

# The generalised plug-in algorithm for the diffeomorphism kernel estimate

Molka Troudi

Département Méthodes quantitatives  
IHEC Carthage  
2016, Carthage, Tunisia  
molkaghorbel@gmail.com

Faouzi Ghorbel

Laboratoire Cristal, ENSI  
Campus Sciences universitaires de la Manouba  
2010, La Manouba, Tunisia  
Faouzi.ghorbel@ensi.rnu.tn

Received: August 7, 2019. Revised: August 25, 2021. Accepted: September 29, 2021. Published: November 27, 2021.

**Abstract**— The optimal value of the smoothing parameter of the Kernel estimator can be obtained by the well known Plug-in algorithm. The optimality is realised in the sense of Mean Integrated Square Error (MISE). In this paper, we propose to generalise this algorithm to the case of the difficult problem of the estimation of a distribution which has a bounded support. The proposed algorithm consists in searching the optimal smoothing parameter by iterations from the expression of MISE of the kernel-diffeomorphism estimator. By some simulations applied to some distribution having a support bounded and semi bounded, we show that the support of the pdf estimator respects the one of the theoretical distribution. We also prove that the proposed method minimizes the Gibbs phenomenon.

**Keywords**—Probability density function, Diffeomorphism Kernel Estimate, Kernel estimate, non parametric estimator, generalised plug-in Algorithm, smoothing parameter, smoothing parameter.

## I. INTRODUCTION

The application of a reliable pdf estimator gives improvements for the systems performances. For examples, the optimal scalar quantification which is based on the pdf estimates is an important step in Signal and image coder. The advanced hashing procedure which is known as an essential task in content image basis indexing improves its performances when the pdf's of image features are well estimated. When the application of the Bayesian classification rule in pattern recognition systems become possible. The determination of the conditional pdf and the mixture one is needed.... The coder parameters, the used features in data base index systems or the shape descriptors could be confined to a bounded or a semi bounded intervals. Densities estimated using the classical estimators beyond their natural support. The orthogonal series method provides an interesting solution to this problem [1]. However, the pdf estimate of such bounded or semi bounded attributes which are modeled by a set of random variables, have some convergence problems in its border values known by the Gibbs phenomenon. For these raisons, some authors have recently developed new non parametric pdf estimate methods taking account of the data support. The Diffeomorphism kernel estimate is one of this pdf estimate kinds[2][3][4]. In the present work, we propose an improvement of such method by optimizing its smoothing

parameter value in the mean of the Mean Integrate Square Error (MISE).

Several methods have been developed in the literature to optimize the smoothing parameter of the standard kernel estimator [5], [6],[7],[8],[9].

We focus in this paper on the plug-in method [7] which gives a good approximation of the optimal smoothing parameter in the mean integrated square error (MISE) sense. We propose to generalize this method to the diffeomorphism kernel estimate. The standard plug-in algorithm becomes a special case when the chosen diffeomorphism is the identity. The proposed estimator is tested on two kinds of distributions: distributions on an interval and those on a semi bounded support.

The present paper is organized as follows. Section 2 is devoted to recall briefly the Kernel pdf estimate method. In Section 3, the theoretical principles of the modified KDE which is adapted to the probability density functions with a bounded support are presented. The convergence according to mean square error criterion gives a sufficient condition so that the estimator converges in terms of the integrated mean square error (IMSE). An asymptotic study is developed in section 4. So, the expression of the optimal smoothing parameter is presented according to IMSE criterion. In section 5, we describe the different steps of the iterative plug-in algorithm witch converges to the optimal smoothing parameter. Therefore, section 6 is devoted to evaluate the performances of the suggested. The results are illustrated by two distributions: A beta distribution for the case of a distribution with bounded support and exponential distribution for the case distribution having semi-bounded support. The conclusion of this paper is presented at the last section where some perspectives are also presented.

## II. BRIEF OVERVIEW OF THE KERNEL PDF ESTIMATE

Let denote by  $(X_i)_{1 \leq i \leq N}$  a sample of absolute continue random variable  $X$  with size  $N$ . Let denote by  $f$  the probability density function (pdf) of  $X$ . The well known Kernel pdf Estimator of  $f$ [10] can be written as follow:

$$\hat{f}_N(x) = \frac{1}{Nh_N} \sum_{i=1}^N K\left(\frac{x - X_i}{h_N}\right) \quad (1)$$

$h_N$  is generally called the smoothing parameter and  $K$  the Kernel.  $K$  is a pdf function and it is assumed to be an even regular function with unit variance and zero mean.

The performance evaluation of estimators is usually based on a distance between the theoretical pdf and its estimator  $\hat{f}_N$ . The Mean Integrated Square Error (MISE) is one of the most important criterions which can be expressed as:

$$MISE(f, \hat{f}_N) = D_2(\hat{f}_N, f) = E\left[\int_{-\infty}^{+\infty} |\hat{f}_N(x) - f(x)|^2 dx\right]$$

where  $E$  is the expectation operator.

The minimization of MISE with respect to the smoothing parameter, for a fixed size  $N$  of the sample, implies the following asymptotic study.

$$\begin{aligned} E\left[|\hat{f}_N - f|^2\right] &= \text{var}(\hat{f}_N) + \left(f - E[\hat{f}_N]\right)^2 \\ &= \frac{1}{Nh_N} \int K^2(u) f(x - h_N u) du \\ &\quad + \left[\int K(u) (f(x - uh_N) - f) du\right]^2 \\ &\quad - \frac{1}{N} \left(\int K(u) f(x - h_N u) du\right)^2 \end{aligned}$$

By consider the Taylor pdf expansion:

$$\begin{aligned} f(x - h_N u) &= f(x) - h_N u f'(x) + \frac{u^2}{2} h_N^2 f''(x) - \\ &\quad \frac{u^3 h_N^3}{6} f^{(3)}(x - \theta h_N u) \end{aligned}$$

where  $0 < \theta < 1$ .

Let denote by

$$M(K) = \int_{-\infty}^{+\infty} K^2(u) du \quad (2)$$

and

$$J(f) = \int_{-\infty}^{+\infty} (f''(x))^2 dx \quad (3)$$

where  $f''$  is the second derivative of the function  $f$ .  $\Delta(h_N)$  which is the Taylor expansion of the MISE (and consequently an approximation of MISE) is given by:

$$MISE \approx \Delta(h_N) = \frac{M(K)}{nh_N} + \frac{J(f)h_N^4}{4}$$

The minimum value of  $\Delta(h_N)$  is obtained for the following value

$$h_N^* = n^{-\frac{1}{5}} \cdot (J(f))^{-\frac{1}{5}} \cdot (M(K))^{\frac{1}{5}} \quad (4)$$

Thus the minimum value of the MISE is given by the following expression:

$$MISE = \frac{5}{4} N^{-\frac{4}{5}} (M(K))^{\frac{4}{5}} (J(f))^{\frac{1}{5}} \quad (5)$$

### III. PRINCIPLES OF THE DIFFEOMORPHISM KERNEL ESTIMATE

The Diffeomorphism Kernel Density Estimate (DKDE) is introduced in [4] [5] [6]. It is based on appropriately diffeomorphism. Let consider now random variables which have a bounded support  $]a, b[$  and  $\phi$  a  $C^1$ -diffeomorphism from  $\mathbb{R}$  to  $\mathbb{R}$ . The following estimator:

$$\hat{f}_N(x) = \frac{|\phi'(x)|}{Nh_N} \sum_{i=1}^N K\left(\frac{\phi(x) - \phi(X_i)}{h_N}\right) \quad (6)$$

is asymptotically unbiased when  $h_N$  tends towards 0 and  $\phi'(x)$  tends towards infinity when  $x$  tends towards  $a$  or  $b$  which are the bounds of the interval  $]a, b[$ . The expectation of the suggested estimator is estimated by:

$$E[\hat{f}_N(x)] = \frac{|\phi'(x)|}{Nh_N} \sum_{i=1}^N E\left[K\left(\frac{\phi(x) - \phi(X_i)}{h_N}\right)\right]$$

Using the following change of variable,  $y = \frac{\phi(x) - \phi(u)}{h_N}$

the expression of variance becomes:

$$\begin{aligned} E[\hat{f}_N(x)] &= |\phi'(x)| \\ &\quad \int_{\mathbb{R}} K(y) f \circ \phi^{-1}(\phi(x) - uh_N) \underbrace{\left|(\phi^{-1})'(\phi(x) - yh_N)\right|}_{g(x,y)} dy \end{aligned}$$

Let us compute the variance of this estimator by using the same change of variable:

$$\begin{aligned} \text{var}[\hat{f}_N(x)] &= \frac{|\phi'(x)|^2}{Nh_N} \\ &\quad \int_{\mathbb{R}} K^2(y) f \circ \phi^{-1}(\phi(x) - uh_N) \underbrace{\left|(\phi^{-1})'(\phi(x) - yh_N)\right|}_{g(x,y)} dy \end{aligned}$$

$$- \frac{1}{N} \left\{ E[\hat{f}_N(x)] \right\}^2$$

The mean square error (MSE) is

IV. ASYMPTOTIC STUDY

The quality of the pdf estimation depends on the choice of the optimal smoothing parameter or bandwidth  $h_N$ . The MSE expression can be written as following:

$$E \left[ \left| \hat{f}_N(x) - f(x) \right|^2 \right] = \text{var} \left[ \hat{f}_N(x) \right] + \left\{ E \left[ \hat{f}_N(x) \right] \right\}^2 - 2f(x)E \left[ \hat{f}_N(x) \right] + f^2(x)$$

The MSE becomes:

$$E \left[ \left| \hat{f}_N(x) - f(x) \right|^2 \right] = \frac{|\phi'(x)|^2}{Nh_N} \left\{ \int_R K^2(y) g(x, y) dy - h_N \left[ \int_R K(y) g(x, y) dy \right]^2 \right\} + \left\{ \int_R K(y) (|\phi'(x)| g_N(x, y) - f(x)) dy \right\}^2$$

As  $(\phi^{-1})'$  is bounded on  $R$ ,  $f$  is assumed to be bounded on  $[a, b]$  and  $K^2$  is integrable on  $R$ , the Lebesgue convergence theorem can be easily applied, then the MSE, for a large value of  $N$ , becomes equivalent to:

$$E \left[ \left| \hat{f}_N(x) - f(x) \right|^2 \right] = \frac{|\phi'(x)| f(x)}{Nh_N} \int_R K^2(y) dy - \frac{f^2(x)}{N} + o(h_N)$$

Because of the continuity of  $\phi'$  on  $R$  and  $f$  on  $[a, b]$ , this estimator converges in IMSE for all compact of  $[a, b]$ . To obtain this convergence according to the IMSE criterion, the function  $\phi'(f)$  have to be integrable on  $[a, b]$  because:

$$\int_a^b E \left[ \left| \hat{f}_N(x) - f(x) \right|^2 \right] dx = \frac{1}{Nh_N} \int_a^b |f(x) \phi'(x)| dx \int_R K^2(y) dy - \frac{1}{N} \int_a^b f^2(x) dx + (b-a) o(h_N)$$

Saoudi and al. [3] shows that the logarithmic diffeomorphism allows a better convergence of the estimator.

$$\phi_{a,b} : ]a, b[ \rightarrow R$$

$$x \rightarrow \text{Log} \left( \frac{x-a}{b-x} \right)$$

$$E \left[ \left| \hat{f}_N(x) - f(x) \right|^2 \right] = A_N(x) + B_N(x) - C_N(x)$$

With

$$A_N(x) = \frac{|\phi'(x)|^2}{Nh_N} \int_R K^2(y) g(x, y) dy$$

$$B_N(x) = \left\{ \int_R K(y) [|\phi'| g(x, y) - f(x)] dy \right\}^2$$

$$C_N(x) = \frac{|\phi'(x)|^2}{N} \left\{ \int_R K(y) g(x, y) dy \right\}^2$$

We consider Taylor expansion of the function  $H_y$  defined as following in the neighborhood of  $\phi(x)$ :

$$\phi(x) \xrightarrow{H_y} f \circ \phi^{-1}(\phi(x) - yh_N) \left| (\phi^{-1})'(\phi(x) - yh_N) \right|$$

It implies that there exists a positive number  $\theta$  less than 1 such that:

$$H_y(\phi(x) - yh_N) = H_y(\phi(x)) - yh_N H'_y(\phi(x)) + \frac{y^2 h_N^2}{2} H''_y(\phi(x)) - \frac{y^3 h_N^3}{6} H'''_y(\phi(x) - \theta y h_N)$$

The following approximations are deduced from the computation of the successive derivatives of the function  $H_y$  in  $\phi(x)$ :

$$A_N(x) \approx \frac{|\phi'(x)| f(x)}{Nh_N} M(K)$$

$$B_N(x) = \frac{h_N^4}{4[\phi'(x)]^8} F^2(x)$$

$$C_N(x) = \frac{[f(x)]^2}{N}$$

with

$$F(x) = \left[ f(x) \left[ 3\phi''(x)^2 - \phi'(x)\phi'''(x) \right] \right. \\ \left. - 3f'(x)\phi'(x)\phi''(x) + f''(x) \left[ \phi'(x) \right]^2 \right] \quad (7)$$

The asymptotical study of IMSE gives:

$$D^2(\hat{f}_N, f) = \int_R [A_N(x) + B_N(x) - C_N(x)] dx \\ \approx \frac{M(K)}{Nh_N} \int_R |\phi'(x)| f(x) dx + \frac{h_N^4}{4} \int_R \frac{F^2(x)}{[\phi'(x)]^8} dx$$

If  $M_\phi$  and  $J_\phi$  exists, we have:

$$M_\phi(K) = M(K) \int_R |\phi'(x)| f(x) dx \quad (8)$$

and

$$J_\phi(f) = \int_R \frac{F^2(x)}{[\phi'(x)]^8} dx \quad (9)$$

The optimal value of  $h_N$  noted by  $h_N^*$  can be deduced by minimization of IMSE.

$$h_N^* = [M_\phi(K)]^{\frac{1}{5}} [J_\phi(f)]^{-\frac{1}{5}} N^{-\frac{1}{5}} \quad (10)$$

#### V. GENERALIZED PLUG-IN ALGORITHM

Several methods are proposed in the literature for selecting optimal smoothing parameter. The best known of these include rules of thumb, oversmoothing, least squares cross-validation, direct plug-in methods, solve-the-equation plug-in method, and the smoothed bootstrap [5],[6],[7]. A method developed in 2004 combines the technique of Least Square Cross Validation and the contrasts with the contrast methods [8]. Recently, Botev presents a new adaptive kernel density estimator based on linear diffusion processes [9]. We focus in this paper on the direct plug-in method applied to the kernel diffeomorphism application. Such a method is an iterative algorithm which converges to the optimal bandwidth.

Following, let's detail the steps of the generalized plug-in algorithm.

- Step 1: Initialization of  $M_\phi(K)$  and  $J_\phi^{(0)}(f)$ . These values can be set arbitrarily. However, we propose to choose  $M(K)$  as first value for  $M_\phi(K)$
- Step 2: It consists of the determinate of  $h_N^{(0)}$  given by the equation (10)
- Step 3: Estimation of the pdf  $f^{(0)}$  according the value  $h_N^{(0)}$  computed by the step 2 task.

- Step 4: At the  $k^{\text{th}}$  iteration, we approximate  $M_\phi(K)$  given by the (equation (8)),  $(f^{(k)})'$  and  $(f^{(k)})''$
- Step 5: Estimation of  $J_\phi(f^{(k)})$  (equation (9)) and deduction of  $h_N^{(k)}$  (equation (10))
- Step 6: Estimation of  $f^{(k)}$  (equation (6))
- Step 7: Stopping the algorithm is conditional on a low relative difference between  $h_N^{(k)}$  and  $h_N^{(k-1)}$  (less than 1%).

#### VI. PERFORMANCES OF THE GENERALIZED PLUG-IN KERNEL DIFFEOMORPHISM ESTIMATE

In this section, we intend to evaluate the performances of the generalized kernel plug-in diffeomorphism estimate. The study was conducted on two kinds of distributions:

- Distributions defined on a semi bounded support illustrated by an exponential distribution defined on  $R^+$  with  $E(X) = 1$ .
- Distributions defined on an interval illustrated by a beta distribution (parameters = (1.5,1)) which is defined on  $[0, 1]$

##### A. Semi-bounded distributions case

Figure 1 represents the estimate of the simulated exponential density by the standard kernel Density estimate (KDE) with adjustment of the smoothing parameter by the plug-in algorithm. Figure 2 represents this estimation by the generalized plug-in diffeomorphism kernel estimate (GKDE). The estimate is obviously improved by GKDE. Indeed, the estimated density approaches so nearly perfectly the theoretical density and the Gibbs phenomenon is significantly reduced. Similarly, the disturbances observed on the estimate of figure 1 are greatly reduced in Figure 2. Mean Integrated Square Error (MISE) values which are shown in the table 1 confirm these observations. Thus, for a sample size of 1000, the observed MISE value is  $37.9 \cdot 10^{-6}$  for the classical kernel estimate whereas it is  $7.76 \cdot 10^{-6}$  the generalized kernel diffeomorphism estimate.

##### B. Bounded distributions case

Figures 3 and 4 shows that the estimation of beta distribution pdf by the generalized plug-in diffeomorphism kernel estimate gives better results than those obtained by usual plug-in kernel pdf estimate. The Gibbs phenomenon is significantly reduced and the smoothing seems to be better. The MISE values versus the sample size presented in table 1 confirm these observations.

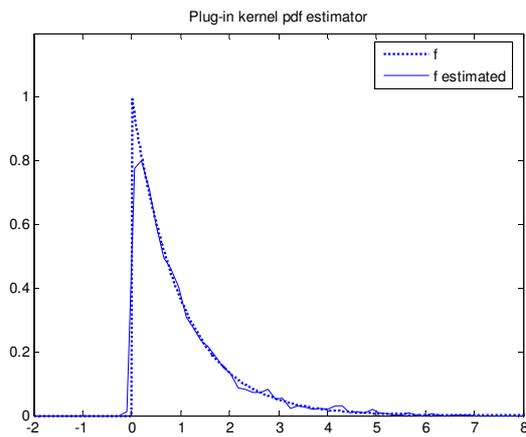


Fig. 1. pdf estimation of an exponential distribution by Plug-in Kernel Density estimator (KDE)

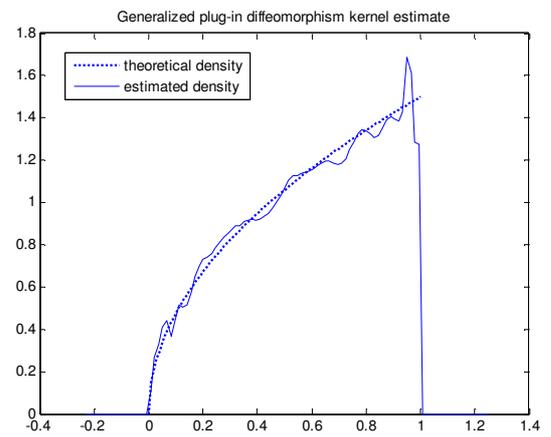


Fig. 4. pdf estimation of a beta distribution by Plug-in Diffeomorphism Kernel Density estimator (DKDE)

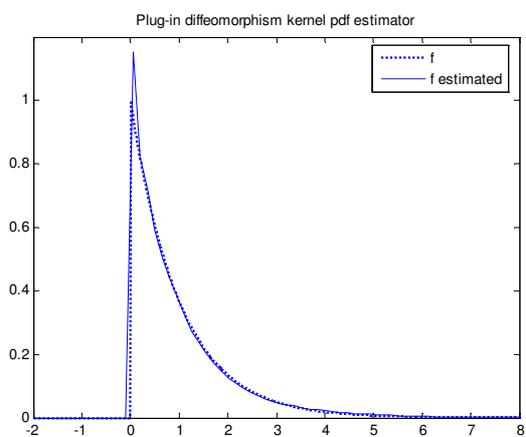


Fig. 2. pdf estimation of an exponential distribution by Plug-in Diffeomorphism Kernel Density Estimator (DKDE)

TABLE I. MISE VALUES VERSUS SAMPLE SIZE

MISE for exponential pdf		
Sample size	Plug-in KDE*10 <sup>-6</sup>	GKDE*10 <sup>-6</sup>
1000	37.9	7.76
2000	26.8	6.01
3000	20.0	5.51
4000	9.06	3.13
5000	7.80	2.81
MISE for beta pdf		
Sample size	Plug-in KDE	Plug-in DKDE
1000	0.3016	0.2487
2000	0.2896	0.2511
3000	0.2827	0.2484
4000	0.2820	0.2530
5000	0.2818	0.2533

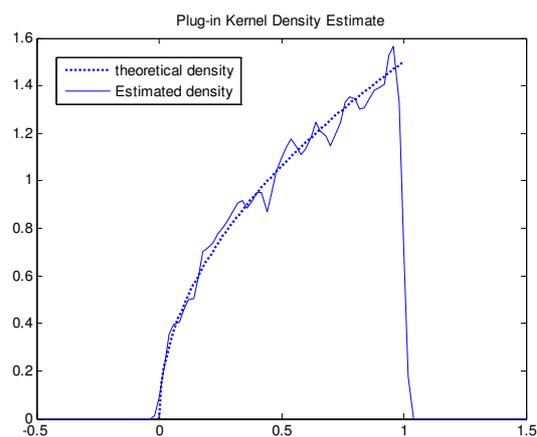


Fig. 3. pdf estimation of a beta distribution by Plug-in Kernel Density estimator (KDE)

## VII. CONCLUSION

In this work, we have generalized the plug-in algorithm which adjusts the smoothing parameter of the kernel pdf estimate, to the diffeomorphism kernel estimate version. Such modified plug-in algorithm comes from the optimization of the MISE of this estimate. This generalization gives a more complicated iterative algorithm since the values of two parameters depending on the unknown pdf have to be approximated along iterations instead of only one parameter on the classical plug-in. It is important to note that the convergence is obtained for the proposed algorithm. By simulations concerning different kinds of distributions confined to bounded or semi bounded supports, we illustrate the better performance of the proposed Plug-in Diffeomorphism Kernel pdf estimate in the sense of MISE.

In our future works, we intend to study the case of multivariate bounded support distributions. We also test this well performance estimate in real data.

REFERENCES

- [1] P. Hall, "Comparison of two orthogonal series methods of estimating a density and its derivatives on interval". *J. Multivariate anal.*, vol. 12, pp. 432 – 449, 1982.
- [2] S. Saoudi, F. Ghorbel, A. Hillion, "Non parametric probability density function estimation on a bounded support: applications to shape classification and speech coding". *Applied Stochastic Models and Data Analysis*, vol. 10, pp. 215-231, 1994.
- [3] S. Saoudi, F. Ghorbel, A. Hillion, "Some statistical properties of the Kernel-diffeomorphism estimator", *Applied Stochastic Models and Data Analysis*, vol. 10, pp. 39-58, 1997.
- [4] F. Ghorbel, *Vers une approche unifiée des aspects géométriques et statistiques de la reconnaissance des formes planes*, s<sup>d</sup> ed. Arts-Pi editions, Tunis, 2011.
- [5] M.C. Jones, J.S Marron, S.J. Seather, "A brief survey of bandwidth selection for density estimation". *J. Amer. Stat. Assoc.* , vol.91, pp. 401 – 407, 1996.
- [6] A.W. Bowman, A. Azzalini, *Applied Smoothing Techniques for Data Analysis*. Oxford University Press, 1997.
- [7] P. Hall, J.S. Marron, " Estimation of integrated squared density derivatives", *Statistics & Probability letters*, vol. 6, pp. 109 – 115, 1987.
- [8] A.R. Mugadi, I.A. Ahmad , "A bandwidth selection for kernel density estimation of functions of random variables", *Computational statistics and data analysis*, vol. 47, pp. 49-62, 2004.
- [9] Z. I . Botev, J.F. Grotowski ,D.P. Kroese , "Kernel density estimation via diffusion", *The annals of statistics*, vol.38, n° 5, pp. 2916 – 2957, 2010.
- [10] E. Parzen, "On estimation of a probability density function and mode", *Annals of mathematical statistics*, vol. 33, pp. 1065-1076, 1962.
- [11] Ghorbel, F., Derrode, S., Alata, O. (ed.): *Récents avancées en reconnaissance de forme statistique*. Arts-Pi editions, Tunis (2012)
- [12] Troudi, M., Alimi, A.M., Saoudi, S.: Analytical Plug-in Method for Kernel Density Estimator Applied to Genetic Neutrality Study. *Eurasip Journal of advances in Signal Processing (Eurasip-JASP)*, Volume 2008, Article ID 739082, 8 pages doi: 10.1155/2008/739082.

**Creative Commons Attribution License 4.0  
(Attribution 4.0 International, CC BY 4.0)**

This article is published under the terms of the Creative Commons Attribution License 4.0  
[https://creativecommons.org/licenses/by/4.0/deed.en\\_US](https://creativecommons.org/licenses/by/4.0/deed.en_US)