# Developmental changes of fundamental frequency and temporal characteristics in Estonian adolescent speech

Einar Meister, Lya Meister

Tallinn University of Technology, Estonia
einar.meister@taltech.ee, lya.meister@taltech.ee

## ABSTRACT

Age- and gender-related anatomical changes and the development of the speech-motor system during childhood and adolescence affect speech acoustics. Development trends of F0 and temporal features in the speech of Estonian adolescents are investigated as a function of age and gender, using speech samples read by 309 Estonian-speaking subjects (175 girls and 134 boys) aged 9–18 years. The results show the expected age-dependent patterns: (1) a decline in F0 with a sharp drop of about 100 Hz in boys aged 12–15 years and a gradual decline in girls, (2) an increase in speech and articulation rates up to 15 years of age. Gender differences in temporal characteristics are not significant, although boys tend to speak slightly faster than girls. The results are in line with the findings reported for several other languages and can be considered reference data for Estonian-speaking subjects aged 9–18 with typical language development.

**Keywords**: adolescent speech, age- and gender-related variations, acoustic analysis, fundamental frequency, speech tempo

## 1. INTRODUCTION

Age- and gender-related anatomic changes of the vocal apparatus during childhood and adolescence have a direct impact on speech acoustics manifested in a sudden drop of fundamental frequency (F0) during puberty in boys and in a gradual decrease of the acoustic vowel space and formant frequencies in both genders [13, 15, 26, 27]. In parallel with anatomical changes, the development of speech-motor control as well as cognitive and linguistic processing takes place, which manifests itself in the increase of speech and articulation rates [14, 24].

In infants (under 1 year old) the length of the vocal folds is 4–5 mm, at the age of 20 the vocal folds have reached a length of 11–15 mm in women and 17–25 mm in men [7, 20]. Consistent with vocal fold growth, changes in F0 have been documented, e.g. mean F0 in boys at the age of 7 is 266 Hz, at the age of 12 years 226 Hz, and lowers to 127 Hz by the age of 15 with marginal changes thereafter. The average F0 of 7-year-old girls is 275 Hz and gradually decreases to 231 Hz until the age of 12, after which there is no significant change [13]. However, another study reports a further decline of girls' mean F0 from 230 Hz to 218 Hz between the ages of 13 and 19 [5].

Studies of speech-motor development have shown that children's articulatory movements are slower and more variable than that of adults. A study [24] examined native English-speaking children and adults aged 4–22 years by recording upper lip, lower lip, and jaw movements while reading different sentences. The results showed that as age increases, the variability of articulatory movements decreases and the time required to form a sentence shortens, and speech-motor control in both boys and girls becomes similar to adults only after the age of 14. The study concludes that the age-related increase in speaking rate is due to improvements in cognitive and linguistic processing and motor control of speech.

According to various studies, the articulation rate in the spontaneous speech of English-speaking children aged 3–6 years varies from 2.9 to 4.3, and the speech rate from 2.3 to 2.6 syllables per second; for 7- to 12-year-olds, 4.5 to 5.6 and 2.4 to 2.9 syllables per second, respectively [14]. The speaking rates increase equally with age in boys and girls, and the differences within age groups are not statistically significant (however, males tend to speak slightly faster) e.g. [10, 11, 12, 19]. Articulation rates of 3.3–5.8 and the speaking rates of 3.7–4.7 syllables per second have been documented in adults [12].

When reading written texts aloud, the speaking rate is affected by the text length as confirmed by several studies e.g. [1, 4, 22]. This is known as anticipatory shortening, whereby in the speech planning process, the speaker adjusts his average syllable duration to the expected length of the phrase [2]. The study [22] explored the relationship between text length and speaking rate in a comparison of children (aged 5–16) and adults (aged 20–23) and found that the duration of a test phrase read in isolation is longer than when read in a frame sentence both in adults and 9-16-year-olds, but not in 5-7-year-old children. The authors hypothesize that children and adults use different motor planning strategies when reading more complex sentences: younger children plan their speech in smaller units (words or syllables), while older children and adults plan speech in longer units (phrases).

The present study aims to document the acoustic characteristics of Estonian adolescent speech related to the speaker's age and gender and gain a better understanding of these variations from a cross-linguistic perspective. In particular, we explore (1) the acoustic variations of the fundamental frequency F0, and (2) the measures of speaking rate.

## 2. MATERIALS AND METHOD

### 2.1. Speech corpus

The Estonian adolescent speech corpus [16] consists of speech samples from 309 native-speaking subjects (175 girls and 134 boys) aged 9–18 years. The corpus contains cross-sectional speech data of different age groups recorded in ten schools across Estonia with high-quality recording equipment. 60 read and 10 spontaneous items were recorded from each speaker, averaging about 15 minutes of speech per speaker. In total, the corpus contains approximately 70 hours of speech.

In the present study, acoustic analyses were performed on a sub-corpus of read speech samples consisting of 20-25 utterances (average duration of utterance 7.8 seconds) per subject, totalling 6964 utterances. All utterances were manually segmented at the word and phoneme levels using Praat [3]. Syllable boundaries and types were added using a custom Praat script.

### 2.2. Acoustic analysis

In the study, fundamental frequency (F0) and speech and articulation rates were explored depending on gender and age.

#### *2.2.1. F0*

For the F0 analysis, a custom Praat script was compiled using the two-step procedure recommended in [8]. First, the F0 values of each utterance were found in the frequency range 75–600 Hz, then the range was narrowed according to the speaker's F0 variation as follows: F0max = 1.5 x 3rd quartile value, F0min = 0.75 x 1st quartile value. This approach takes into account the subject's individual F0 range and thus provides more reliable results. However, about 20% of the utterances required manual F0 validation in cases where the automatically found F0 values seemed unlikely, e.g. when boys had maximum F0 values above 400 Hz or girls had minimum F0 values below 150 Hz. There were also more F0 detection errors in the speech of boys with a pubertal voice mutation period, where F0 variations were larger and sometimes reached the falsetto register.

#### *2.2.2. Speaking rate*

To calculate the speaking rate measures, the number of syllables in each utterance, and utterance duration with intra-sentence pauses and without pauses (pauses longer than 250 ms were excluded) were found. Two speaking rate measures for each utterance were calculated:

- the speaking rate = the number of syllables in an utterance / the duration of an utterance including pauses,
- the articulation rate = the number of syllables in an utterance / the duration of an utterance without pauses.

### 2.3. Statistical analysis

For statistical data processing RStudio [21] was used with the *mgcv* package [29] for Generalized Additive Mixed Models (GAMM) and the *itsadug* package [25] for model validation and visualization of the results).

## 3. RESULTS

### 3.1. F0

The histograms (Figure 1) show the distribution of mean F0 in boys and girls. In the case of boys, the F0 distribution is binomial with clearly distinguished peaks at 110 Hz and 215 Hz. The first peak represents the most frequent mean F0 value of boys who have undergone pubertal voice change (boys aged 15–18), and the second peak corresponds to the most frequent mean F0 value of younger boys (aged 9–12) before pubertal voice change. As expected, there is only one peak in the histogram for girls, representing the median F0 value for all girls (226 Hz).

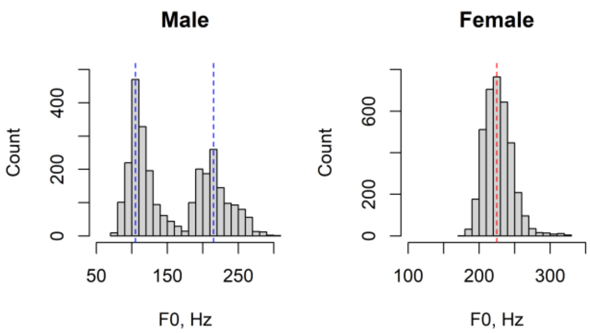


**Figure 1**: Histograms of F0 mean values, males on the left, females on the right.

Figures 2 and 3 represent the boxplots of the measured F0 means for each age group in boys and girls, respectively. In boys, the general developmental pattern of F0 mean values shows a gradual decrease between the ages 9 and 12 by 17 Hz, a more prominent decline (by 28 Hz) occurs between the ages 12–13 followed by the largest drop (by 61 Hz) between the ages 13–14, a further decrease (by 19 Hz) continues till age 15. During the ages of 15–18, the mean F0 stabilizes around 110 Hz (in pairwise comparison, the differences between the age groups are still statistically significant). However, individual F0 developmental paths can be different from the general pattern, as the outliers in Figure 3 manifest. There are two 13-years old (F0 means 101 and 113 Hz), three 14-years old (F0 means 228, 194, and 164 Hz), two 15-years old (F0 means 149 and 141 Hz), and one 16-years old (F0 mean 201 Hz) boys whose F0 mean values deviate significantly ($p<0.001$) from the other speakers in the respective age group (F0 means 199, 114, 105, and 111 Hz, respectively). We suggest that these deviating F0 mean values reveal an early (in the case of 13-year-olds) and late (in the case of 14–16-year-olds) onset of the pubertal voice mutation period.

In girls, F0 development shows an almost linear decline pattern between the ages 9 and 18, with the decrease of mean F0 from 245 Hz to 212 Hz. The outliers occurring in several female age groups could be attributed to the individual characteristics of the subjects' voice source.

Gender difference is significant ($p<0.001$) already from age 9 (F0 means 235 and 245 Hz for boys and girls, respectively).

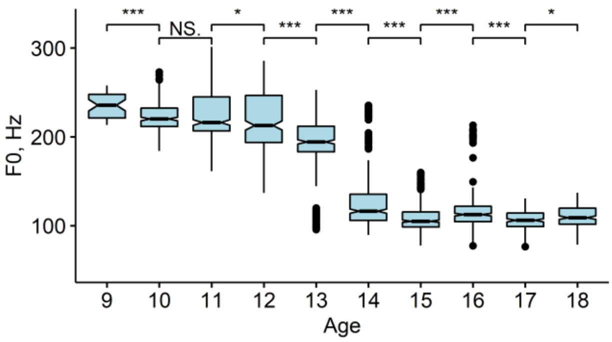


**Figure 2**: Male speakers' boxplots of F0 means by age with significance levels of pairwise t-test (NS. p > 0.05, * p ≤ 0.05, ** p ≤ 0.01, *** p ≤ 0.001).

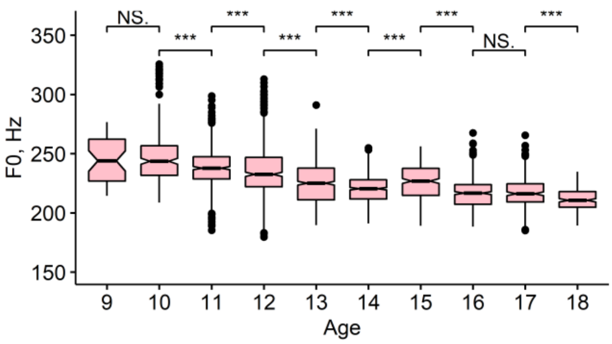


**Figure 3**: Female speakers' boxplots of F0 means by age with significance levels of pairwise t-test.

The measured F0 mean values for each read utterance were used to fit the GAMM with the smooth term *age,* the explanatory factor *gender,* and *subject* as the random effect. The model-predicted F0 values for both genders are presented in Figure 4.

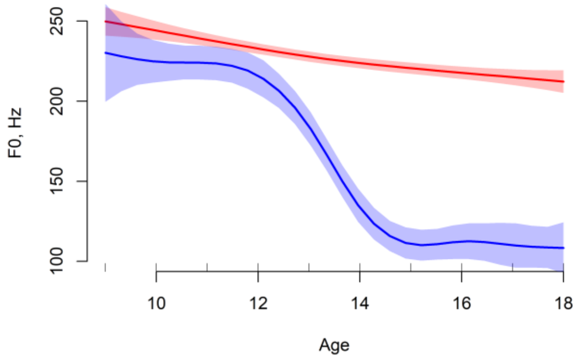


**Figure 4**: The age-dependent change of F0 mean with 95% confidence bands for girls (top) and boys (bottom) as predicted by the GAMM.

According to the GAMM's prediction, in girls between the ages of 9 and 18 years, the F0 mean gradually decreases from 250 Hz to 212 Hz. In boys between the ages 9 – 12, the mean F0 decreases by 17 Hz, followed by a rapid drop of F0 (by 105 Hz) between the ages 12–15, followed by minor F0 variations in further ages.

**Table 1.** The GAMM-predicted mean F0 values and the standard errors (SE) for each age group by gender (in Hz).

| | | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 | 17 | 18 |
|---|---|---|---|---|---|---|---|---|---|---|---|
| F | F0 | 250 | 244 | 238 | 233 | 228 | 224 | 221 | 218 | 215 | 212 |
| | SE | 4,4 | 2,7 | 1,6 | 1,2 | 1,1 | 1,1 | 1,3 | 1,6 | 2,2 | 3,4 |
| M | F0 | 230 | 225 | 224 | 216 | 184 | 134 | 111 | 112 | 110 | 108 |
| | SE | 15,6 | 6,2 | 4,9 | 5,5 | 4,8 | 4,8 | 4,4 | 5,1 | 6,7 | 7,8 |

### 3.1. Speaking rate characteristics

GAMMs were fitted for the speech and articulation rates with the smooth terms *age* and *text length*, and the explanatory factor *gender*; *subject* was added as an independent random variable. The GAMM-predicted speech and articulation rates are presented in Table 2 and the plots of the smooth terms are shown in Figure 5.

**Table 2.** The GAMM-predicted articulation and speech rates and the standard errors (SE) for each age group (in syllables per second).

| | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 | 17 | 18 |
|---|---|---|---|---|---|---|---|---|---|---|
| Art. rate | 4.0 | 4.3 | 4.6 | 4.9 | 5.0 | 5.2 | 5.3 | 5.3 | 5.2 | 5.2 |
| SE | 0.13 | 0.07 | 0.04 | 0.04 | 0.04 | 0.04 | 0.04 | 0.05 | 0.06 | 0.09 |
| Sp. rate | 3.8 | 4.2 | 4.5 | 4.8 | 5.0 | 5.1 | 5.2 | 5.2 | 5.2 | 5.1 |
| SE | 0.14 | 0.07 | 0.05 | 0.04 | 0.04 | 0.04 | 0.04 | 0.05 | 0.06 | 0.09 |

The data in Table 2 confirm similar age-related development patterns of both speaking rate characteristics (Figure 5, left): at the age 9–15 both rates increase significantly (pairwise comparison of age groups is statistically significant, $p<0.001$) and reach a maximum at 15–16 years of age, in ages 17–18 both rates slightly decrease ($p<0.05$). The effect of text length on the speaking rates has similar patterns (Figure 5, right): in sentences with 8–12 syllables speaking rates increase with the text length (differences are significant at least at the $p < 0.05$ level), in sentences with 12–18 syllables the speaking rates are stable, followed by a decreasing trend (differences are insignificant) in longer sentences (19–20 syllables). There is no gender difference in both rates ($p>0.5$).

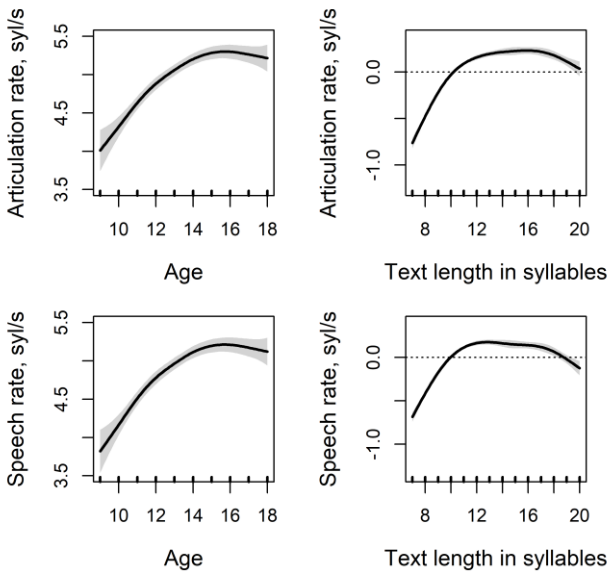


**Figure 5.** The partial effect of features fitted with the GAMM models for articulation rate (top) and speech rate (bottom) with 95% confidence bands.

## 4. DISCUSSION

In the study, we explored the developmental characteristics of F0 and two speaking rate measures in Estonian adolescent speech. Due to the cross-sectional nature of the corpus, the findings reveal general patterns of the age- and gender-related development of the studied acoustic characteristics.

According to GAMM prediction, the largest drop in boys' F0 (by 105 Hz) occurs between age 12 (F0 = 216 Hz) and age 15 (F0 = 111 Hz). The finding is in line with [13], which suggests that pubertal voice change in male speakers starts between ages 12 and 13, and ends around age 15. Similar to previous studies [9, 28], individual differences in 13–16-year-old males have been found in our corpus, with mean F0 values that significantly differ from those of age-matched peers suggesting that the onset of puberty varies among speakers.

Both articulation and speech rates increase between the ages of 9 and 15 and become stable at further ages. The difference between articulation rate and speech rate in the analyzed speech material is rather small as there are considerably fewer pauses in read speech compared to spontaneous speech. The effect of text length on speech rate characteristics was significant for sentences up to 12 syllables long. In texts longer than this, the articulation rate shows a further rising trend up to 18-syllable sentences followed by a slight decrease in 19–20-syllable sentences, while the speech rate decreases in 13–20-syllable sentences because speakers make more pauses when reading longer texts. Such patterns of speaking rate variations may be related to motor planning in text reading.

According to previous research, the temporal characteristics of children's speech are acoustically more variable up to the age of 12 years [13] and articulatory up to the age of 14 years [24]. These results have been interpreted as evidence of the achievement of adult-like speech-motor skills at ages 12–14 [18]. On the other hand, it has been argued that the development of motor patterns continues into late adolescence [23]. In line with the latter, it has been reported that the speech rate of 13–14-year-old British children has not yet reached adult levels [6], and the speech rate of 13–17-year-old Hebrew-speaking children continues to increase [1]. The average speech rate of 14–18-year-old Estonian speakers (5.1 syllables per second) compared well to the read speech rate of Estonian young adults of 4.9–5.3 syllables per second [17].

Based on the discussion above and our results, we suggest that the speech-motor control of Estonian adolescents achieve adult-like skills between the ages of 14 and 15 with further improvement of proficiency in ages 15–18.

## 5. SUMMARY

The study explored the changes in F0 and two speaking rate measures in Estonian adolescent speech as a function of age and gender. The discovered developmental patterns – (1) a decline in F0 with a sharp drop of about 100 Hz in boys aged 12–15 years and a gradual decline in girls, (2) an increase in speech and articulation rates up to 15 years of age without gender-specific differences – are similar to the findings reported for several other languages.

The results of the study will further the knowledge about the age- and gender-related variability of the acoustic properties of adolescent speech and can be considered reference data that are typical for Estonian-speaking individuals aged 9–18 years with normal language development.

# 6. ACKNOWLEDGEMENTS

This study was supported by the ERDF (the project "Centre of Excellence in Estonian Studies") and by the national programme “Estonian Language Technology 2018-2027” (the project “Speech Recognition”).


# 7. REFERENCES


[1] Amir, O., Grinfeld, D. 2011. Articulation Rate in Childhood and Adolescence: Hebrew Speakers. *Language and Speech*, 54(2), 225–240.

[2] Bishop, J., Kim, B. 2018. Anticipatory shortening: Articulation rate, phrase length, and lookahead in speech production. *Proceedings of 9th International Conference on Speech Prosody 2018*, 235–239.

[3] Boersma, P., Weenink, D. 2022. Praat: doing phonetics by computer [Computer program]. Version 6.3., retrieved 15 November 2022 from http://www.praat.org//

[4] Darling-White, M., Banks, S. W. 2021. Speech Rate Varies with Sentence Length in Typically Developing Children. *Journal of Speech, Language, and Hearing Research*, 64 (6S), 2385–2391.

[5] Draxler, C., Schiel, F., Ellbogen, T. 2008. F0 of adolescent speakers: First results for the German Ph@ttSessionz Database. *Proceedings of the 6th International Conference on Language Resources and Evaluation*. European Language Resources Assoc. http://www.lrec-conf.org/proceedings/lrec2008/

[6] Hazan, V.; Pettinato, M. 2014. The emergence of rhythmic strategies for clarifying speech: Variation of syllable rate and pausing in adults, children and teenagers. *10th International Speech Production Seminar*, Köln, Germany.

[7] Hirano, M., Kurita, S., Toh, Y. 1981. Growth, Development and Aging of the Vocal Fold, *Practica Oto-Rhino-Laryngologica*, 74(8), 1791–180.

[8] Hirst, D. 2007. A Praat plugin for Momel and INTSINT with improved algorithms for modelling and coding intonation. *Proceedings of the XVIth International Conference of Phonetic Sciences*, 1233–1236.

[9] Hollien, H., Green, R., and Massey, K. 1994. Longitudinal research on adolescent voice change in males. *J. Acoust. Soc. Am*. 96, 2646–2654.

[10] Jacewicz, E., Fox, R. A., O’Neill, C. 2009. Articulation rate across dialect, age, and gender. *Language Variation and Change*, 21(2), 233–256.

[11] Jacewicz, E., Fox, R. A, Wei, L. 2010. Between-speaker and within-speaker variation in speech tempo of American English. *Journal of the Acoustical Society of America*, 128(2), 839–850.

[12] Lee, A., Doherty, R. 2017. Speaking rate and articulation rate of native speakers of Irish English. *Speech, Language and Hearing*, 20(4), 206–211.

[13] Lee, S., Potamianos, A., Narayanan, S. 1999. Acoustics of children’s speech: Developmental changes of temporal and spectral parameters. *J. Acoust. Soc. Am*. vol. 105, no. 3, 1455–1468.

[14] Logan, K. J., Byrd, Courtney T., Mazzocchi, E. M., Gillam, R. B. 2011. Speaking rate characteristics of elementary-school-aged children who do and do not stutter. *Journal of Communication Disorders*, 44(1), 130–147.

[15] Meister, E., Meister, L. 2021. Developmental changes of vowel acoustics in adolescents. *Interspeech 2021: Brno, Czech Republic.* International Speech Communication Association, 4014−4018.

[16] Meister, L., Meister, E. 2014. Development of the corpus of Estonian Adolescent Speech. *Human Language Technologies – the Baltic Perspective: Proceedings of the Sixth International Baltic Conference, Baltic HLT 2014.* Amsterdam: IOS Press, 206−209.

[17] Meister, E., Meister, L. 2022. Estonian adolescent speech III: analysis of speech tempo and syllable durations [in Estonian]. Keel ja Kirjandus, 3, 226−244.

[18] Redford, M. A.; Oh, G. E. 2017. The representation and execution of articulatory timing in first and second language acquisition. *Journal of Phonetics*, 63(1) 127–138.

[19] Robb, M. P., Maclagan, M. A., Chen, Y. 2004. Speaking rates of American and New Zealand varieties of English. *Clinical Linguistics & Phonetics*, 18(1), 1–15.

[20] Rogers, D. J., Setlur, J., Raol, N., Maurer, R., Hartnick, C.J. 2014. Evaluation of true vocal fold growth as a function of age. *Otolaryngol Head Neck Surg*. 151(4), 681–687. doi: 10.1177/0194599814547489.

[21] RStudio Team. 2020. RStudio: Integrated Development for R, RStudio, PBC, Boston, MA. http://www.rstudio.com/.

[22] Sadagopan, N., Smith, A. 2008. Developmental Changes in the Effects of Utterance Length and Complexity on Speech Movement Variability. *Journal of Speech, Language, and Hearing Research*, 51, 1138–1151.

[23] Smith, A. 2006. Speech motor development: Integrating muscles, movements, and linguistic units. *Journal of Communication Disorders*, 39(5), 331–349.

[24] Smith, A., Zelaznik, H. N. 2004. Development of functional synergies for speech motor coordination in childhood and adolescence. *Developmental Psychobiology*, 45(1), 22–33.

[25] van Rij, J., Wieling, M., Baayen, R., van Rijn, H. 2022. itsadug: Interpreting Time Series and Autocorrelated Data Using GAMMs. R package version 2.4.1.

[26] Vorperian, H. K., Kent, R. D. 2007. Vowel acoustic space development in children: A synthesis of acoustic and anatomic data. *Journal of Speech, Language, and Hearing Research*, vol. 50, no. 6, 1510–1545.

[27] Whiteside, S.P. 2001. Sex-specific fundamental and formant frequency patterns in a cross-sectional study. *J. Acoust. Soc. Am*., vol. 110, no. 1, 464–478.

[28] Whiteside, S. P, Hodgson, C., Tapster, C. 2002. Vocal characteristics in pre-adolescent and adolescent children: a longitudinal study. *Logopedics Phoniatrics Vocology,* 27(1), 12–20.

[29] Wood, S. N. 2017. Generalized Additive Models: An Introduction with R (2nd ed.), CRC Press.