset{\sim}{\varepsilon}^{(1)}\\ \underset{\sim}{\varepsilon}^{(2)}\\ \vdots\\ \underset{\sim}{\varepsilon}^{(R)}\end{bmatrix}$$

$$\underset{\sim}{\theta}^{(r)} = \left[\underset{\sim}{\beta}_0^{(r)} \quad \underset{\sim}{\beta}_1^{(r)} \quad \underset{\sim}{\beta}_2^{(r)} \quad \cdots \quad \underset{\sim}{\beta}_p^{(r)} \vdots \quad \underset{\sim}{\gamma}_1^{(r)} \quad \underset{\sim}{\gamma}_2^{(r)} \vdots \quad \cdots \quad \vdots \underset{\sim}{\gamma}_q^{(r)}\right]$$

Errors random vectors of each response variable $\underset{\sim}{\varepsilon}^{(1)},\underset{\sim}{\varepsilon}^{(2)},...,\underset{\sim}{\varepsilon}^{(R)}$ correlate with each other. For estimating of parameters $\underset{\sim}{\theta}$, we use weighed least square (WLS) optimization method. In the WLS method, we minimize weighted sum square errors. The weight is invert of variance-covariance of errors matrix. The variance-covariance matrix structure of the multiresponse semiparametric regression model is as follows:

$$E\left(\underset{\sim}{\varepsilon}\underset{\sim}{\varepsilon}^T\right) = E\begin{bmatrix}E\left(\underset{\sim}{\varepsilon}^{(1)}\underset{\sim}{\varepsilon}^{(1)T}\right) & E\left(\underset{\sim}{\varepsilon}^{(1)}\underset{\sim}{\varepsilon}^{(2)T}\right) & \cdots & E\left(\underset{\sim}{\varepsilon}^{(1)}\underset{\sim}{\varepsilon}^{(R)T}\right)\\ E\left(\underset{\sim}{\varepsilon}^{(2)}\underset{\sim}{\varepsilon}^{(1)T}\right) & E\left(\underset{\sim}{\varepsilon}^{(2)}\underset{\sim}{\varepsilon}^{(2)T}\right) & \cdots & E\left(\underset{\sim}{\varepsilon}^{(2)}\underset{\sim}{\varepsilon}^{(R)T}\right)\\ \vdots & \vdots & \ddots & \vdots\\ E\left(\underset{\sim}{\varepsilon}^{(R)}\underset{\sim}{\varepsilon}^{(1)T}\right) & E\left(\underset{\sim}{\varepsilon}^{(R)}\underset{\sim}{\varepsilon}^{(2)T}\right) & \cdots & E\left(\underset{\sim}{\varepsilon}^{(R)}\underset{\sim}{\varepsilon}^{(R)T}\right)\end{bmatrix}$$

$$= \begin{bmatrix}W_{11} & W_{12} & \cdots & W_{1R}\\ W_{21} & W_{22} & \cdots & W_{2R}\\ \vdots & \vdots & \ddots & \vdots\\ W_{R1} & W_{R2} & \cdots & W_{RR}\end{bmatrix} = W \quad (3)$$

The steps to estimate $\underset{\sim}{\theta}$ by using WLS method are as follows:

1. Defining the function of Q

$$Q(\underset{\sim}{\theta}) = (\underset{\sim}{y} - C\underset{\sim}{\theta})^T W^{-1}(\underset{\sim}{y} - C\underset{\sim}{\theta}) \quad (4)$$

and then taking derivation of $Q(\underset{\sim}{\theta})$ with respect to $\underset{\sim}{\theta}$

2. Minimize equation Q by solving the following equation:

$$\frac{\partial Q(\underset{\sim}{\theta})}{\partial \underset{\sim}{\theta}} = \underset{\sim}{0}$$

International Journal of Academic and Applied Research (IJAAR)
ISSN: 2643-9603
Vol. 4 Issue 1, January – 2020, Pages: 14-18

$$\frac{\partial Q(\underset{\sim}{\theta})}{\partial \underset{\sim}{\theta}} = \frac{\partial}{\partial \underset{\sim}{\theta}}\left\{(\underset{\sim}{y}-C\underset{\sim}{\theta})^T W^{-1}(\underset{\sim}{y}-C\underset{\sim}{\theta})\right\}$$
$$= \frac{\partial}{\partial \theta}\left\{(\underset{\sim}{y}^T W^{-1}-\underset{\sim}{\theta}^T C^T W^{-1})(\underset{\sim}{y}-C\underset{\sim}{\theta})\right\}$$
$$= \frac{\partial}{\partial \theta}\left\{(\underset{\sim}{y}^T W^{-1}\underset{\sim}{y}-2\underset{\sim}{\theta}^T C^T W^{-1}\underset{\sim}{y}+\underset{\sim}{\theta}^T C^T W^{-1}C\underset{\sim}{\theta})\right\} \quad (5)$$
$$= -2C^T W^{-1}\underset{\sim}{y} \quad + 2C^T W^{-1}C\theta$$
$$\hat{\underset{\sim}{\theta}} = (C^T W^{-1}C)^{-1}C^T W^{-1}\underset{\sim}{y}$$

Next, we design the shortest parameter confidence interval $(1-\alpha)100\%$ for $\underset{\sim}{\theta}_v$ where σ^2 is unknown, presumably with mean square error (MSE) by using pivotal quantity as follows:

$$T_v(x_1, x_2, ..., x_p, t_1, t_2, ..., t_q, y) = \frac{\hat{\underset{\sim}{\theta}}_v - \underset{\sim}{\theta}_v}{\sqrt{MSE(C^T W^{-1}C)^{-1}_{vv}}},$$

where $\text{MSE} = \frac{\underset{\sim}{y}^T A\underset{\sim}{y}}{nR-R^*}$, A=1-D, $D = C(C^T W^{-1}C)^{-1}C^T W^{-1}$

and $\theta_v, \quad v = 1, 2, ..., R^*, R^* = \left(R + Rp + qR + R\sum_{\ell=1}^{q} S_\ell\right)$

Furthermore, we determine the confidence interval $(1-\alpha)$ by solving the probability equation, $P(a_v \le T_v(x_1, ..., x_p, t_1, ..., t_q, y) \le b_v) = 1-\alpha$. where a_v and b_v an element is a real number $a_v \le b_v$, so the length of the shortest parameter confidence interval is as follows:

$$P\left(\hat{\underset{\sim}{\theta}}_v - t_{\left(\frac{\alpha}{2}, nR-R^*\right)}\sqrt{\frac{\underset{\sim}{y}^T A\underset{\sim}{y}}{nR-R^*}V} \le \underset{\sim}{\theta}_v \le \hat{\underset{\sim}{\theta}}_v - t_{\left(\frac{\alpha}{2}, nR-R^*\right)}\sqrt{\frac{\underset{\sim}{y}^T A\underset{\sim}{y}}{nR-R^*}V}\right) = 1-\alpha$$

where $V = (C^T W^{-1}C)^{-1}_{vv}$

4. SIMULATION STUDY

Simulation study aims to show the ability of the weighted truncated spline estimator in homoscedasticty case in multiresponse semiparametric regression model. while the functions are as follows:

$$y_1^{(1)} = \beta_0^{(1)} + \beta_1^{(1)}x_1^{(1)} + \gamma_1^{(1)}t_1^{(1)} + \gamma_2^{(1)}(t_1^{(1)} - K_1^{(1)})_+$$
$$y_2^{(2)} = \beta_0^{(2)} + \beta_1^{(2)}x_1^{(2)} + \gamma_1^{(2)}t_1^{(2)} + \gamma_2^{(2)}(t_1^{(2)} - K_1^{(2)})_+$$
$$\vdots$$
$$y_n^{(R)} = \beta_0^{(R)} + \beta_1^{(R)}x_1^{(R)} + \gamma_1^{(R)}t_1^{(R)} + \gamma_2^{(R)}(t_1^{(R)} - K_1^{(R)})_+$$

Simulation data generated for large samples and small samples, each of them was repeated with 100 times and 300 replications with two treatments, namely choose one using weighted and unweighted. The estimation of multiresponse semiparametric regression model by using weighted truncated spline involve invert of covariance matrix of $\underset{\sim}{\varepsilon}$. The generated data consist of three responses, one predictor variable of the parametric component (x), one predictor variable of nonparametric component (t), and $\sigma_1^2 = 26.75$ $\sigma_2^2 = 31.81, \sigma_3^2 = 16.56, \sigma_{12} = 26.25$ $\sigma_{13} = 17.9, \sigma_{23} = 16.56$, as follows:

$$y_1^{(1)} = 9^{(1)} + 2^{(1)}x_1^{(1)} + 4^{(1)}t_1^{(1)} + 1^{(1)}(t_1^{(1)} - 1^{(1)})_+$$
$$y_2^{(2)} = 15^{(2)} + 3^{(2)}x_1^{(2)} + 1^{(2)}t_1^{(2)} + 6^{(2)}(t_1^{(2)} - 1^{(2)})_+$$
$$y_3^{(3)} = 11^{(3)} + 1^{(3)}x_1^{(3)} + 6^{(3)}t_1^{(3)} + 2^{(3)}(t_1^{(3)} - 1^{(3)})_+$$

The result of the simulation study for estimating confidence interval of multiresponse semiparametric regression model parameters by creating R-code is obtained in Table 1:

Table 1. Summary of simulation results with n=25 u=100 and u=300 repetitions

No	$\underset{\sim}{\theta}$	n=25 u=100 repetitions				n=25 u=300 repetitions			
		weighted		unweighted		weighted		unweighted	
		$\hat{\underset{\sim}{\theta}}$	Accuracy	$\hat{\underset{\sim}{\theta}}$	Accuracy	$\hat{\underset{\sim}{\theta}}$	Accuracy	$\hat{\underset{\sim}{\theta}}$	Accuracy
1	9	8.85	100%	8.64	100%	9.00	100%	9.03	99%
2	2	2.03	100%	2.04	100%	2.00	100%	2.00	99%
3	4	4.03	100%	4.08	100%	4.01	100%	4.01	99%
4	1	0.94	100%	0.90	98%	0.98	100%	1.02	99%
5	15	15.31	100%	15.94	97%	14.95	100%	14.67	96%
6	3	2.91	100%	2.83	97%	3.04	100%	3.06	97%
7	1	1.04	100%	0.88	99%	1.00	100%	1.06	95%
8	6	5.90	100%	5.96	98%	6.02	100%	5.93	96%
9	11	11.20	100%	10.41	85%	10.82	100%	10.45	83%
10	1	0.99	100%	1.03	82%	1.01	100%	1.03	81%
11	6	6.01	100%	6.21	83%	6.08	100%	6.19	83%
12	2	2.04	100%	1.75	86%	1.91	100%	1.65	85%

In simulation study for small samples (n = 25) with repetition of 100 times and 300 times, Table 1 shows by using significance level ($\alpha = 0.05$), the accuracy of estimated confidence interval of parameters in multiresponse semiparameter model based on weighted spline estimator is in average of 100% but it based on unweighted spline estimator of 93.6%. It means that for small samples the accuracy of estimated confidence interval of parameters in multiresponse semiparameter model based on weighted spline estimator is better than unweighted spline estimator.

Table 2. Summary of simulation results with n=100 u=100 and u=300 repetitions

No	$\underset{\sim}{\theta}$	n=100 u=100 repetitions				n=100 u=300 repetitions			
		Weighted		unweighted		weighted		unweighted	
		$\hat{\underset{\sim}{\theta}}$	Accuracy	$\hat{\underset{\sim}{\theta}}$	Accuracy	$\hat{\underset{\sim}{\theta}}$	Accuracy	$\hat{\underset{\sim}{\theta}}$	Accuracy
1	9	8.95	100%	8.81	99%	9.02	100%	8.96	100%
2	4	4.00	100%	4.04	100%	4.01	100%	4.01	100%
3	2	2.00	100%	2.00	100%	2.01	100%	1.99	100%
4	1	1.01	100%	0.95	100%	1.00	100%	0.99	99%
5	15	15.04	100%	14.98	97%	14.94	100%	14.88	90%
6	3	3.01	100%	3.06	98%	3.00	100%	2.98	92%
7	1	0.96	100%	0.93	95%	1.02	100%	1.05	90%

8	6	6.07	100%	6.14	96%	5.98	100%	5.94	99%
9	11	11.02	100%	11.17	88%	10.90	100%	10.96	84%
10	1	1.00	100%	1.00	86%	1.00	100%	1.00	85%
11	6	5.98	100%	5.92	90%	6.02	100%	6.00	83%
12	2	2.03	100%	2.13	94%	2.00	100%	2.08	83%

Based on Table 2, for large samples (n = 100) with repetition of 100 times and 300 times shows by using significance level ($\alpha = 0.05$) in multiresponse semiparameter model, the accuracy of estimated confidence interval of parameters based on weighted spline estimator is in average of 100% and based on unweighted spline estimator of 94 %. This means it can be concluded that for large samples the accuracy of estimated confidence interval of parameters in multiresponse semiparametric model based on weighted spline estimator is better than unweighted spline estimator.

5. CONCLUSION

In the simulation study, the accuracy of estimated confidence interval of parameters based on weighted spline estimator is better than unweighted spline estimator in multiresponse semiparametric regression model for homoscedasticity cases.

6. REFERENCES

[1] Hardle, W., (1994), *Applied Nonparametric Regression*, Humboldt-Universitat Zu Berlin.

[2] Budiantara, I.N. (2009). S*pline dalam regresi nonparametrik dan semiparametrik: sebuah pemoelan statistika masa kini dan masa mendatang*. Pidato Pengukuhan Untuk Jabatan Guru Besar dalam Bidang Ilmu: Matematika Statistika dan Probabilitas, Pada Jurusan Statistika FMIPA, ITS, Surabaya.

[3] Bandyopadhyay, S., & Maity, A., (2011). Analysis of sabine river flow data using semiparametric spline modeling. *Journal of Hydrology* 399 : 274–280.

[4] Tong, T., Wu., & He, X., (2012). Coordinate ascent for penalized semiparametric regression on high-dimensional panel count data. *Journal of Computational Statistics and Data Analysis* 56 : 23-33

[5] Yang. J., & Yang. H., (2016). A robust penalized estimation for identification in semiparametric additive models. *Statistics and Probability Letters* 110 : 268-277.

[6] Kim, Y, J., (2013). A partial spline approach for semiparametric estimation of varying-coefficient partially linear models. *Journal of Computational Statistics and Data Analysis* 62:181-187

[7] Chen, M., & Song, Q., (2016). Semiparametric estimation and forecasting for exogenous log-GARCH models. *Journal of TEST* 25: 93–112.

[8] Ramadan W, Chamidah N, Zaman B, Muniroh L, and Lestari B 2019 Standard Growth Chart of Weight for Height to Determine Wasting Nutritional Status in East Java Based on Semiparametric Least Square Spline Estimator *IOP Conf. Series: Materials Science and Engineering* 546 052063052063. doi:10.1088/1757-899X/546/5/052063

[9] Chamidah, N., and Rifada. M., (2016). Local linier estimator in bi-response semiparametric regression model for estimating median growth charts of children. *Far East Journal of Mathematical Sciences (FJMS)* 99(8):1233-1244. SJR:0.24,Q4, SCOPUS, ISSN:09720871

[10] Chamidah, N., and Eridani., (2015). Designing of growth reference chart by using bi-response semiparametric regression approach based on p-spline estimator. *International Journal of Applied Mathematics and Statistics*, Int. J. Appl. Math. Stat.; Vol. 53; Issue No. 3.

[11] Chamidah, N., Kurniawan, K., Zaman, B., Muniroh, L. (2018). Least Square-Spline Estimator In Multi-Response Semiparametric Regression Model For Estimating Median Growth Charts Of Children In East Java, Indonesia. *Far East Journal of Mathematical Sciences*, Vol.107 (2), pp.295-307

[12] Lo'pez, B.P. and Manteiga, W.G., (2006). Multivariate partially linier models, *Statistics and Probability Letters*, 76, 1543-1549

[13] Wibowo, W., Haryatmi, S., & Budiantara, I, N., (2012). On multiresponse semiparametric regression model. *Journal of Mathematics and Statistics* 8 (4): 489-499.

[14] Wibowo, W., Haryatmi, S., & Budiantara, I.N. 2013. Modeling of regional banking activitas using spline multiresponse semiparametric regression. *Journal of Applied Mathematics and Statistics*. 23 : 102-110.

[15] Chamidah, N., Kurniawan, K., Zaman, B., Muniroh, L. (2018). Least Square-Spline Estimator In Multi-Response Semiparametric Regression Model For Estimating Median Growth Charts Of Children In East Java, Indonesia. Far East Journal of Mathematical Sciences, Vol.107 (2), pp.295-307

[16] Hidayati, L., Chamidah, N., and Budiantara, I.N., (2019). Spline truncated estimator in multiresponse semiparametric regression model for computer based national exam in west nusa tenggara *Proceeding* 9th Annual Basic Science International Conference. IOP Conf. Series : Material Science and Engineering 546 (2019a) 052029 doi: 10.1088/1757-899X/546/5/052029

[17] Hidayati, L., Chamidah, N., and Budiantara, I.N., (2019b). Bi-response semiparametric regression model based on spline truncated for estimating computer based

national exam In west nusa tenggara. *Proceeding* 1st International Conference on Mathematics and Islam (ICMIs 2018), SCITEPRESS – Science and Technology Publications, ISBN: 978-989-758-407-7, Depósito Legal: 465131/19. :357-361

[18] Prawanti, D.D., Budiantara, I.N., and Purnomo, J.D.T., (2019) Parameter interval estimation of semiparametric spline truncated regression model for longitudinal data. *Proceeding* 9th Annual Basic Science International Conference. IOP Conf. Series : Material Science and Engineering 546 (2019) 052053 doi: 10.1088/1757-899X/546/5/052053