Stability of attractor local dimension estimates in non-Axiom A dynamical systems

Flavio Maria Emanuele Pons, Gabriele Messori, Davide Faranda

Stability of attractor local dimension estimates in non-Axiom A dynamical systems

Flavio Maria Emanuele Pons *[1], Gabriele Messori[2,3], and Davide Faranda[1,4]

[1] *LSCE-IPSL, CEA Saclay l'Orme des Merisiers, CNRS UMR 8212 CEA-CNRS-UVSQ, Université Paris-Saclay, 91191 Gif-sur-Yvette, France*

[2] *Department of Earth Sciences and Centre of Natural Hazards and Disaster Science (CNDS), Uppsala University, Uppsala, 752 36, Sweden*

[3] *Department of Meteorology and Bolin Centre for Climate Research, Stockholm University, 106 91, Stockholm, Sweden*

[4] *London Mathematical Laboratory, 14 Buckingham Street, London, WC2N 6DF, UK*

flavio.pons@lsce.ipsl.fr

Abstract

We study different extreme value theory (EVT)-based estimators for the local Hausdorff dimension (also known as local attractor dimension) of dynamical systems. The attractor dimension is an important quantity related to the number of effective degrees of freedom of the underlying dynamical system, and its estimation has been a central topic in the dynamical systems literature since the 80s. The framework considered here combines the analysis of recurrences in phase space with EVT to estimate the local attractor dimension in the neighborhood of a state of interest. While the EVT framework enables the analysis of high-dimensional complex systems, such as the Earth's climate, its applicability relies on robust statistical parameter estimation for the assumed extreme value distribution. In this study,

*Corresponding author
email: flavio.pons@lsce.ipsl.fr

we provide a critical review of several EVT-based local dimension estimators, analyzing and comparing their performance across a range of data-generating systems. Our results provide insights for researchers using the EVT-based estimates of the local dimension, aiding in the selection of an appropriate estimator for their specific application.

Dynamical systems are characterized by an attractor, i.e. a compact region embedded in the space of the physical variables, hosting all the system's trajectories. The attractor is a fractal object, thus its dimension is not integer. In real-world systems, such as the Earth's climate, the local dimension of the attractor around a point corresponding to a state of interest provides insight on the predictability of the said state. Recent applications of statistical extreme value theory have led to developing estimators of the attractor dimension. These have been used in several studies, especially in geophysics, sometimes making strong simplifying assumptions that may not be met in the real world. In this article, we compare the performance of several of these estimators, providing some general guidelines for their use in applied studies.

1 Introduction

Dynamical systems theory is a powerful mathematical framework that provides information on the behavior of time-evolving systems, from something as simple as the motion of a pendulum to as complex as the global climate system. The underlying idea is to characterize all possible trajectories of a system, including their density, recurrence properties and persistence. The object which hosts all the information on these trajectories is called the attractor of the system. It lives in the phase space, i.e. the space spanned by all the variables of the system. For natural chaotic systems originating from forced-dissipative dynamics, the attractor is a compact object in phase space: the trajectories have finite energy and cannot escape from the preferential region defined by the attractor.

The geometrical properties of the attractor of a dynamical system provide a number of insights on the system's behaviour. Here, we specifically consider the Hausdorff dimension, also known

simply as attractor dimension, D. The Hausdorff dimension is an important quantity related to the number of effective degrees of freedom of the underlying dynamical system, whose estimation has been a central topic in the dynamical systems literature since the 80s. Early efforts to estimate attractor dimensions were based on the embedding approach (Grassberger, 1988). This consists of performing a nearest-neighbours search in spaces of dimensions $k \ll n$, where n indicates the number of variables (or, equivalently, the dimension of the full phase space). While the embedding approach is very successful when applied to low-dimensional attractors, it shows severe limitations for high-dimensional systems, where the choice of the embedding dimension k appears to be crucial. It was shown that, for $k > 6$, the so-called curse of dimensionality strongly reduces the capability to identify neighbors in phase space, resulting in biased estimates of $D \simeq k$ (Grassberger and Procaccia, 1984; Nicolis and Nicolis, 1984; Grassberger, 1986; Lorenz, 1991). For several years, estimating D in high-dimensional systems remained a major numerical challenge. Recently, a technique combining the analysis of recurrences in phase space with extreme value theory (EVT), has allowed to estimate D without relying on embedding (Lucarini et al., 2016). The key idea is that the limiting distribution of suitably re-scaled recurrences of the system around a state of interest ζ can be modelled by the generalized extreme value distribution or by the generalized Pareto distribution. In other words, recurrences in phase space can be viewed as extreme events and follow an extreme value law. The scaling of the hypersphere centered around ζ and containing the recurrences serves as an estimate of the *local* attractor dimension around the state ζ, $d(\zeta)$. When a sufficient number of states ζ is considered, the hyperspheres densely cover the attractor and, by averaging d over several states ζ, one obtains an estimate for D. Therefore, the EVT approach offers a way to measure D, while providing valuable information on the dynamical system in a local (in phase space) and instantaneous (in time) sense. More specifically, the local dimension informs on the geometry of the system's trajectories in the neighborhood of the state of interest. In general terms, a larger dimension is associated to a lack of predictability and a lower dimension to higher predictability.

The above EVT-based technique also suffers from some limitations linked to the curse of dimen-

sionality (Pons et al., 2020). Nonetheless, its applicability to high dimensional complex systems, and in particular the Earth system, has provided a new dimensionality reduction strategy and enabled new research directions. These include dynamical characterisations of global climate states (Faranda et al., 2017a; Buschow and Friederichs, 2018; Brunetti et al., 2019), key mechanisms of climate change (Faranda et al., 2019a), evaluation of climate models (Rodrigues et al., 2018; Falasca and Bracco, 2022), and investigations of atmospheric predictability (Messori et al., 2017; Scher and Messori, 2018; Hochman et al., 2019a,b, 2021b, 2022), atmospheric dynamics (Faranda et al., 2017b, 2019b; Messori et al., 2021; Hochman et al., 2021a), dynamics and attribution of climate extremes (De Luca et al., 2020; Giamalaki et al., 2021; Faranda et al., 2023; Ginesta et al., 2022), palaeoclimates (Messori and Faranda, 2021), slow earthquakes (Gualandi et al., 2020) and more. These studies chiefly leveraged the local information provided by d to discriminate between different states (or, more properly, of special Poincaré sections) of the climate system. As a concrete example, low dimensions of sea-level pressure maps in the North Atlantic region were associated to frequent extratropical storms, and high dimensions to so-called blocked flows, namely times when the climatological west to east atmospheric flow over the region is diverted by a persistent high pressure system (Faranda et al., 2017b).

Given the broad range of applications of d, it is important to choose a robust statistical estimator. Indeed, a range of different estimators, all based on extreme value laws, can be used to obtain d, yet a systematic comparison is lacking in the literature. The goal of this paper is to present different EVT-based estimators for d, and provide a critical review of the drawbacks and advantages of each of them, by comparing their performances on a range of data-generating systems.

Our study is structured as follows: in Section 2 we introduce the concept of local and global (Hausdorff) attractor dimension and its estimation via the Generalized Pareto distribution, and Section 3 provides an overview of different statistical estimators proposed in the literature. In Section 4 we describe the generated and real-world data that we use for the statistical evaluation of all the estimators, whose results are described in Section 5. Finally, Section 6 contains our

conclusions and recommendations on the estimation of the local dimension of dynamical systems.

2 Theoretical Framework

2.1 Hausdorff and local attractor dimension via EVT

The Hausdorff dimension is the extension of the classical concept of topological dimension to non-integer values, introduced in 1918 by Hausdorff (Hausdorff, 1918) as a way to measure the "size" of fractal objects.

For example, the Hausdorff dimension of a straight line is 1, of a volume is 3, while for a fractal object like the Koch curve it is greater than 1, but less than 2. Local dimensions are a way to measure the dimension of a fractal object at a specific point in the phase space, i.e. around a specific instantaneous state.

A classical method to estimate the Hausdorff dimension and local dimensions of a fractal object is box counting. This involves covering the fractal object with boxes of a given size and counting the number of boxes needed to cover the object. By repeating this process for boxes of different sizes, one obtains a curve that represents the relationship between the box size and the number of boxes needed to cover the object. The slope of this curve can be used to estimate the Hausdorff dimension and local dimensions of the fractal object.

Despite its usefulness in estimating the attractor dimension of low-dimensional systems, the box counting method is limited by the fact that it only provides information about the behavior of the fractal object at a finite range of scales, and therefore it is heavily affected by the so-called curse of dimensionality. As a result, estimates of the attractor dimensions for complex systems such as Earth's climate are unrealistically small (Nicolis and Nicolis, 1984; Grassberger, 1986; Lorenz, 1991).

More recently, a new methodology based on EVT has been introduced to estimate attractor dimensions. This framework allows to extrapolate the behavior of the fractal object to very large or very small scales by modeling the extreme values of the box counting curve. In other terms, the probability distribution of the distance between a reference point and nearby points in the attractor

is modeled as a power law. The exponent of this power law is related to the local dimension of the attractor at that point. EVT thus provides a more accurate estimate of the Hausdorff dimension and local dimensions of the fractal object than the box counting approach, especially for high-dimensional systems. However, it does not entirely overcome the curse of dimensionality (Pons et al., 2020).

Let y_j be $j = 1, \ldots, N$ realizations in time of a dynamical system Y, and let ζ be a particular realization taken as a reference state. Following Freitas et al. (2010); Faranda et al. (2011), we consider the distance observable defined as

$$g_j = -\ln(\mathrm{dist}(y_j, \zeta)), \tag{1}$$

where $\mathrm{dist}(y_j, \zeta)$ is the Euclidean distance between each state y_j and the reference state ζ. We define the *recurrences* of the reference state ζ as $\{y_j : \mathrm{dist}(y_j, \zeta) < r\}$, i.e. the ensemble of states of the system located in the neighbourhood of ζ defined by the hypersphere $B(\zeta, r)$. Let us consider the sequence of N values of g, and let g_q be an extreme quantile of the probability density function of g, $f_G(g)$, corresponding to the probability q. We define the Peaks Over Threshold (POTs) as the ensemble of $n = [N(1-q)]$ values such that $g > g_q$, where $[\cdot]$ denotes the closest integer operator.

The observable g in Eq. 1 is defined in such a way that the recurrences, states characterized by minima of the distance function, correspond to POTs of the observable itself. Given an n-dimensional sequence of POTs, we define the *exceedances* X as the POTs shifted by subtracting the threshold g_q, so that $X = \{X_\ell\}_{\ell=1}^{n} = \{g - g_q : g \geq g_q\}$, where $\ell = 1, \ldots, n$ and $\min X = 0$.

The choice of the Euclidean distance to define g guarantees that the threshold g_q used to defined the exceedances is in direct relation to the scaling of the hypersphere $B(\zeta, r)$, since $r = \exp\{-g_q\}$ (Faranda et al., 2011). In other words, the asymptotic frequency with which the system enters a hypersphere of radius r centred on ζ can now be expressed as the probability of exceeding a threshold corresponding to a high quantile g_q of the distribution of g.

In the limit $N \to \infty$, the shifted exceedances follow a two-parameter Generalized Pareto Distribution $GPD(\sigma, \xi)$. In this setting, the local dimension around the state ζ is given by (Faranda

et al., 2011):

$$d(\zeta) = \frac{1}{\sigma}. \tag{2}$$

By averaging d over a sufficiently large ensemble of states ζ on the attractor, one then obtains an estimate the Hausdorff dimension D (Caby et al., 2018).

For Axiom A systems in the asymptotic case $n \to \infty$, $\xi \to 0$, and the two-parameter $GPD(\sigma, \xi)$ converges to an exponential distribution, $Exp(\sigma)$. In this case, the local dimension estimator is the inverse of the exceedances sample mean

$$\hat{d} = \frac{1}{\overline{x}}. \tag{3}$$

Many of the existing applied studies leveraging on EVT to estimate d use Eq. 3; however, real-world systems do not realistically satisfy the Axiom A assumption, and the exponential model for the exceedances distribution is misspecified, likely leading to biased estimates.

2.2 Generalized Pareto Distribution

A random variable X follows a generalised Pareto distributon $GPD(\mu, \xi, \sigma)$ if its cumulative distribution function (CDF), giving the probability $P(X > x)$, is defined as

$$F(x|\xi, \mu, \sigma) = \begin{cases} 1 - \left(1 + \xi \frac{x-\mu}{\sigma}\right)^{-1/\xi}, & \xi \neq 0 \\ 1 - \exp\left(-\frac{x-\mu}{\sigma}\right), & \xi = 0 \end{cases} \tag{4}$$

and the corresponding probability density function (pdf) reads:

$$f(x|\xi, \mu, \sigma) = \begin{cases} \frac{1}{\sigma}\left(1 + \xi \frac{x-\mu}{\sigma}\right)^{-1/\xi - 1}, & \xi \neq 0 \\ \frac{1}{\sigma}\exp\left(-\frac{x-\mu}{\sigma}\right), & \xi = 0 \end{cases} \tag{5}$$

where $\xi \in \mathbb{R}$ is the shape parameter, $\mu \in \mathbb{R}$ is the location parameter and $\sigma > 0$ is the scale parameter. The support of X depends on the shape parameter: if $\xi \leq 0$, $\mu \leq x \leq \mu + \frac{\sigma}{\xi}$; if $\xi > 0$,

$\mu \leq x \leq \infty$. The moment generating function is:

$$M_X(t) = E_X[e^{tx}] = e^{t\mu} \sum_{j=0}^{\infty} \left[\frac{(t\sigma)^j}{\prod_{k=0}^{j}(1-k\xi)} \right], \quad \text{for } k\xi < 1$$

where $E_X[\cdot]$ denotes expectation computed with respect to the density of X. The r-th moment around zero can be obtained as:

$$E[X^r] = \left. \frac{d^r M_X(t)}{dt^r} \right|_{t=0} = r! \frac{\sigma^r}{(-\xi)^{r+1}} \frac{\Gamma\left(-\frac{1}{\xi} - r\right)}{\Gamma\left(1 - \frac{1}{\xi}\right)}, \quad \text{for } \xi < \frac{1}{r} \tag{6}$$

In the special case $\mu = 0$ the three-parameter $GPD(\mu, \xi, \sigma)$ reduces to a two-parameter $GPD(\sigma, \xi)$. In the following, we will consider this specific case, since the procedure described in Section 3 requires to shift the peak-over-thresholds (POTs) so that their minimum is 0.

There are several methods to estimate the parameters of the GPD distribution, ranging from frequentist techniques such as Maximum likelihood (ML), the method of moments (MM), probability-weighted method of moments (PWM), linear combination of order statistics and others, to Bayesian methods and approaches based on entropy (de Zea Bermudez and Kotz, 2010a,b).

Here, we focus on several frequentist methodologies, including ML, MM and PWM estimation, plus an alternative formulation of ML, which we will call pseudo-ML (PML). We choose these estimators based on their computational simplicity (MM and PWM), or because an effective implementation is readily available in a `R` package.

3 Estimating the attractor dimension

In this section, we introduce five estimators of the parameter $\sigma = \frac{1}{d}$, namely the inverse local dimension. The first one (given in Eq. 19), is constructed assuming the degenerate case of POTs following a $GPD(\xi = 0, \sigma) \equiv Exp(\sigma)$. The others (given in Eq.s 20-23) are constructed under the assumptions of POTs following a $GPD(\sigma, \xi)$.

3.1 Estimation based on $Exp(\sigma)$

As already mentioned, for Axiom A systems in the asymptotic case $n \to \infty$ $\xi \to 0$, so that the distribution of the exceedances converges to an exponential distribution.

In such a case, a ML estimator for the local dimension is readily available in close form, given by $\hat{\sigma}_m = \overline{x}$. Using this estimator, Pons et al. (2020) have shown that the estimated attractor dimension is biased, with a bias growing nonlinearly with the topological dimension of the system, as a particular case of curse of dimensionality (Bellman, 1957), due to the concentration of the Euclidean norm. This effect was analytically proven for very long time series of random vectors, and shown to equally affect real-world complex data, such as sea-level pressure fields or multivariate time series of financial indices.

However, for both real-world and simulated systems displaying chaotic behaviour, the asymptotic case may not be met as easily as for purely stochastic systems. Much longer sampling is required to observe the invariant distribution of chaotic systems compared to random vectors, and assuming exponential exceedances may constitute a serious model misspecification, which may introduce further bias.

3.2 Estimation based on $GPD(\sigma, \xi)$

Maximum Likelihood (ML)

Maximum Likelihood is one of the most popular methods to estimate the parameters of a statistical model, and it has been widely used since its first introduction Fisher (1912). It is beyond our scope to give an overview of ML, which can be found in many classic statistics textbooks (Casella and Berger, 2002). In brief, given a sample x consisting of n independent realizations of a random variable with probability density function $f(x|\theta)$, where θ denotes the vector of parameters, we call likelihood function the quantity

$$L(\theta|x) = \prod_{i=1}^{n} f(x_i|\theta). \tag{7}$$

The ML estimator of the parameter vector is the point $\hat{\theta}$ of the parameter space at which $L(\theta|x)$ attains its maximum. Finding the ML parameter estimates can be a simple analytical task or a challenging numerical procedure, depending on the smoothness of the likelihood function. In the case of the GPD, it is not possible to obtain close-form estimators analytically, so that alternative strategies must be found to maximize $L(\theta|x)$. To obtain ML estimates for the GPD, we use the `gpd.mle()` function from the `R` package `'mev'` (Belzile et al., 2022), based on the profile likelihood maximization method Grimshaw (1993).

Pseudo Maximum Likelihood (PML)

An alternative way to obtain ML estimates of the GPD parameters has been introduced by Villaseñor-Alva and González-Estrada (2009) and implemented in the `R` package `gPdtest` (Estrada and Alva, 2012). The estimator is constructed differently depending on the sign of ξ; given this composite strategy and the passage through intermediate steps not based on the likelihood function, we will refer to this method as "pseudo maximum likelihood". Here we only provide a sketch and refer the reader to the aforementioned article for full details.

For $\xi \geq 0$, `gPdtest` implements a maximization of the asymptotic likelihood, based on the logarithm of the k largest values in the ordered sample $x_{1:n}, x_{2:n}, \ldots, x_{n:n}$: $W_j = \log x_{j:n}$ with $j = n-k+1, n-k+2, \ldots, n$. For $n \to \infty$ the log-likelihood can be expressed as a function of W_j, and the two ML estimators follow:

$$\hat{\xi}_k = -\left(W_{n-k+1} - \frac{1}{k}\sum_{j=1}^{k} W_{n-j+1}\right) \tag{8}$$

$$\hat{\sigma}_k = \hat{\xi}_k \exp\left\{W_{n-k+1} + \hat{\xi}_k \log\frac{k}{n}\right\}. \tag{9}$$

The case $\xi < 0$ requires a composite strategy, combining a moment equation step and a maximum likelihood estimator of the upper boundary of the GPD. Recalling that we consider the case $\mu = 0$, the variable $U = [1 - F(X|\xi,\sigma)]^{-\xi} = 1 + \frac{\xi}{\sigma}X$ follows a Beta$(-\frac{1}{\xi}, 1)$ distribution. The first sample moment is $\overline{u} = \frac{1}{n}\sum_{i=1}^{n}\left(1 + \frac{\xi}{\sigma}x_i\right) = 1 + \frac{\xi}{\sigma}\overline{x}$, to be equated to the expected value of the Beta distribution $E[U] = \frac{1}{1-\xi}$, resulting in $\xi = 1 - \frac{\sigma}{\overline{x}}$. Let us now recall that, for $\xi < 0$, $0 \leq x \leq \sigma/\xi$:

then, the sample maximum $x_{n:n}$ is a ML estimator for the upper bound σ/ξ. Combining the two steps, one obtains:

$$\hat{\sigma}_{PML} = -\hat{\xi} x_{n:n} \tag{10}$$

$$\hat{\xi}_{PML} = \frac{\overline{x}}{\overline{x} - x_{n:n}} \tag{11}$$

Notice that the fact that the estimation strategy depends on the sign of the shape parameter, a prior knowledge on the value of ξ must be used. Results discussed in this paper from PML estimation are based on the sign of $\hat{\xi}_{PWM}$, introduced later. We have also tested the possibility to use $\hat{\xi}_{MM}$ (not shown in the article) obtaining equivalent results.

Method of Moments (MM)

The method of moments is another common parameter estimation method. It consists of expressing the theoretical moments as a function of the parameters and solving the resulting equations for the parameter estimates using the values of the sample moments.

Let X be a random variable $X \sim GPD(\sigma, \xi)$, $x = x_1, x_2, \ldots, x_n$ a sample of n observations from X. The first two GPD moments, required to estimate the two parameters, can be obtained using Eq. 6:

$$\begin{aligned} E[X] &= \frac{\sigma}{1+\xi}, && \text{for } \xi < 1 \\ Var[X] = E[X^2] - E[X]^2 &= \frac{\sigma^2}{(1+\xi)^2(1+2\xi)}, && \text{for } \xi < \frac{1}{2} \end{aligned} \tag{12}$$

and the MM parameter estimators follow by equating sample and theoretical moments, and solving for σ, ξ:

$$\begin{aligned} \hat{\xi}_{MM} &= \frac{1}{2}\left(\frac{\overline{x}^2}{s^2} - 1\right) \\ \hat{\sigma}_{MM} &= \frac{\overline{x}}{2}\left(\frac{\overline{x}^2}{s^2} + 1\right) \end{aligned} \tag{13}$$

where $\overline{x} = n^{-1}\sum x_i$ is the sample mean and $s^2 = n^{-1}\sum(x_i - \overline{x})^2$ is the sample variance. Given the condition $\xi < \frac{1}{r}$ in Eq. 6, the estimators in Eq. 13 only hold for $\xi < 0.5$.

Notice that $\hat{\sigma} = \overline{x}(\hat{\xi} + 1)$. If $\xi = 0$, the GPD reduces to a standard exponential distribution, $X \sim Exp(\sigma)$, and the MM estimator of σ is the sample mean. On the other hand, if it is erroneously

assumed that $\xi = 0$, the estimate of σ is affected by a bias equal to $-\overline{x}\hat{\xi}$.

Probability Weighted Moments (PWM)

Finally, we consider a variation of the MM, based on probability weighted moments (PWM). This method is particularly suited for distributions that can be uniquely expressed in terms of their inverse CDF in a simple way, but it is also popular in hydrological modeling due to its computational simplicity. For a random variable X with CDF $F(X)$, the PWM are defined as

$$M_{p,r,s} = E[X^p (F(X))^r (1 - F(X))^s], \qquad p, r, s \in \mathbb{R} \tag{14}$$

For the estimation of $GPD(\sigma, \xi)$, is is suggested (de Zea Bermudez and Kotz, 2010a) to use the moments defined by $p = 1$, $r = 0$:

$$\begin{aligned} \alpha_s = M_{1,0,s} &= E[X(1 - F(X))^s] \\ &= \frac{\sigma}{(s+1)(s+1+\xi)}, \qquad\qquad \xi < 1, \quad s = 0, 1, 2, \ldots \end{aligned} \tag{15}$$

so that the PWM expression of the GPD parameters is:

$$\begin{aligned} \xi &= 2 - \frac{\alpha_0}{\alpha_0 - 2\alpha_1} \\ \sigma &= \frac{2\alpha_0\alpha_1}{\alpha_0 - 2\alpha_1}. \end{aligned} \tag{16}$$

Given a sample x of size n from $X \sim GPD(\sigma, \xi)$, the sample version of the PWM α_s is

$$a_s = \frac{1}{n} \sum_{i=1}^{n} x_{i:n} (1 - p_{i:n})^s, \tag{17}$$

where $x_{i:n}$ is the i-th element of the ordered sample and $p_{i:n}$ are the corresponding plotting positions, such that $(1-p_{i:n})$ is a sample estimate of $(1-F(X))$. Notice that $a_0 = \overline{x}$. There are several possible ways to estimate the plotting positions. Here, we adopt the simple expression recommended in the

literature (Landwehr et al., 1979; Hosking et al., 1985):

$$p_{i:n} = \frac{i - 0.35}{n}.$$

The two estimators can be obtained by plugging Eq. 17 in Eq.s 16. Once again, we recall that the upper limit of $GPD(\sigma, \xi)$ is given by $-\sigma/\xi$ and therefore estimated by $-\hat{\sigma}/\hat{\xi}$; however, if $-\hat{\sigma}/\hat{\xi} < x_{n:n}$, the sample maximum $x_{n:n}$ itself is a better estimator of the upper limit. The final expression of the PWM estimators follows:

$$\begin{aligned} \hat{\sigma}_{PWM} &= \frac{2a_0a_1}{a_0 - 2a_1}, \quad \hat{\xi}' = 2 - \frac{a_0}{a_0 - 2a_1} \\ \hat{\xi}_{PWM} &= \begin{cases} \hat{\xi}' & -\hat{\sigma}/\hat{\xi}' \geq x_{n:n} \\ -\frac{\hat{\sigma}}{x_{n:n}} & -\hat{\sigma}/\hat{\xi}' < x_{n:n} \end{cases} \end{aligned} \tag{18}$$

The PWM estimators are unbiased, and for large samples are normally distributed with mean (σ, ξ) and known covariance matrix.

To summarize, we will consider the following five expressions of the inverse attractor dimension:

$$\hat{d}_m^{-1} = \hat{\sigma}_m = \overline{x} \tag{19}$$

$$\hat{d}_{ML}^{-1} = \hat{\sigma}_{ML} = \arg\max\{L(\sigma, \xi|x)\} \tag{20}$$

$$\hat{d}_{PML}^{-1} = \hat{\sigma}_{PML} = -\frac{\overline{x}x_{(n)}}{\overline{x} - x_{(n)}} \tag{21}$$

$$\hat{d}_{MM}^{-1} = \hat{\sigma}_{MM} = \frac{\overline{x}}{2}\left(\frac{\overline{x}^2}{s^2} + 1\right) \tag{22}$$

$$\hat{d}_{PWM}^{-1} = \hat{\sigma}_{PWM} = \frac{2\overline{x}a_1}{\overline{x} - 2a_1} \tag{23}$$

with $a_1 = \frac{1}{n}\sum_{i=1}^{n} x_{1:n}(\frac{i-0.35}{n})$. We give no analytical expression available for $\hat{\sigma}_{ML}$, as this is obtained by numerical optimization of the likelihood function.

4 Data

In order to evaluate and compare the statistical performance of the estimators, we use data from three sources. First, we perform a Monte Carlo study, so that we can compare estimates to known true values of σ, as a function of ξ and of the sample size. Then, we generate a very long trajectory from the well-known Lorenz attractor, whose Hausdorff dimension is known, and which is not affected by the curse of dimensionality. Finally, we compare results obtained from real-world climatological data, that are the main motivation for this study.

4.1 Monte Carlo experiment

As a first step, we perform a Monte Carlo experiment to assess estimator performance as a function of the true values of σ, ξ and of the number of exceedances n. In order to do so, for every combination (σ, ξ, n) we generate $M = 5000$ random samples from a $GPD(\sigma, \xi)$ with $\xi \in [-1, 1]_{0.2}$; $\sigma = 1/d$, with $d \in [1, 5]_{0.2} \cup [5, 100]_5$; $n \in [10, 100]_{10} \cup \{150, 200, 500, 1000\}$, where $[a, b]_\delta$ denotes a partition of the closed a, b interval with step δ. Notice that n is the number of available exceedances, so that the original sample size is implied, and in real cases it depends on the chosen POT threshold: for example, $n = 100$ and a threshold at the 99th percentile implies an initial time series of 10^5 observations.

4.2 Lorenz 1963

Moving further beyond Monte Carlo data, we consider a simple chaotic dynamical system: the well-known Lorenz63 model (hereinafter L63; Lorenz, 1963). This is a three-variable system meant to describe Rayleigh-Bénard convection, and consisting of the following ordinary differential equa-

tions:

$$\begin{aligned} \frac{dX}{dt} &= \sigma(Y - X) \\ \frac{dY}{dt} &= X(\rho - Z) - Y \\ \frac{dZ}{dt} &= XY - \beta Z \end{aligned} \tag{24}$$

Here X represents convective overturning rate, Y horizontal temperature variations, and Z departure from a linear vertical temperature gradient. We use the standard parameter configuration to obtain chaotic behaviour (the well-known "butterfly"), namely $\sigma = 10$, $\rho = 28$ and $\beta = 8/3$. We use a timestep of 0.001 s and generate $4 \cdot 10^5$ iterations. We discard the first 2000 iterations and sub-sample the trajectory taking one every 4 iterations.

4.3 ERA5 climate reanalysis

As a last step, we consider climate reanalysis data. This data is issued from a full-complexity atmospheric global model, constrained by a very large amount of in situ and remotely sensed data.

We use the ERA5 reanalysis dataset (Hersbach et al., 2020), produced by the European Centre for Medium-Range Weather Forecasts (ECMWF). We use 500 hPa geopotential height data on a $0.25°$ horizontal resolution, daily temporal resolution and the underlying climate model includes accurate physical parameterizations. We consider data over the period 1979–2018 and a domain spanning $26°$ – $70°$ North and $22°$ West – $46°$ East. This provides a lat-lon grid of 273×177 points covering Europe, and parts of the Eastern Atlantic, Russia and North Africa.

5 Results

5.1 Monte Carlo experiment

As described in Section 4.1, we specify a set of combinations (σ, ξ, n), for each combination we generate 5000 Monte Carlo samples, and for each sample we estimate the dimension with the five

estimators described by Eq.s 19-23. All results discussed in this section are obtained as the average over the 5000 Monte Carlo samples for each (σ, ξ, n) combination, and refer to the local dimension $d = 1/\sigma$. The dependence of the estimator performance on ξ is important for understanding the actual applicability of each estimator to particular classes of systems.

In the following, we will consider two error measures to evaluate the performance of the different estimators: the root mean square error (RMSE) and the relative error. The RMSE is defined as the square root of the average squared deviation:

$$RMSE = \sqrt{E[(\hat{d} - d)^2]} = \sqrt{Bias(\hat{d})^2 + Var(\hat{d})} \tag{25}$$

where the expectation is taken over the M Monte Carlo replicates, and it is thus a function of n, ξ and the true value d. Since the mean square error can be decomposed in terms of bias and variance of the estimator as shown in Eq. 25, the RMSE allows to find the most efficient estimator, in the sense of the best bias-variance trade-off. Figures 1-4 show values of RMSE as a function of d and n, for $\xi = -1, -0.5, 0, 0.5$. The complete set of results for all ξ is shown in supplementary Figures 19-39.

In general, larger sample sizes correspond to smaller RMSE, and larger RMSE are found for large values of the true dimension d. For the mean estimator, however, the positive effect of the sample size is observed only at $\xi \sim 0$, where the underlying assumption of an exponential distribution of the exceedances if met. The MM estimator also produces large RMSE for large n when $\xi \gtrapprox 0.5$, which is expected as the second moment of the GPD only exists for $\xi < 0.5$. Notice that the PML shows a qualitatively similar dependence of RMSE on n, d, but the RMSE magnitude is much larger compared to the other estimators.

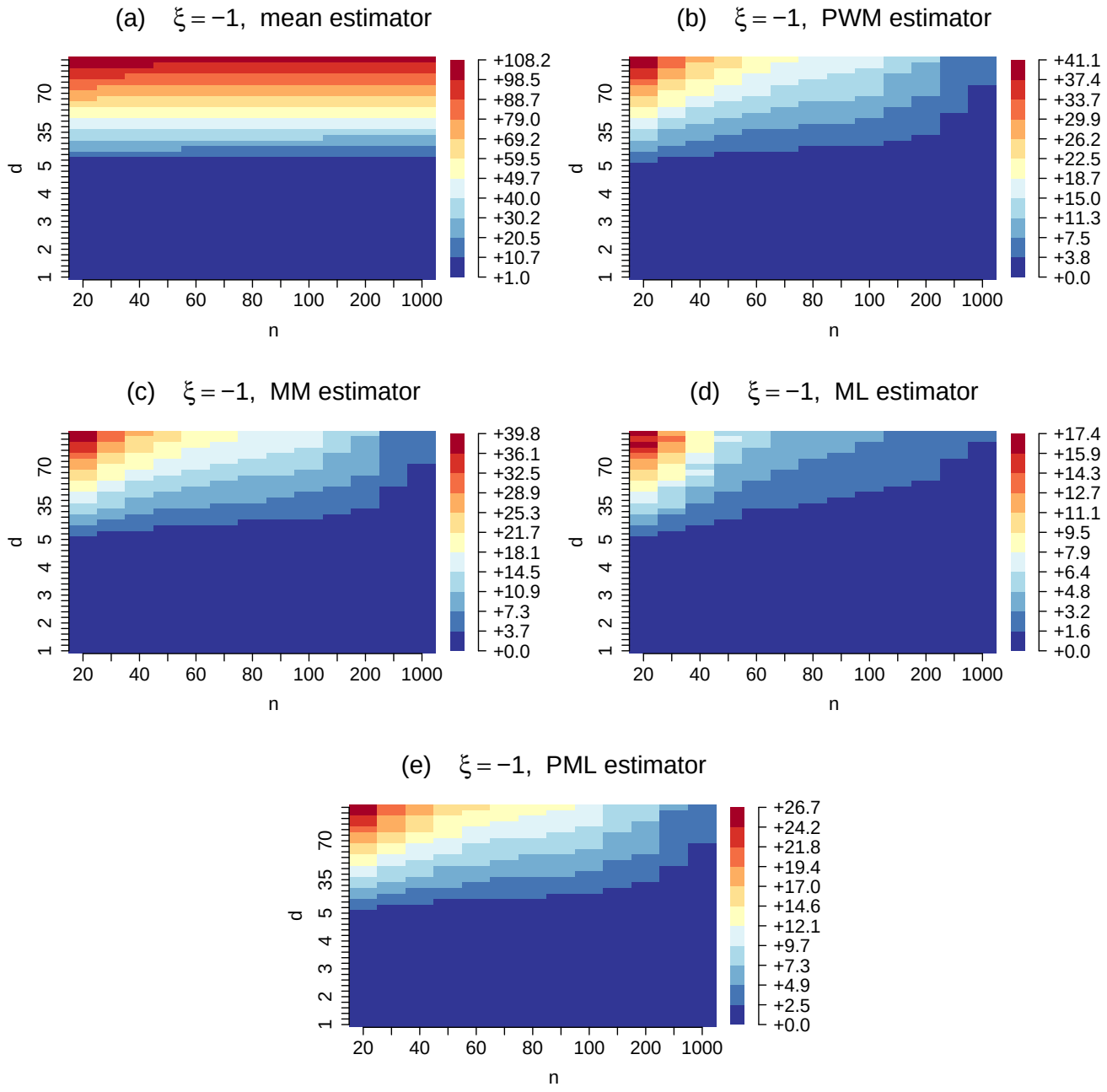

Figure 1: Root mean square error of the mean estimator (a), PWM (b), MM (c), ML (d) and PML (e) on d as a function of d and n, for $\xi = -1$.

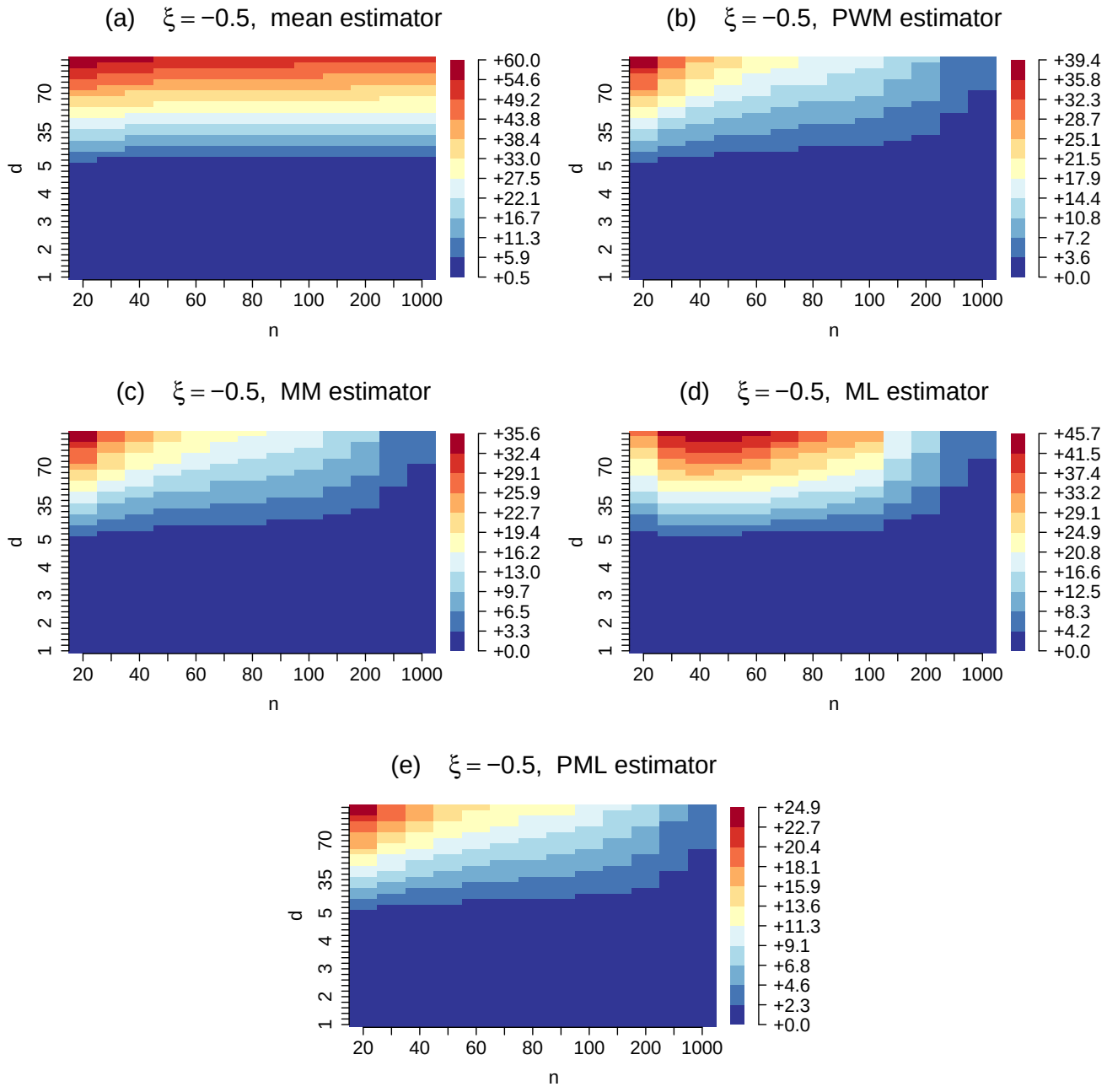

Figure 2: Root mean square error of the mean estimator (a), PWM (b), MM (c), ML (d) and PML (e) on d as a function of d and n, for $\xi = -0.5$.

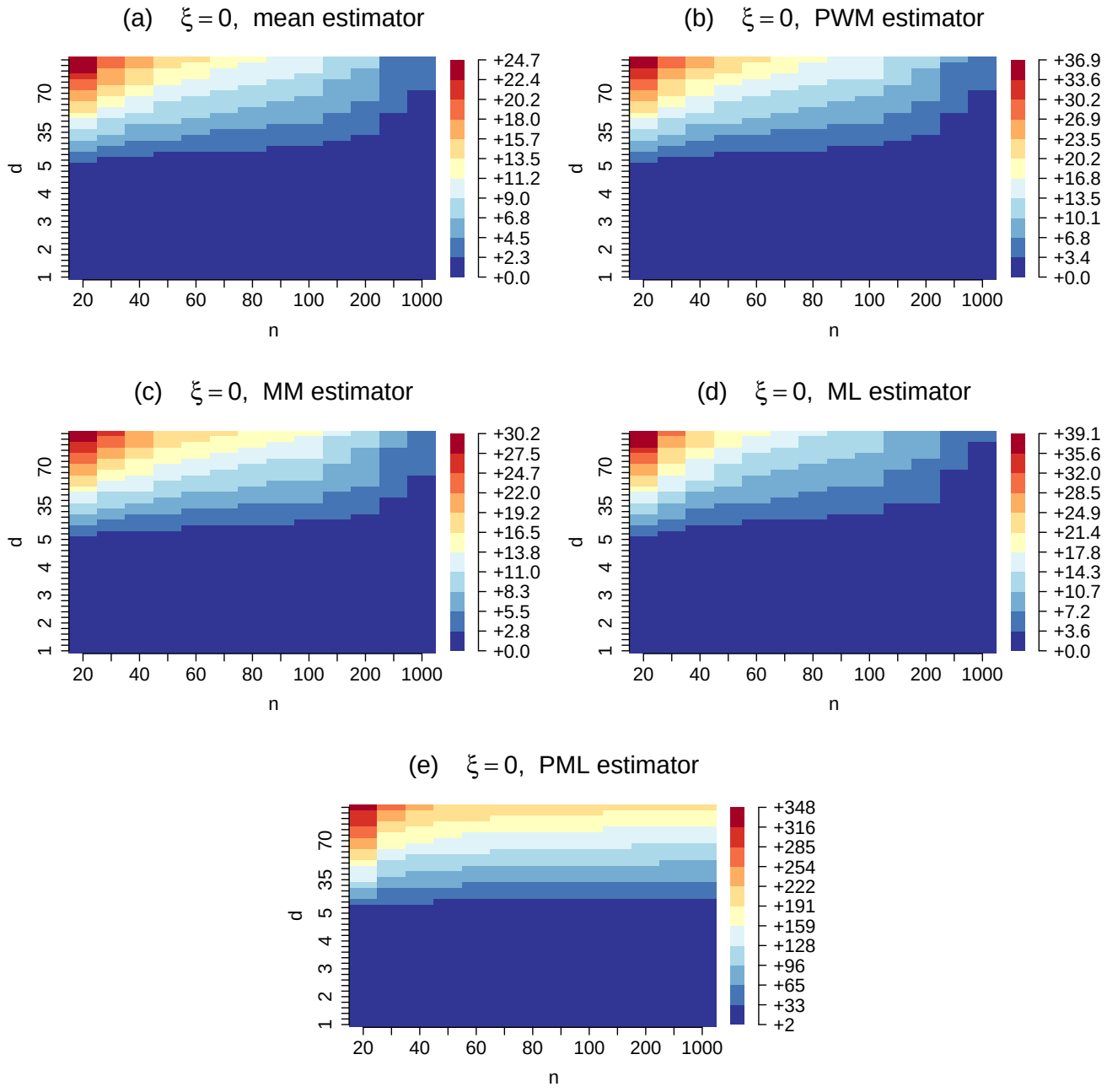

Figure 3: Root mean square error of the mean estimator (a), PWM (b), MM (c), ML (d) and PML (e) on d as a function of d and n, for $\xi = 0$.

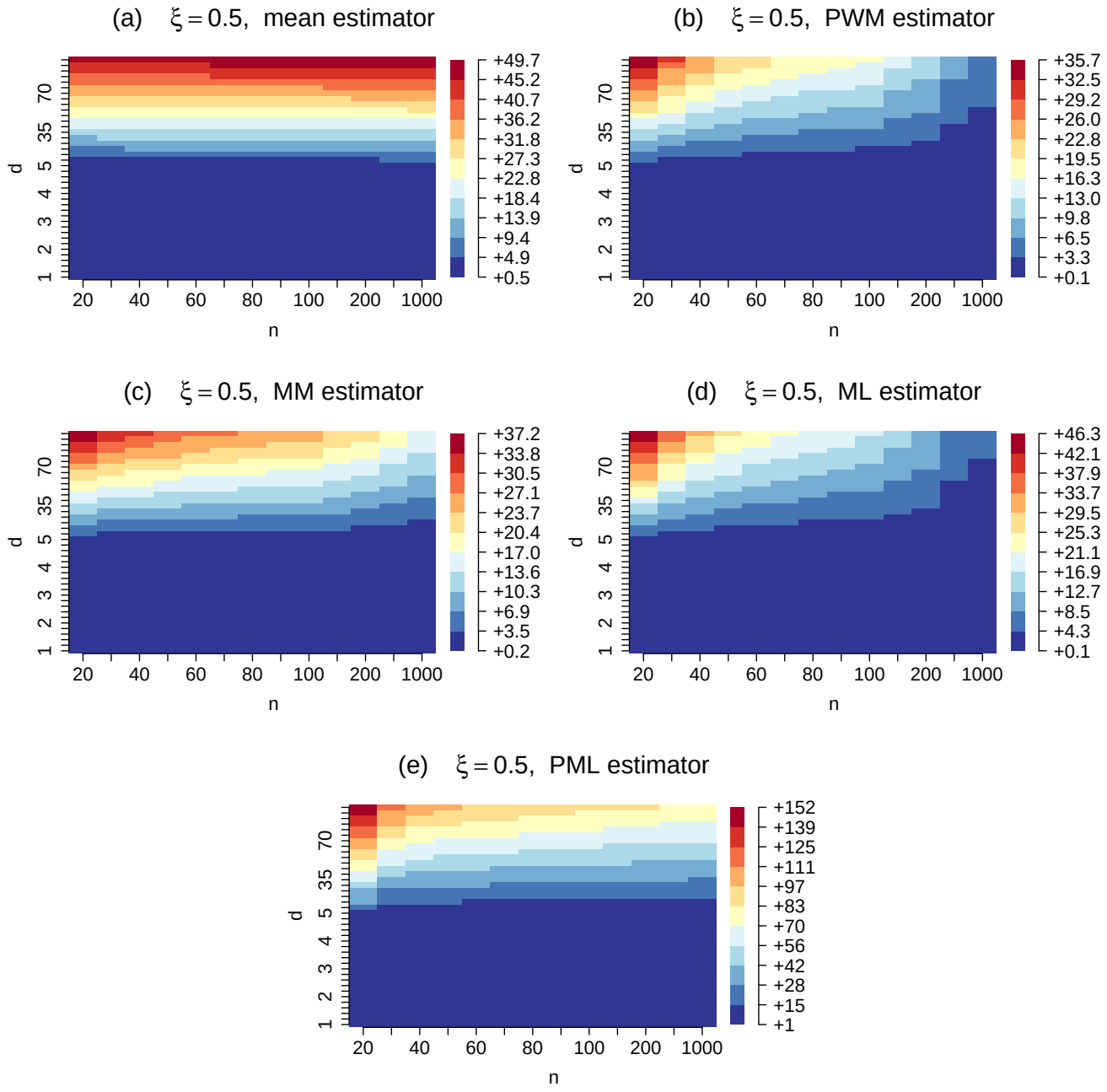

Figure 4: Root mean square error of the mean estimator (a), PWM (b), MM (c), ML (d) and PML (e) on d as a function of d and n, for $\xi = 0.5$.

Table 1 and Figure 5 summarize the effect of the true value of ξ on RMSE for $n = 20$, 100, 1000, averaged over the true dimension. As expected, the mean estimator has the smallest RMSE when $\xi \simeq 0$, but it loses efficiency when this assumption is not met. The PML is the most efficient over a restricted range of negative values of ξ, however for $\xi \gtrapprox -0.1$ it produces by far the largest RMSE values. The efficiency of the ML as a function of ξ changes visibly from small to large

ξ	-1.0	-0.8	-0.6	-0.4	-0.2	0.0	0.2	0.4	0.6	0.8	1.0
	n = 20										
mean	29.72	24.31	19.10	14.13	9.71	6.98	7.85	11.03	14.50	17.55	20.06
MM	16.68	15.89	15.25	14.61	13.78	12.98	12.52	13.86	17.58	22.08	26.42
PWM	17.33	16.84	16.56	16.42	16.06	15.83	15.72	15.11	15.14	15.81	17.70
ML	7.27	9.35	12.97	15.25	16.16	16.85	17.93	19.04	20.70	22.01	23.50
PML	11.14	10.65	10.38	10.55	11.49	146.81	103.16	75.64	59.71	54.74	61.16
	n = 100										
mean	28.09	22.60	17.10	11.67	6.45	2.84	5.94	10.91	15.68	19.61	22.52
MM	6.91	6.66	6.38	6.12	5.95	5.91	6.08	8.07	14.49	22.23	28.92
PWM	6.96	6.86	6.75	6.65	6.64	6.62	6.70	6.91	7.21	8.02	10.92
ML	1.50	5.20	8.77	16.42	10.21	5.63	6.67	7.33	7.85	8.28	8.80
PML	4.92	4.76	4.69	5.04	6.29	85.97	60.54	43.70	32.95	26.70	24.06
	n = 1000										
mean	27.74	22.20	16.67	11.14	5.64	0.87	5.58	11.06	16.39	20.97	24.23
MM	2.13	2.06	1.98	1.91	1.87	1.88	2.44	4.05	12.06	23.20	32.24
PWM	2.14	2.11	2.08	2.06	2.06	2.04	2.10	2.29	2.84	3.58	6.20
ML	0.75	1.50	1.89	2.93	8.45	1.66	2.09	2.24	2.41	2.54	2.69
PML	1.55	1.50	1.49	1.74	2.95	80.94	56.75	40.24	28.89	21.12	15.95

Table 1: Root mean square error of the five considered estimators on the local dimension, averaged over d and shown as a function of ξ, for sample sizes $n = 20$, 100, 1000.

sample sizes, where it shows larger RMSE for $-0.5 < \xi < 0$. The MM estimator shows a good performance for $\xi \lessapprox 0.5$, but it is affected by large RMSE values at larger values of ξ, which does not improve with sample size. Finally, the PWM is the most stable estimator: for small n it is not the most efficient, but its performance is independent of the true value of ξ. At larger sample sizes the RMSE is overall small, and with the smallest variations as a function of ξ, with only a slightly decreased efficiency for ξ close to 1.

Overall, it is difficult to point out the best estimator in terms of efficiency when the sample size is small. For larger samples, the PWM is the best choice; however, if it can be safely assumed that $\xi < 0$, MM and PWM are in close competition in terms of efficiency.

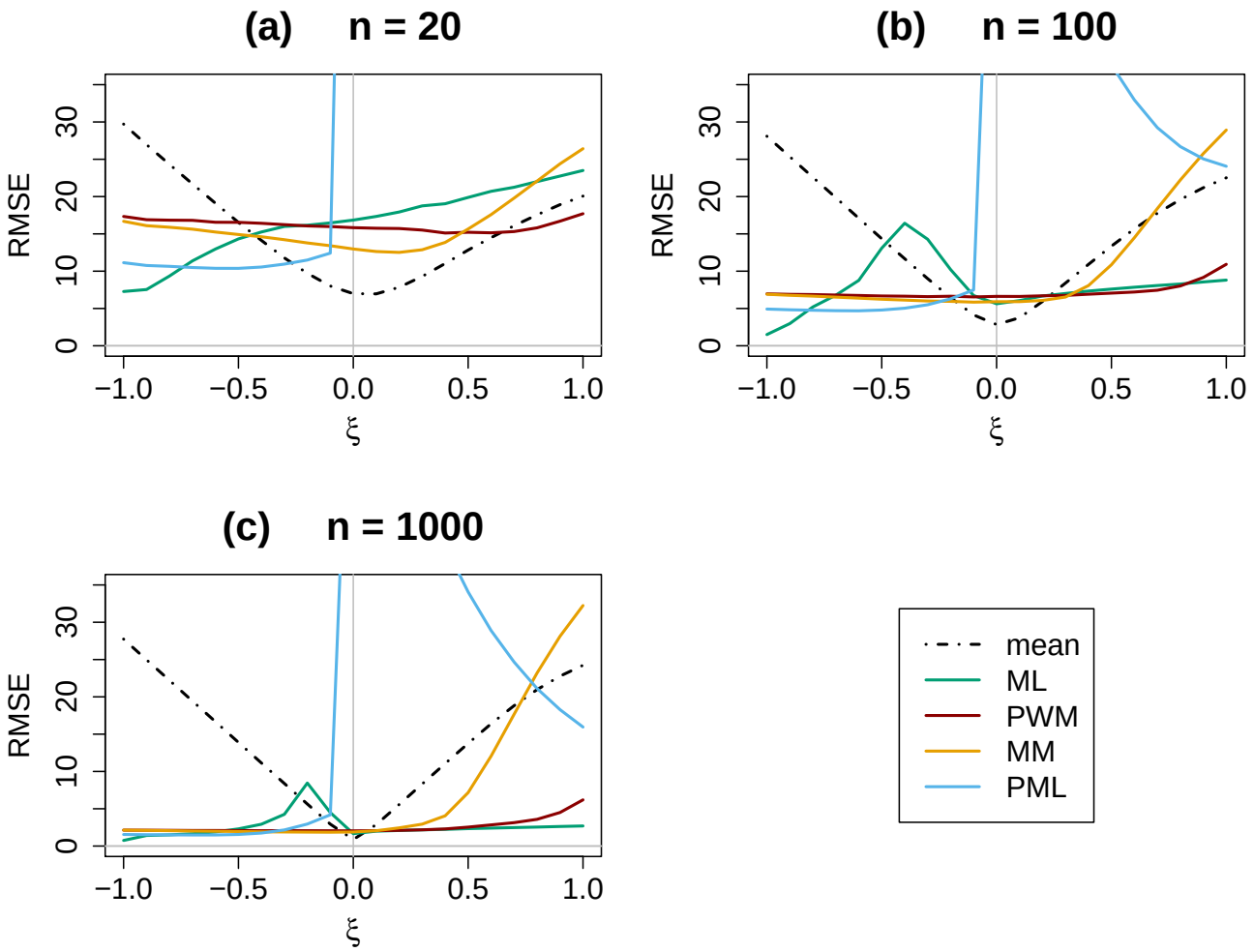

Figure 5: Graphic visualization of the results shown in Table 1. Root mean square error of the five considered estimators on the local dimension, averaged over d and shown as a function of ξ, for sample sizes $n = 20$ (a), 100 (b), 1000 (c).

As a further way to characterize the performance of the estimators, we consider the average relative error $E[(\hat{d} - d)/d]$, where the expectation is computed over the M Monte Carlo replicates as for the RMSE. This measure allows to evaluate the magnitude of the estimation error relative to the true value, while also assessing the sign of the bias.

Figures 6-9 show the values of the relative error for the five considered estimators, as a function of the sample size n and of the true dimension d. The Figures refer to $\xi = -1, -0.5, 0, 0.5$; the complete set of results for all ξ is shown in supplementary Figures 40-60.

For all values of ξ and for all estimators, results improve with the sample size, with overall convergence at values around $n \sim 200$. Results appear to be stable with respect to the true

dimension d for all methods, except maximum likelihood. For $\xi > 0$, the error of the ML estimator seems to be independent on d, or shows a weak dependence for $0 < \xi \lessapprox 0.2$. However, for $\xi < 0$ the ML estimator shows a strong dependence on d, especially for sample sizes $50 \lessapprox n \lessapprox 200$: for $-1 \lessapprox \xi \lessapprox -0.5$ the relative error becomes more negative with increasing d, while for $-0.5 \lessapprox \xi \lessapprox 0$ it becomes more positive.

For the special case $\xi = 0$ (Fig. 8), when the GPD converges to an exponential distribution, the performance of the mean estimator shows a great improvement, with relative errors $\sim 0.5\%$ at $n = 100$ and $\sim 0.01\%$ for $n > 200$. This is expected, since in this case the sample mean is also the exact maximum likelihood estimator of σ.

The PML estimator has the overall worst performance, with relative errors comparable to other methods for $\xi < 0$ but jumping to errors of $100 - 200\%$ for $\xi \geq 0$. Even though the package authors recommend the use of `gPdtest` for all values of ξ, Eq.s 8-9 show that an unbiased estimation is not possible for $\xi = 0$, since if ξ was correctly estimated, it would necessarily result in $\hat{\sigma} = 0$, corresponding to a degenerate distribution with 0 variance, and to an infinite estimated dimension.

Since the performance of the estimators is mostly independent of the true value of σ (or d), and convergence with n is reached for $n \lesssim 200$, we can fix n and average over d to obtain a synthetic representation on the relationship between ξ and the relative estimation error affecting $\hat{d}$. As we pointed out earlier, the ML estimator error shows some dependence on d for $\xi \lessapprox 0.2$. However, the range of relative value errors is small, in particular compared to the mean and the PML estimators. We still include ML in this analysis for a global comparison. Results are summarized in Table 2 and shown in Fig. 10, where the relative error averaged over d for $n = 20, 100, 1000$ is plotted as a function of ξ.

The performance of PML shows large relative errors for $\xi > -0.2$, and a non-monotonic, discontinuous behaviour as a function of ξ. The ML estimator also shows a non-monotonic dependence of the relative error for negative values of ξ. This irregularity is likely due to the aforementioned dependence of the ML estimator on d, overlooked in this analysis. All the other estimators produce a relative error that is monotonically decreasing with the true value of ξ. The two moment-based

estimators have a similar performance, with relative errors $\approx 0.1\%$ for $\xi \leq 0$. The performance of the MM estimator degrades for values larger than $0 < \xi \lessapprox 0.5$, depending on the sample size with large samples allowing for accurate estimations close to the limit of the domain of existence $\xi < 0.5$. Similarly, the PWM exist for $\xi < 1$ and the corresponding estimator shows a tendency to under-estimate σ when $\xi \to 1$, with smaller errors at larger sample size.

The performance of the mean estimator is particularly interesting. We recall that this is a well-specified unbiased estimator only under the assumption of an Axiom A system in the asymptotic case, condition that can be rarely satisfied in real-world data; on the other hand, it is very simple to compute, and it has been used in several studies, providing physically coherent results. Indeed, while the relative error of this estimator is large compared to the others, it has the peculiar property of being $(\hat{\sigma} - \sigma)/\sigma = -\xi$, except for $\xi \approx 1$, where it slightly deviates from this relationship.

Overall, for systems where it is expected that $\xi \leq 0$, the MM and PWM perform the best. ML performs slightly worse for $\xi \leq 0$, but it is the only estimator providing unbiased estimates for $\xi > 0$. All three estimators are quite stable with the sample size, providing fairly good estimates already for $n = 20$. The mean estimator is severely biased, but its error is monotonic and approximately linear with ξ, so that conclusions obtained comparing local dimension of different states or systems still hold.

ξ	-1.0	-0.8	-0.6	-0.4	-0.2	0.0	0.2	0.4	0.6	0.8	1.0
	n = 20										
mean	1.035	0.836	0.639	0.441	0.246	0.052	-0.134	-0.308	-0.461	-0.589	-0.691
MM	0.061	0.055	0.051	0.042	0.026	-0.011	-0.078	-0.185	-0.314	-0.445	-0.564
PWM	0.082	0.075	0.070	0.065	0.057	0.043	0.025	-0.011	-0.068	-0.151	-0.250
ML	0.084	-0.038	-0.091	-0.085	-0.054	-0.032	-0.010	0.007	0.023	0.036	0.049
PML	-0.017	-0.040	-0.071	-0.116	-0.177	2.834	1.920	1.312	0.894	0.617	0.449
	n = 100										
mean	1.007	0.808	0.607	0.408	0.209	0.010	-0.187	-0.378	-0.552	-0.698	-0.807
MM	0.012	0.012	0.010	0.008	0.006	-0.004	-0.040	-0.140	-0.308	-0.498	-0.663
PWM	0.016	0.016	0.014	0.012	0.012	0.009	0.004	-0.007	-0.036	-0.098	-0.200
ML	0.023	-0.088	-0.121	0.034	0.085	0.017	0.000	0.001	0.005	0.008	0.010
PML	-0.003	-0.012	-0.029	-0.060	-0.112	1.997	1.388	0.972	0.686	0.489	0.353
	n = 1000										
mean	1.001	0.801	0.601	0.401	0.201	0.001	-0.199	-0.397	-0.589	-0.755	-0.873
MM	0.001	0.001	0.001	0.001	0.001	0.000	-0.008	-0.076	-0.275	-0.537	-0.754
PWM	0.002	0.001	0.002	0.001	0.001	0.001	0.000	-0.001	-0.010	-0.045	-0.131
ML	0.009	-0.015	-0.024	-0.033	0.078	0.005	0.000	0.000	0.001	0.001	0.001
PML	0.000	-0.002	-0.008	-0.023	-0.060	1.905	1.334	0.943	0.672	0.483	0.351

Table 2: Relative error of the five considered estimators on the local dimension, averaged over d and shown as a function of ξ, for sample sizes $n = 20$, 100, 1000.

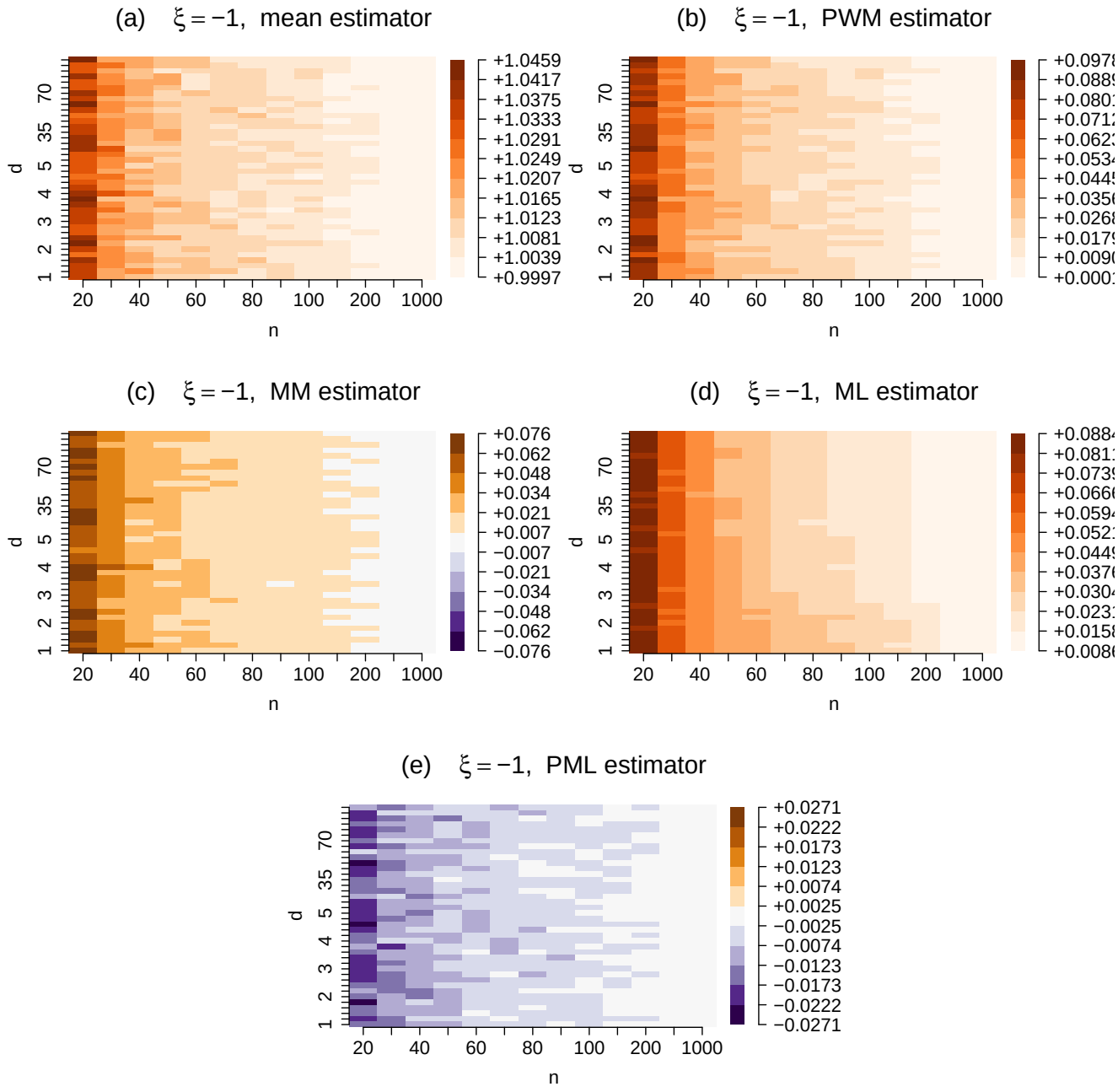

Figure 6: Relative error of the mean estimator (a), PWM (b), MM (c), ML (d) and PML (e) on d as a function of d and n, for $\xi = -1$.

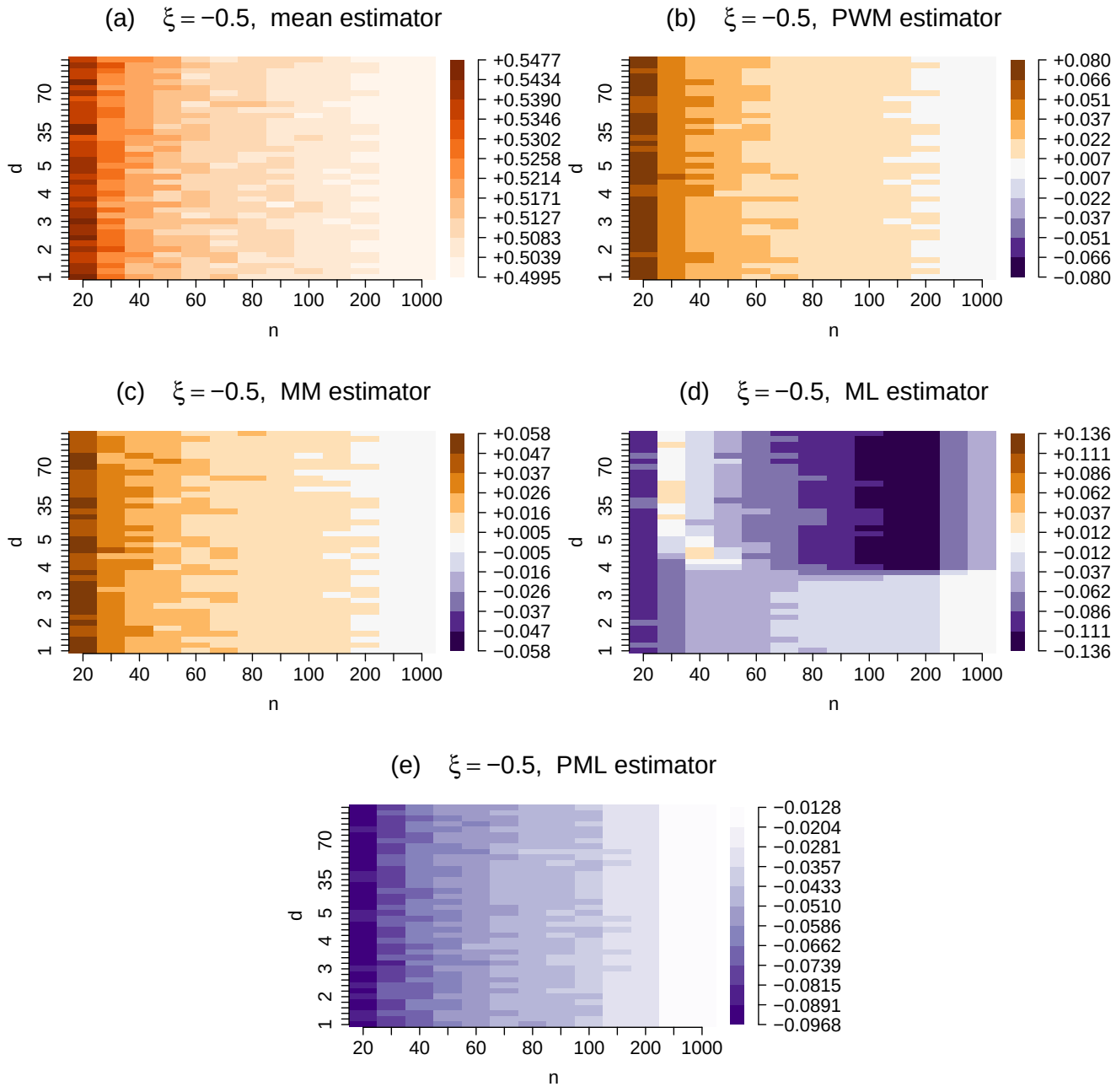

Figure 7: Relative error of the mean estimator (a), PWM (b), MM (c), ML (d) and PML (e) on d as a function of d and n, for $\xi = -0.5$.

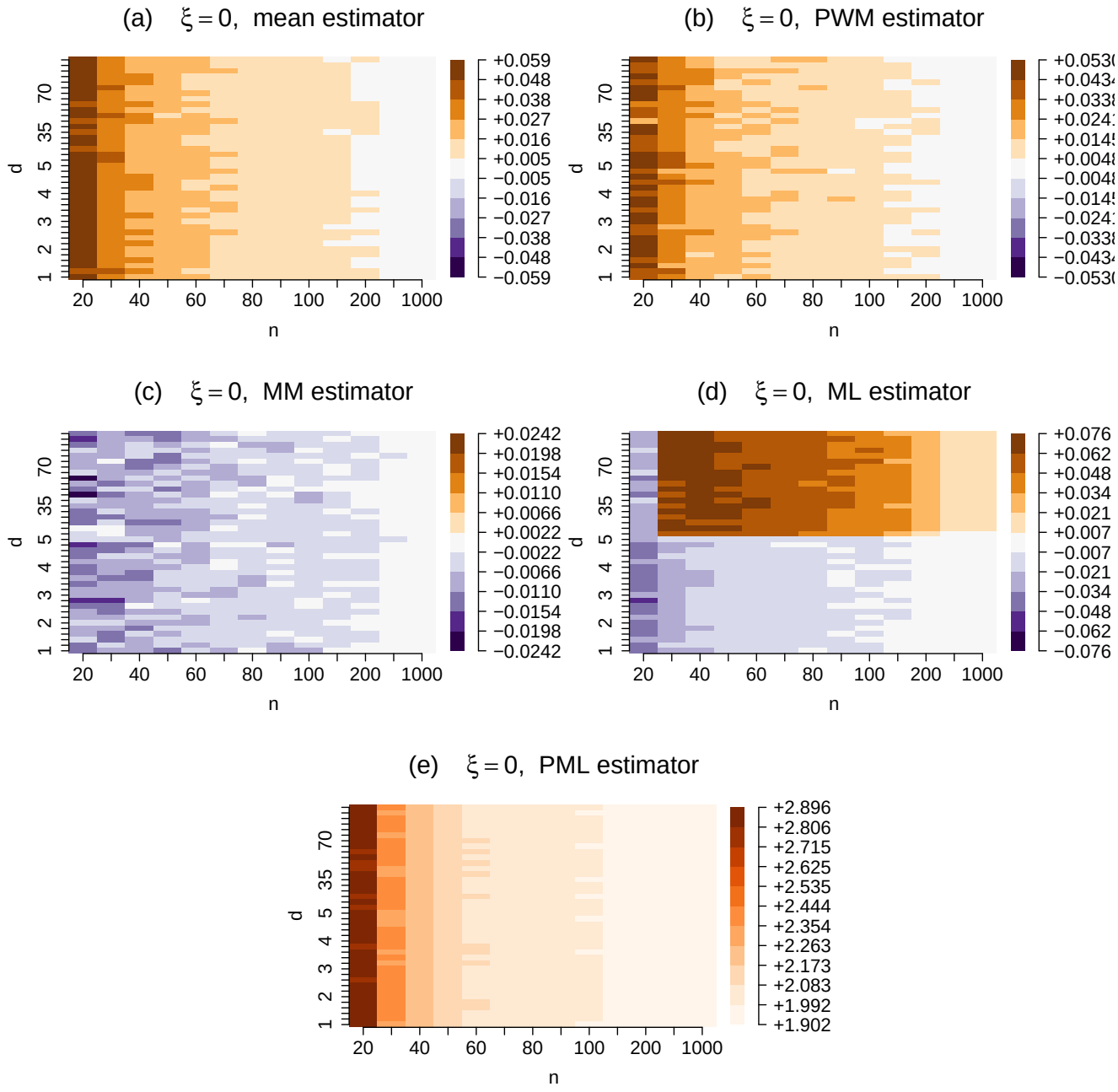

Figure 8: Relative error of the mean estimator (a), PWM (b), MM (c), ML (d) and PML (e) on d as a function of d and n, for $\xi = 0$.

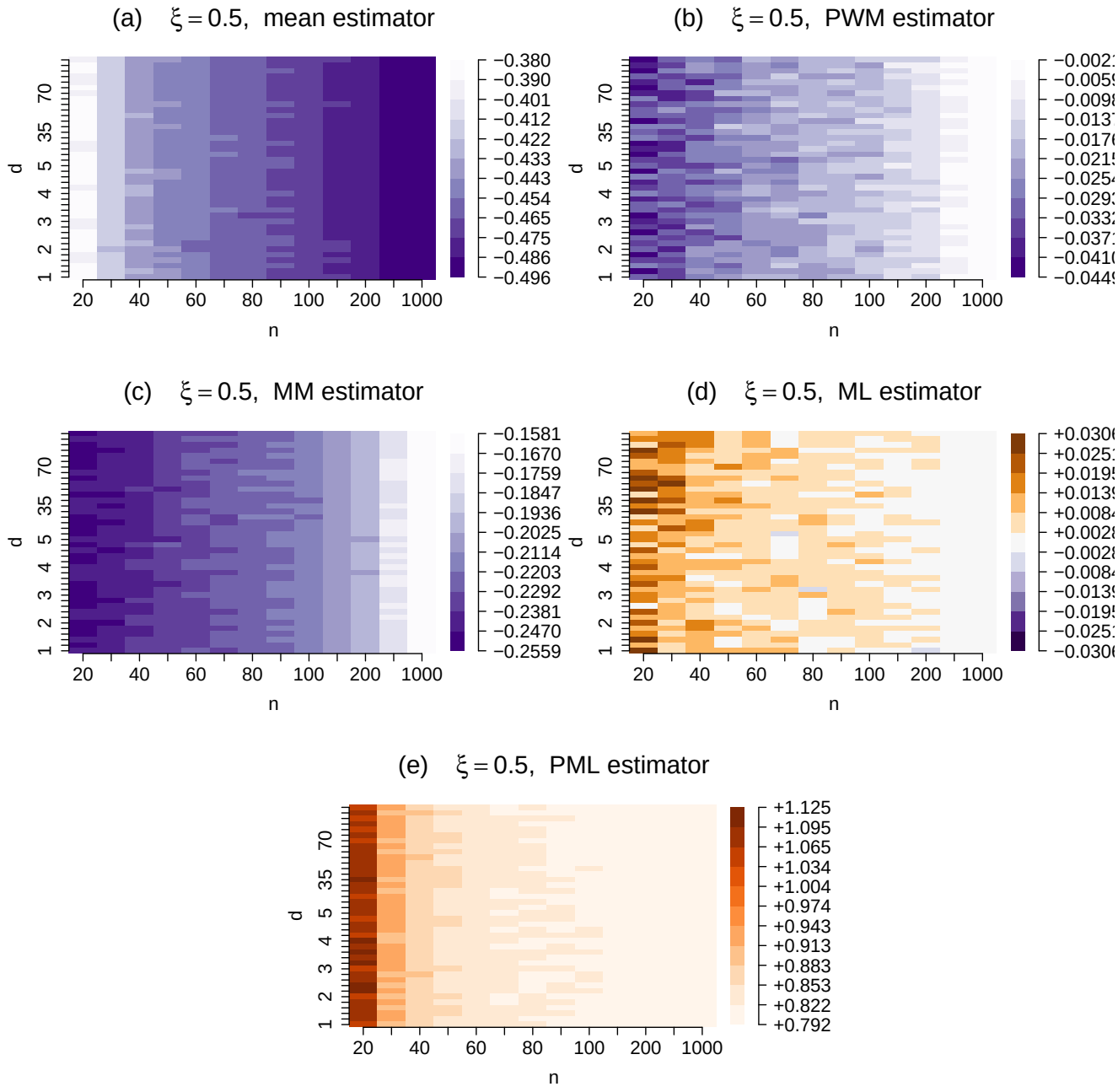

Figure 9: Relative error of the mean estimator (a), PWM (b), MM (c), ML (d) and PML (e) on d as a function of d and n, for $\xi = 0.5$.

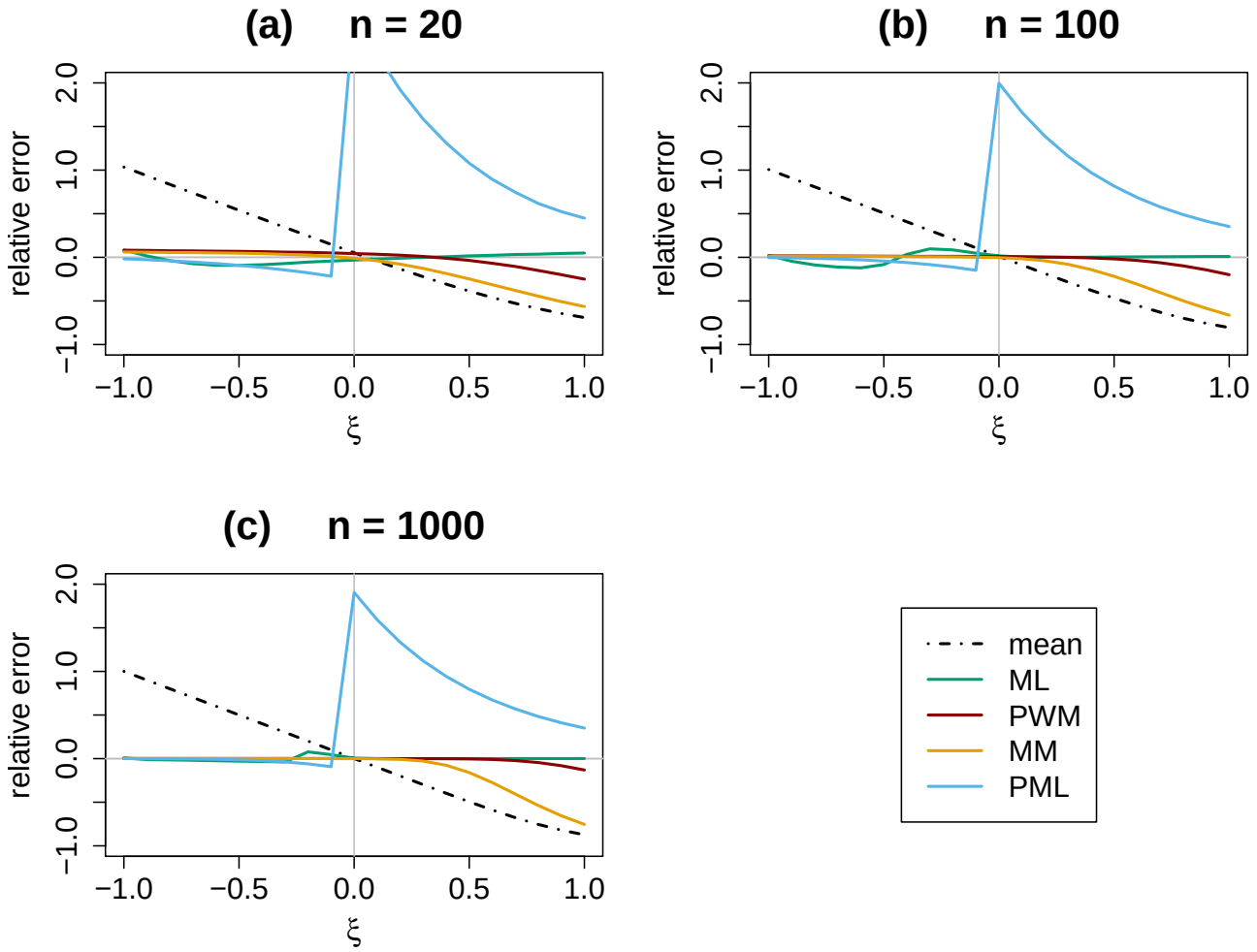

Figure 10: Graphic visualization of the results shown in Table 1. Relative error of the five considered estimators on the local dimension, averaged over d and shown as a function of ξ, for sample sizes $n = 20$ (a), 100 (b), 1000 (c).

5.2 Lorenz 1963 model

Next, we analyze the performance of the estimators on the local dimension of the L63 system. Its Hausdorff dimension is known to be $D = 2.06 \pm 0.01$ since the early work of Grassberger and Procaccia (1983), and more recently it has been estimated to be $D = 2.0627160$ (Viswanath, 2004). We choose the L63 system not only because its D is known, but also because in this case the estimation of the local dimension via Euclidean distance is not affected by the curse of dimensionality documented for higher-dimensional systems (Pons et al., 2020), since the phase space of L63 has dimension 3.

Before looking at the estimates of d and D, we consider the performance of the different

estimators on the shape parameter ξ. For the MM, PWM and PML estimators, $\hat{\sigma}$ can be written as a function of $\hat{\xi}$, and analysing ξ may bring insights on the source of estimation errors on d. Figure 11 shows the probability density functions of $\hat{\xi}$ for all estimators, except the mean estimator which assumes $\xi = 0$. We can notice marked differences: PML results in a roughly symmetric distribution, centered around $\hat{\xi}_{PML} \sim -0.15$, while regular ML produces a much wider density with its peak just below zero. The MM and PWM estimators both produce a bi-modal distribution, with a peak very close to zero and another one between -0.1 and -0.15. To better characterize this result, we used the break-point search based on segmented regression implemented in the `R` package `segmented` (Muggeo, 2003) to separate the two modes of each distribution. The change points are located at -0.058 for the MM and at -0.049 for the PWM estimator. Fig. 12 shows the two distributions, and the L63 attractor colored based on the local value of $\hat{\xi}$ obtained with the two estimators. While the two modes of each distribution seem to separate similar regions of the attractor, the values are inverted, so that the regions producing the lowest mode of $\hat{\xi}_{MM}$ correspond to the highest peak of $\hat{\xi}_{PWM}$, and vice versa. This is reflected in the negative correlation coefficient of -0.42 between $\hat{\xi}_{MM}$ and $\hat{\xi}_{PWM}$.

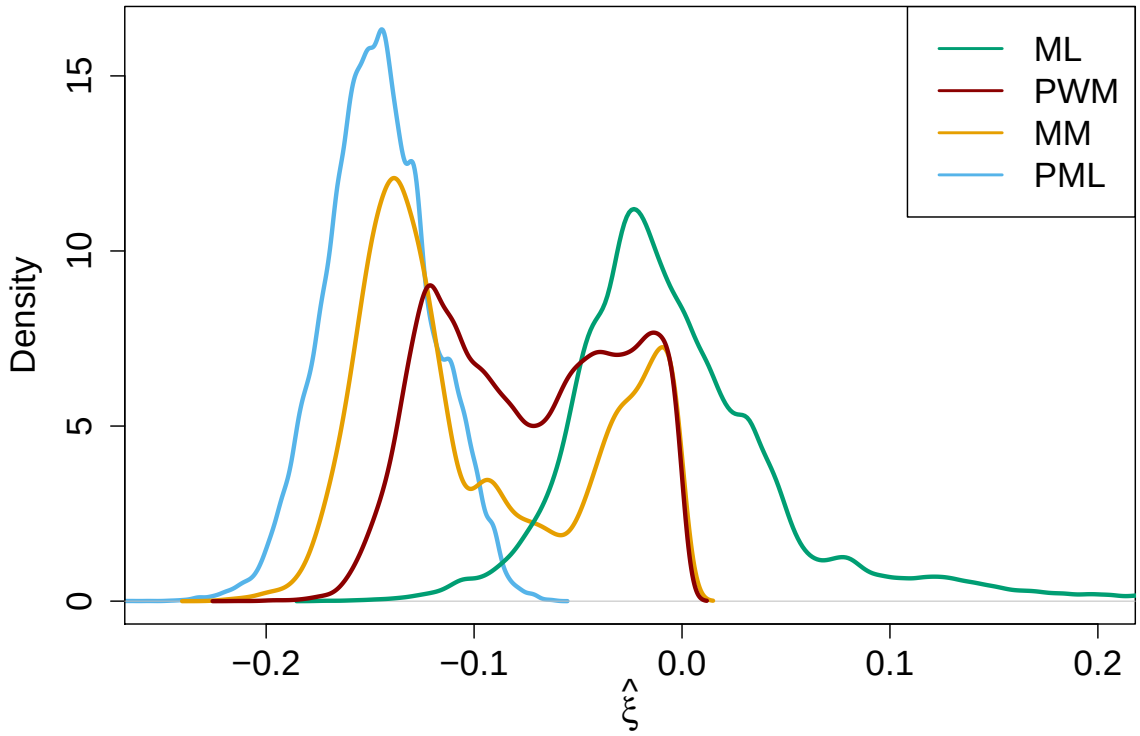

Figure 11: Probability density functions of the estimated local dimension of L63.

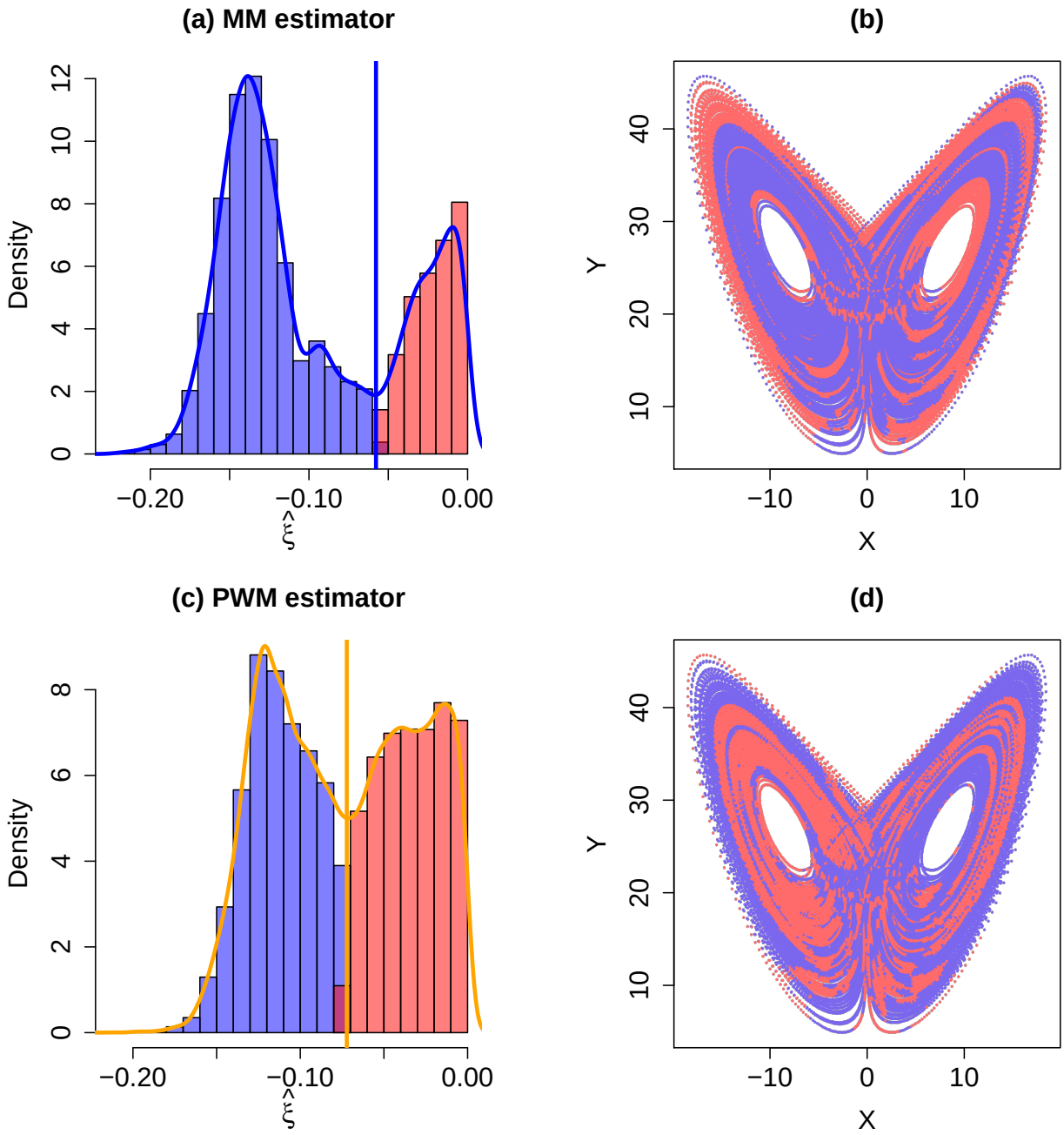

Figure 12: Attractor of the L63 system colored according to the shape parameter ξ estimated with the MM and PWM methods. The bimodal distributions of $\hat{\xi}$ have been separated with a breakpoint search algorithm.

In Fig. 13 we show the probability density function of the local dimension d for all the estimators, in logarithmic scale to reduce the visual effect of the large skewness. As expected from the Monte Carlo results, the PML estimator has by far the worst performance. To obtain the shown results, we have used $\hat{\xi}_{PWM}$ as initial estimate. Using $\hat{\xi}_{MM}$ (not shown) provided almost identical results. The resulting Hausdorff dimension estimate is $\hat{D}_{PML} = 4.259003$, not only very

different from the true value, but also larger than the phase-space dimension of the system, and thus non-physical.

The estimates of D obtained with the other estimators are overall much closer to the true value. We apply a Student's t test at the 5% level with Bonferroni correction to each mean, under the null hypothesis that $\hat{D} = 2.0627160$ against the bilateral alternative. All estimates are significantly different, except for $\hat{D}_{MM}$. The MM estimator also has the smallest RMSE, meaning that it provides not only the most accurate estimate, but also the best trade-off between bias and variance. However, it is worth to mention that variability in the estimates of the local dimension in this experiment is partially due to the structure of the attractor, so that this metric is less important than in the case of the Monte Carlo experiment.

Figure 14 shows the $X - Y$ plot of the L63 attractor, colored according to the local dimension obtained with the four remaining estimators. Values are saturated at 3, since larger dimensions are non-physical, but possible due to finite sampling effects. The four estimators show comparable performances, with largely superimposed high- and low-dimensional regions. This is confirmed by the large positive correlations between time series of estimated local dimension, ranging from 0.89 for the correlation between $\hat{d}_m$ and $\hat{d}_{PWM}$ to 0.99 for the correlation between $\hat{d}_{ML}$ and both $\hat{d}_{MM}$ and $\hat{d}_{PWM}$. Thus, the differences in the estimation of ξ, especially concerning the bimodality of the MM and PWM estimates discussed above, do not translate in equally visible differences in the estimation of d.

These findings corroborate the result already emerged in the Monte Carlo experiment that, for $\xi \leq 0$, the MM estimator provides the smallest bias and the best bias-variance trade-off.

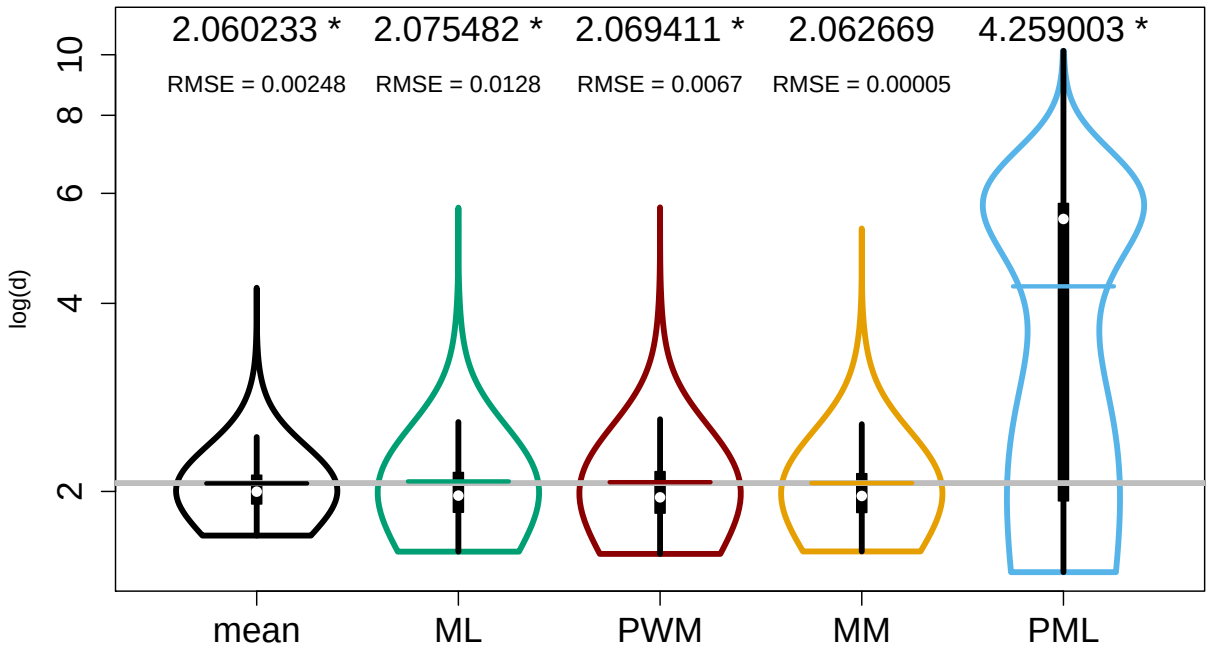

Figure 13: Probability density functions of the estimated local dimension for the L63 system in logarithmic scale. The horizontal grey line shows the value of the L63 Hausdorff dimension D_{L63}; horizontal coloured segments show estimates $\hat{D}$ obtained as the average of all local dimensions. The values are shown above each plot; stars indicates a statistical difference with respect to the theoretical value of D_{L63}. Values of root mean square error are also shown, with the exclusion of PML.

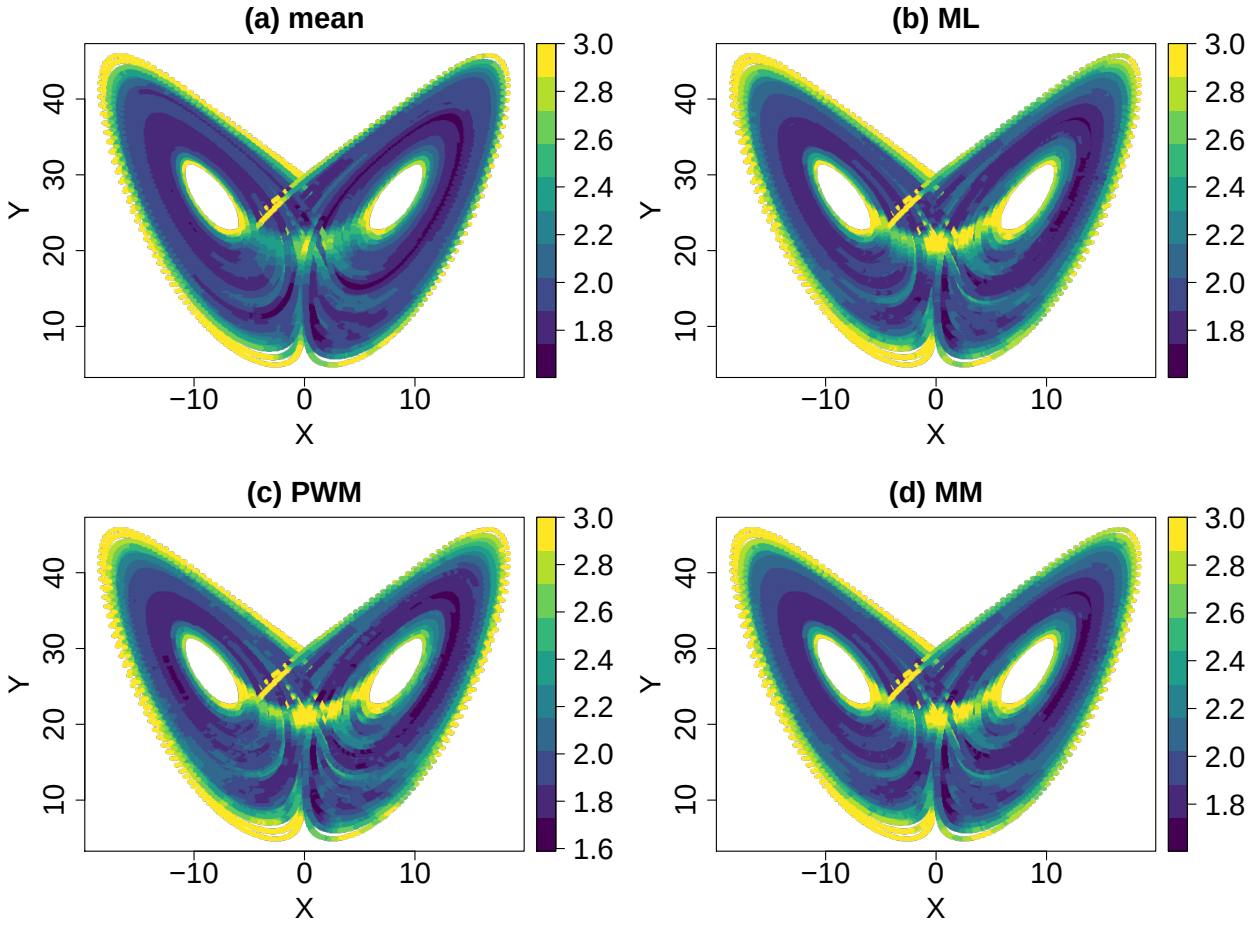

Figure 14: Attractor of the L63 system colored according to the estimated local dimension obtained with the estimator based on the mean (a), ML (b), PWM (c) and MM (d).

5.3 Climate Data

Finally, we replicate the previous analysis for the ERA5 dataset. In this case, it is not possible to compare estimates to the true value, as in the Monte Carlo experiment, or to previous knowledge about the attractor dimension, as for the L63 system. The goal of considering real-world data is rather to check the stability of local and attractor dimension estimates when using different estimators. This is particularly relevant because many existing climatological studies have been based on the mean estimator, while the climate system is expected to behave as a non-Axiom A system, characterized by $\xi < 0$.

Figure 15 shows the probability density functions of estimated values of ξ with all methods, except the mean estimator which assumes $\xi = 0$. Qualitatively, the distributions of ξ_{MM} and ξ_{PML} present the highest resemblance, with the main difference due to a peak of estimated values

close to 0 for the MM estimator. However, all possible combinations of two-sample Kolmogorov-Smirnov test reject the null hypothesis of identical distribution at the 5% level. The PWM and ML distributions are shifted towards larger values, with a large peak around 0 in the distribution of $\hat{\xi}_{ML}$.

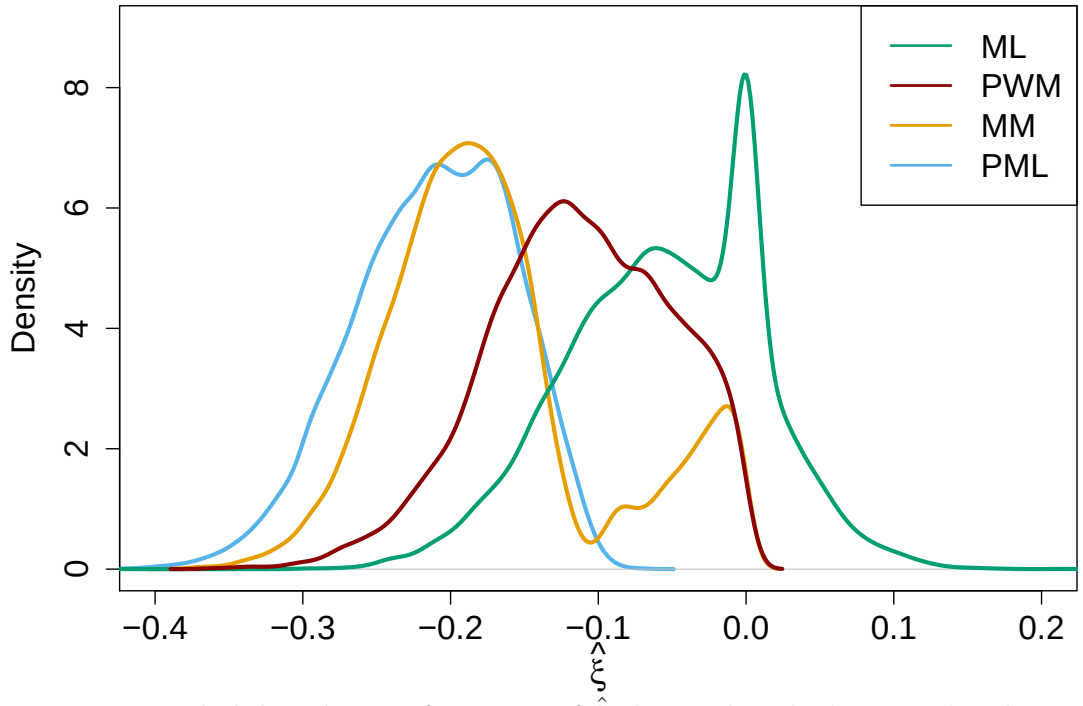

Figure 15: Probability density functions of $\hat{\xi}$ obtained with the considered estimators.

However, as in the case of L63, the differences in the estimation of ξ do not reflect on the local dimension estimates. All the estimators produce values $8 \leq \hat{D} \leq 9$ (see Fig. 16), with the exception of PML which gives an estimate $\hat{D}_{PML} = 23.4$, with values of the local dimension up to $\hat{d}_{PML} = 41.6$.

Also in this case, all possible combinations of Kolmogorov-Smirnov test reject the null hypothesis of identical distributions at the 5% level. We also test the differences among the means, i.e. the estimates of D, through pairwise Student's t test with Bonferroni correction for multiple test comparisons, without assuming equal variances in the distributions of d. All comparisons reject the null hypothesis of equal means at the 5% level.

While these differences are statistically significant, from the point of view of many climatological

applications, the estimates are equivalent. As an example of this type of application, we mention a study (Fery et al., 2022) that uses a machine learning algorithm to decompose the sea level pressure (slp) anomaly fields over the North Atlantic in a number of coherent structures. Each structure corresponds to a low or high pressure anomaly pattern, and can be seen as a degree of freedom of slp variability. To decide the optimal number of patterns, the authors construct a few statistics that allow to draw scree plots, and choose 28. They note that this value is compatible with the range of local dimensions estimated for the same atmospheric fields using the mean estimator (Faranda et al., 2017b). In this type of analysis, where an integer approximation of $\hat{D}$ is used to determine an appropriate order of magnitude, any of the four estimates shown in Fig. 16 can be considered compatible and would provide consistent results.

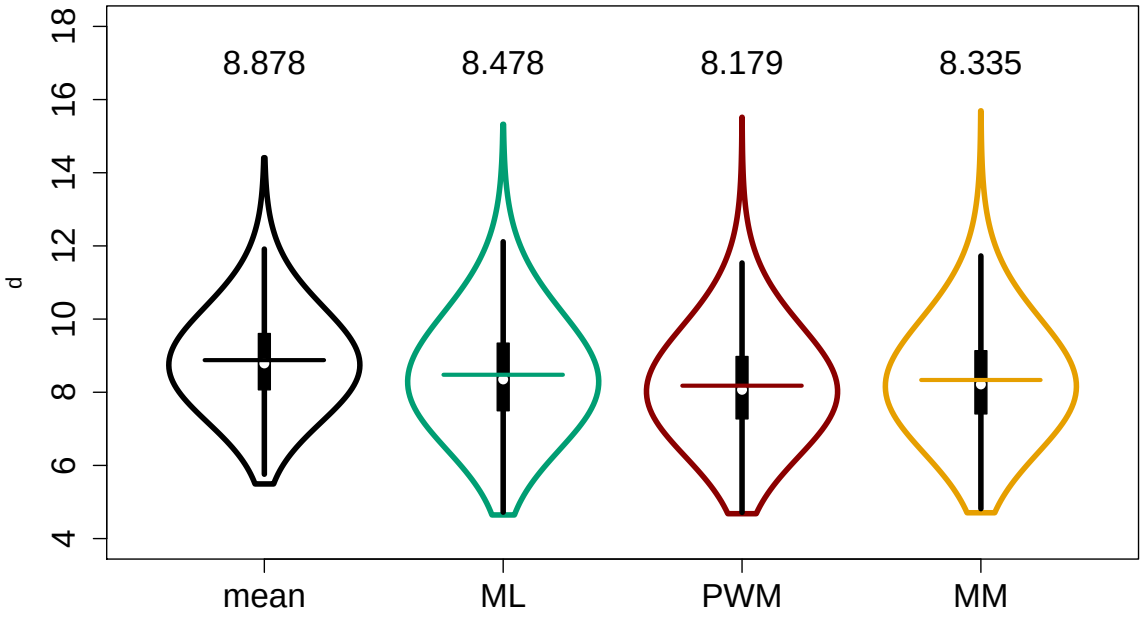

Figure 16: Probability density functions of $\hat{d}$ obtained with different estimators. The horizontal color segments and the numbers above the violin plots indicate the estimated Hausdorff dimension $\hat{D}$.

Other previous applications have used estimates of the (instantaneous) local dimension. For example, past studies have investigated temporal properties of $\hat{d}_m$, such as its seasonality and relation to a dynamical measure of persistence (Faranda et al., 2017b,a), or long-term trends in

reanalysis and climate projections (Faranda et al., 2019a). A simple way to check the similarity between time series is Pearson's correlation, as it is insensitive to mean shifts or scaling factors, but it attains small values if two time series are not synchronized. The upper triangle of the correlation plot in Fig. 17 shows values of Pearson's correlation between pairs of estimated local dimension time series. Corresponding scatterplots with regression lines are depicted in the lower triangle. Both low correlation values and scatterplots show that the differences between $\hat{d}_{PML}$ and the other estimators are not simply due to a scale factor, but rather to the bimodality in the distribution of $\hat{d}_{PML}$. We will not investigate the issues of this particular estimator, that may be due to the implementation of the `R` package `gPdtest`. At this point, it is clear that the other estimators constitute better alternatives, both because they are more accurate, and because they do not require a prior knowledge of the sign of ξ.

The correlations among the other estimators are positive and large, ranging between 0.88 and 0.97. This shows that, even under the erroneous assumption of $\xi = 0$, the mean estimator produces time series of $\hat{d}_m$ well synchronized with the ones obtained by its competitors that assume a GPD as the exceedances distribution. To ensure that seasonality and modes of interannual variability are consistently reflected by all the estimators, we also inspect the autocorrelation function (ACF) and the spectral density function (SDF) of the corresponding time series of $\hat{d}$, both shown in Fig. 18. We show the ACF up to 730 daily lags, corresponding to 2 years. The four best estimators follow a similar behavior, with a clear yearly cycle. PML estimates show a faster decay of the correlation and very weak seasonal signal, with values mostly inside the confidence interval constructed under the white noise null hypothesis. The SDF shows a dominant peak at 2 years, and two secondary peaks at 1 and 3 years. All estimators (except PML which has an overall flat spectrum, not shown) reflect these dominant frequencies, although the mean estimator shows a lower peak at the yearly frequency. This suggest that the mean estimator may display a subdued seasonal cycle. However, the overall performance of the four methods is largely comparable.

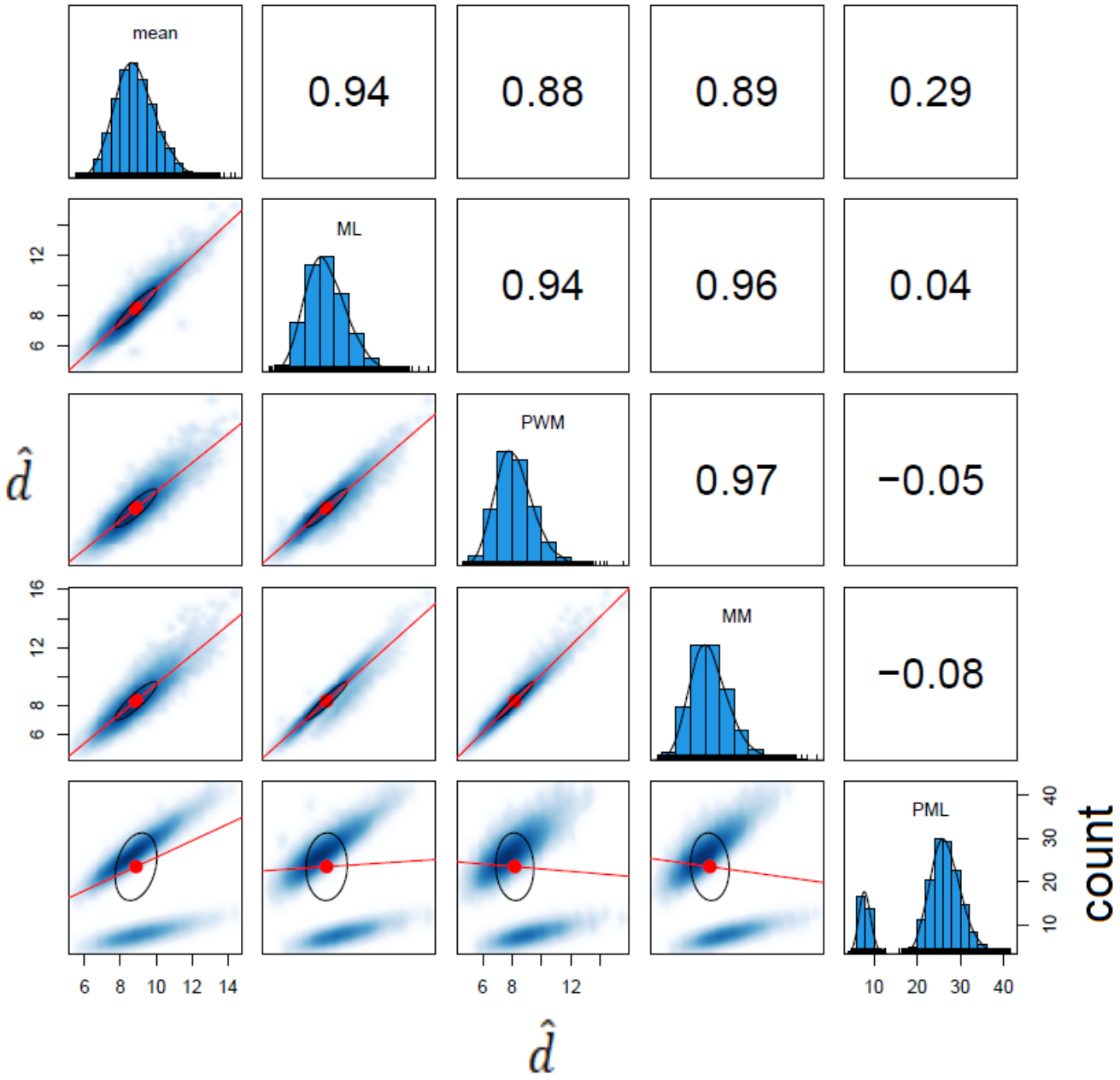

Figure 17: Correlation plot among the time series of $\hat{d}$ obtained with all estimators. Panels in the upper triangle show pairwise Pearson's correlation coefficient. Panels in the lower triangle show pairwise scatterplots, with superimposed correlation ellipse and linear regression line. The red dots denote the scatterplot centers of mass. Panels along the diagonal show the histograms of the marginal distributions of $\hat{d}$.

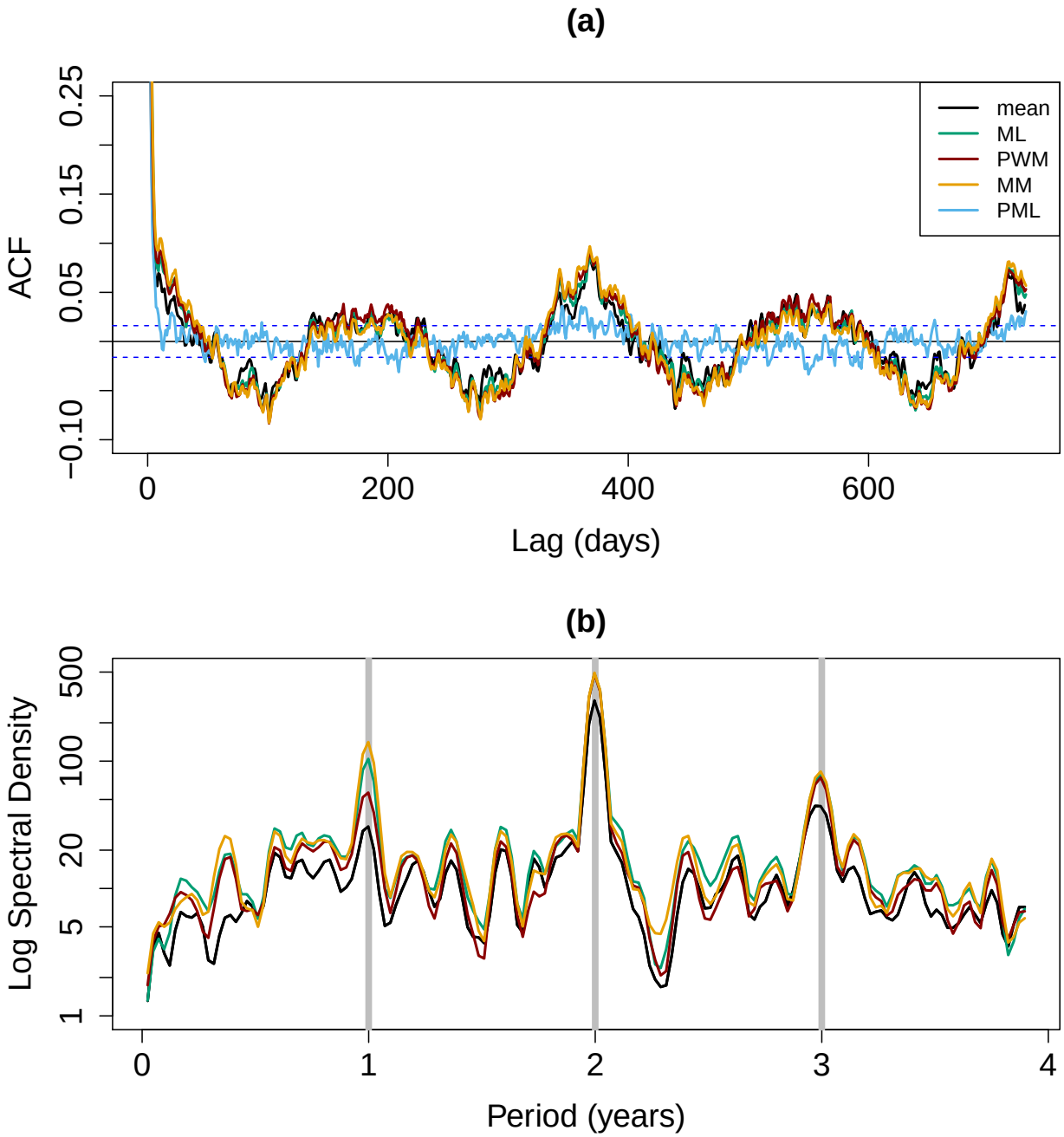

Figure 18: (a) Global autocorrelation function of $\hat{d}$ for all estimators, at lags up to 730 days. Dashed blue lines show the 95% confidence interval. (b) Logarithm of the spectral density function of $\hat{d}$ for all estimators except PML, at periods up to 4 years. Vertical grey lines mark the main frequencies, showing annual and interannual variability.

6 Discussion

We have assessed different estimators of the local attractor dimension of a dynamical system, by using extreme value theory for peaks over threshold (POT). This method also provides a

way to estimate the Hausdorff dimension, which can be obtained as the average over all local dimension, provided that the observed trajectories are long enough. We have considered four methods assuming that POTs follow a Generalized Pareto Distribution (GPD) with parameters ξ, σ: the method of moments (MM), probability weighted method of moments (PWM), maximum likelihood (ML) and pseudo-ML (PML) estimators. Notice that the MM estimator is only defined for $\xi < 0.5$, and that PML estimation requires two different procedures depending on the sign of ξ. We also added a naive estimator, namely the inverse of the POTs mean, based on the simplified assumption of exponentially distributed POTs, which is valid in the asymptotic case for Axiom A dynamical systems.

We have performed a Monte Carlo experiment by simulating the POTs directly from a GPD over a realistic range of true parameter values. In this way, we were able to measure the bias of each estimator as a function of the true parameter space. We assessed the performance in terms of the root mean square error and of the relative error on the local attractor dimension, given by the inverse of the GPD scale parameter σ.

In terms of statistical efficiency as measured by the root mean square error (RMSE), the overall best estimates are provided by the PWM. However, for dynamical systems whose POTs are characterized by $\xi < 0$, the MM estimator produces the best estimates and it is very easy to compute, without any tuning parameter such as the $c = 0.35$ coefficient in the definition of the plotting positions for the PWM. Mean estimator, ML and PML produce large RMSE over some range of $\xi < 0$, indicating that their performance is less optimal in terms of statistical efficiency.

Results also show that the relative error does not depend on the true value of σ, and convergence towards a fixed value is overall reached for a POT sample size $\sim$200. The PWM estimator shows the overall best performance, and the MM estimator performs equally well in its domain of existence $\xi \lessapprox 0.5$, despite the theoretical superiority of ML estimators under well-specified statistical models. The PML performance degrades rapidly for growing $\xi \geq -0.2$. The average relative errors of the naive estimator display a smooth, monotonic relationship with the true dimension value as a function of the true value of ξ. The relationship is linear for the mean estimator (with slope equal to -1).

We further test the five estimators when applied to the Lorenz 1963 system – a 3-equation system in a 3-dimensional phase space – and some high complexity climate data from the ERA5 reanalysis. For the Lorenz system, the Hausdorff dimension has been previously estimated, so we can use that as a reference value. The results are consistent with those for the Monte Carlo data, with the MM estimator showing the best performance for $\xi < 0$. For the climate data we have no prior estimate of the dimension, yet we know that the data is characterized by a negative estimated ξ. We thus consider the MM estimator as our baseline. We are particularly interested in testing the performance of the mean estimator, as it has been repeatedly used in the literature but there has been no systematic comparison of how it performs on complex systems relative to alternative estimators. The correlations between the mean and the MM estimator are very close to or even equal to 1, corroborating the idea that, even far from the asymptotic sampling of Axiom A systems, the simple mean estimator can be effectively used. Correlations of MM with the other estimators, with the exception of PML, are also very high.

Since for the MM, PWM and PML estimators $\hat{\sigma}$ can be written as a function of $\hat{\xi}$, we have also inspected the performance of the different methods in estimating the shape parameter, to check whether it could be a source of estimation error on σ. Results suggest the opposite: in both cases, the PML estimates are the only one characterized by a unimodal distribution centered on a realistic value, while PML estimates of the local and Hausdorff dimension are both inaccurate and characterized by large variability, compared to the competitors. ML is centered around zero for the L63 system and has a steep mode around zero for climate data, and both MM and PWM show negative values with a second mode near zero. In the L63 case, MM and PWM show a very well defined bimodality, but the values belonging to the two modes tend to correspond to well defined but different regions of the attractor. Despite this, estimates of the local dimension have the highest correlation among all pairs, with a coefficient of 0.97. Overall, we find that if one is interested only in the estimation of the local attractor dimension (or simply the σ parameter in a different setting), the performance of the chosen estimator on ξ does not seem to be crucial.

To conclude, our results can be used to draw simple, general guidelines for the choice of attractor dimension estimators. If it is known a priori that $\xi < 0$, the MM estimator is the best choice.

However, if no prior assumption can be made on the true value of ξ, the mean estimator could be the most suitable choice. In contrast, despite the popularity of maximum likelihood in statistical estimation and testing, it appears that the considered PML estimator only performs well in a narrow interval of values of ξ, which does not include ξ values usually observed in climate data. Moreover, the different estimators (with the exception of PML) show a high correlation. When one is interested in relative values of the dimension, the choice of estimator thus does not appear crucial. This is often the case when studying the behaviour of a system through the attractor dimension, as the relationship between dimension values of different states is more important than the value itself, since dimensions are usually compared to one another to detect changes in system behaviour, rather than interpreted in terms of their absolute values.

GM acknowledges the support of the European Union's H2020 research and innovation programme under ERC grant no. 948309 (CENÆ). FP and DF received support from the European Union's Horizon 2020 research and innovation programme under grant agreement No. 101003469 (XAIDA). DF and GM received further support from the European Union's Horizon 2020 Marie Sklodowska-Curie grant agreement No. 956396 (EDIPI) and the LEFE-MANU-INSU-CNRS grant "CROIRE".

Data Availability Statement

The ERA5 reanalysis data (Hersbach et al., 2018) that support the findings of this study are openly available on the Climate Data Store of the Copernicus Climate Change Service (http://doi.org/10.24381/cds.ad

A Supplementary Figures

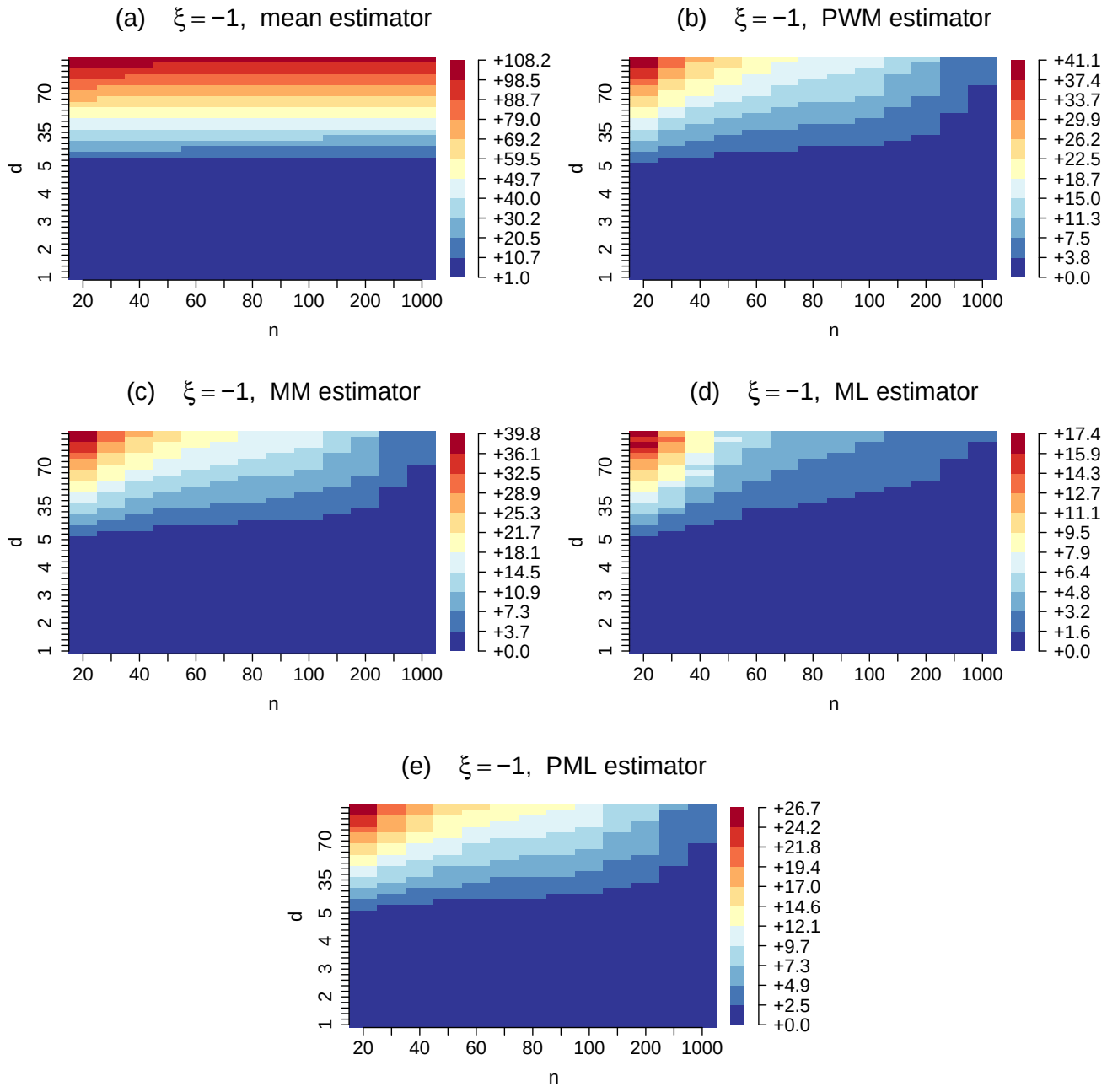

Figure 19: Root mean square error of the mean estimator (a), PWM (b), MM (c), ML (d) and PML (e) on d as a function of d and n, for $\xi = -1$.

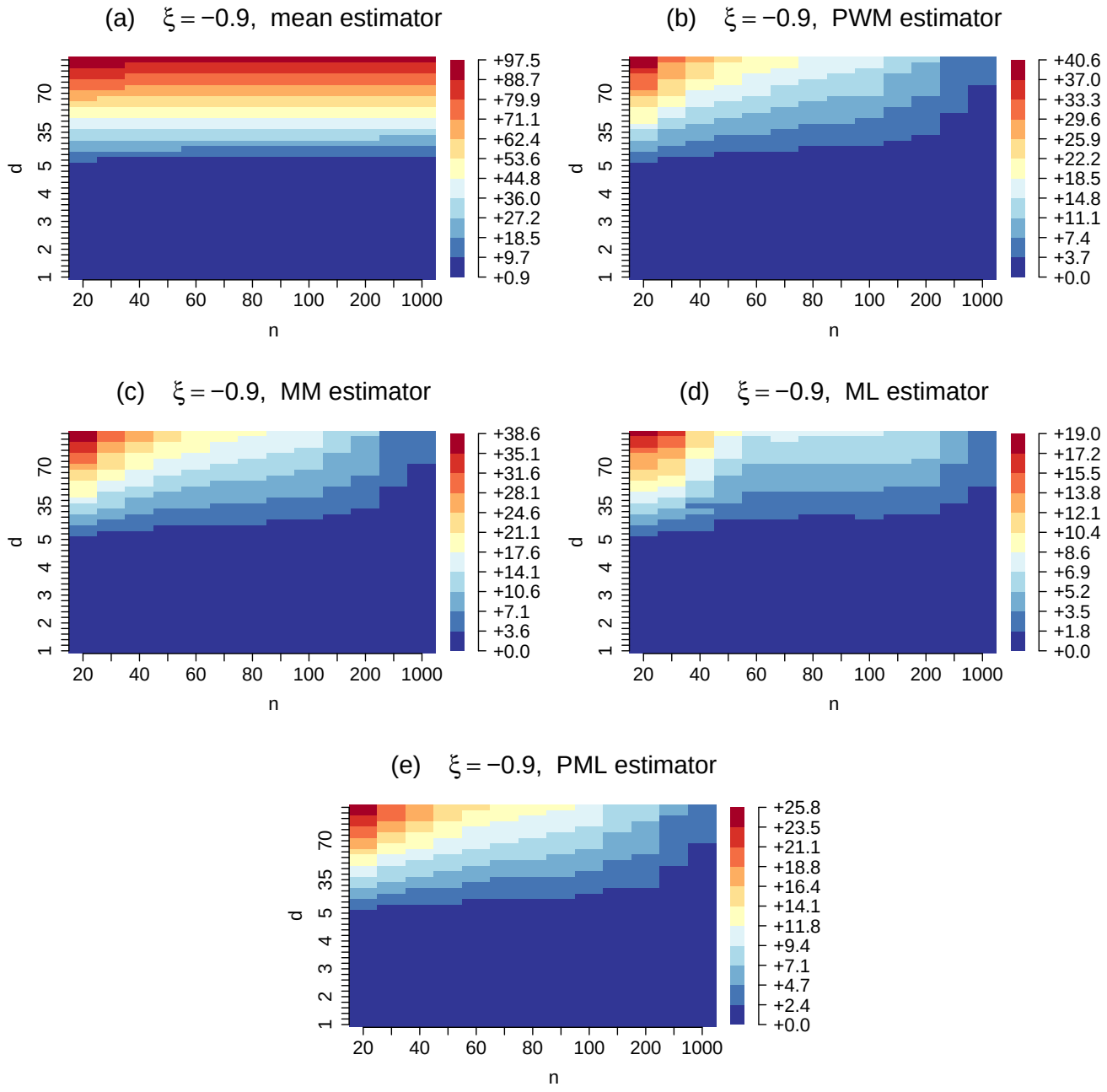

Figure 20: Root mean square error of the mean estimator (a), PWM (b), MM (c), ML (d) and PML (e) on d as a function of d and n, for $\xi = -0.9$.

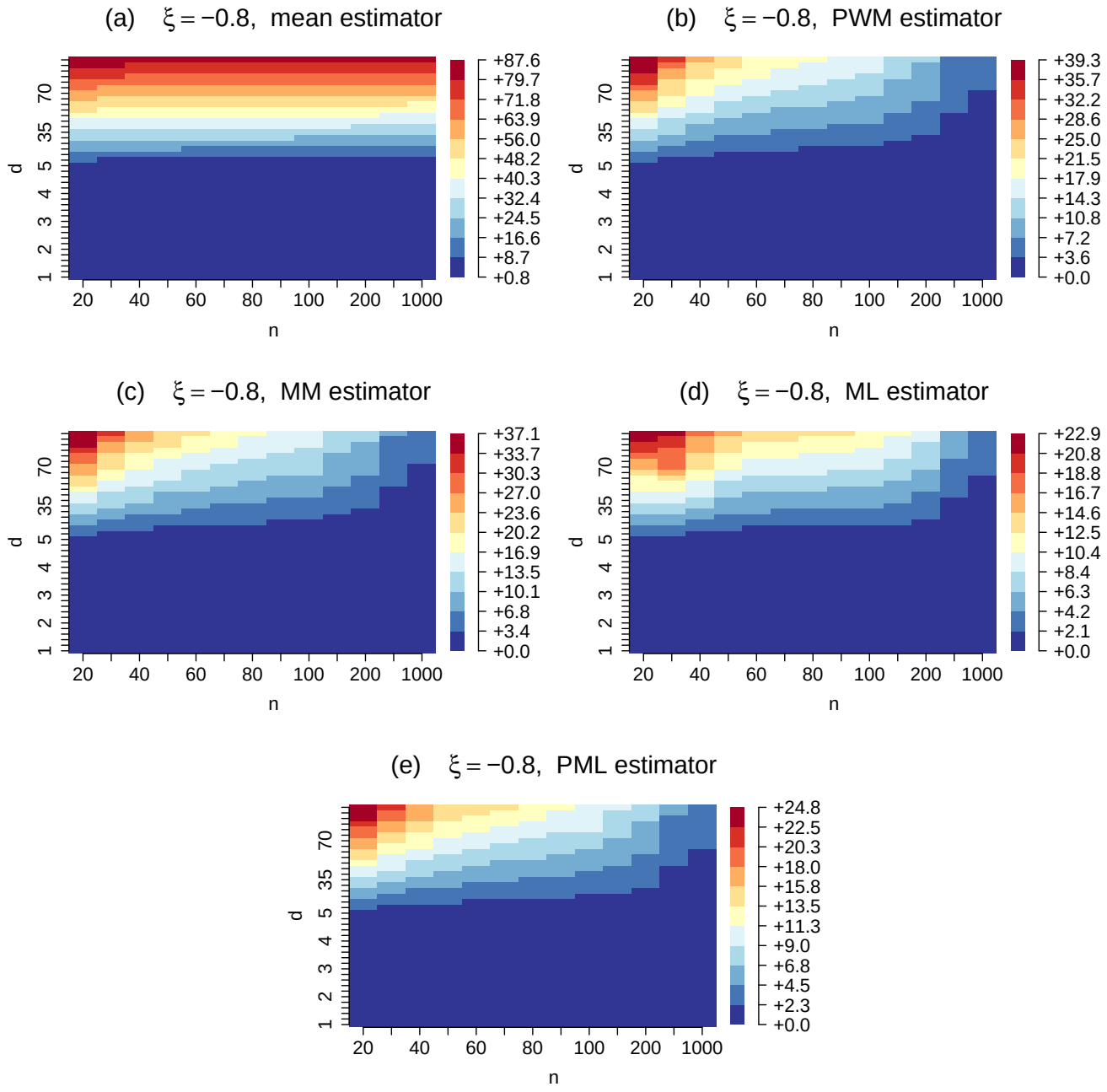

Figure 21: Root mean square error of the mean estimator (a), PWM (b), MM (c), ML (d) and PML (e) on d as a function of d and n, for $\xi = -0.8$.

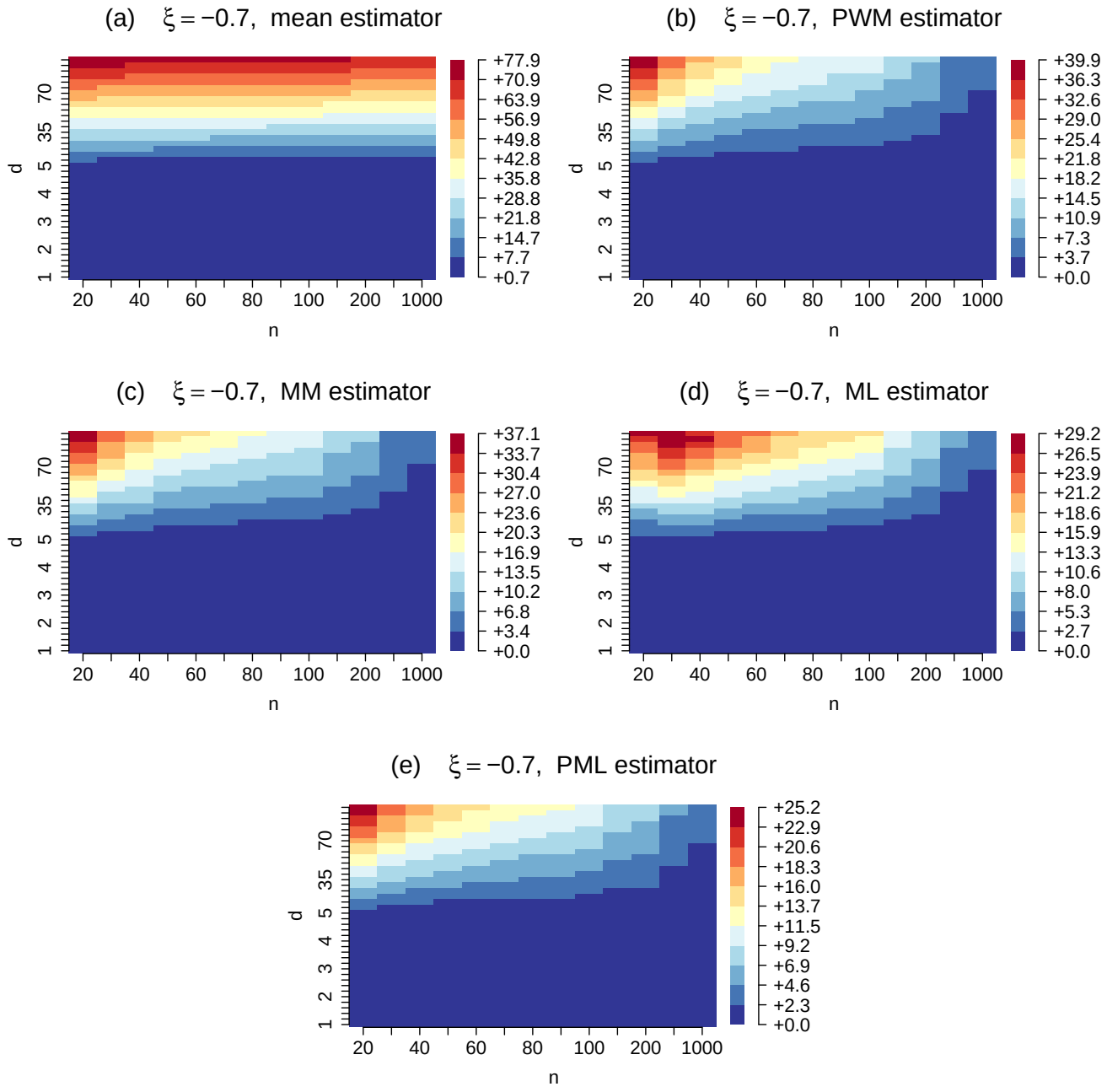

Figure 22: Root mean square error of the mean estimator (a), PWM (b), MM (c), ML (d) and PML (e) on d as a function of d and n, for $\xi = -0.7$.

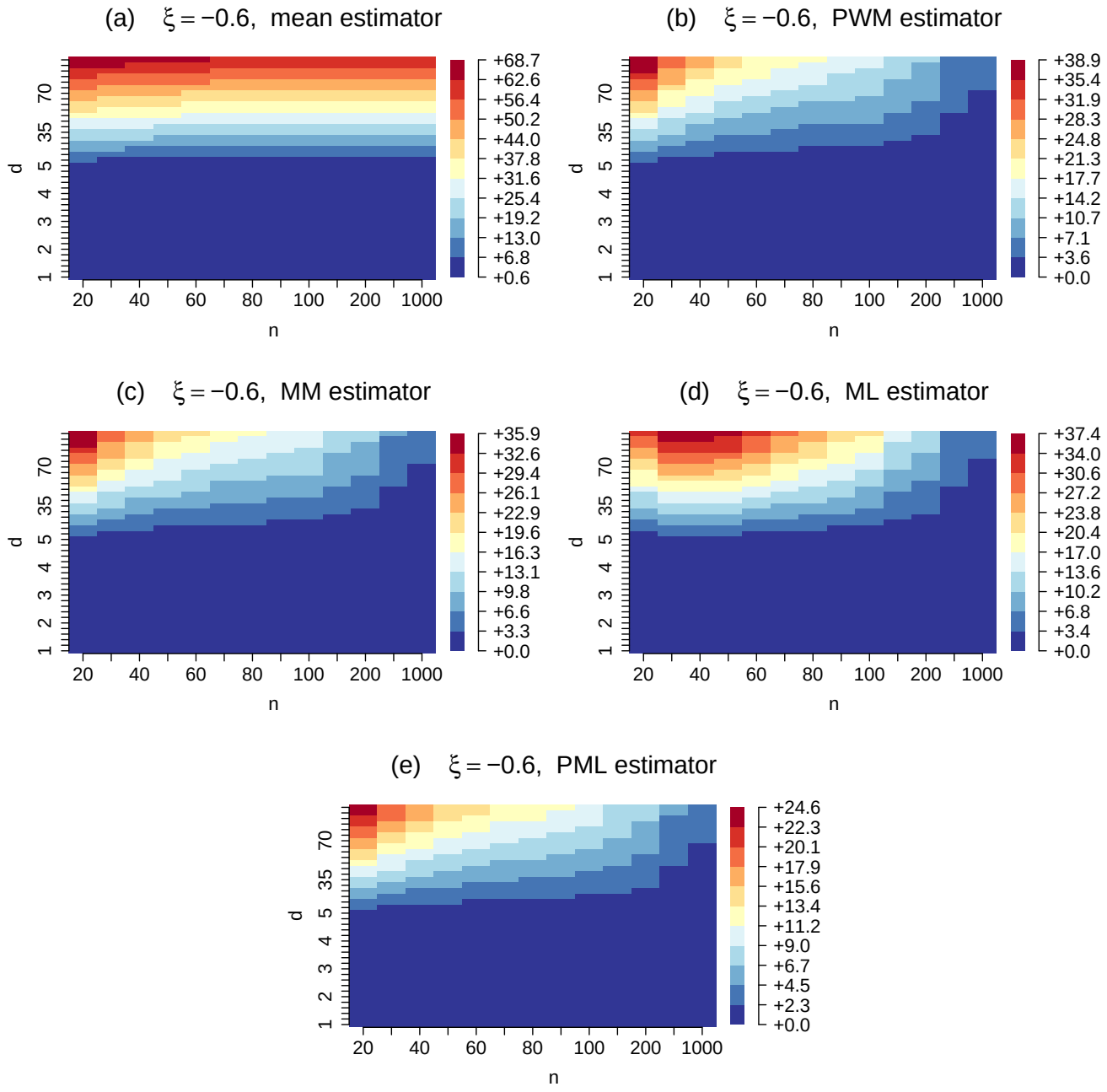

Figure 23: Root mean square error of the mean estimator (a), PWM (b), MM (c), ML (d) and PML (e) on d as a function of d and n, for $\xi = -0.6$.

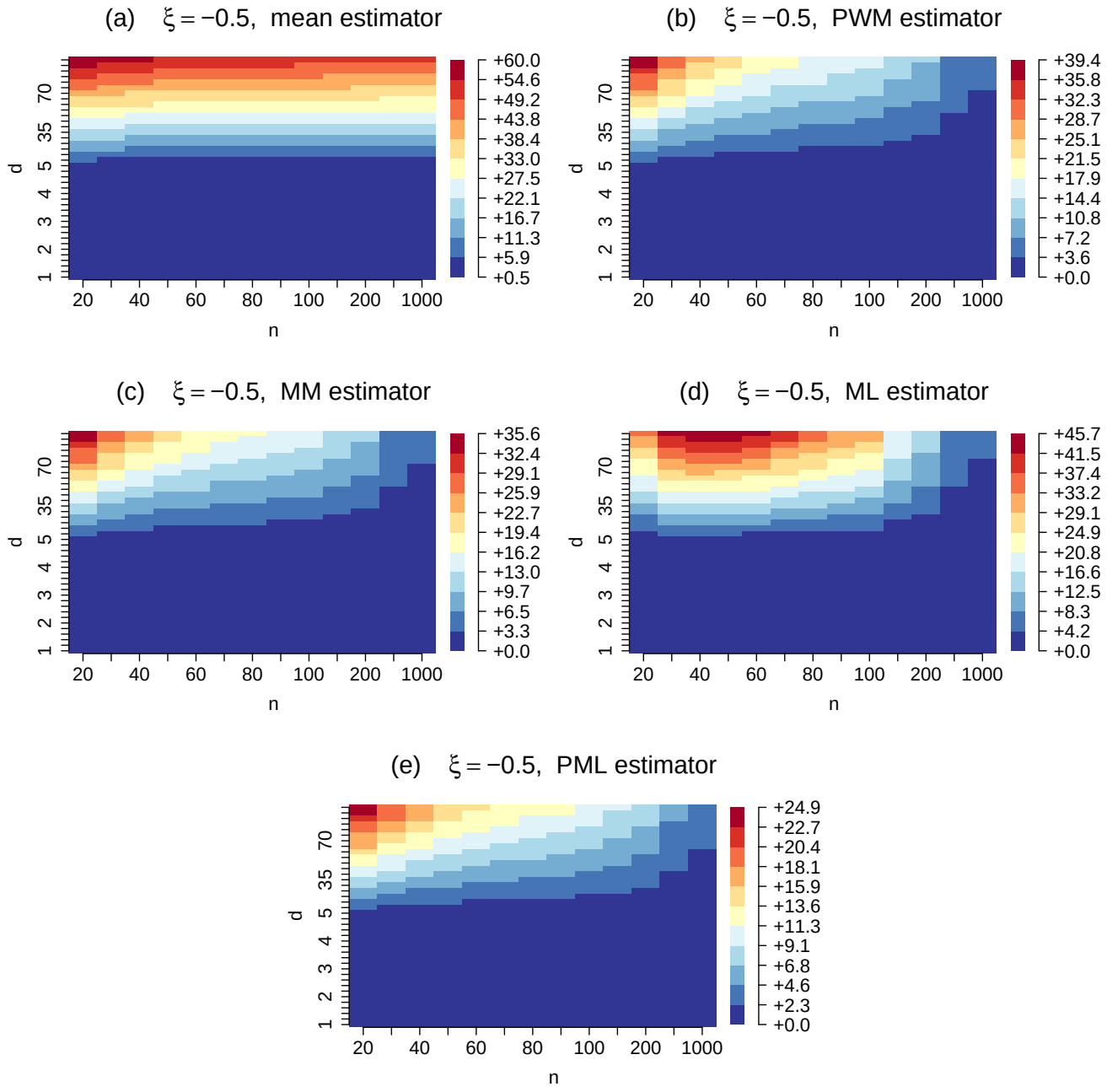

Figure 24: Root mean square error of the mean estimator (a), PWM (b), MM (c), ML (d) and PML (e) on d as a function of d and n, for $\xi = -0.5$.

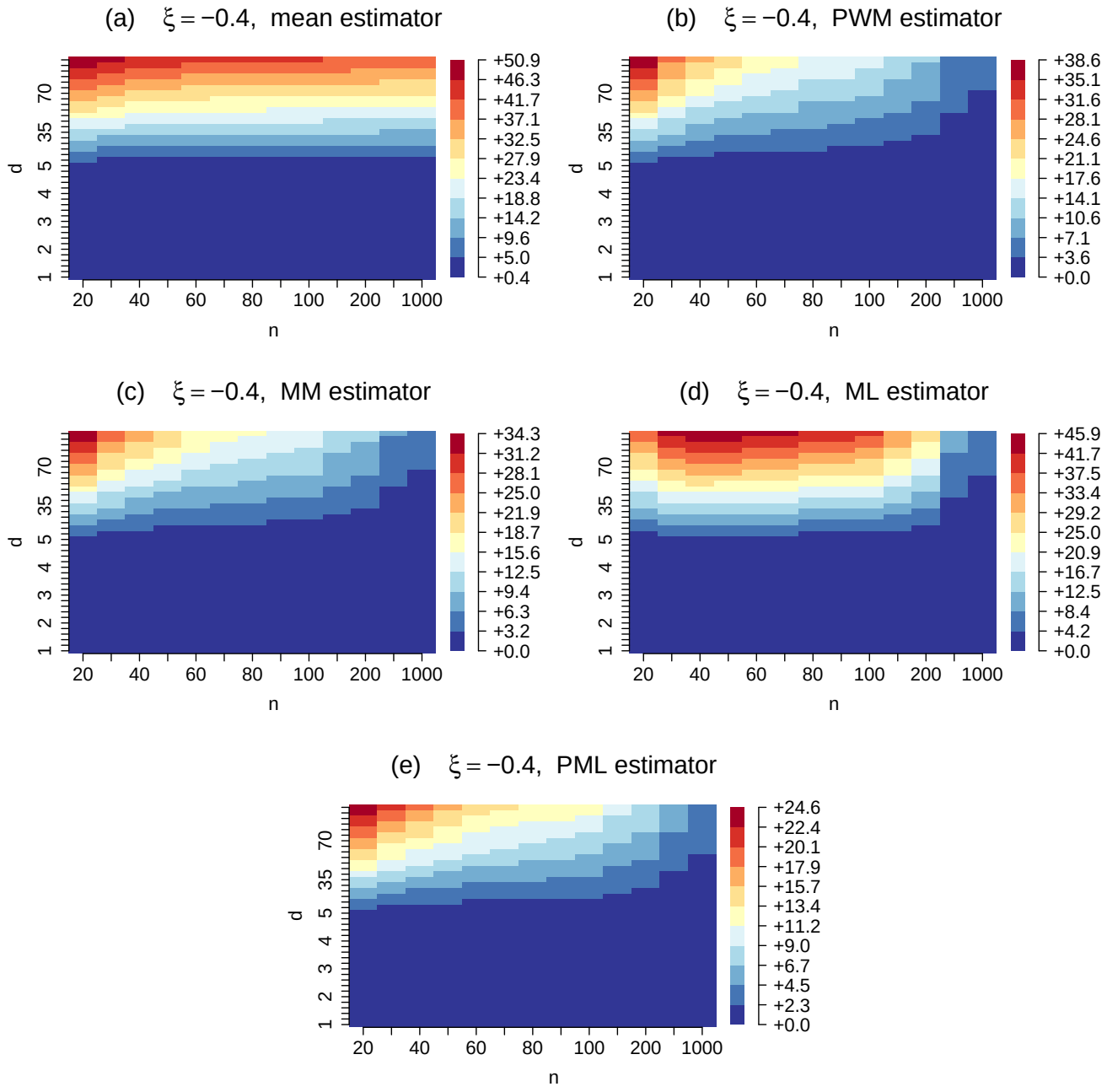

Figure 25: Root mean square error of the mean estimator (a), PWM (b), MM (c), ML (d) and PML (e) on d as a function of d and n, for $\xi = -0.4$.

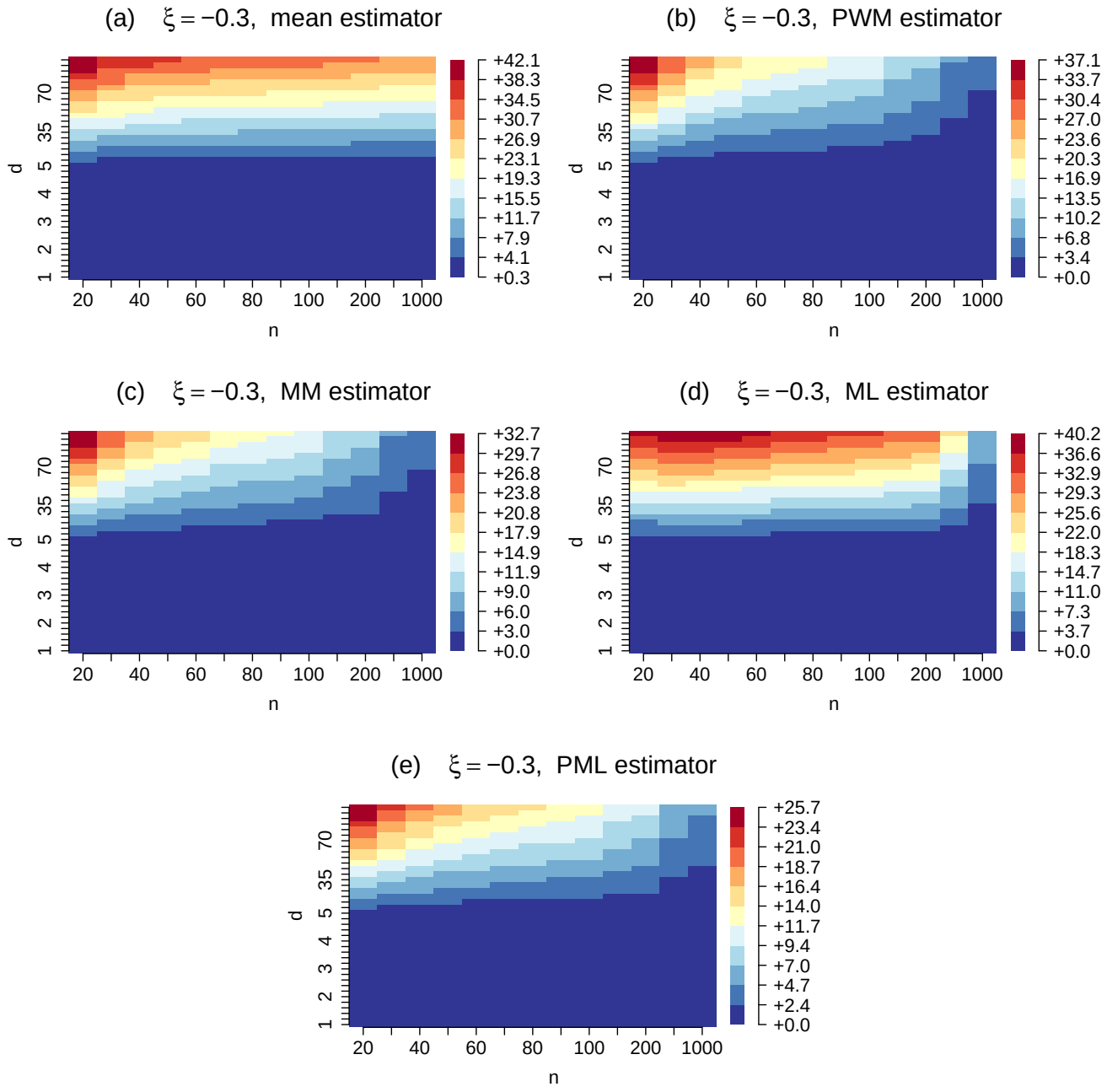

Figure 26: Root mean square error of the mean estimator (a), PWM (b), MM (c), ML (d) and PML (e) on d as a function of d and n, for $\xi = -0.3$.

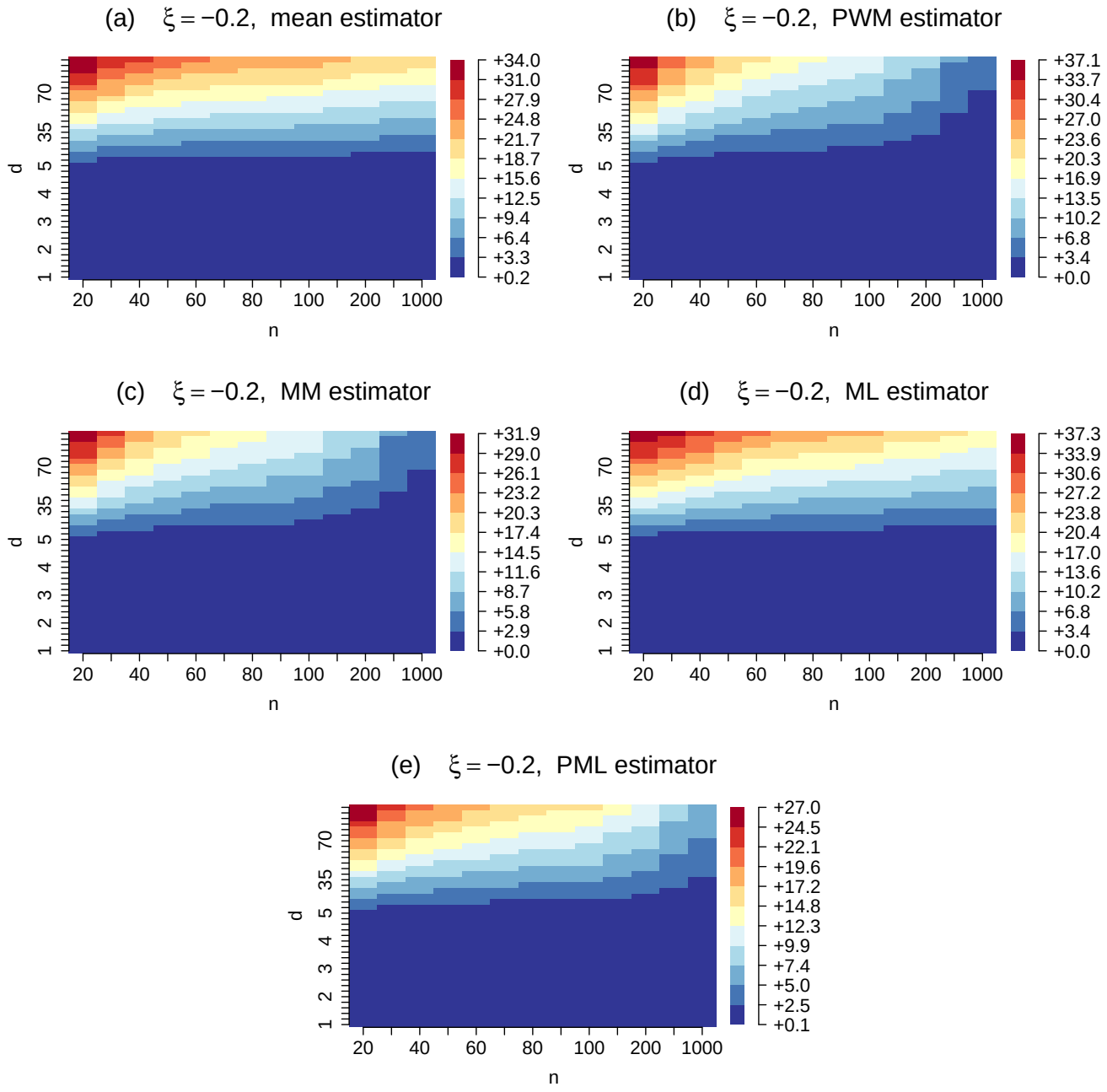

Figure 27: Root mean square error of the mean estimator (a), PWM (b), MM (c), ML (d) and PML (e) on d as a function of d and n, for $\xi = -0.2$.

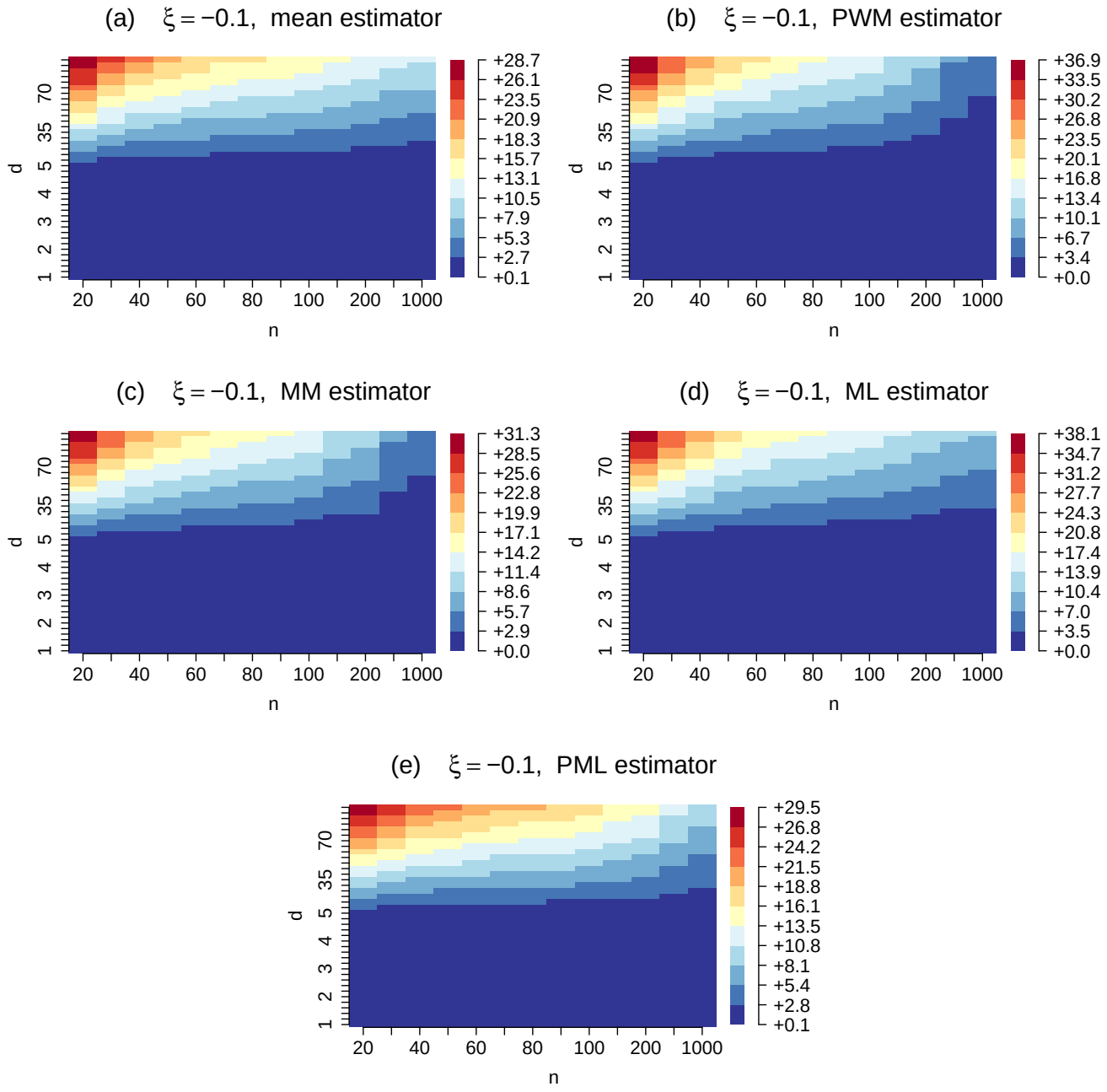

Figure 28: Root mean square error of the mean estimator (a), PWM (b), MM (c), ML (d) and PML (e) on d as a function of d and n, for $\xi = -0.1$.

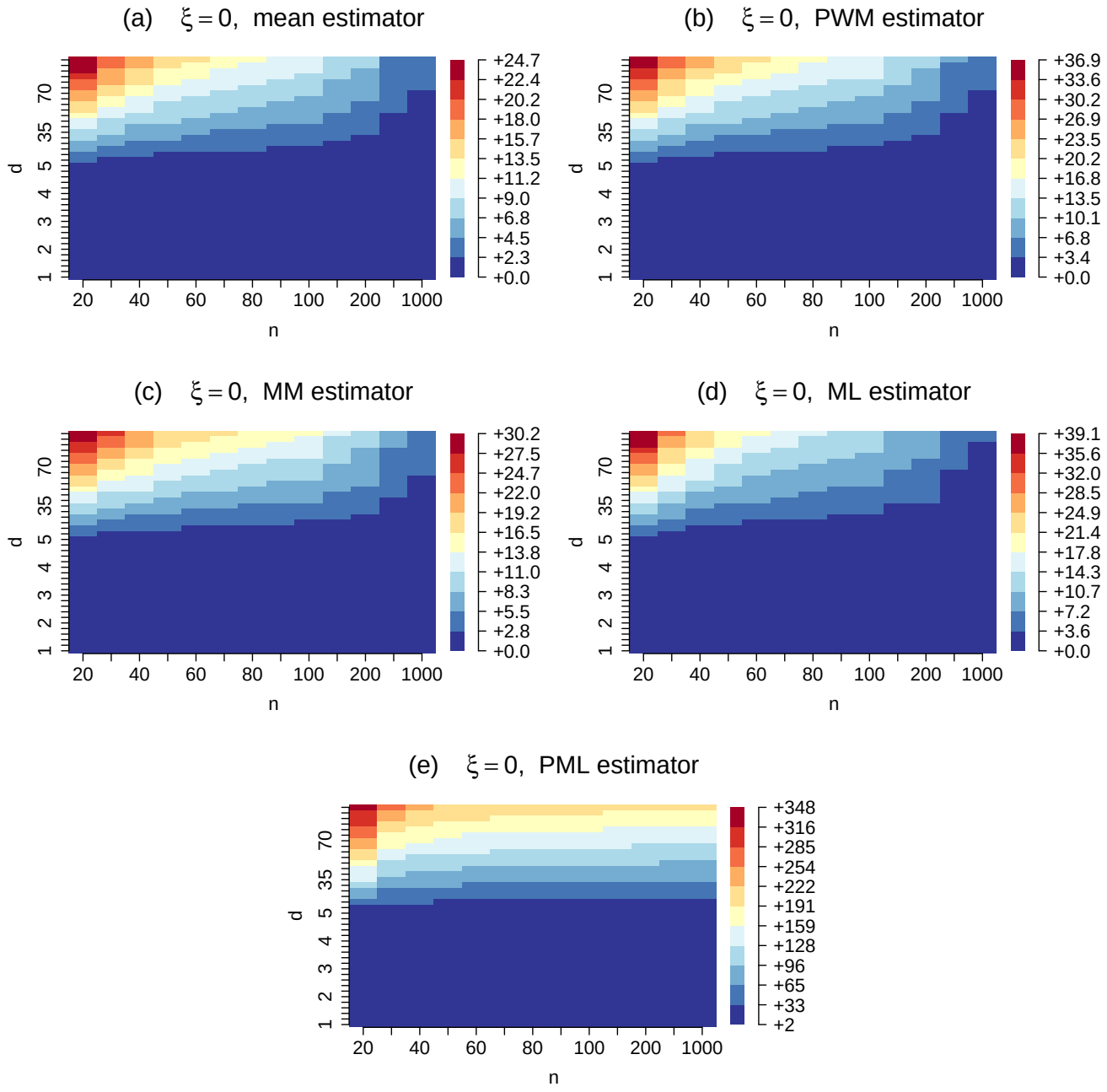

Figure 29: Root mean square error of the mean estimator (a), PWM (b), MM (c), ML (d) and PML (e) on d as a function of d and n, for $\xi = 0$.

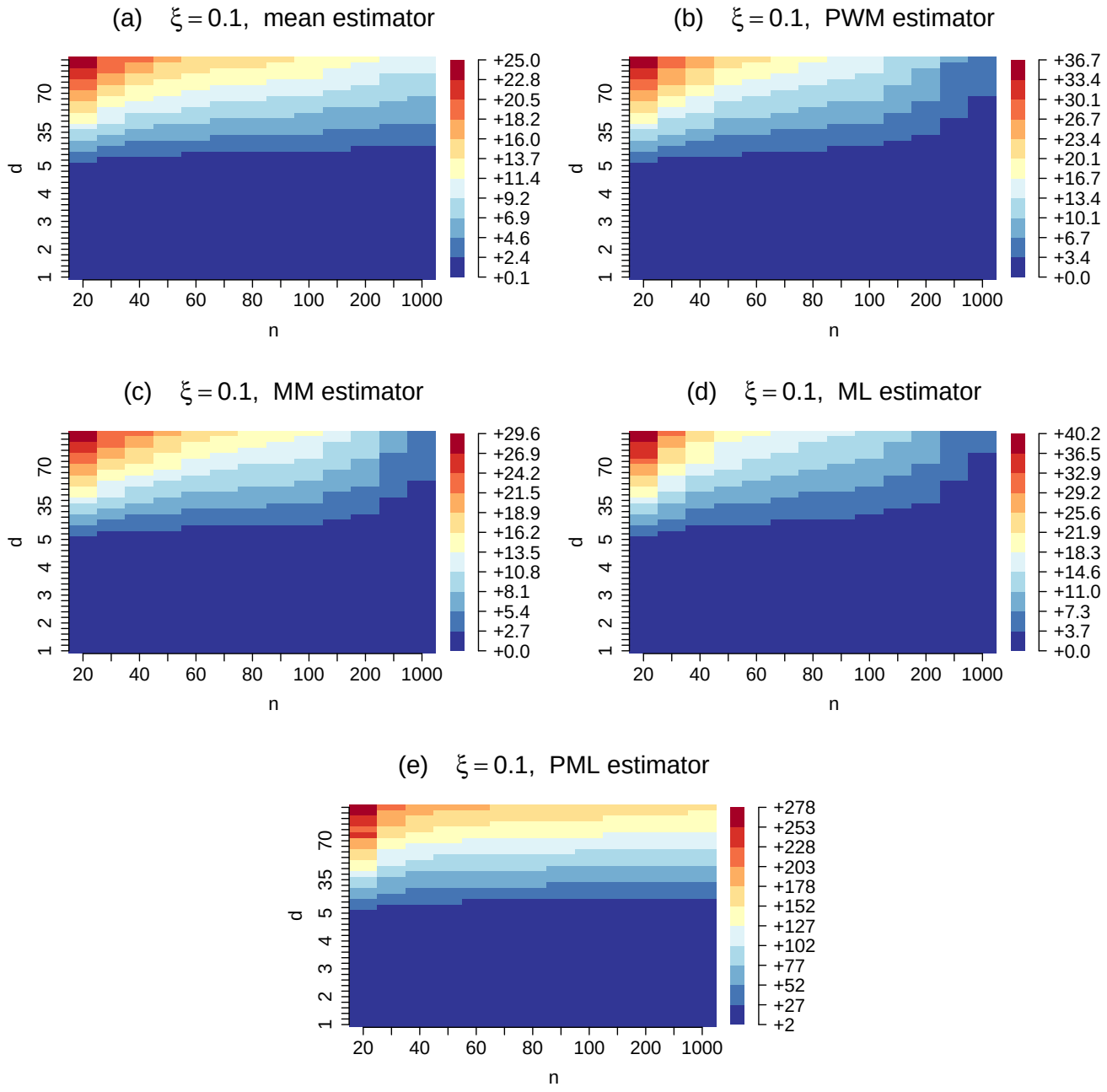

Figure 30: Root mean square error of the mean estimator (a), PWM (b), MM (c), ML (d) and PML (e) on d as a function of d and n, for $\xi = 0.1$.

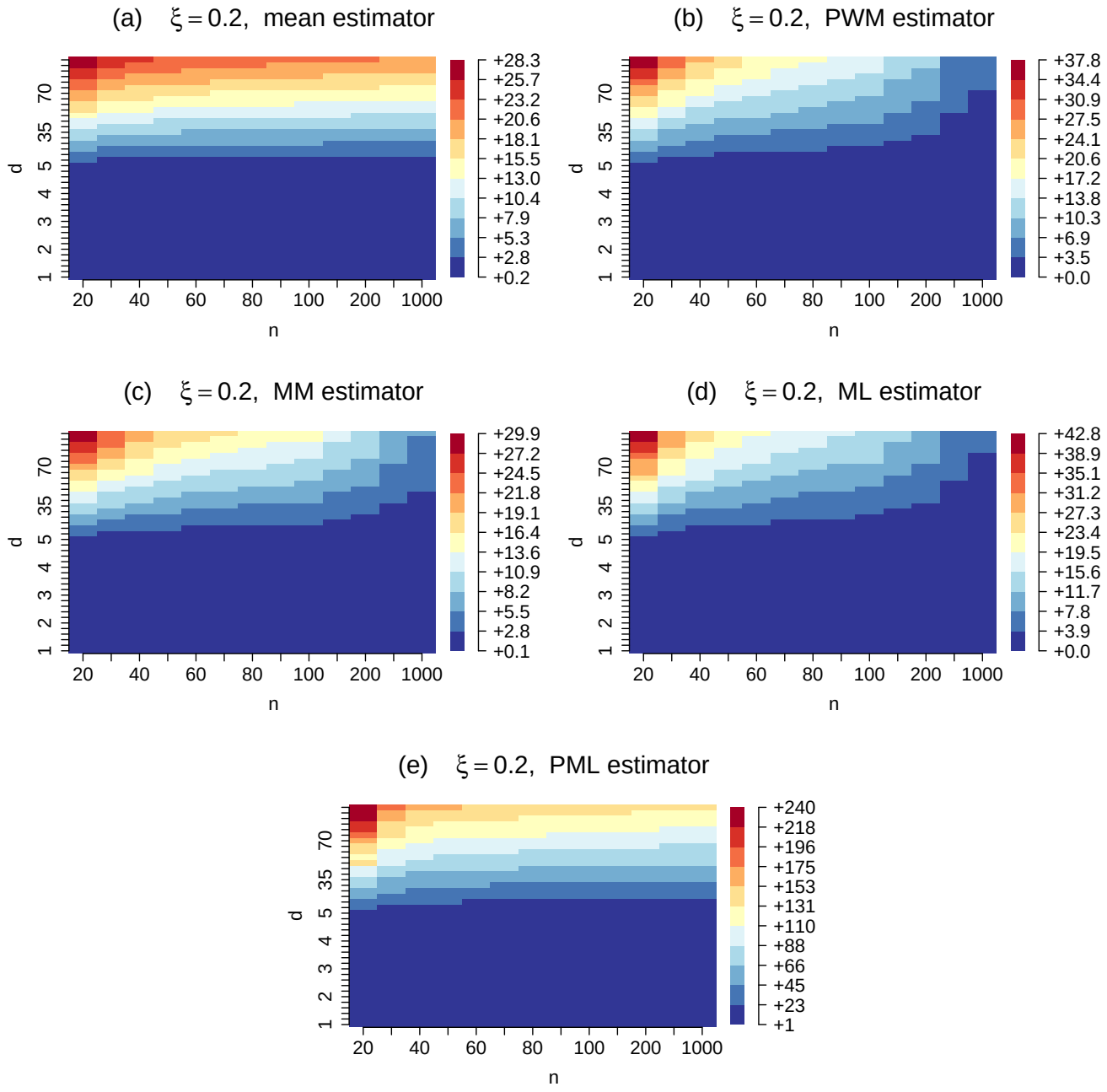

Figure 31: Root mean square error of the mean estimator (a), PWM (b), MM (c), ML (d) and PML (e) on d as a function of d and n, for $\xi = 0.2$.

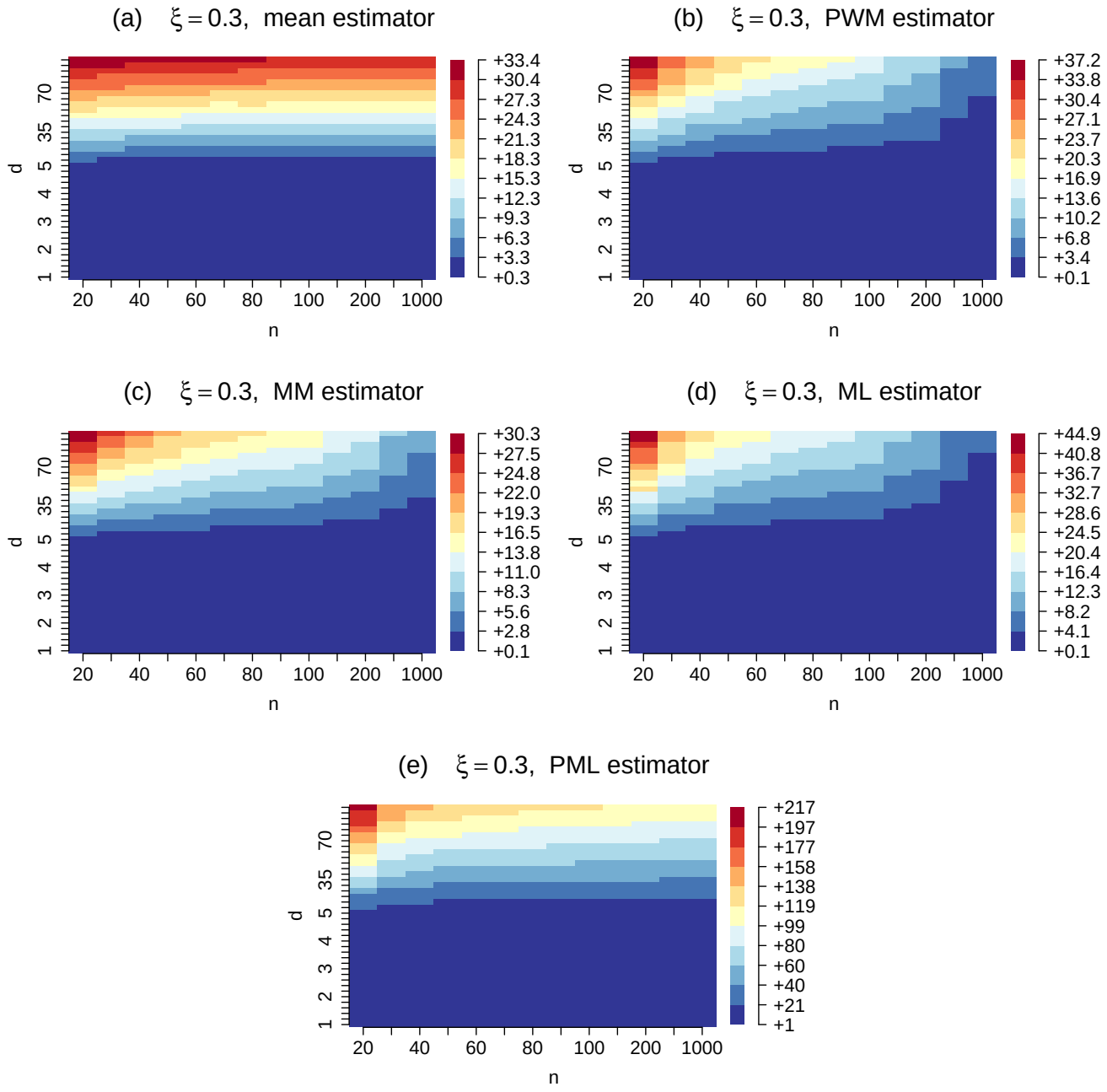

Figure 32: Root mean square error of the mean estimator (a), PWM (b), MM (c), ML (d) and PML (e) on d as a function of d and n, for $\xi = 0.3$.

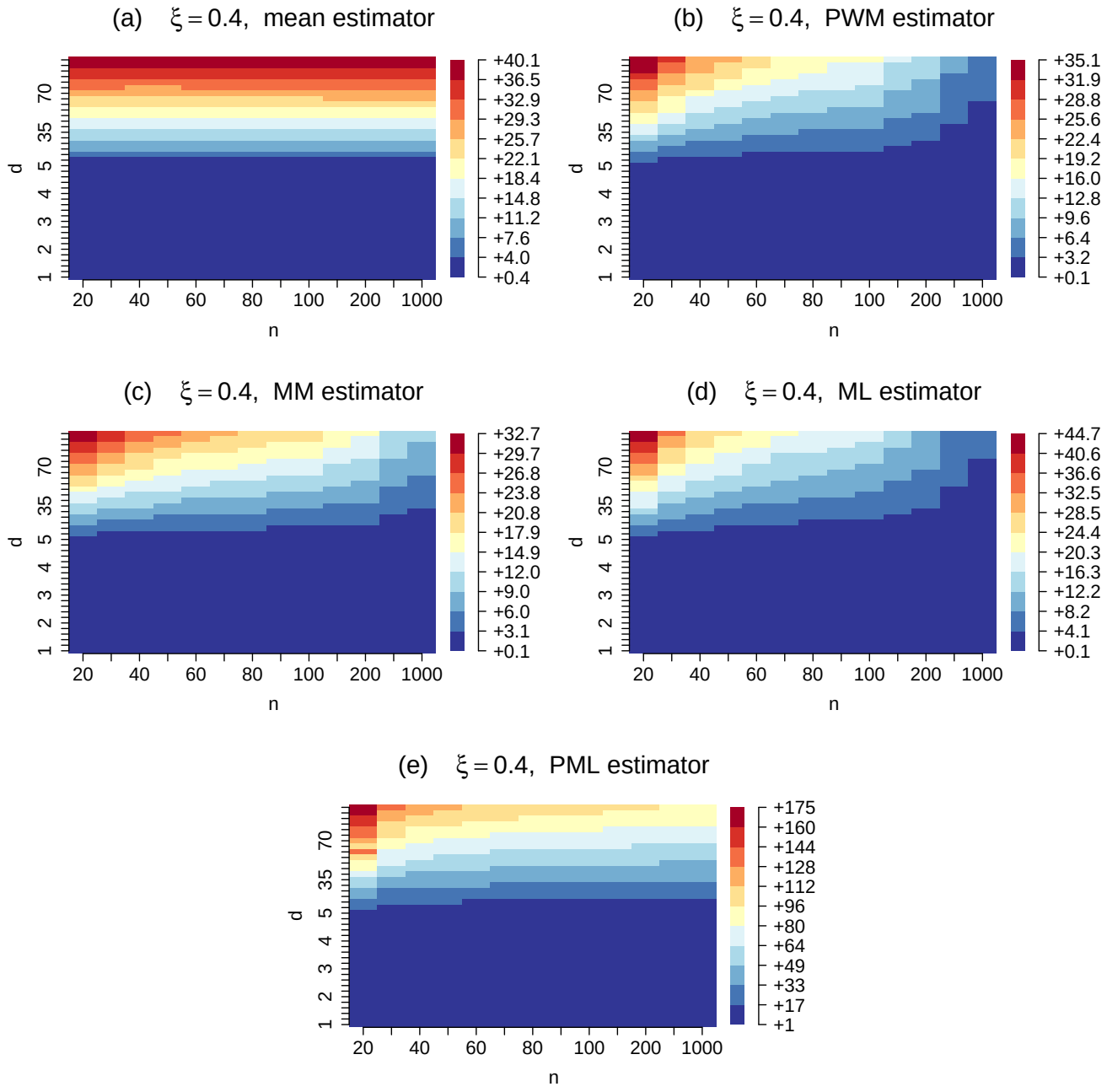

Figure 33: Root mean square error of the mean estimator (a), PWM (b), MM (c), ML (d) and PML (e) on d as a function of d and n, for $\xi = 0.4$.

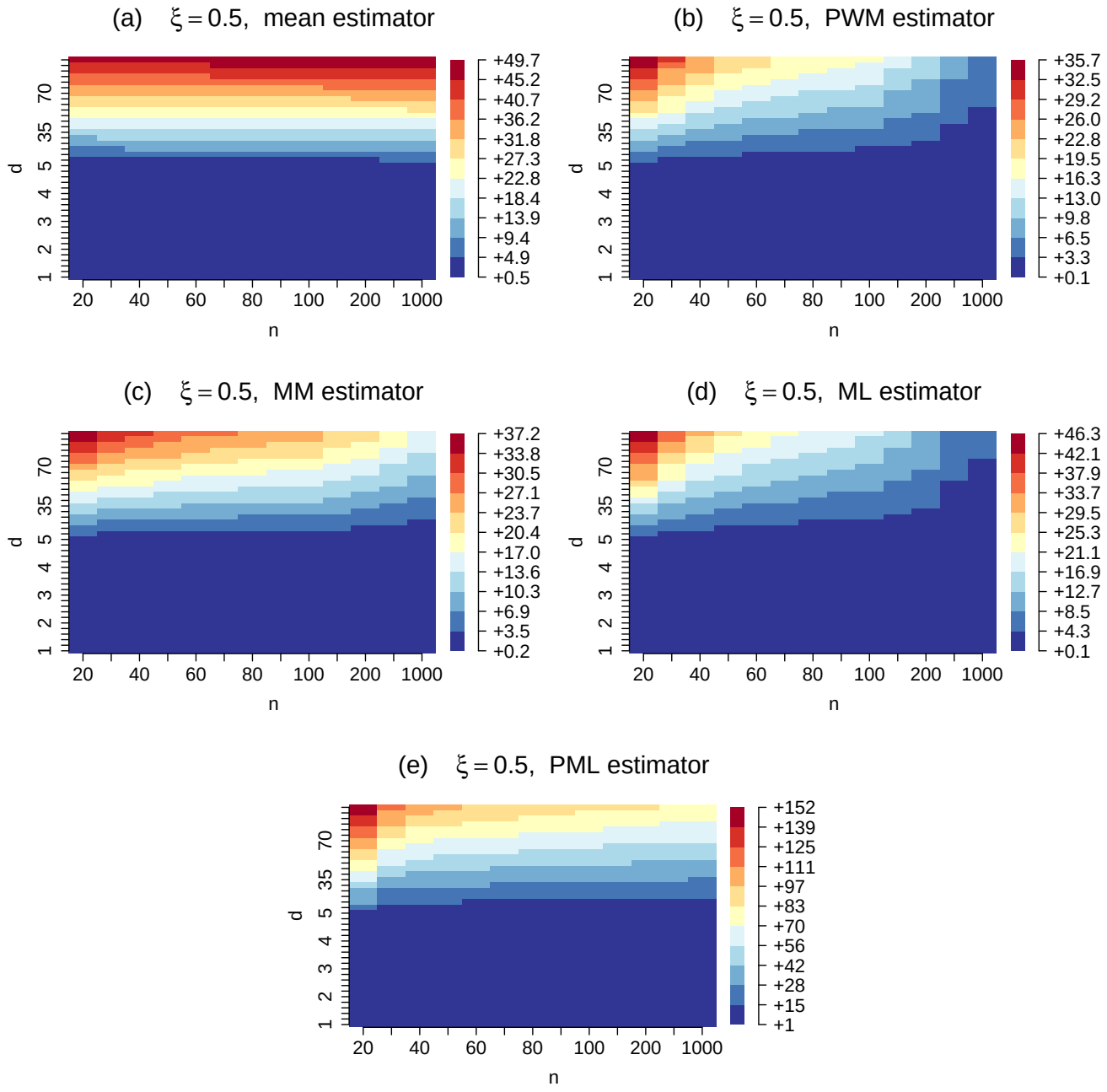

Figure 34: Root mean square error of the mean estimator (a), PWM (b), MM (c), ML (d) and PML (e) on d as a function of d and n, for $\xi = 0.5$.

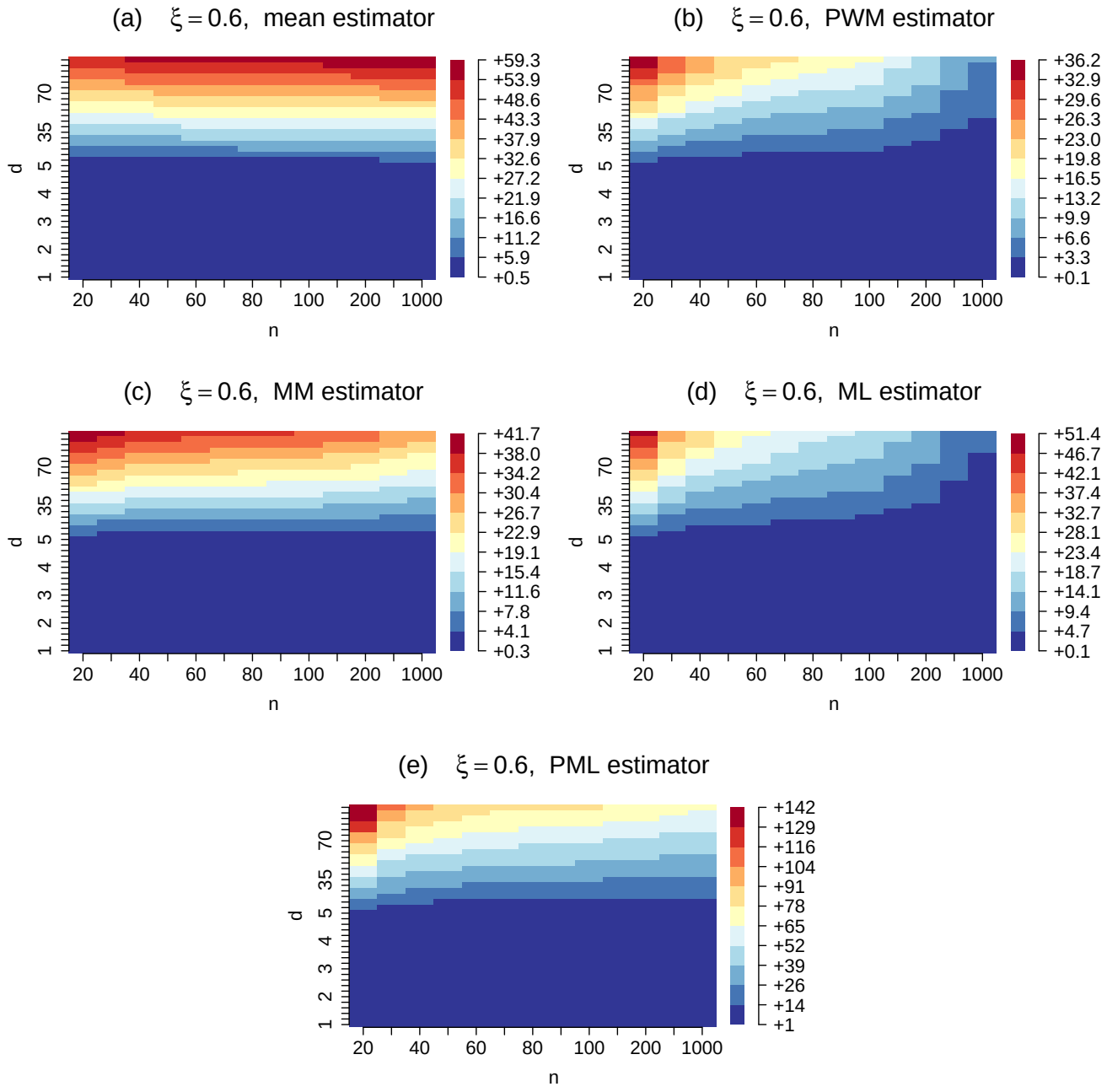

Figure 35: Root mean square error of the mean estimator (a), PWM (b), MM (c), ML (d) and PML (e) on d as a function of d and n, for $\xi = 0.6$.

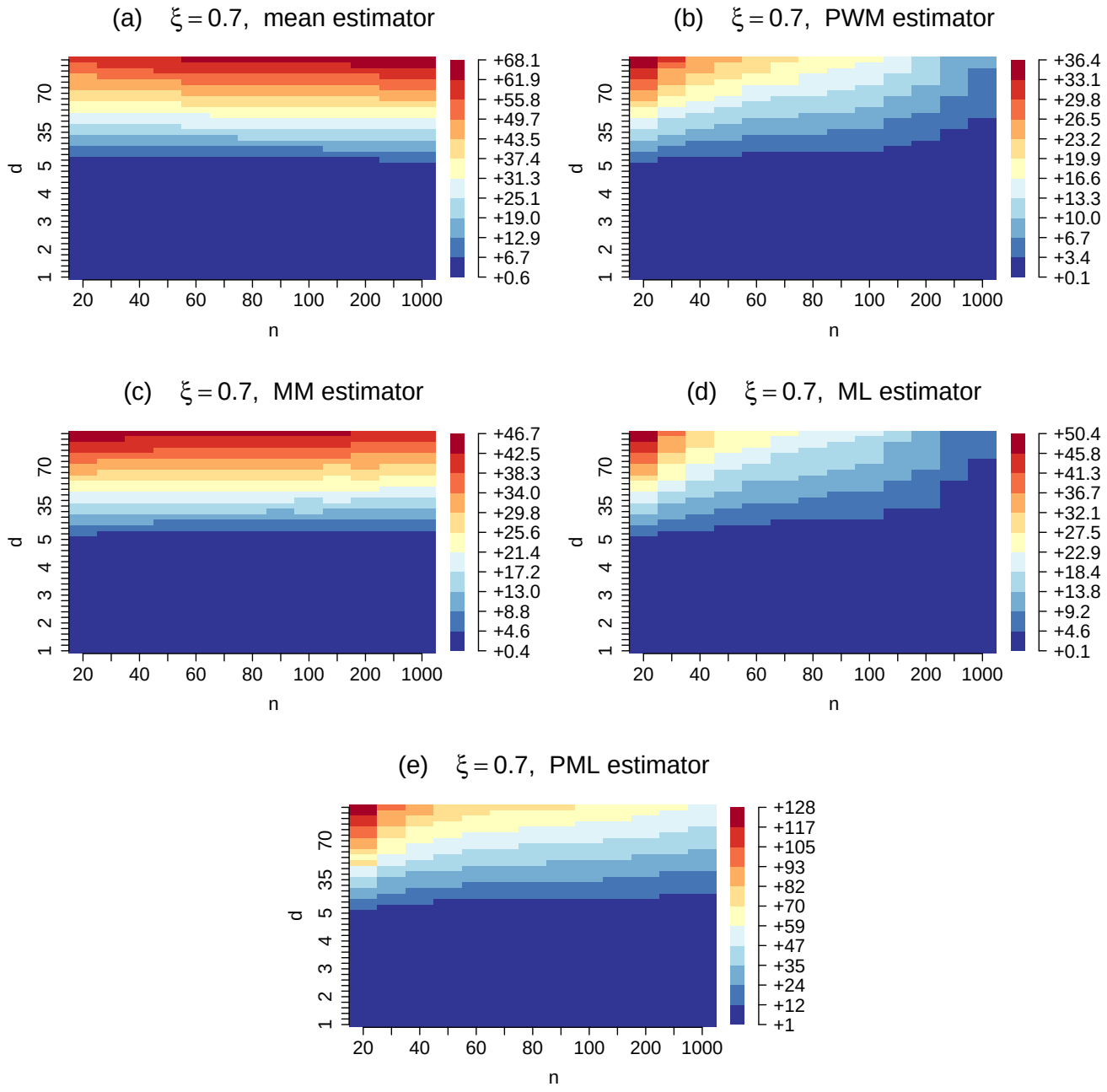

Figure 36: Root mean square error of the mean estimator (a), PWM (b), MM (c), ML (d) and PML (e) on d as a function of d and n, for $\xi = 0.7$.

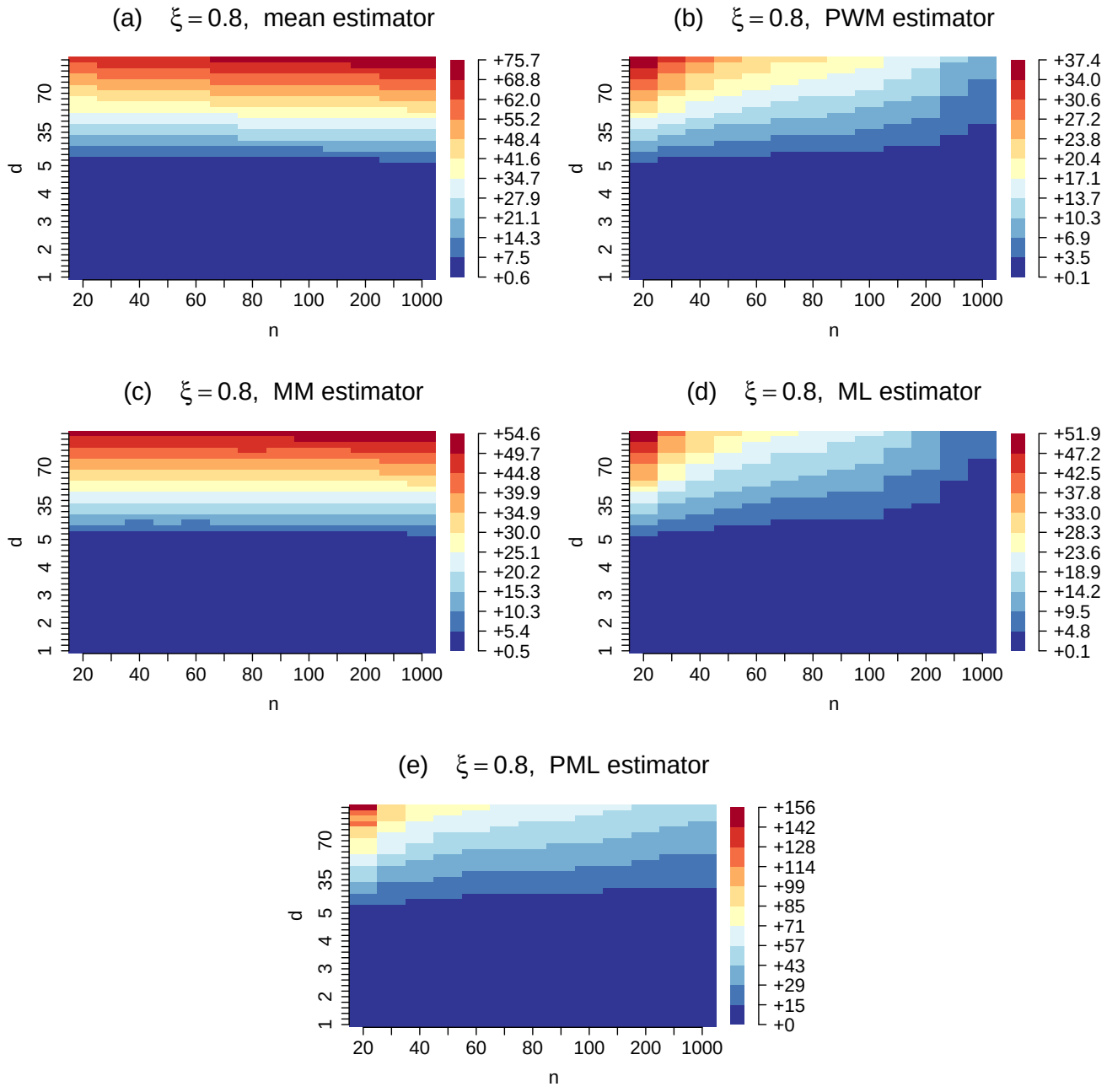

Figure 37: Root mean square error of the mean estimator (a), PWM (b), MM (c), ML (d) and PML (e) on d as a function of d and n, for $\xi = 0.8$.

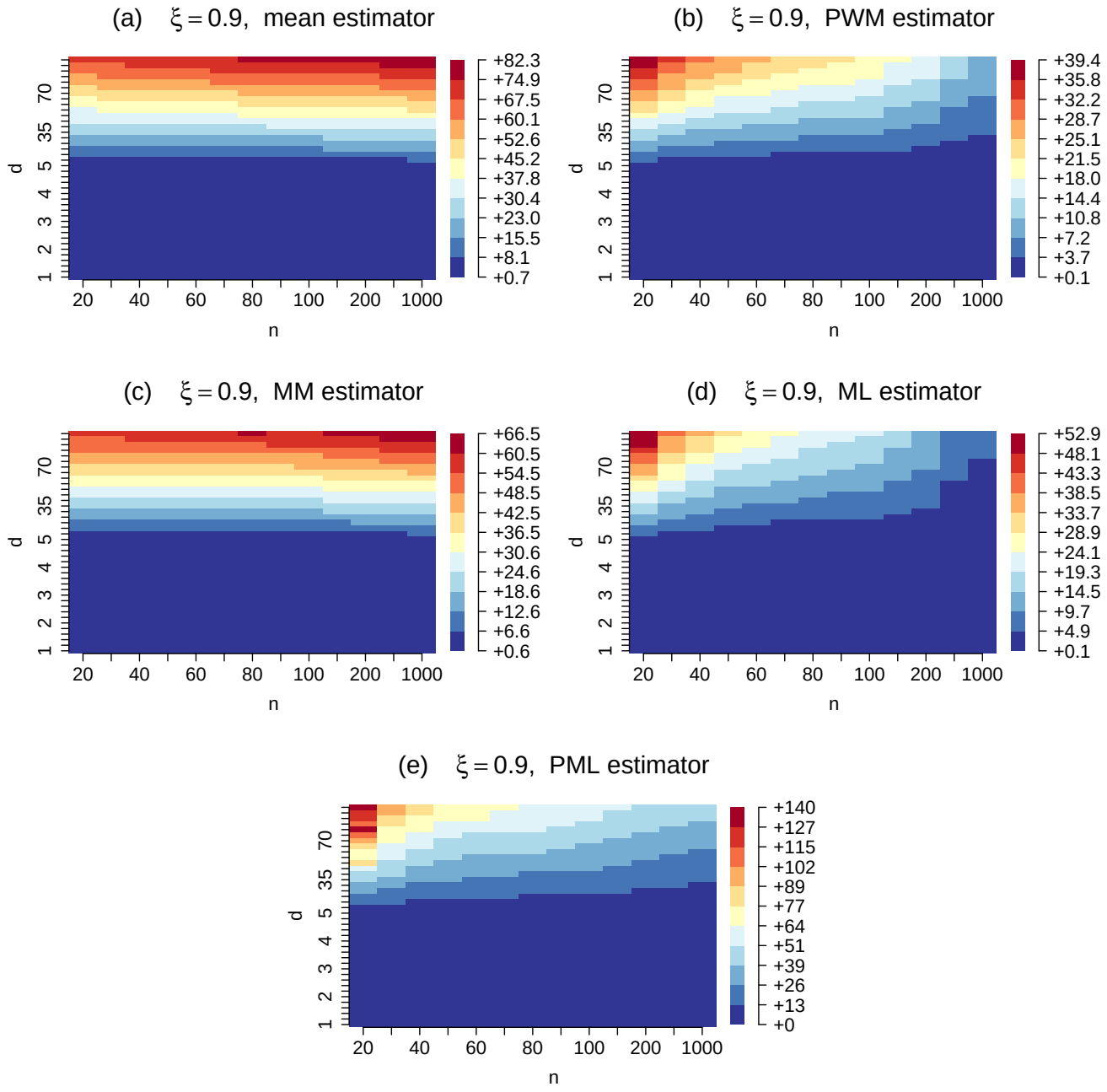

Figure 38: Root mean square error of the mean estimator (a), PWM (b), MM (c), ML (d) and PML (e) on d as a function of d and n, for $\xi = 0.9$.

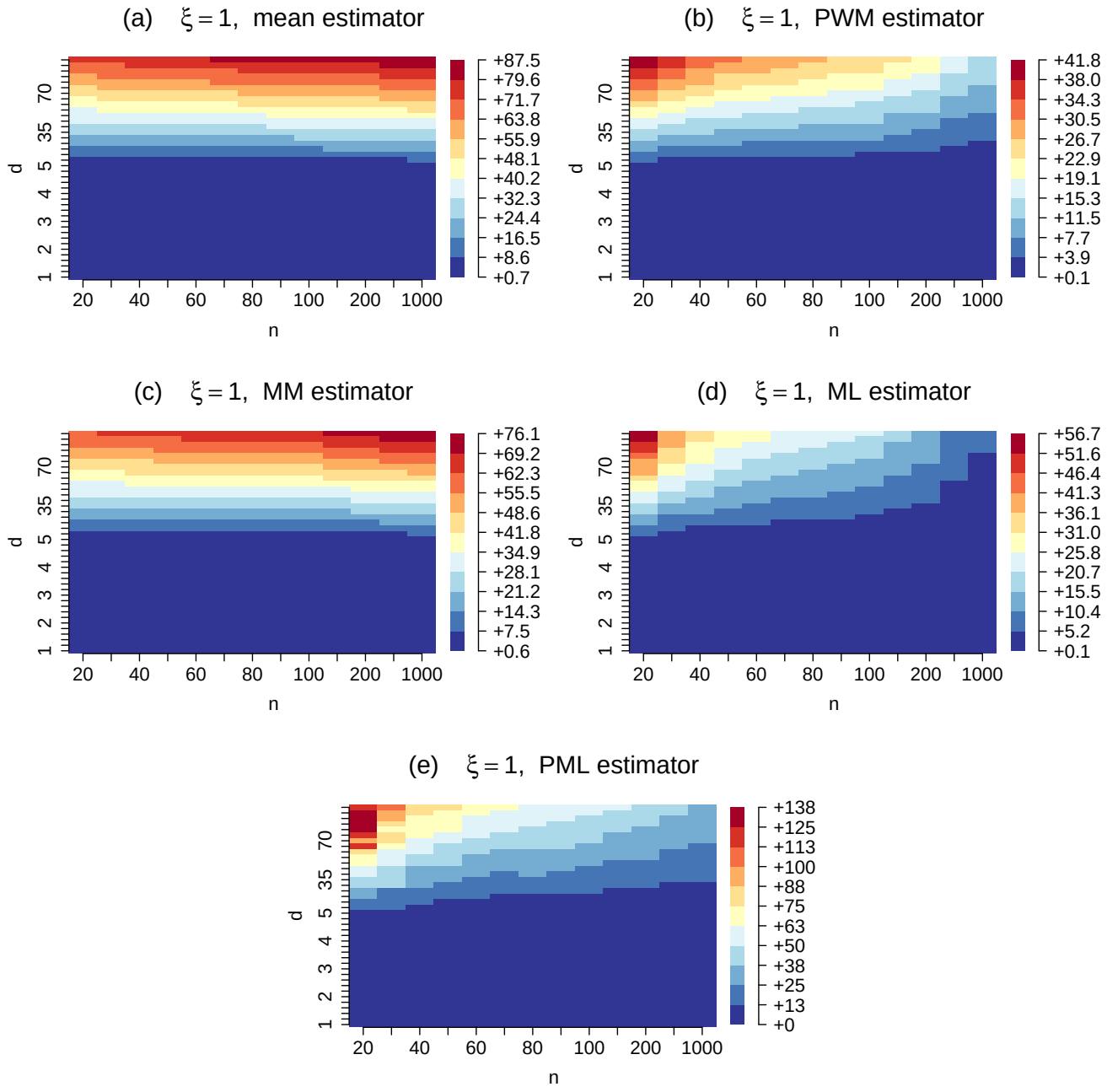

Figure 39: Root mean square error of the mean estimator (a), PWM (b), MM (c), ML (d) and PML (e) on d as a function of d and n, for $\xi = 1$.

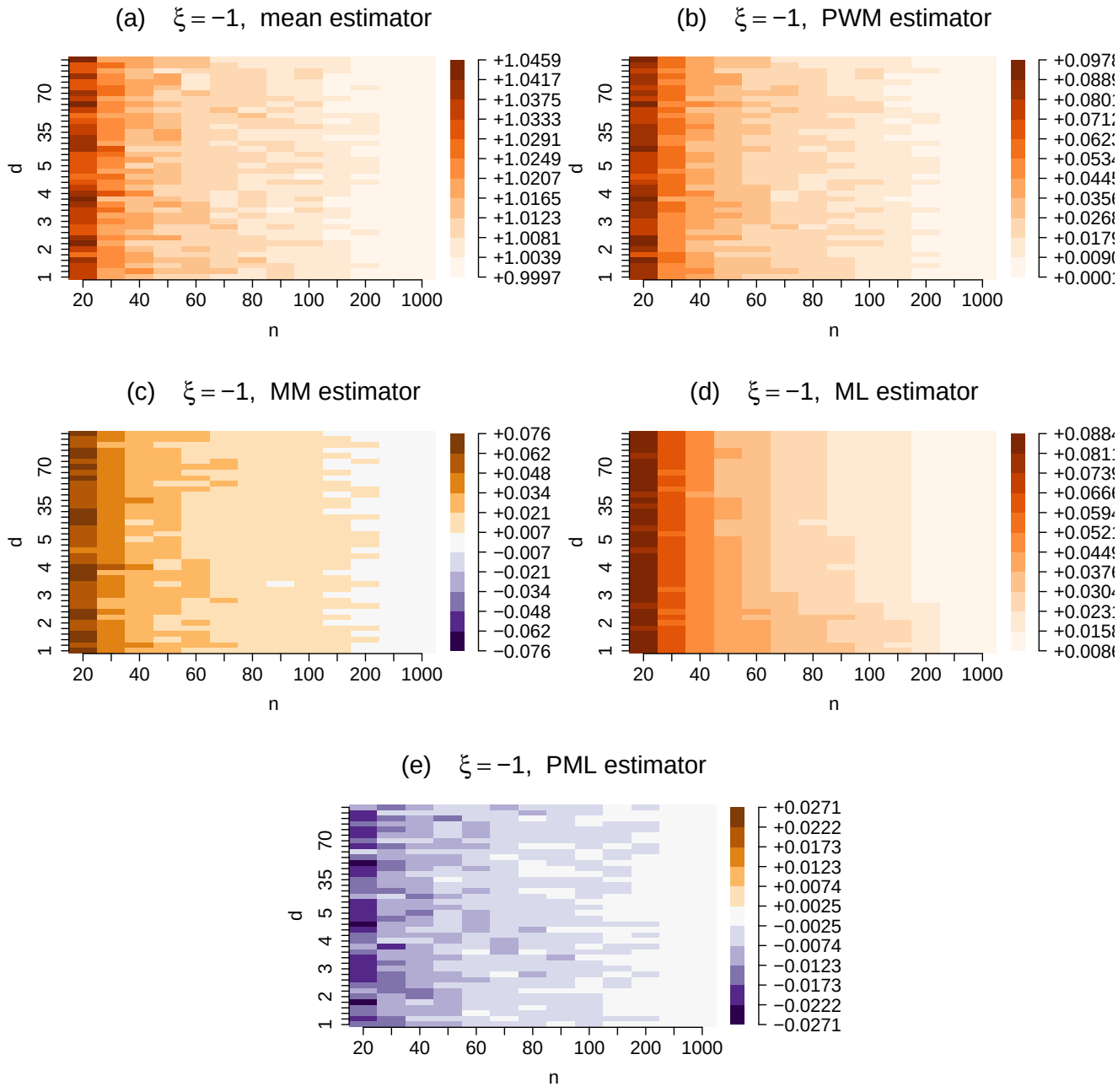

Figure 40: Relative error of the mean estimator (a), PWM (b), MM (c), ML (d) and PML (e) on d as a function of d and n, for $\xi = -1$.

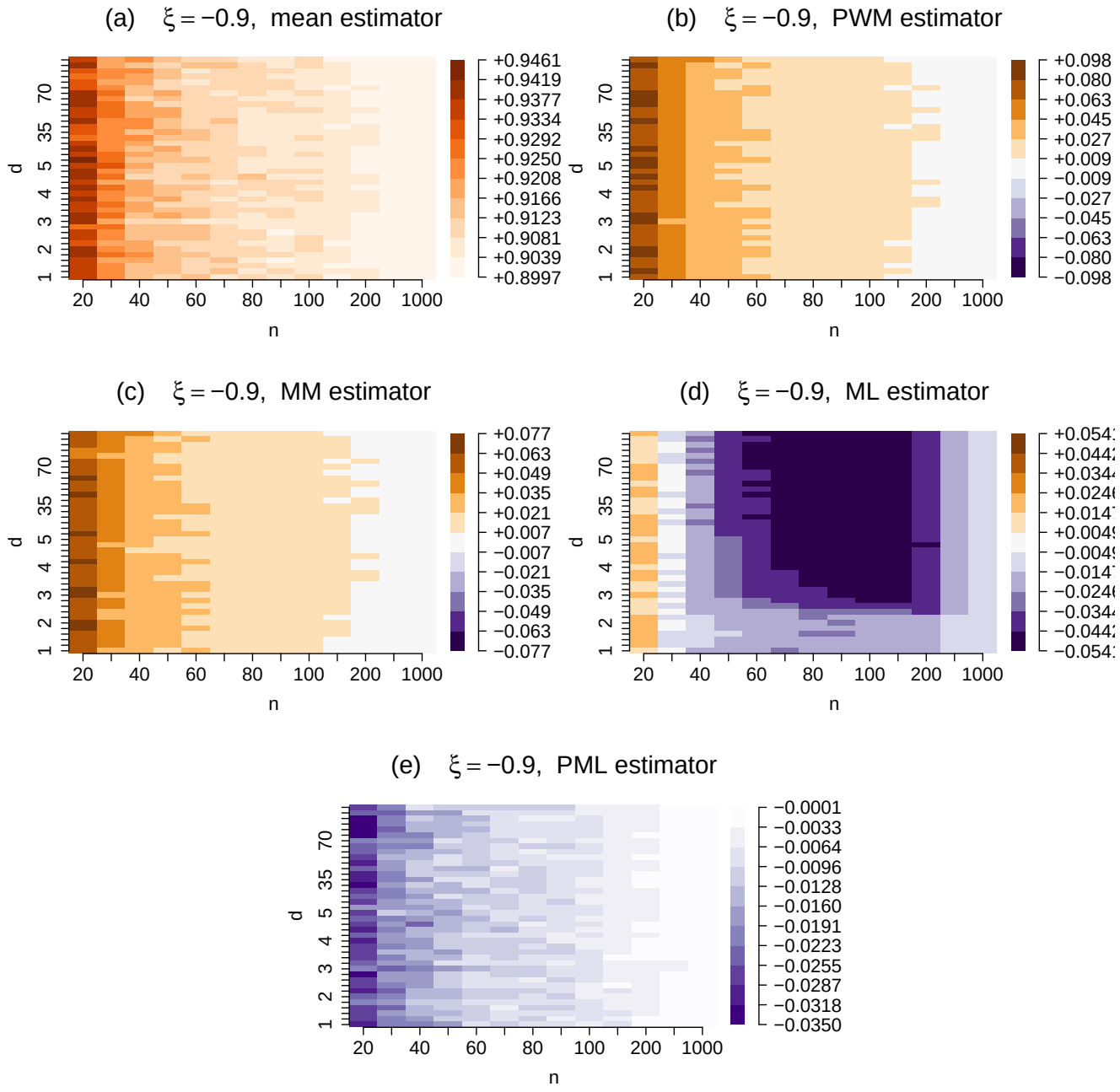

Figure 41: Relative error of the mean estimator (a), PWM (b), MM (c), ML (d) and PML (e) on d as a function of d and n, for $\xi = -0.9$.

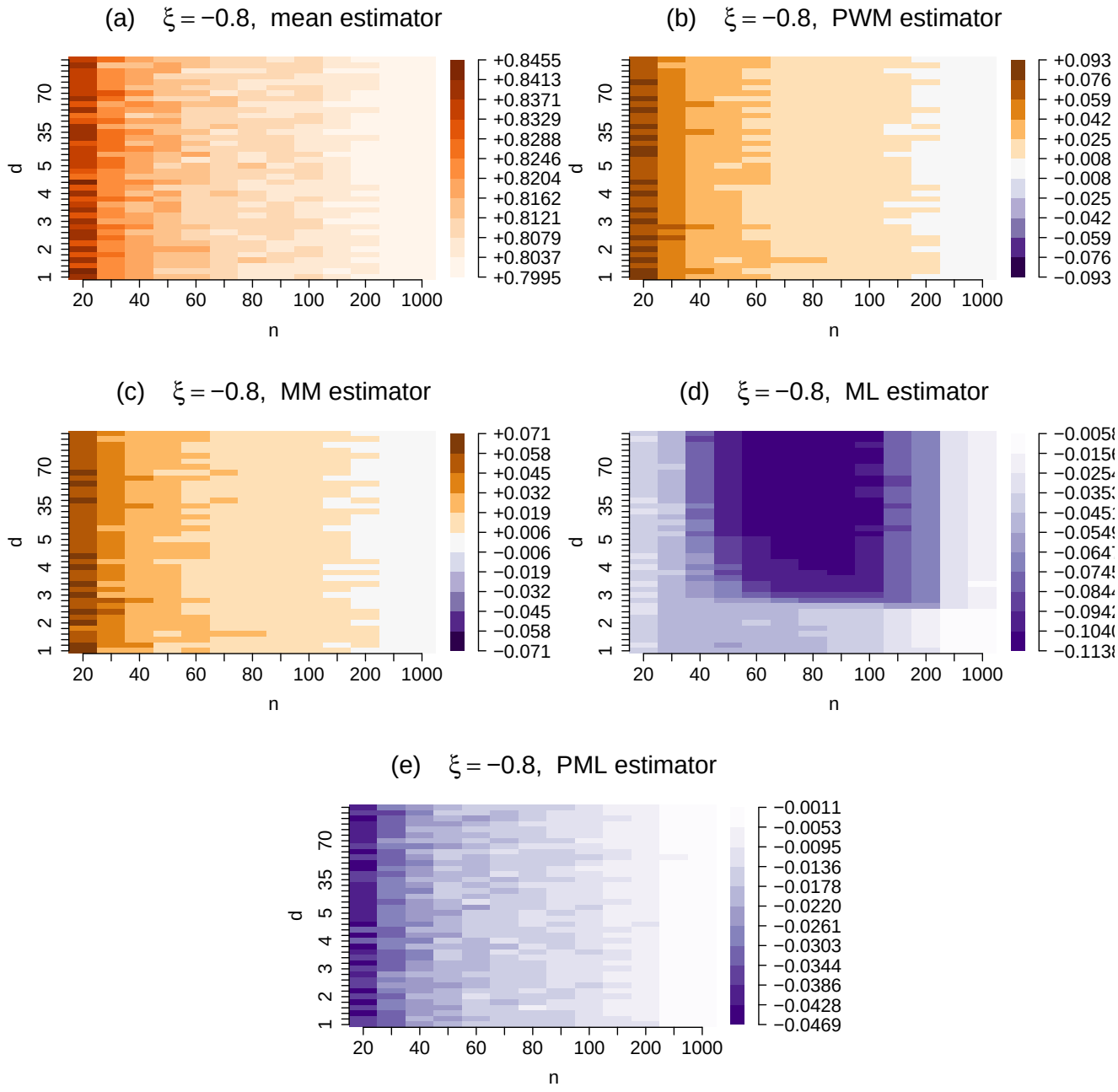

Figure 42: Relative error of the mean estimator (a), PWM (b), MM (c), ML (d) and PML (e) on d as a function of d and n, for $\xi = -0.8$.

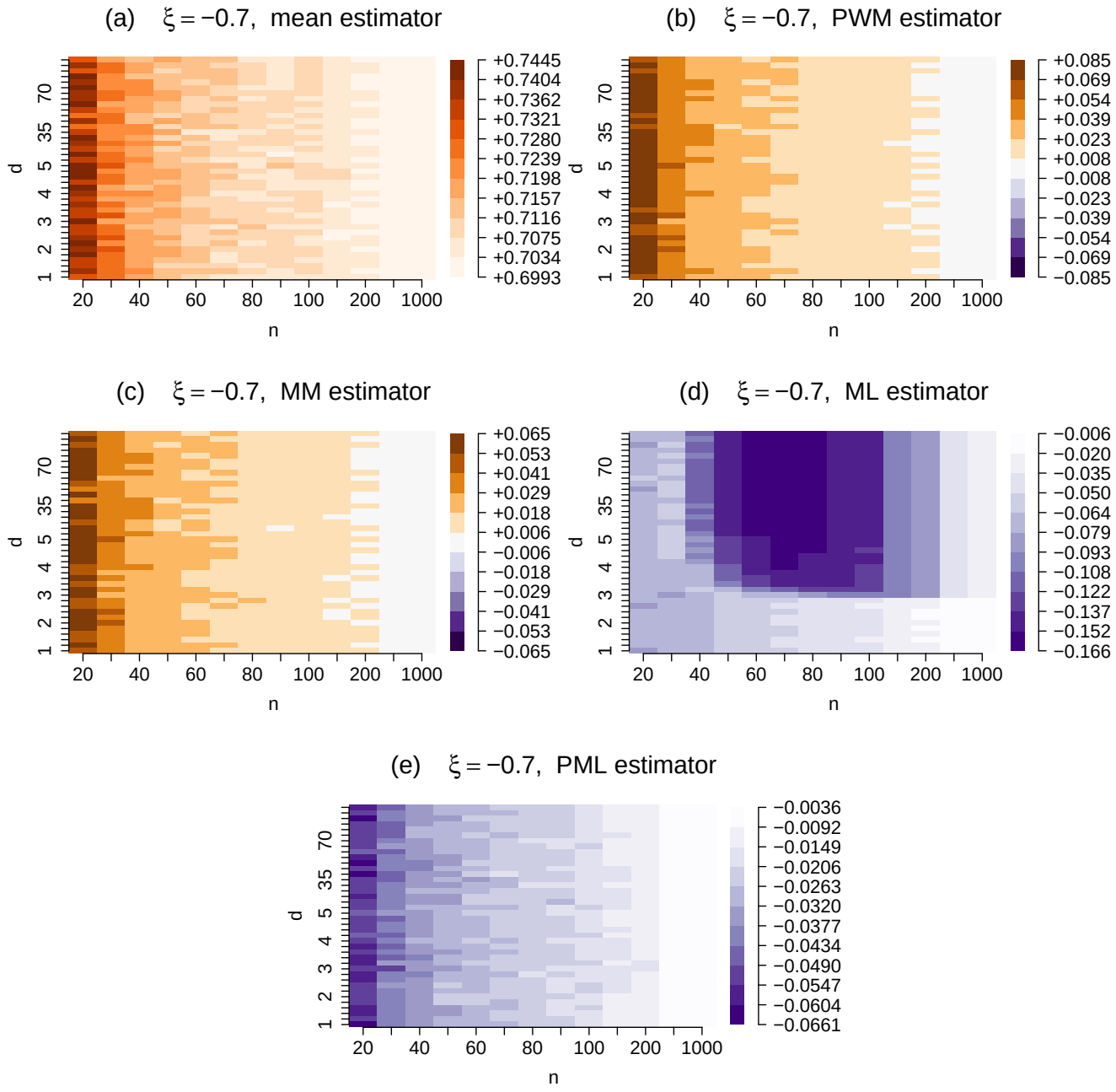

Figure 43: Relative error of the mean estimator (a), PWM (b), MM (c), ML (d) and PML (e) on d as a function of d and n, for $\xi = -0.7$.

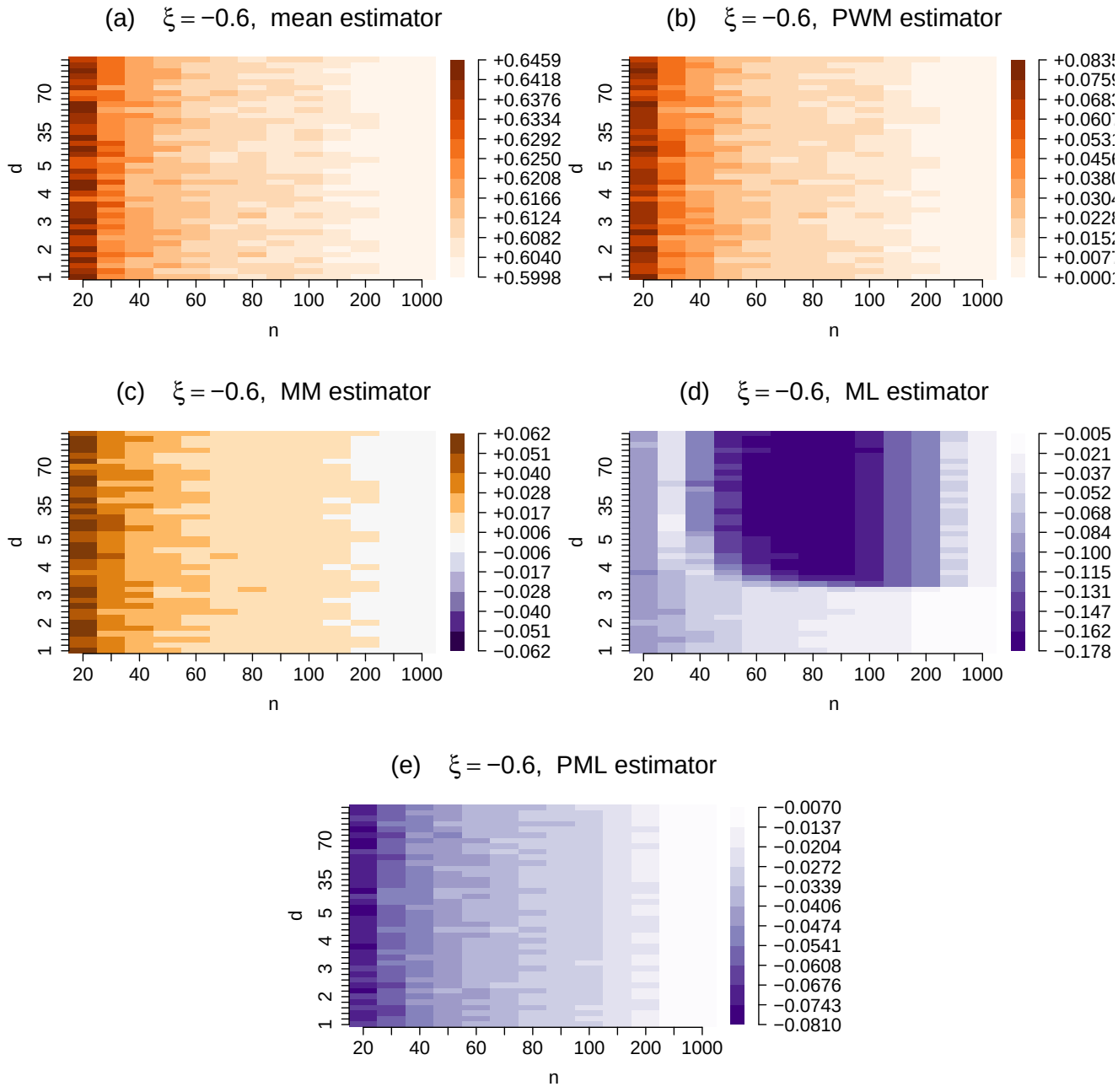

Figure 44: Relative error of the mean estimator (a), PWM (b), MM (c), ML (d) and PML (e) on d as a function of d and n, for $\xi = -0.6$.

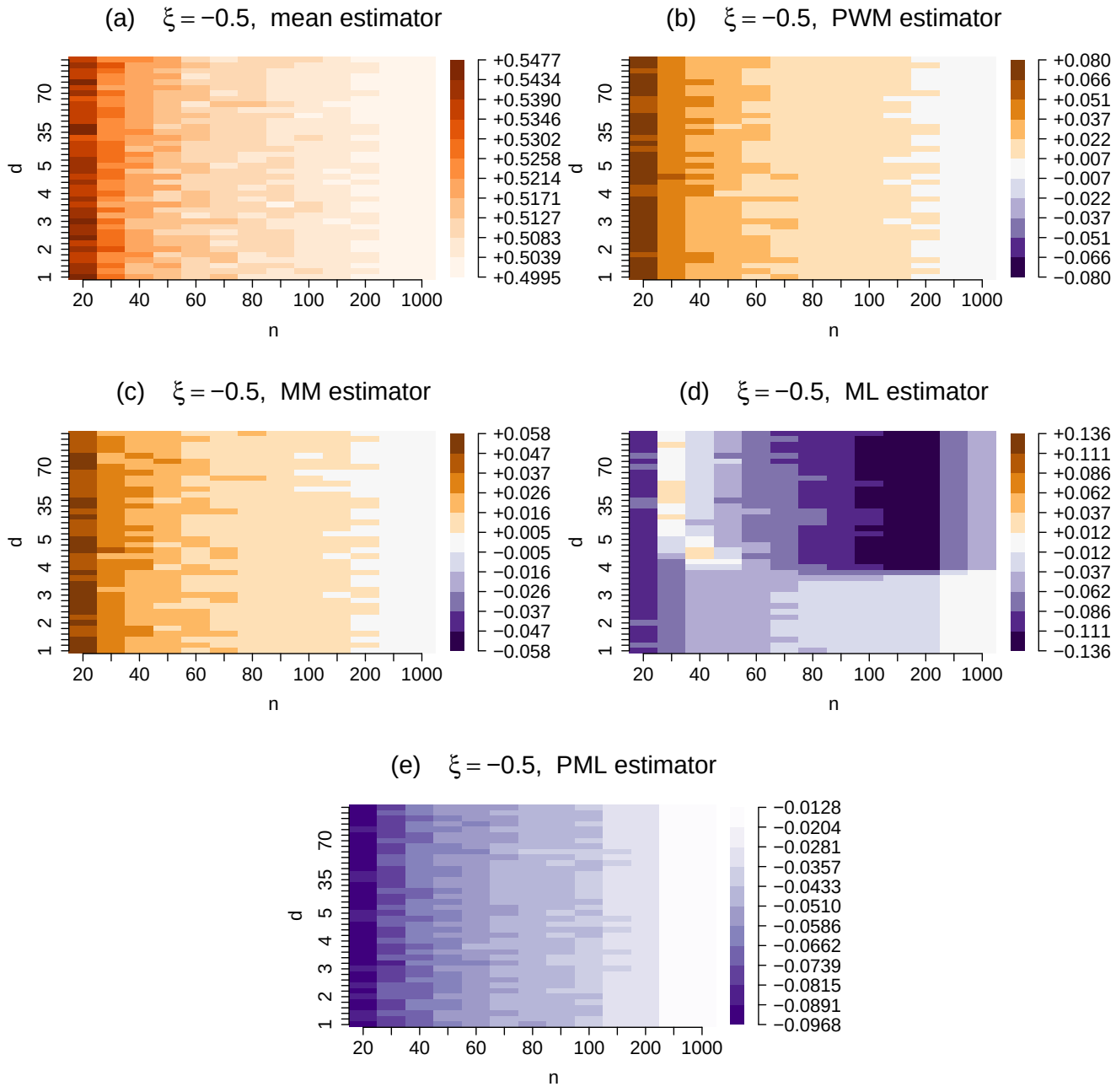

Figure 45: Relative error of the mean estimator (a), PWM (b), MM (c), ML (d) and PML (e) on d as a function of d and n, for $\xi = -0.5$.

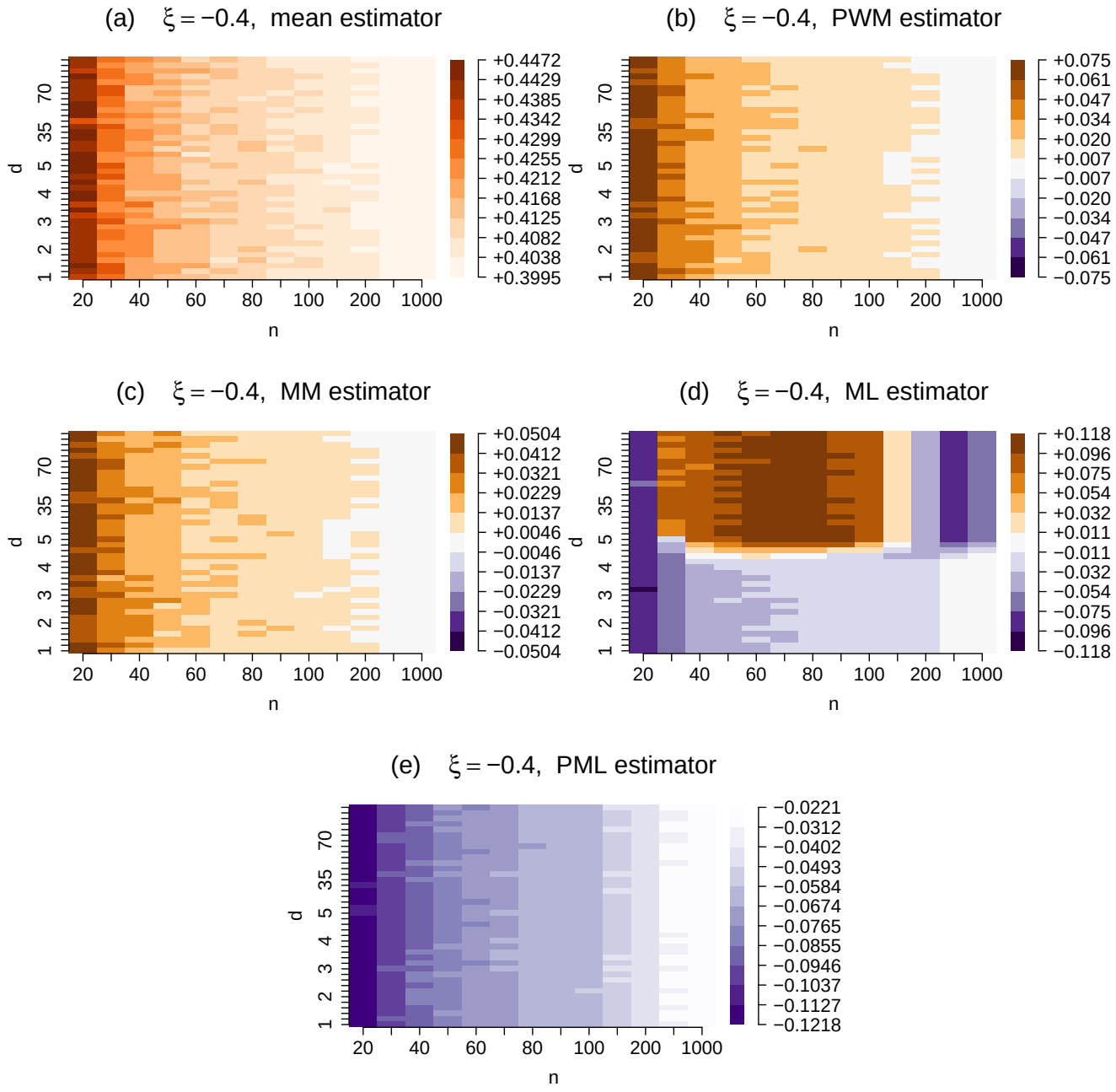

Figure 46: Relative error of the mean estimator (a), PWM (b), MM (c), ML (d) and PML (e) on d as a function of d and n, for $\xi = -0.4$.

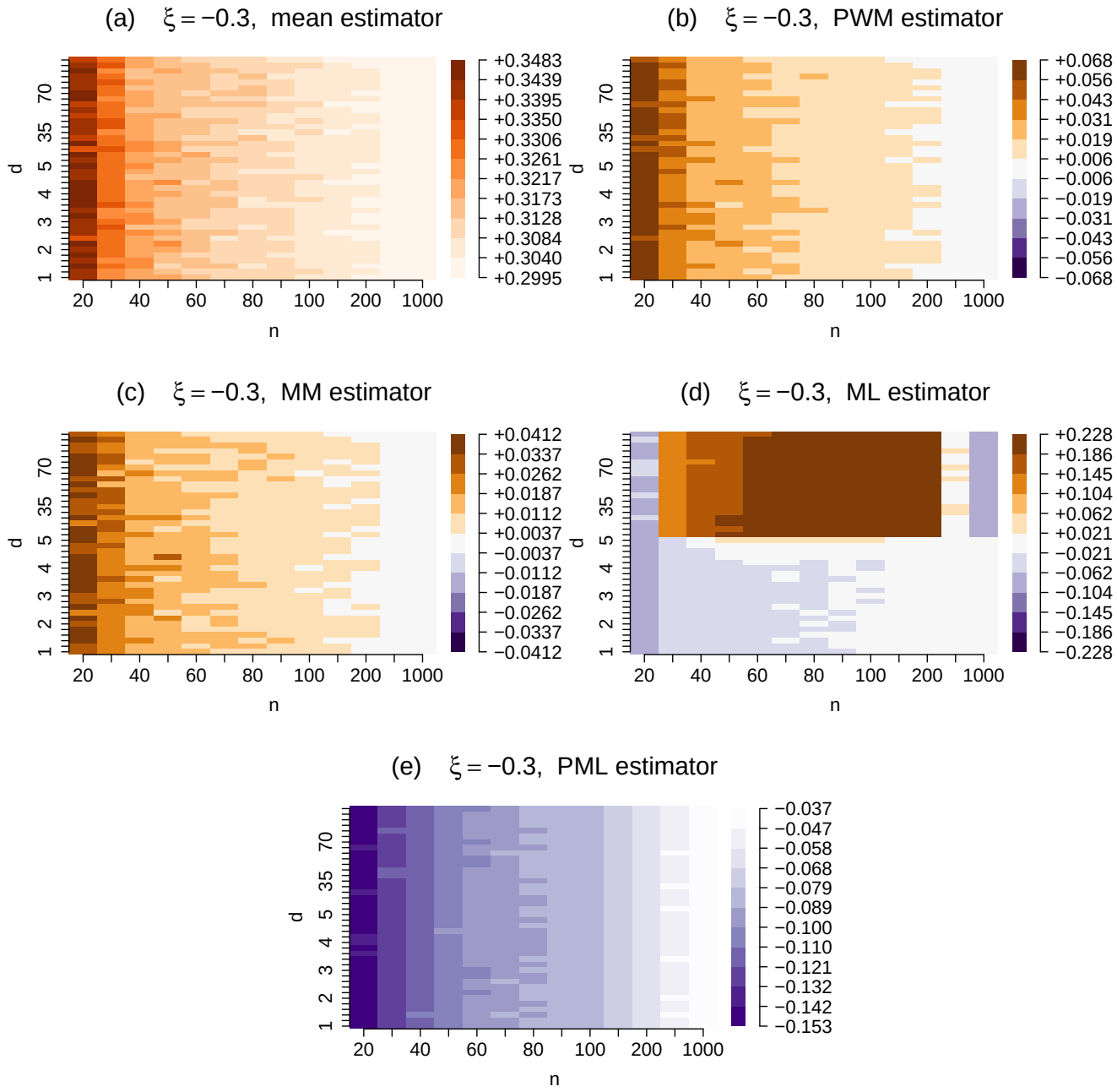

Figure 47: Relative error of the mean estimator (a), PWM (b), MM (c), ML (d) and PML (e) on d as a function of d and n, for $\xi = -0.3$.

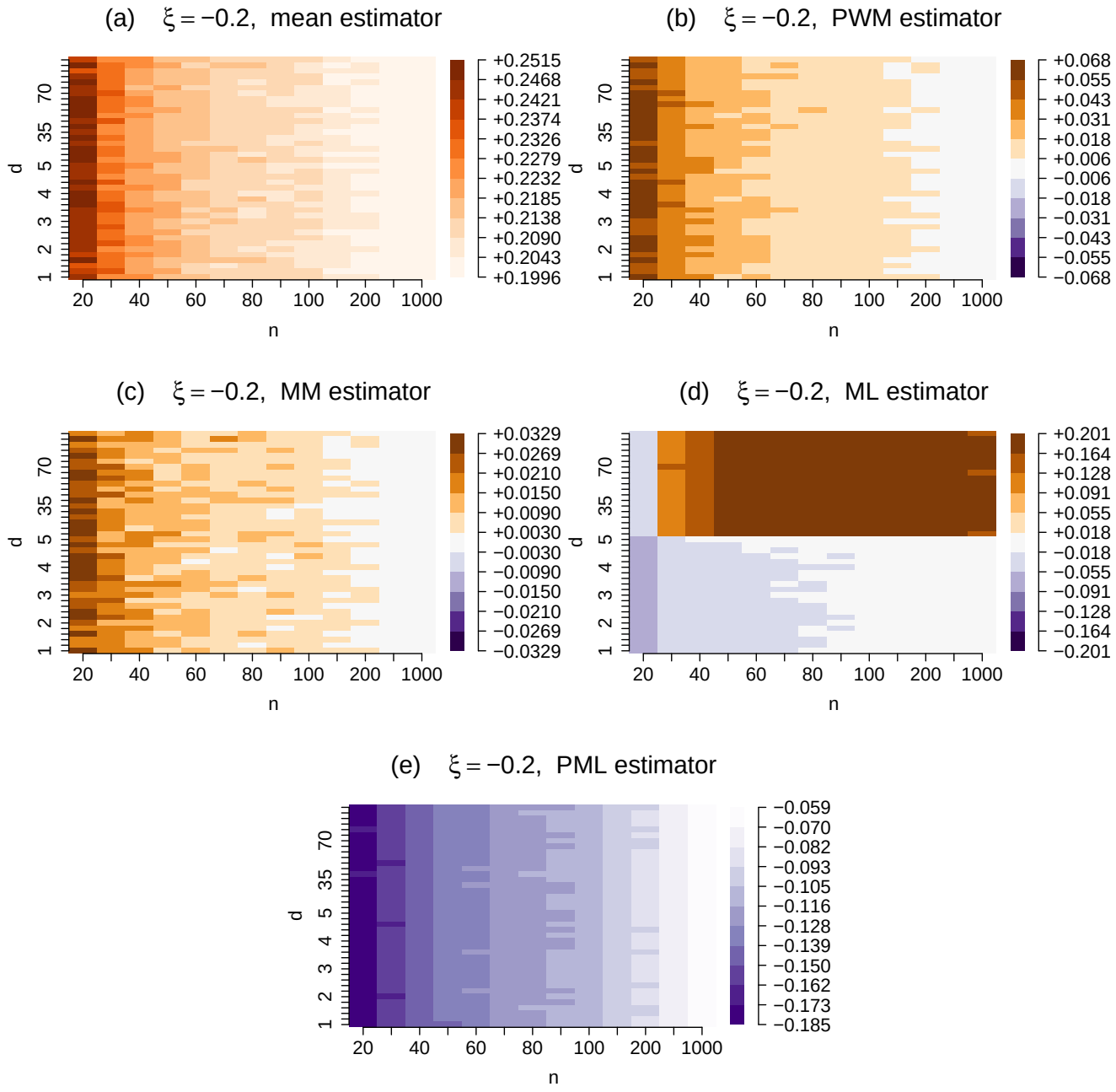

Figure 48: Relative error of the mean estimator (a), PWM (b), MM (c), ML (d) and PML (e) on d as a function of d and n, for $\xi = -0.2$.

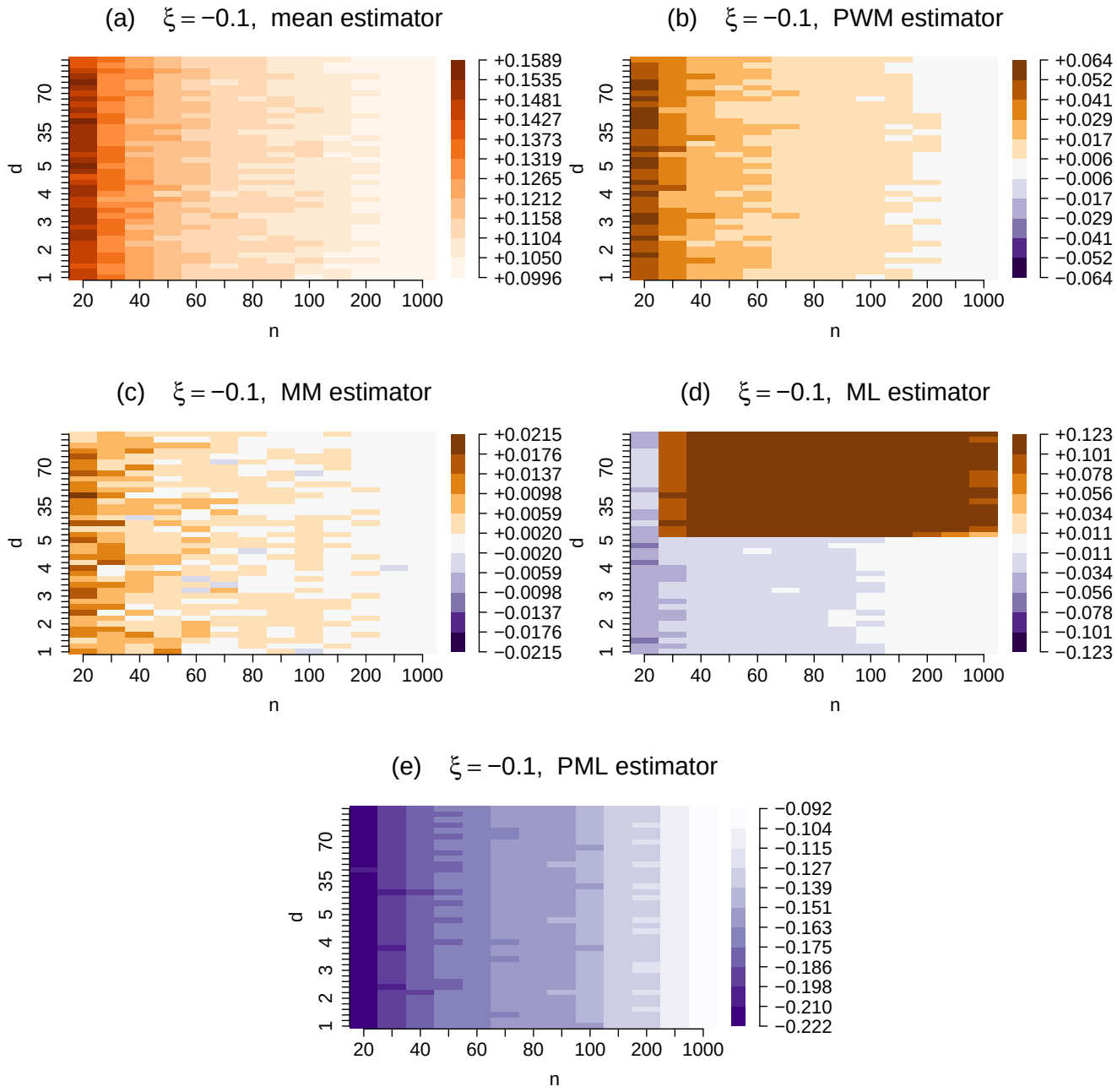

Figure 49: Relative error of the mean estimator (a), PWM (b), MM (c), ML (d) and PML (e) on d as a function of d and n, for $\xi = -0.1$.

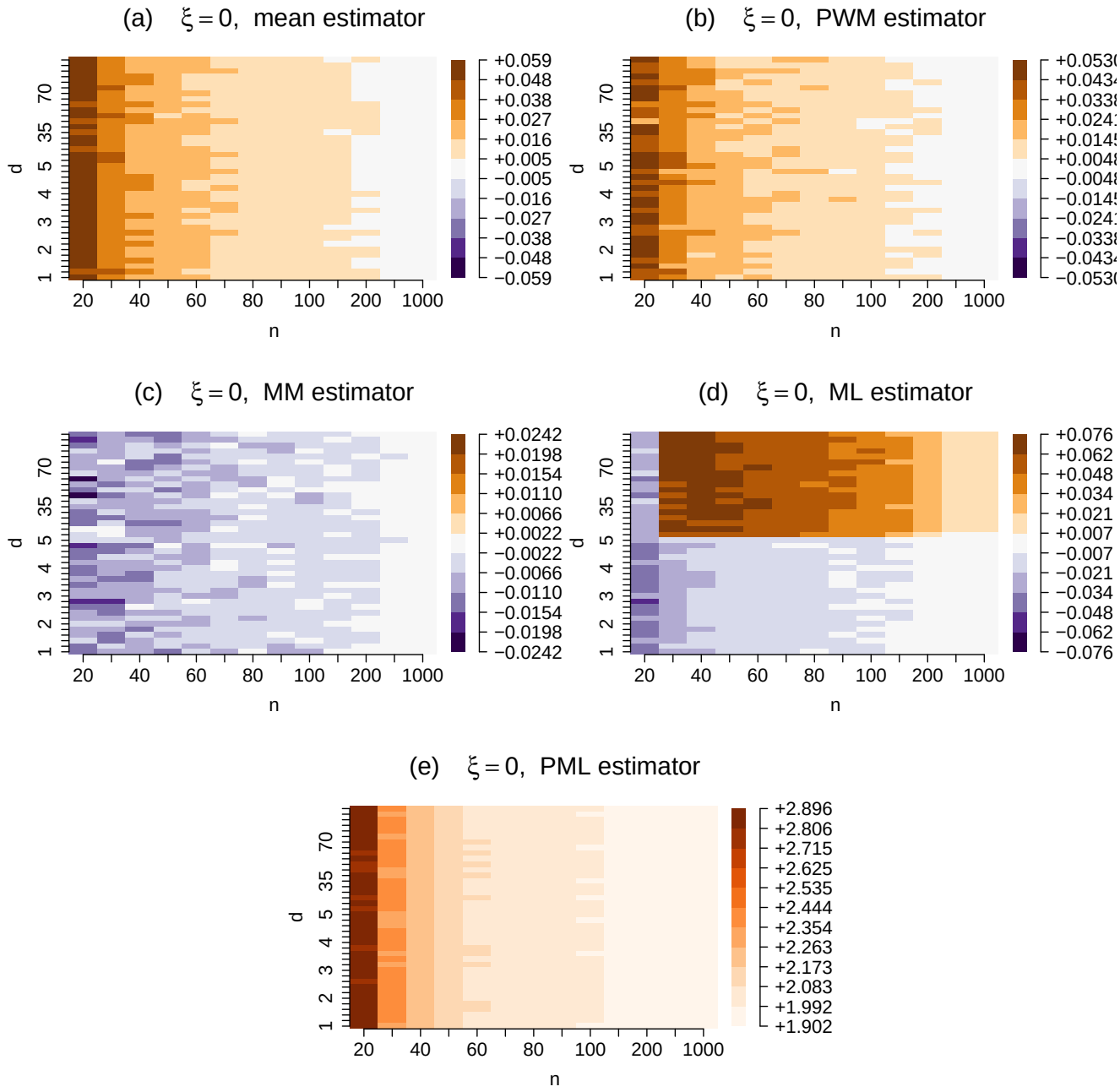

Figure 50: Relative error of the mean estimator (a), PWM (b), MM (c), ML (d) and PML (e) on d as a function of d and n, for $\xi = 0$.

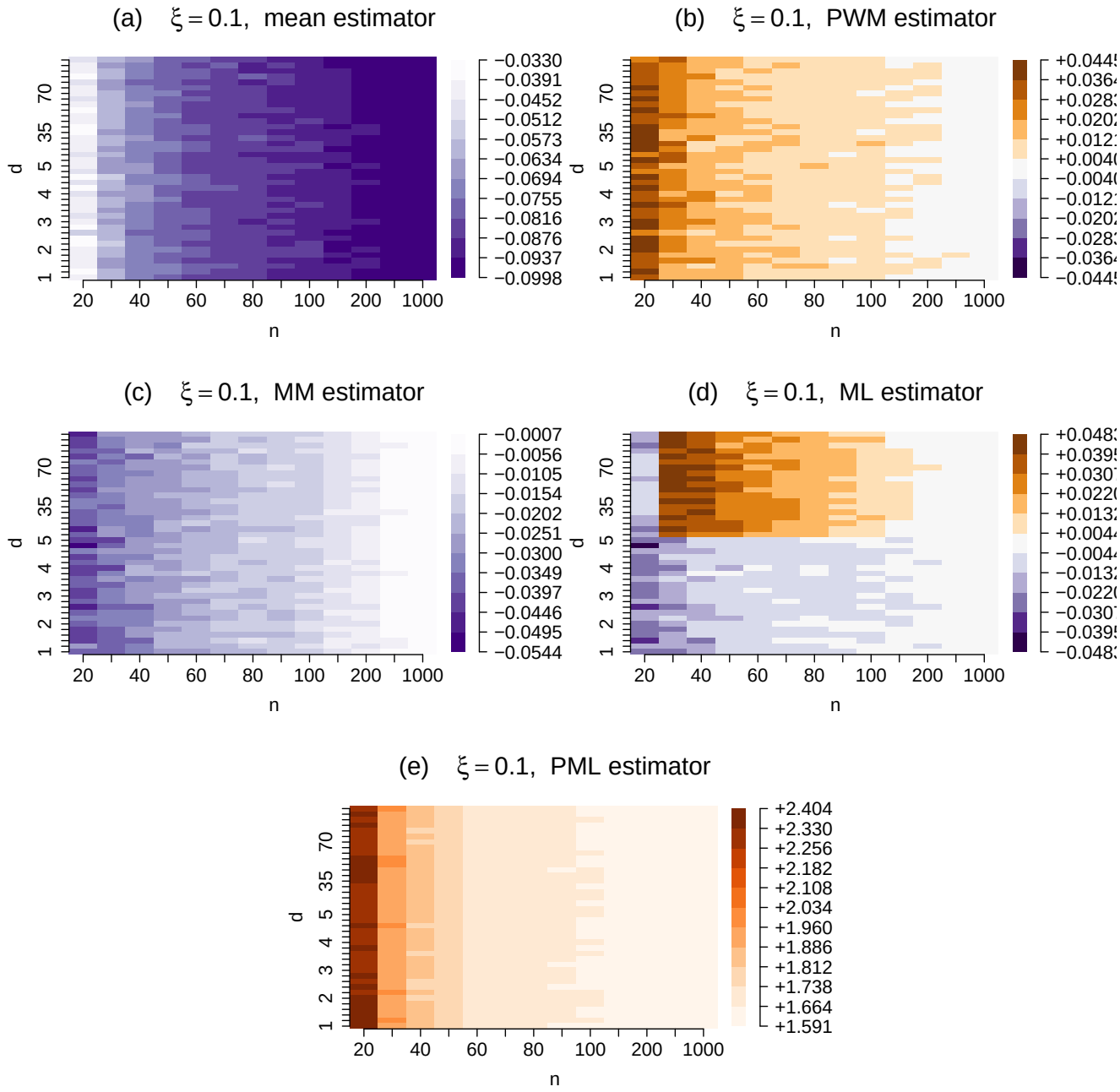

Figure 51: Relative error of the mean estimator (a), PWM (b), MM (c), ML (d) and PML (e) on d as a function of d and n, for $\xi = 0.1$.

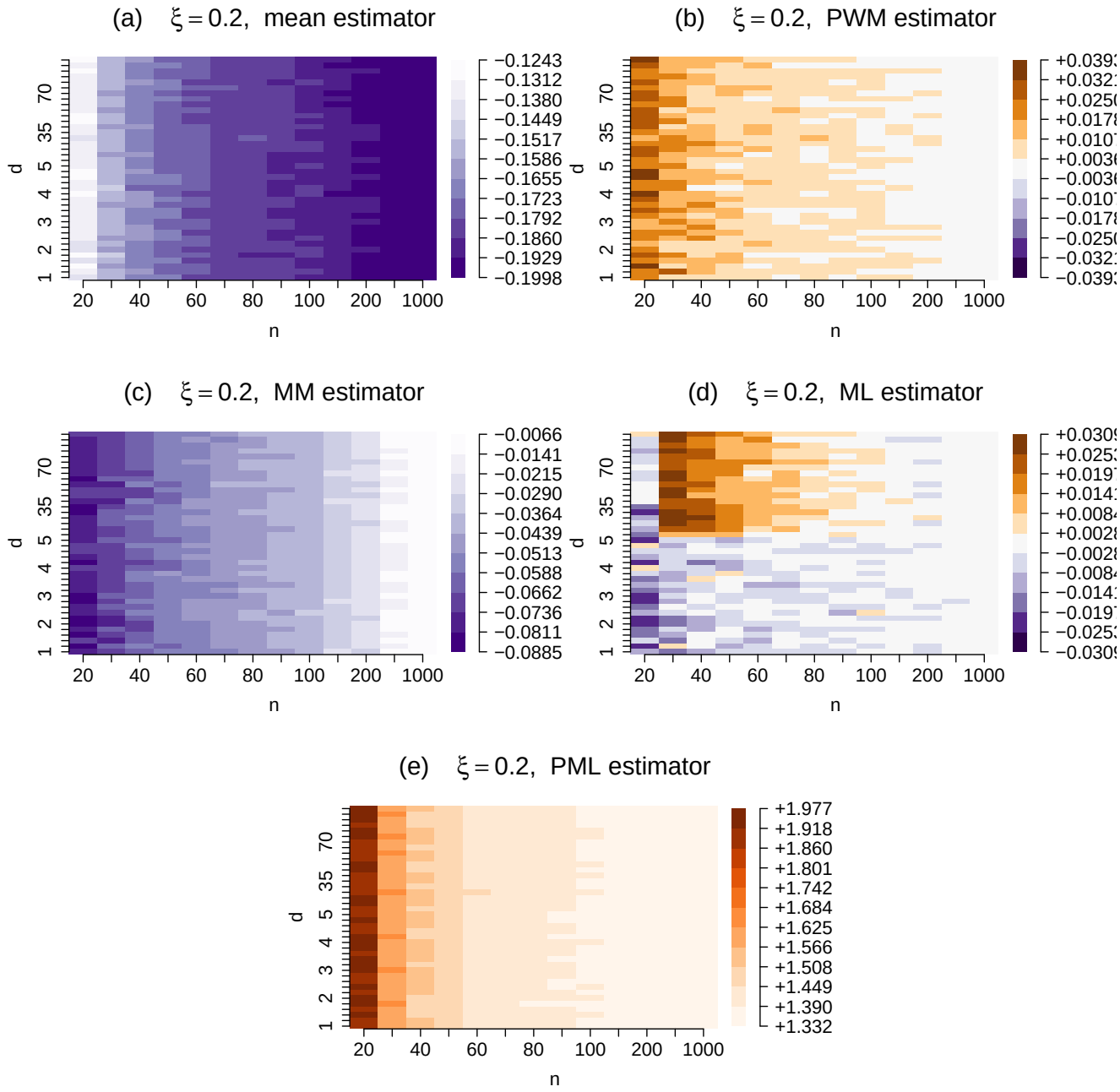

Figure 52: Relative error of the mean estimator (a), PWM (b), MM (c), ML (d) and PML (e) on d as a function of d and n, for $\xi = 0.2$.

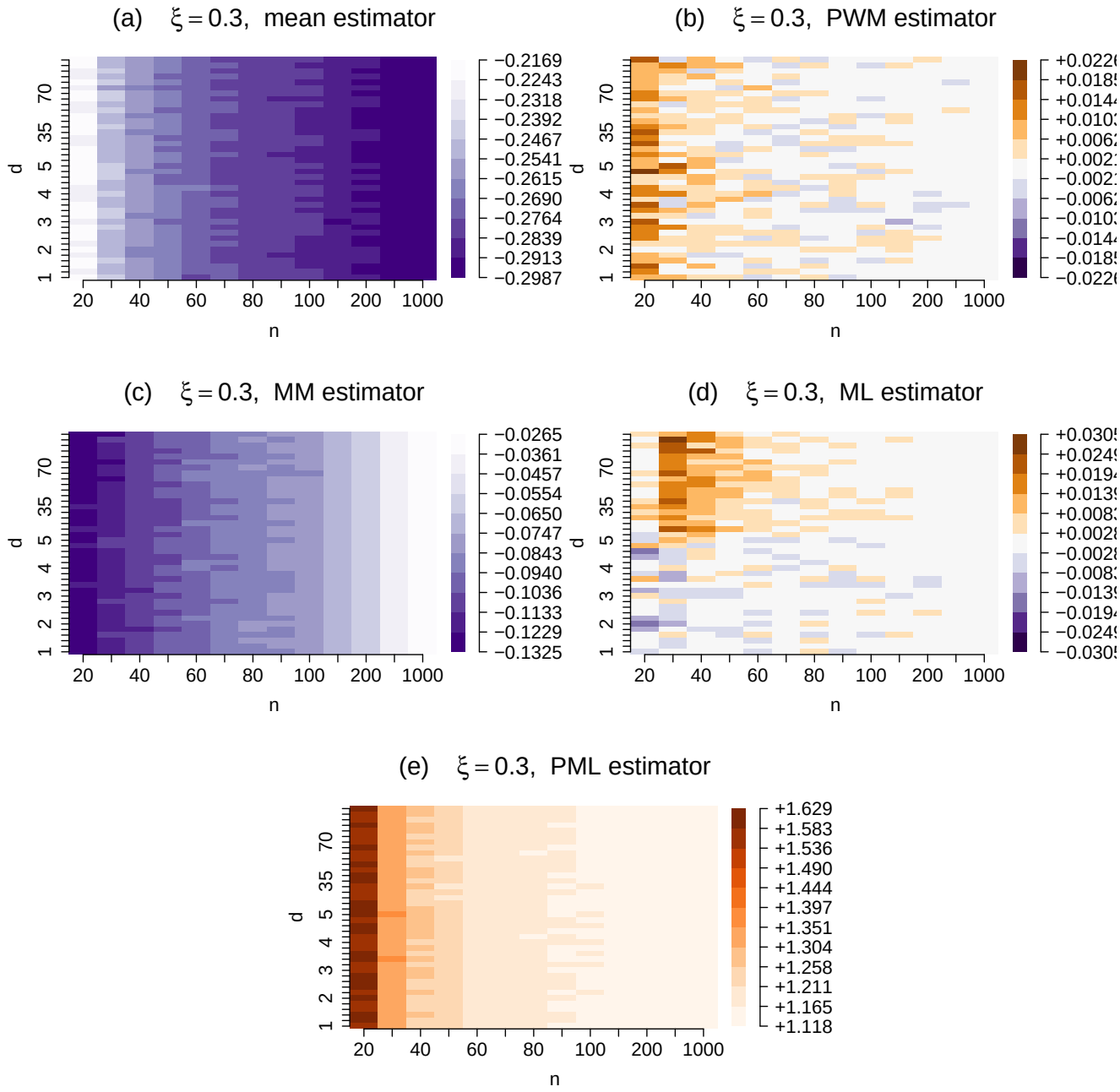

Figure 53: Relative error of the mean estimator (a), PWM (b), MM (c), ML (d) and PML (e) on d as a function of d and n, for $\xi = 0.3$.

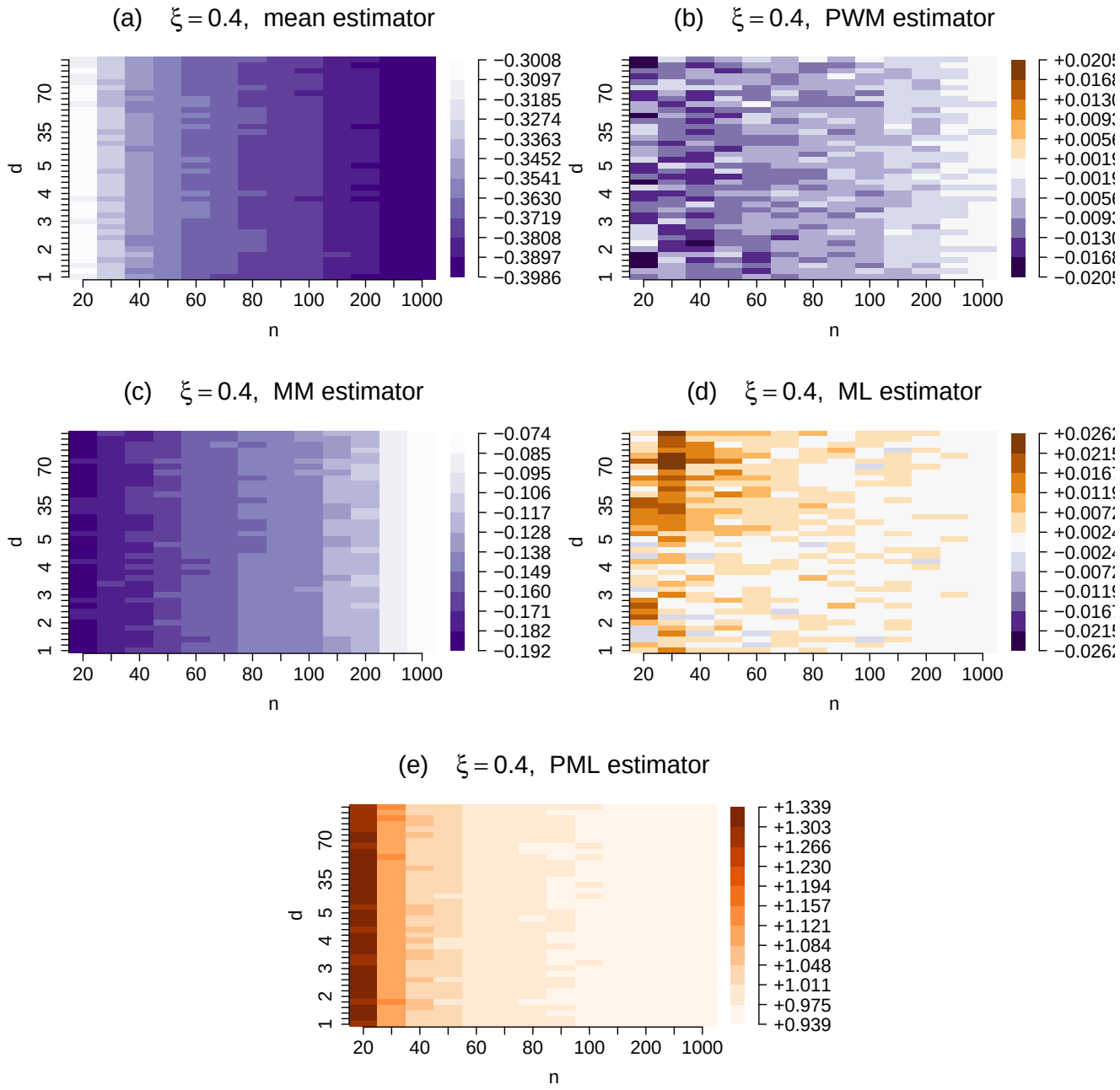

Figure 54: Relative error of the mean estimator (a), PWM (b), MM (c), ML (d) and PML (e) on d as a function of d and n, for $\xi = 0.4$.

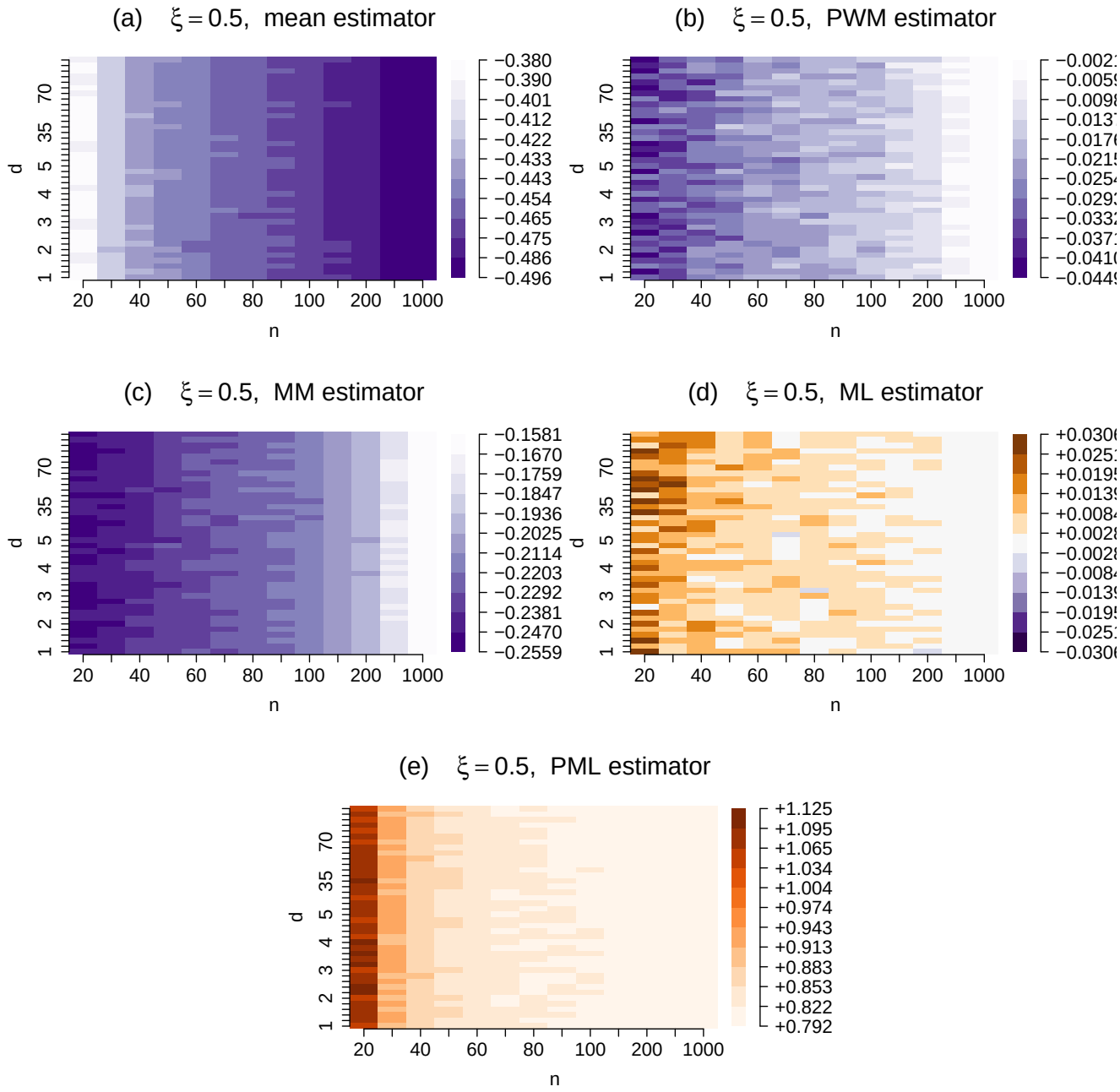

Figure 55: Relative error of the mean estimator (a), PWM (b), MM (c), ML (d) and PML (e) on d as a function of d and n, for $\xi = 0.5$.

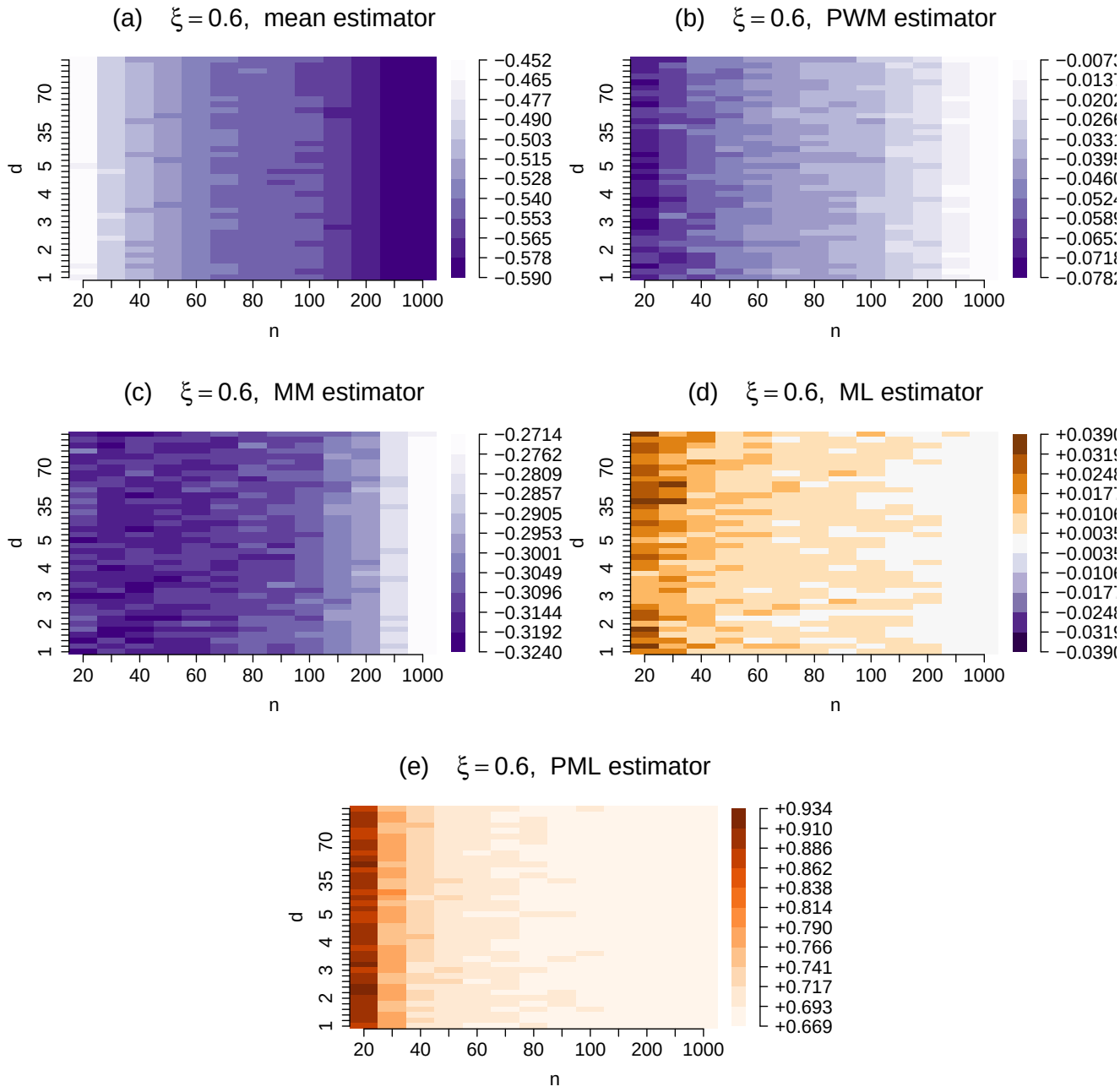

Figure 56: Relative error of the mean estimator (a), PWM (b), MM (c), ML (d) and PML (e) on d as a function of d and n, for $\xi = 0.6$.

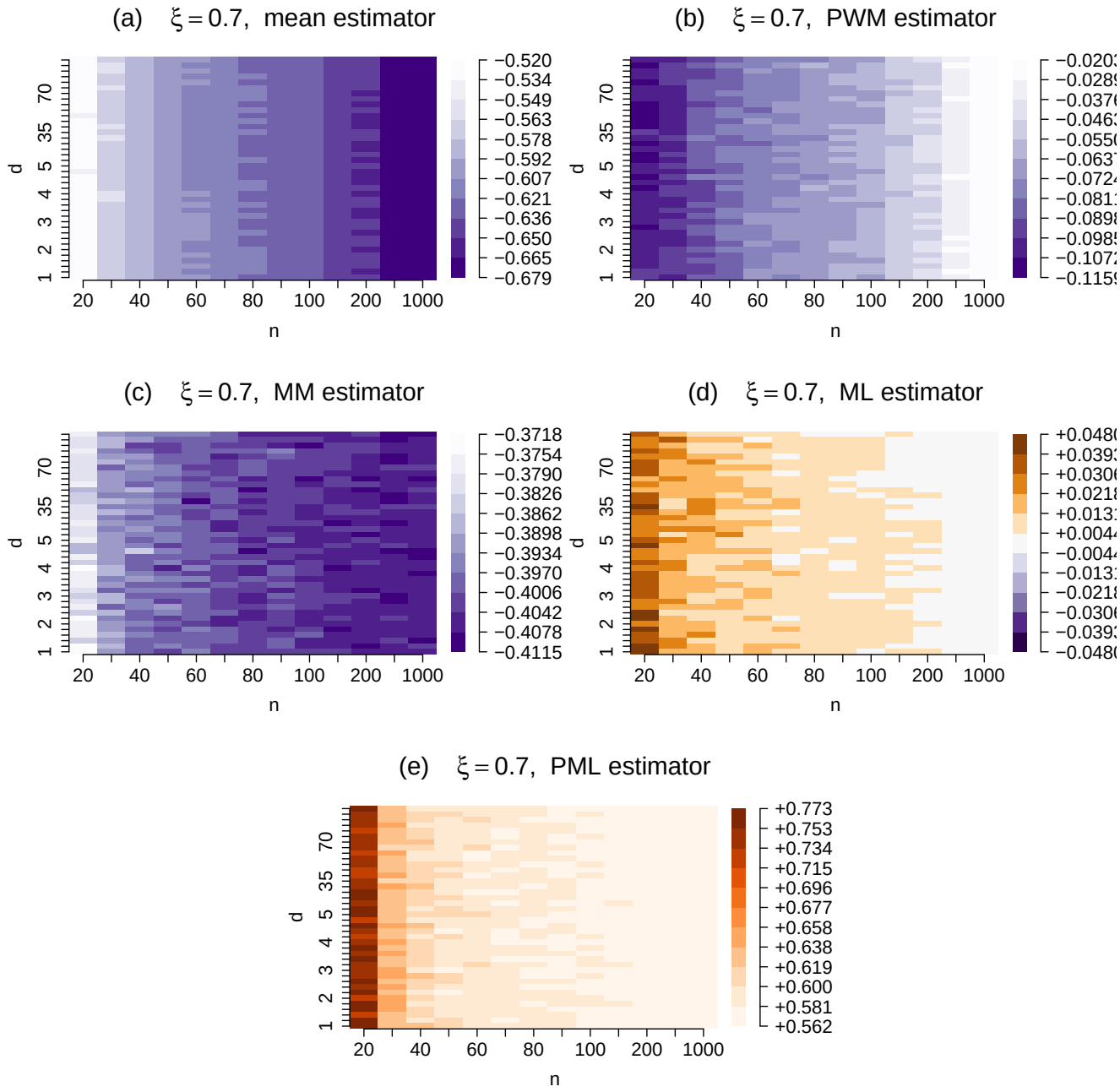

Figure 57: Relative error of the mean estimator (a), PWM (b), MM (c), ML (d) and PML (e) on d as a function of d and n, for $\xi = 0.7$.

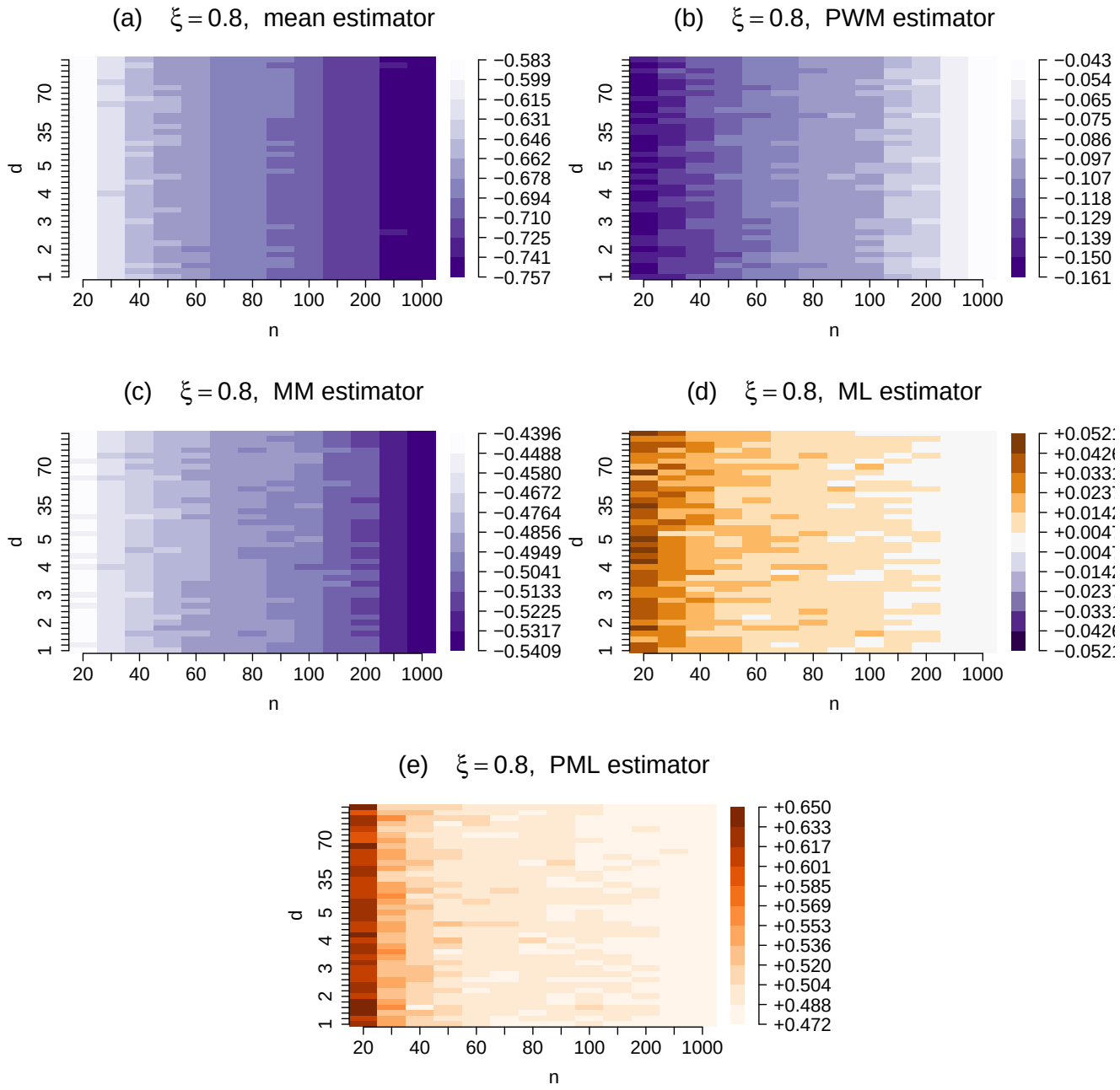

Figure 58: Relative error of the mean estimator (a), PWM (b), MM (c), ML (d) and PML (e) on d as a function of d and n, for $\xi = 0.8$.

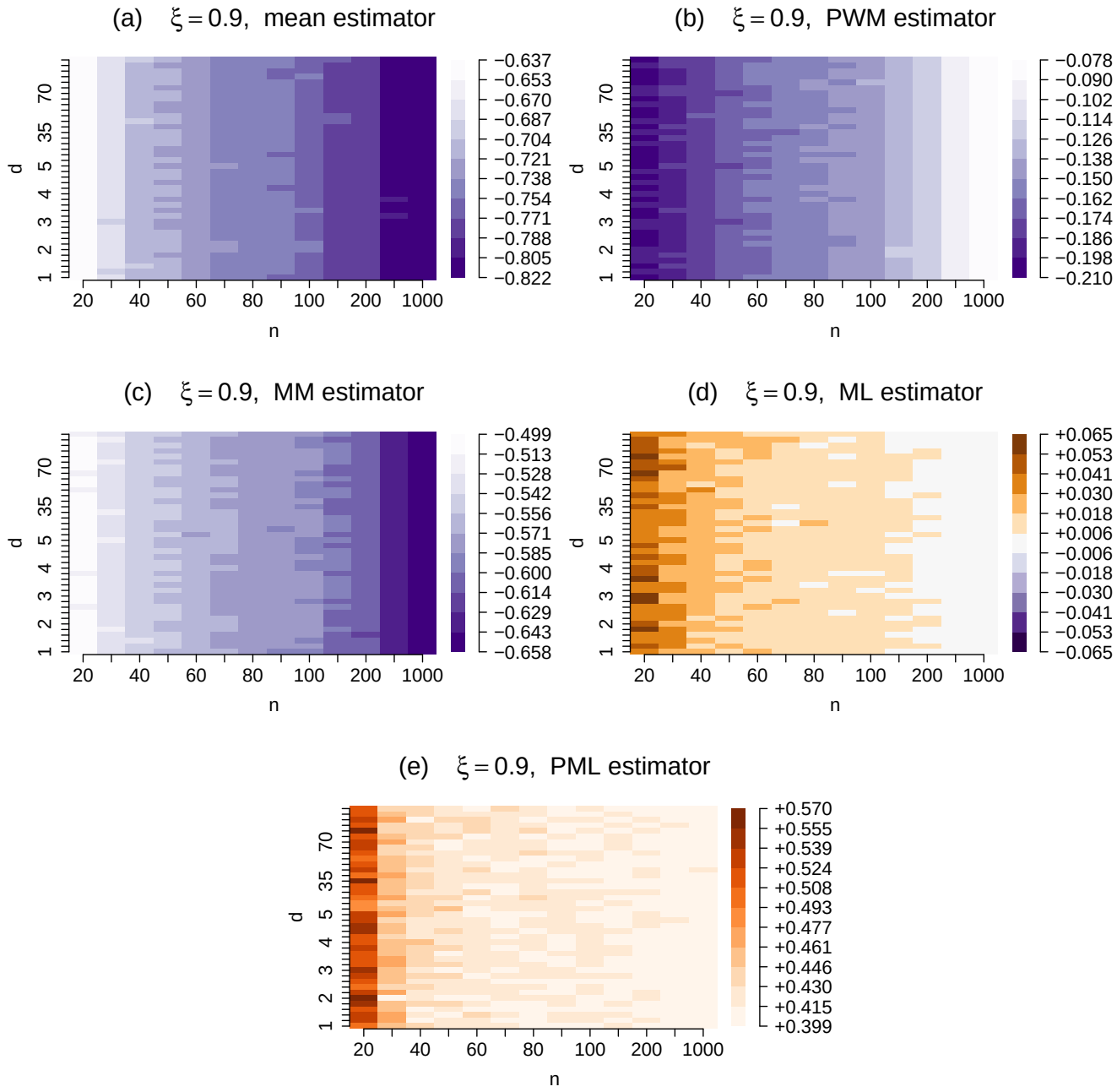

Figure 59: Relative error of the mean estimator (a), PWM (b), MM (c), ML (d) and PML (e) on d as a function of d and n, for $\xi = 0.9$.

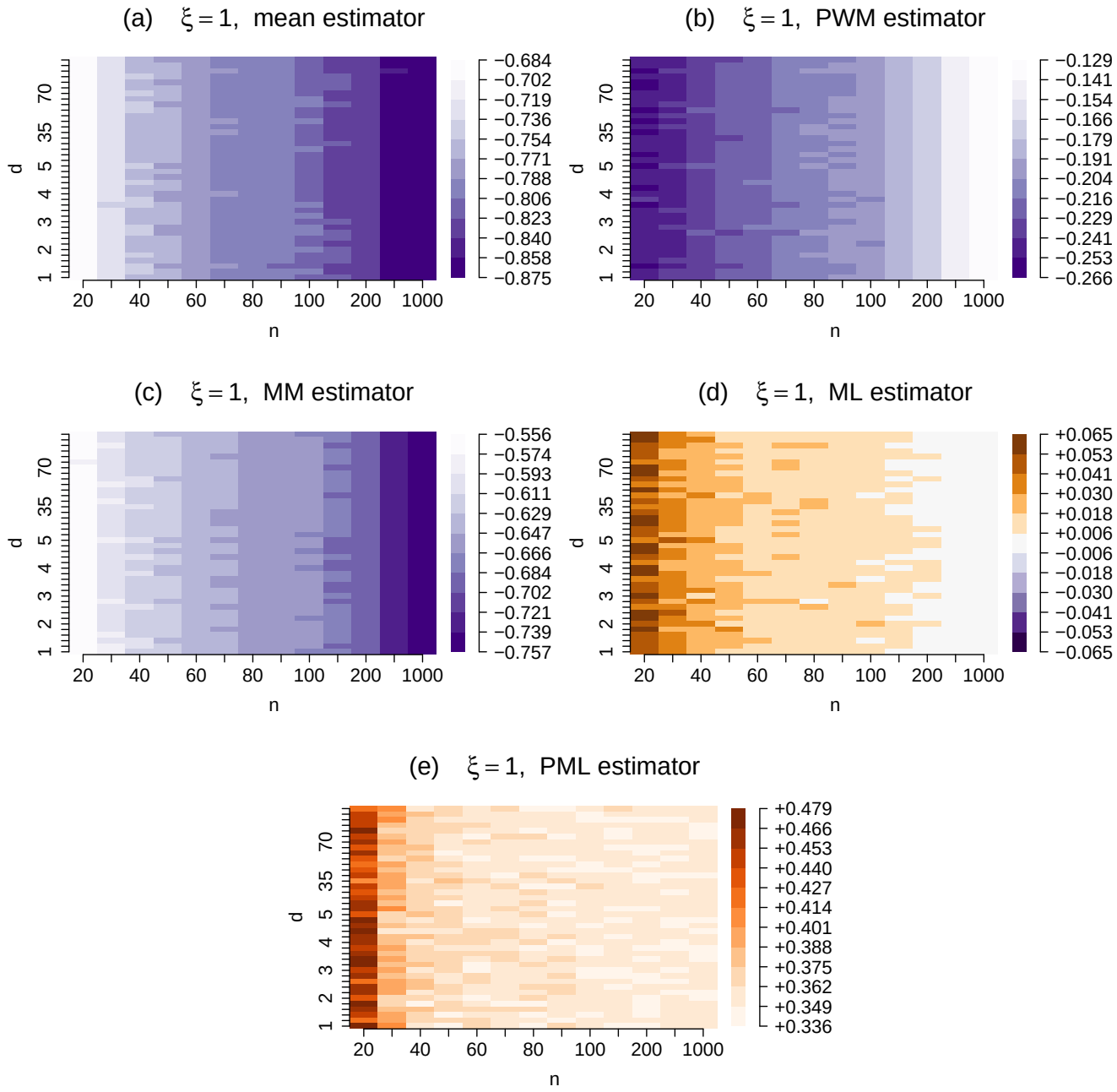

Figure 60: Relative error of the mean estimator (a), PWM (b), MM (c), ML (d) and PML (e) on d as a function of d and n, for $\xi = 1$.

References

Bellman, R. (1957). Dynamic programming. *Princeton university press, Princeton*.

Belzile, L. et al. (2022). *mev: Modelling Extreme Values*. R package version 1.14.

Brunetti, M., Kasparian, J., and Vérard, C. (2019). Co-existing climate attractors in a coupled aquaplanet. *Climate Dynamics*, 53:6293–6308.

Buschow, S. and Friederichs, P. (2018). Local dimension and recurrent circulation patterns in long-term climate simulations. *Chaos: An Interdisciplinary Journal of Nonlinear Science*, 28(8):083124.

Caby, T., Faranda, D., Mantica, G., Vaienti, S., and Yiou, P. (2018). Generalized dimensions, large deviations and the distribution of rare events. *arXiv preprint arXiv:1812.00036*.

Casella, G. and Berger, R. L. (2002). *Statistical inference*, volume 2. Duxbury Pacific Grove, CA.

De Luca, P., Messori, G., Pons, F., and Faranda, D. (2020). Dynamical systems theory sheds new light on compound climate extremes in europe and eastern north america. *Quarterly Journal of the Royal Meteorological Society*.

de Zea Bermudez, P. and Kotz, S. (2010a). Parameter estimation of the generalized pareto distribution—part i. *Journal of Statistical Planning and Inference*, 140(6):1353–1373.

de Zea Bermudez, P. and Kotz, S. (2010b). Parameter estimation of the generalized pareto distribution—part ii. *Journal of Statistical Planning and Inference*, 140(6):1374–1388.

Estrada, E. G. and Alva, J. A. V. (2012). *gPdtest: Bootstrap goodness-of-fit test for the generalized Pareto distribution*. R package version 0.4.

Falasca, F. and Bracco, A. (2022). Exploring the tropical pacific manifold in models and observations. *Physical Review X*, 12(2):021054.

Faranda, D., Alvarez-Castro, M. C., Messori, G., Rodrigues, D., and Yiou, P. (2019a). The hammam effect or how a warm ocean enhances large scale atmospheric predictability. *Nature communications*, 10(1):1316.

Faranda, D., Lucarini, V., Turchetti, G., and Vaienti, S. (2011). Numerical convergence of the block-maxima approach to the generalized extreme value distribution. *J. Stat. Phys.*, 145(5):1156–1180.

Faranda, D., Messori, G., Alvarez-Castro, M. C., and Yiou, P. (2017a). Dynamical properties and extremes of northern hemisphere climate fields over the past 60 years. *Nonlinear Processes in Geophysics*, 24(4):713–725.

Faranda, D., Messori, G., and Yiou, P. (2017b). Dynamical proxies of north atlantic predictability and extremes. *Scientific reports*, 7:41278.

Faranda, D., Messori, G., Yiou, P., Thao, S., Pons, F., and Dubrulle, B. (2023). Dynamical footprints of hurricanes in the tropical dynamics. *Chaos: An Interdisciplinary Journal of Nonlinear Science*, 33(1):013101.

Faranda, D., Sato, Y., Messori, G., Moloney, N. R., and Yiou, P. (2019b). Minimal dynamical systems model of the northern hemisphere jet stream via embedding of climate data. *Earth System Dynamics*, 10(3):555–567.

Fery, L., Dubrulle, B., Podvin, B., Pons, F., and Faranda, D. (2022). Learning a weather dictionary of atmospheric patterns using latent dirichlet allocation. *Geophysical Research Letters*, 49(9):e2021GL096184.

Fisher, R. A. (1912). On an absolute criterion for fitting frequency curves. *Messenger of Mathematics*, 41:155–156.

Freitas, A. C. M., Freitas, J. M., and Todd, M. (2010). Hitting time statistics and extreme value theory. *Probability Theory and Related Fields*, 147(3):675–710.

Giamalaki, K., Beaulieu, C., Henson, S., Martin, A., Kassem, H., and Faranda, D. (2021). Future intensification of extreme aleutian low events and their climate impacts. *Scientific reports*, 11(1):18395.

Ginesta, M., Yiou, P., Messori, G., and Faranda, D. (2022). A methodology for attributing severe extratropical cyclones to climate change based on reanalysis data: the case study of storm alex 2020. *Climate Dynamics*, pages 1–25.

Grassberger, P. (1986). Do climatic attractors exist? *Nature*, 323(6089):609.

Grassberger, P. (1988). Finite sample corrections to entropy and dimension estimates. *Physics Letters A*, 128(6-7):369–373.

Grassberger, P. and Procaccia, I. (1983). Characterization of strange attractors. *Phys. Rev. Lett.*, 50(5):346.

Grassberger, P. and Procaccia, I. (1984). Dimensions and entropies of strange attractors from a fluctuating dynamics approach. *Physica D*, 13(1):34–54.

Grimshaw, S. D. (1993). Computing maximum likelihood estimates for the generalized pareto distribution. *Technometrics*, 35(2):185–191.

Gualandi, A., Avouac, J.-P., Michel, S., and Faranda, D. (2020). The predictable chaos of slow earthquakes. *Science advances*, 6(27):eaaz5548.

Hausdorff, F. (1918). Dimension und äußeres maß. *Mathematische Annalen*, 79(1-2):157–179.

Hersbach, H., Bell, B., Berrisford, P., Biavati, G., Horányi, A., Muñoz Sabater, J., Nicolas, J., Peubey, C., Radu, R., Rozum, I., et al. (2018). Era5 hourly data on single levels from 1979 to present. *Copernicus climate change service (c3s) climate data store (cds)*, 10(10.24381).

Hersbach, H., Bell, B., Berrisford, P., Hirahara, S., Horányi, A., Muñoz-Sabater, J., Nicolas, J., Peubey, C., Radu, R., Schepers, D., et al. (2020). The era5 global reanalysis. *Quarterly Journal of the Royal Meteorological Society*, 146(730):1999–2049.

Hochman, A., Alpert, P., Harpaz, T., Saaroni, H., and Messori, G. (2019a). A new dynamical systems perspective on atmospheric predictability: Eastern mediterranean weather regimes as a case study. *Science advances*, 5(6):eaau0936.

Hochman, A., Alpert, P., Kunin, P., Rostkier-Edelstein, D., Harpaz, T., Saaroni, H., and Messori, G. (2019b). The dynamics of cyclones in the twentyfirst century: the eastern mediterranean as an example. *Climate Dynamics*, pages 1–14.

Hochman, A., Messori, G., Quinting, J. F., Pinto, J. G., and Grams, C. M. (2021a). Do atlantic-european weather regimes physically exist? *Geophysical Research Letters*, 48(20):e2021GL095574.

Hochman, A., Scher, S., Quinting, J., Pinto, J. G., and Messori, G. (2021b). A new view of heat wave dynamics and predictability over the eastern mediterranean. *Earth System Dynamics*, 12(1):133–149.

Hochman, A., Scher, S., Quinting, J., Pinto, J. G., and Messori, G. (2022). Dynamics and predictability of cold spells over the eastern mediterranean. *Climate Dynamics*, 58(7-8):2047–2064.

Hosking, J. R. M., Wallis, J. R., and Wood, E. F. (1985). Estimation of the generalized extreme-value distribution by the method of probability-weighted moments. *Technometrics*, 27(3):251–261.

Landwehr, J. M., Matalas, N., and Wallis, J. (1979). Probability weighted moments compared with some traditional techniques in estimating gumbel parameters and quantiles. *Water resources research*, 15(5):1055–1064.

Lorenz, E. N. (1963). Deterministic nonperiodic flow. *J. Atmos. Sci.*, 20(2):130–141.

Lorenz, E. N. (1991). Dimension of weather and climate attractors. *Nature*, 353(6341):241–244.

Lucarini, V., Faranda, D., de Freitas, A., de Freitas, J., Holland, M., Kuna, T., Nicol, M., Todd, M., and Vaienti, S. (2016). *Extremes and Recurrence in Dynamical Systems*. Pure and Applied Mathematics: A Wiley Series of Texts, Monographs and Tracts. Wiley.

Messori, G., Caballero, R., and Faranda, D. (2017). A dynamical systems approach to studying midlatitude weather extremes. *Geophysical Research Letters*, 44(7):3346–3354.

Messori, G. and Faranda, D. (2021). Characterising and comparing different palaeoclimates with dynamical systems theory. *Climate of the Past*, 17(1):545–563.

Messori, G., Harnik, N., Madonna, E., Lachmy, O., and Faranda, D. (2021). A dynamical systems characterization of atmospheric jet regimes. *Earth System Dynamics*, 12(1):233–251.

Muggeo, V. M. (2003). Estimating regression models with unknown break-points. *Statistics in Medicine*, 22:3055–3071.

Nicolis, C. and Nicolis, G. (1984). Is there a climatic attractor? *Nature*, 311:529–532.

Pons, F. M. E., Messori, G., Alvarez-Castro, M. C., and Faranda, D. (2020). Sampling hyperspheres via extreme value theory: implications for measuring attractor dimensions. *Journal of statistical physics*, 179(5):1698–1717.

Rodrigues, D., Alvarez-Castro, M. C., Messori, G., Yiou, P., Robin, Y., and Faranda, D. (2018). Dynamical properties of the north atlantic atmospheric circulation in the past 150 years in cmip5 models and the 20crv2c reanalysis. *Journal of Climate*, 31(15):6097–6111.

Scher, S. and Messori, G. (2018). Predicting weather forecast uncertainty with machine learning. *Quarterly Journal of the Royal Meteorological Society*, 144(717):2830–2841.

Villaseñor-Alva, J. A. and González-Estrada, E. (2009). A bootstrap goodness of fit test for the generalized pareto distribution. *Computational statistics & data analysis*, 53(11):3835–3841.

Viswanath, D. (2004). The fractal property of the lorenz attractor. *Physica D: Nonlinear Phenomena*, 190(1-2):115–128.