

基于部分线性变系数模型的飞行训练成绩综合评定

未振军, 刘 嘉, 王昌海

海军航空大学, 山东 烟台

收稿日期: 2025年8月24日; 录用日期: 2025年9月14日; 发布日期: 2025年9月25日

摘 要

针对当前飞行训练成绩综合评定工作中耗费大、效率低、过度依赖人工等问题, 以飞行学员飞行训练成绩数据为研究对象, 构建了部分线性变系数模型, 通过数据拟合, 对参数和非参数分量进行了估计和假设检验, 并完成了变量的有效选择, 最后使用模型评价指标决定系数 R^2 和残差标准差 $\hat{\sigma}$, 对部分线性变系数模型和多元线性模型进行了比较。结果表明, 部分线性变系数模型在对飞行训练成绩综合评定中要明显优于多元线性模型。

关键词

部分线性变系数模型, 飞行训练, 综合评定, 参数估计

Comprehensive Evaluation of Flight Training Performance Based on Partially Linear Varying Coefficient Model

Zhenjun Wei, Jia Liu, Changhai Wang

Naval Aviation University, Yantai Shandong

Received: August 24th, 2025; accepted: September 14th, 2025; published: September 25th, 2025

Abstract

Aiming at the problems of high cost, low efficiency, and excessive reliance on manual work in the comprehensive evaluation of flight training performance, this study takes flight training performance data of cadets as the research object and constructs a partially linear varying coefficient model. Through data fitting, the parameter and non-parameter components were estimated and

subjected to hypothesis testing. Effective variable selection was also performed. Finally, the model evaluation metrics—coefficient of determination (R^2) and residual standard deviation—were used to compare the partially linear varying coefficient model with the multiple linear model. The results show that the partially linear varying coefficient model significantly outperforms the multiple linear model in the comprehensive evaluation of flight training performance.

Keywords

Partially Linear Varying Coefficient Model, Flight Training, Comprehensive Evaluation, Parameter Estimation

Copyright © 2025 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

1. 引言

飞行训练是飞行学员提高飞行能力的一项重要训练内容，而飞行训练成绩评定又是飞行训练工作中的一个重要组成部分，客观合理的对飞行学员飞行训练的质量进行评估，对于衡量飞行学员操作水平、打好飞行基础、提升飞行能力、改善飞行技巧、保障飞行安全、提高飞行训练质量等具有重要的作用。飞行训练成绩评定的目的是让飞行学员充分利用每架次飞行，保证每个架次都能进行高质量的飞行训练，并且通过后续分析飞行训练成绩的变化，对成绩不稳定的学员有针对性地进行训练课目的调配和强化训练，可以显著提升训练质效。

目前，飞行训练评估方法主要是以带飞教官的人工评分为主，这种方式需要花费庞大的人力和物力资源，且效率低下，还会因为不同带飞教官等人为因素造成评估偏差[1]。为了有效提高飞行学员飞行训练评估的准确性，减少带飞教官的工作量，节省在飞行训练评估上的精力和物力等资源，同时降低人为等主观因素所产生的评估误差，近年来，许多专家学者基于各类模型对飞行训练质量的评估进行了大量研究。比如，尹大伟[1]研究了基于飞参数据的飞行学员飞行训练质量评估方法，深入分析了飞行训练质量评估的特点，建立了飞参数据辅助评估流程和基本框架；王昌海等[2]基于多元线性回归分析的方法，通过从飞行训练成绩单数据中提取出了 15 个变量，建立了多元回归模型，进行了飞行训练成绩评定问题研究；罗渝川等[3]通过分析文献和访谈专家收集了训练绩效指标，并通过层次分析法和因子分析法构建了飞行学员训练绩效指标模型。但是上述研究大都停留在理论阶段或定性分析[4]-[6]，缺乏定量的研究，而定量研究对于提升飞行学员飞行能力、提高飞行学员的训练效率具有重要的指导作用，因此研究基于部分线性变系数模型的飞行训练成绩综合评定具有重要的应用价值和现实意义。

2. 部分线性变系数模型

部分线性变系数模型[7] (Partial Linear Variable Coefficient Models, PLVCM)，是近年来发展起来的一种高级数据分析模型，其既含有线性模型部分，又含有变系数模型部分，是线性模型与变系数模型的结合，因此其具有线性模型便于解释和非参数模型稳健的特征。

(一) 模型形式

部分线性变系数模型的基本形式如下：

$$Y = \alpha^T(U)X + \beta^T Z + \varepsilon \quad (1)$$

其中 Y 是响应变量, X , Z 和 U 为协变量, $\beta = (\beta_1, \beta_2, \dots, \beta_q)^T$ 是 $q \times 1$ 维的未知参数向量, $\alpha(\cdot) = (\alpha_1(\cdot), \alpha_2(\cdot), \dots, \alpha_p(\cdot))^T$ 是 $p \times 1$ 维未知的系数函数向量, ε 为模型误差并且满足 $E(\varepsilon|X, Z, U) = 0$, $Var(\varepsilon|X, Z, U) = \sigma^2$ 。为了避免“维数祸根”, 一般假设 U 为单变量。

(二) 参数估计

对 $\beta, \alpha(u)$ 的估计通常采用 profile 最小二乘法[4], 估计结果为:

$$\hat{\beta} = \{Z^T (I - S)^T (I - S) Z\}^{-1} Z^T (I - S)^T (I - S) Y \quad (2)$$

$$\hat{\alpha}(u) = [I_p \quad 0_p] \{D_u^T W_u D_u\}^{-1} D_u^T W_u (Y - Z \hat{\beta}) \quad (3)$$

其中 $K_h(\cdot) = K(\cdot/h)/h$, $K(\cdot)$ 为核函数, h 为带宽, I 为单位阵, $Y = (Y_1, Y_2, \dots, Y_n)^T$, $X = (X_1, X_2, \dots, X_n)^T$,

$$Z = (Z_1, Z_2, \dots, Z_n)^T, U = (U_1, U_2, \dots, U_n)^T, D_u = \begin{pmatrix} X_1^T & \frac{U_1 - u}{h} X_1^T \\ X_2^T & \frac{U_2 - u}{h} X_2^T \\ \vdots & \vdots \\ X_n^T & \frac{U_n - u}{h} X_n^T \end{pmatrix}, S = \begin{pmatrix} [X_1^T \quad 0] \{D_{u_1}^T W_{u_1} D_{u_1}\}^{-1} D_{u_1}^T W_{u_1} \\ \vdots \\ [X_n^T \quad 0] \{D_{u_n}^T W_{u_n} D_{u_n}\}^{-1} D_{u_n}^T W_{u_n} \end{pmatrix},$$

$$W_u = \begin{bmatrix} K_h(U_1 - u) & 0 & 0 & 0 \\ 0 & K_h(U_2 - u) & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & K_h(U_n - u) \end{bmatrix}.$$

(三) 模型检验

由样本计算得到的回归系数 $\hat{\beta}$ 及回归系数函数 $\hat{\alpha}(u)$ 是总体回归系数和 β 回归系数函数 $\alpha(u)$ 的估计值, 如果这些总体回归系数或总体回归系数函数等于 0, 那么部分线性变系数模型将没有任何意义。因此在建立回归方程后还需要对这些回归系数进行检验, 来验证回归方程是否统计显著。

对参数分量检验, 要进行的假设检验问题为:

$$H_0: \beta_1 = \beta_2 = \dots = \beta_l = 0, l \leq q \quad (4)$$

构造广义似然比 PLR 统计量为:

$$T_n = \frac{n}{2} \log \left(\frac{RSS_0}{RSS_1} \right) \approx \frac{n}{2} \frac{RSS_0 - RSS_1}{RSS_1} \quad (5)$$

当原假设 H_0 成立时, 上述 PLR 统计量渐进服从自由度为 l 的卡方分布, 即在原假设下: $2T_n \sim \chi^2(l)$ 。因此可以使用卡方分布对回归模型参数分量的显著性进行假设检验。

对非参数分量检验时, 在得到 $\alpha(\cdot) = (\alpha_1(\cdot), \alpha_2(\cdot), \dots, \alpha_p(\cdot))^T$ 的非参数分量估计之后, 需要检验非参数分量是否可以用参数模型进行拟合, 为此可构造的假设检验问题为:

$$H_0: \alpha_j(u) = a_j, j = 1, 2, \dots, p, a_j \text{ 为常数} \quad (6)$$

构造广义似然比 GLR 统计量为:

$$T_0 = \frac{n}{2} \log \left(\frac{\widetilde{RSS_0}}{RSS_1} \right) \quad (7)$$

当原假设 H_0 成立时，广义似然比 GLR 统计量 T_0 渐进服从卡方分布，因此可以使用卡方分布对回归模型非参数分量的显著性进行假设检验。

(四) 模型评价

统计学中评价回归模型性能最常用的指标之一为 R^2 (R-squared)，又称决定系数(Coefficient of Determination)，是评价估计量与被估计量之间差异程度的一种度量方法，用于衡量模型对样本数据的拟合优度，反映了模型的预测准确性，核心思想是通过比较模型预测值与实际值的差异，量化模型解释目标变量变异的能力，其值取值范围为 0 到 1，越接近与 1，说明模型拟合的越好，解释能力越强，预测精度也越高；反之，则说明模型拟合的效果较差。计算公式为：

$$R^2 = 1 - \frac{SSR}{SST} \tag{8}$$

其中 SSR 为残差平方和，即预测值与实际值之差的平方和； SST 为总平方和，即实际值与其均值的偏差平方和。 R^2 有如下两个特点：① 具有直观的解释性， $R^2 = 0.8$ 表示模型能解释 80% 的目标变量差异，剩余 20% 由其他因素或者随机噪声导致；② 无量纲性，不受目标变量单位影响，便于不同模型间的横向比较。

3. 飞行训练成绩评定

(一) 数据来源及说明

本文选取了 100 名中高级阶段飞行学员高教机型飞行训练成绩数据，经过前期的数据预处理，共得到 315 条有效样本，每个样本包含 10 个变量，分别为：当前架次总成绩、是否为单元首个架次、前一架次总成绩、当前架次动作条目总数、架次动作条目成绩大于标准的总和、架次动作条目成绩小于大纲标准的总和、当前架次大纲要求的标准总数、当前架次难度、架次任务相似度、当前总架次数。其中“当前架次总成绩”为响应变量，其他 9 个变量为解释变量。

为了使用上的方便，定义变量及变量名称见表 1。

Table 1. Variable and variable name
表 1. 变量表示及变量名称

变量	变量名称
Y	当前架次总成绩
X_1	当前架次动作条目总数
X_2	前一架次总成绩
X_3	是否为单元首个架次
X_4	架次动作条目成绩大于大纲标准的总和
X_5	架次动作条目成绩小于大纲标准的总和
X_6	当前架次大纲要求的标准总数
X_7	当前架次难度
X_8	架次任务相似度
X_9	当前总架次数

(二) 构建多元线性回归模型

为了便于将本文提出的模型与王昌海使用的多元线性模型[2]进行比较，先基于上述 10 个变量构建了如下的多元线性回归模型：

$$Y = \beta_0 + \sum_{i=1}^9 \beta_i X_i + \varepsilon$$

通过数据拟合, 得到的变量估计值及对应的 p 值见表 2。

Table 2. Variable estimates and p -values

表 2. 变量估计及对应的 p 值

变量	系数估计值	误差标准差	t-统计量值	p 值
截距项	12.21	4.83	2.53	0.012
X_1	11.99	1.03	11.67	0
X_2	0.66	0.04	16.2	0
X_3	0.12	0.15	0.77	0.44
X_4	0.17	0.07	2.42	0.016
X_5	-0.65	0.26	-2.49	0.013
X_6	0.12	0.03	3.43	0
X_7	-0.23	0.06	-4.06	0
X_8	-0.20	2.42	-0.08	0.93
X_9	-0.01	0.01	-0.83	0.41

并进一步计算得到了多元线性回归模型的决定系数为 $R^2 = 0.6981$, 残差标准差为 $\hat{\sigma} = 5.206$ 。下面, 将使用本文提出的部分线性模型来拟合飞行训练成绩数据。

(三) 构建部分线性变系数模型

众所周知, 随着飞行架次的增多, 飞行水平会逐渐提高。因此, 为了考查“当前总架次数”对其他变量的影响, 故本文选取“当前总架次数” X_9 作为模型非参数分量的自变量, 即取 $U = X_9$ 。

首先, 构建变系数模型[8]:

$$Y = \beta_0 + \sum_{i=1}^8 \alpha_i(U) X_i + \varepsilon$$

本文选取的核函数为 Epanechnikov 核, 即: $K(u) = \frac{3}{4}(1-u^2)I(|u| \leq 1)$, 使用交叉验证法确定最优带宽。通过数据拟合, 可以得到上述变系数模型的决定系数 $R^2 = 0.763$, 残差标准差 $\hat{\sigma} = 4.245$ 。然后对非参数系数函数进行假设检验, 来判断系数函数是否是统计显著的。表 3 展示了每一个系数函数的统计量值及对应的 p 值, 此处取显著性水平为 0.01。

Table 3. Coefficient function test statistic and p -values

表 3. 系数函数检验统计量及对应的 p 值

	$\alpha_1(U)$	$\alpha_2(U)$	$\alpha_3(U)$	$\alpha_4(U)$	$\alpha_5(U)$	$\alpha_6(U)$	$\alpha_7(U)$	$\alpha_8(U)$
GLR 统计量	21.169	151.34	0.805	4.469	2.160	8.313	5.571	0.713
p 值	0	0	0.448	0.012	0.088	0	0	0.636

由上表结果可知, X_3, X_4, X_5, X_8 系数函数的 p 值大于 0.01。可以判断 X_3, X_4, X_5, X_8 的系数函数不随 U 的变化而变化, 因此可以认为 X_3, X_4, X_5, X_8 的系数为常数。

基于上述分析, 在变系数模型的基础上, 可以将 X_3, X_4, X_5, X_8 的系数设为常数, 从而构建如下的部

分线性变系数模型:

$$Y = \alpha_1(U)X_1 + \alpha_2(U)X_2 + \alpha_3(U)X_6 + \alpha_4(U)X_7 + \beta_1X_3 + \beta_2X_4 + \beta_3X_5 + \beta_4X_8 + \beta_0 + \varepsilon$$

再次使用上述模型拟合数据, 此时需要判断 X_3, X_4, X_5, X_8 的系数是否是统计显著的。为了验证上述想法, 需要构建参数分量的假设检验统计量, 假设检验的结果见表 4。

Table 4. Parameter component coefficient test statistic and p -values
表 4. 参数分量系数检验统计量及对应的 p 值

	β_1	β_2	β_3	β_4
PLR 统计量	-0.758	2.924	-2.403	-0.776
p 值	0.449	0.004	0.017	0.44

由上表可知, X_3, X_5, X_8 系数 $\beta_1, \beta_3, \beta_4$ 检验的 p 值大于 0.01, 即 X_3, X_5, X_8 这三个变量的变化对响应变量 Y 几乎没有影响, 可以认为 $\beta_1 = \beta_3 = \beta_4 = 0$ 。可得最终的模型为:

$$Y = \alpha_1(U)X_1 + \alpha_2(U)X_2 + \alpha_3(U)X_6 + \alpha_4(U)X_7 + \beta_2X_4 + \beta_0 + \varepsilon$$

继续使用上述模型来拟合数据, 最后得到参数分量 X_4 的系数 $\beta_2 = 0.146$, 估计标准误差为 0.061。非参数分量系数为函数形式, 其估计值见图 1 黑色曲线所示, 其中阴影部分为非参数系数函数的 95%置信区间。

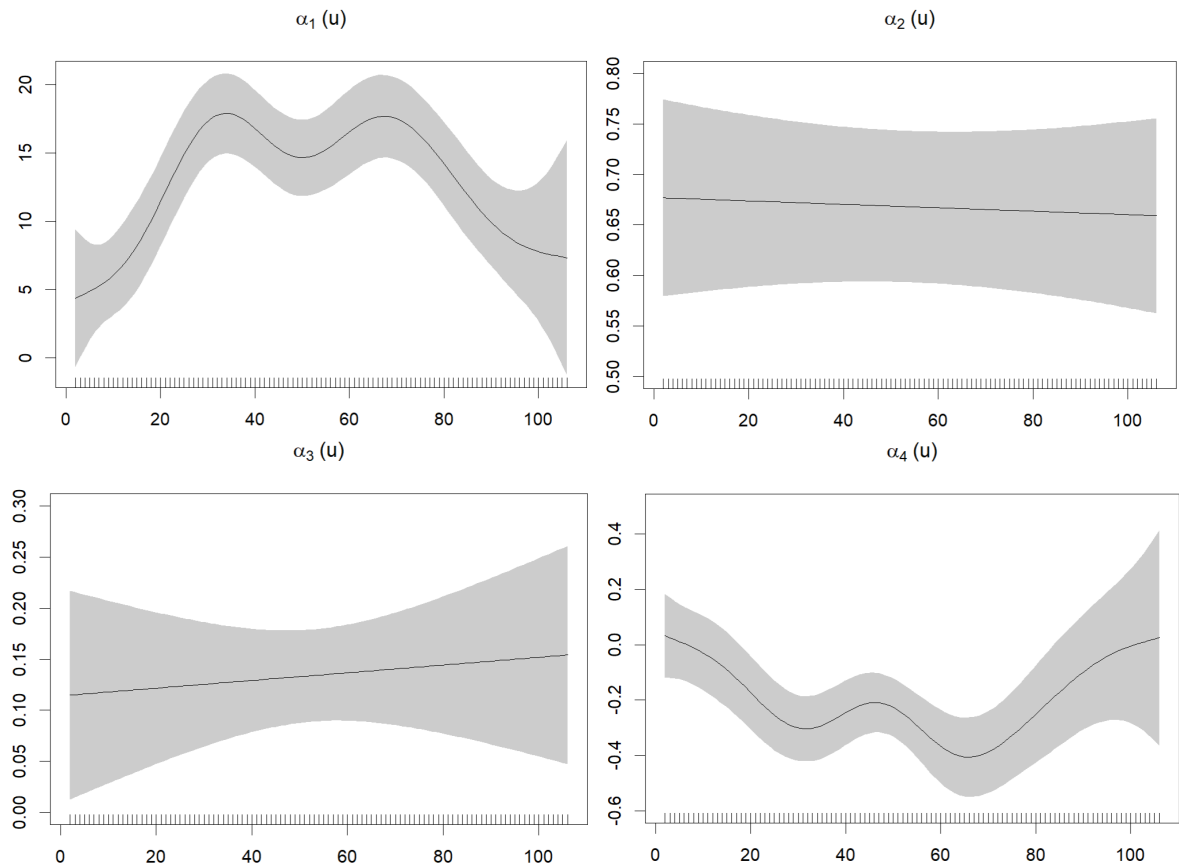


Figure 1. Nonparametric component coefficient function estimate
图 1. 非参数分量系数函数估计值

此外, 可得该部分线性变系数模型的决定系数 $R^2 = 0.759$, 明显大于多元线性回归模型的决定系数 $R^2 = 0.6981$, 且该模型的残差标准差 $\hat{\sigma} = 4.124$, 明显小于多元线性回归模型的残差标准差为 $\hat{\sigma} = 5.206$. 因此, 从上述结果上看, 本文使用的部分线性变系数模型在对飞行训练成绩综合评定中要明显优于多元线性模型。

4. 实际指导意义

首先, 科学评定能够“以数促训”, 实现训练资源的精准投放。传统模式下, 教官通常依据经验判断学员薄弱环节, 容易出现“一刀切”或“重复加码”的低效训练。引入部分线性变系数模型的方法评价后, 模型可自动识别出哪些变量(如前一架次成绩、当前架次难度、任务相似度等)对总成绩影响显著, 哪些变量影响微弱。例如, 数据显示, “前一架次总成绩”的系数显著为正, 而“当前架次动作条目总数”在某些情境下并不显著。这意味着, 当飞行学员成绩波动主要由前一架次表现决定时, 应将重点放在帮助学员快速消化前序经验、强化动作记忆, 而非简单增加飞行架次。通过精准定位关键变量, 可把有限的航油、空域、机务保障资源集中到最能产生训练效益的课目和学员身上, 显著提升单位资源的训练产出比。

其次, 动态评估能够“以数促改”, 实现训练方案的滚动优化。部分线性变系数模型的最大优势在于, 其非参数部分允许变量效应随“当前总架次数”平滑变化, 从而刻画出“学习曲线”的真实形态。图1中 $\alpha_1(u)$ 的估计曲线表明, 随着总架次数增加, 当前架次动作条目总数对成绩的边际贡献先升后降: 初期多动作有助于快速积累经验, 但达到一定阈值后, 过多动作反而分散注意力, 导致成绩提升放缓甚至出现波动。这表明在学员飞行生涯早期应敢于“加课目、加动作”, 促成“陡坡式”成长; 而当飞行小时数累积到拐点后, 应及时转入“精炼动作、强化质量”的阶段, 避免“高原现象”。通过滚动更新模型参数, 训练部门可以每月、每季度生成最新的“学习曲线”档案, 据此微调课目编排和训练节奏, 实现训练方案与学员成长阶段的精准匹配。

最后, 公正评定能够“以数促战”, 夯实飞行能力生成的底层逻辑。飞行学员选拔历来是敏感环节, 若缺乏量化依据, 容易滋生“人情分”“印象分”, 挫伤飞行学员积极性。本文所提模型将所有飞行学员置于同一量化标尺下, 依据客观数据自动生成排名, 让学员对自身水平一目了然。更重要的是, 模型结果可与实际背景结合: 例如, 未来飞行中对低空飞行的精度要求极高, 可将“当前架次动作条目成绩大于大纲标准的总和”权重调高, 引导学员养成“超标”意识。通过实际需求转化为模型变量权重, 训练评定与实际需求实现无缝衔接。

综上, 以部分线性变系数模型为核心的飞行训练成绩综合评定体系, 不仅能解决传统人工评分“高耗低效、主观偏差”的痛点, 更能够通过精准识别关键变量、动态跟踪学习曲线、公正支撑人才选拔, 形成“数据驱动、闭环管理”的训练新范式。对航空部门而言, 这意味着训练资源投入更精准、训练效益评估更科学; 对学员而言, 则意味着成长路径更清晰、短板补齐更及时、个人潜力释放更充分。随着数据积累与模型迭代, 该体系将持续反哺训练大纲修订、课目标准优化, 最终推动飞行学员训练质量呈螺旋式上升。

5. 结束语

本文以飞行学员日常飞行训练成绩数据为研究对象, 构建了包括响应变量在内的10个可以衡量飞行学员飞行训练质量的特征, 通过数据的预处理, 获得了315条有效样本。然后建立了部分线性变系数模型, 并基于样本数据, 完成了对参数分量和非参数函数分量的有效估计, 之后对参数分量和非参数函数分量进行了假设检验, 在假设检验的过程中同时完成了对变量选择, 即保留与模型显著相关的变量, 剔

除与模型不相关的变量。最后使用模型评价指标决定系数 R^2 和残差标准差 $\hat{\sigma}$ ，对部分线性变系数模型和多元线性模型进行了比较，结果表明，无论是决定系数还是残差标准差，部分线性变系数模型在对飞行学员飞行训练成绩综合评定中都要明显优于多元线性模型。研究结果可以有效地辅助带飞教官对学员飞行训练成绩的评定，更加科学、客观、公正地评估飞行学员的训练质量，进一步提高飞行学员人才培养的质量。

参考文献

- [1] 尹大伟. 基于飞参数据的飞行学员飞行训练质量辅助评估研究[J]. 航空标准化与质量, 2023(S1): 116-120.
- [2] 王昌海, 冷旭, 李匡迪. 基于多元线性回归分析的初教机飞行训练成绩评定问题研究[J]. 军事运筹与评估, 2023, 38(4): 20-25.
- [3] 罗渝川, 冯鑫源. 飞行学员训练绩效评估方法研究[J]. 自动化应用, 2024, 65(8): 217-219+223.
- [4] 姚裕盛, 徐开俊. 基于 BP 神经网络的飞行训练品质评估[J]. 航空学报, 2017, 38(S1): 24-32.
- [5] 高文琦, 张复春, 王立波, 等. 飞行训练成绩评估模型的建立与实现[J]. 电子设计工程, 2011, 19(24): 50-52.
- [6] 刘浩, 王昊, 孟光磊, 等. 基于动态贝叶斯网络和模糊灰度理论的飞行训练评估[J]. 航空学报, 2021, 42(8): 250-261.
- [7] Fan, J.Q. and Huang, T. (2005) Profile Likelihood Inferences on Semiparametric Varying-Coefficient Partially Linear Models. *Bernoulli*, **11**, 1031-1057. <https://doi.org/10.3150/bj/1137421639>
- [8] Simon, N. (2012) Wood. Generalized Additive Models an Introduction with R. CRC Press, 478-580.