Evaluation and comparison of copulas

David L. Woodruff, Dominic Yang, Simon Strotmann, Maël Forcier

February 2, 2018

Contents

1 Introduction

This report refers to the script "horse_race_copulas.py", which can be used to evaluate the performance of different copulas in describing dependences in a given data set. It can be also used to make a horse race between copulas by running it with the same data set but different copulas. After that one can evaluate, which copula fits the given data the best.

The 2. section explains the theory behind the implemented algorithm, which is shown in section 3. In section 4 there are given some results of some experiments, which make clear that the script works.

2 Theoretical background

The goal is to get a method for evaluating how accurate a computed copula fits the given data. At the end we want compare different computed copulas and see which one performs better in describing the dependence between the several marginals.

2.1 Unreproducibility

For each day the copula is computed using the specific historical data for that day. But the actuals of each day are observed only once. Thus, for each day there is a different copula computed and only one observation made. Therefore, a statement about the quality of each of these copulas is not possible. Not even a statement about the general way of computing the copulas is possible.

For the purpose of evaluating the quality of the whole method, we use the probability integral transformation[1]:

Theorem 1. *If a random variable X has a continuous distribution function $F(x)$, then the random variable $U = F(X)$ has a uniform distribution on the interval $(0,1)$, that is, U is a $U(0,1)$ random variable.*

We define the random variable $U_j = F_j(O_j)$ for each day j, where O_j is the observation from day j and F_j the computed distribution for day j. Theorem 1 says, that U is uniformly distributed, if F is the real density function for the observation. If you repeat this step for more days, then you get a sample $\boldsymbol{U} = (U_1, U_2, \ldots, U_n)$ which is computed independently. All in all this sample should be independent and uniformly distributed.

Now, to make a statement about the quality of our general procedure for computing the copula, we have to measure the difference between the empirical distribution of the sample U and the uniform distribution. In a perfect case, the empirical distribution would be uniform and the distance would be zero. So, the lower the distance is, the better the general method works.

2.2 Rank histograms

One method to evaluate the distance from the uniform distribution are rank histograms. The idea is to produce a plot of the empirical distribution of U and compare it to the plot of a uniform distribution on the interval $(0,1)$ (i.e. equals 1 between 0 and 1, and equals 0 elsewhere). The theory behind this idea is explained in [2].

The main problems of rank histograms are that they are only useful in one dimension and the question how to decide which copula is closer to the uniform distribution by just seeing the plot. We solve the first issue by projecting the samples onto one dimension [2] and the other by using a specific metric: the Earth Mover's Distance or Wasserstein Distance.

[1] see [1]

[2] see section 2.4 "Diagonal"

2.3 Earth Mover's Distance / Wasserstein Distance

The Earth Mover's Distance or Wasserstein Distance is a very nice tool for measuring the distance between two histograms or more general between two probability density functions. In this approach we are using the Wasserstein Distance from scipy.stats, which is defined as follows [3]:

Definition 1. The first Wasserstein Distance between u and v is:

$$l_1(u,v) = \inf_{\pi \in \Gamma(u,v)} \int_{\mathbb{R}\times\mathbb{R}} |x-y| d\pi(x,y)$$

where $\Gamma(u,v)$ is the set of (probability) distributions on $\mathbb{R}\times\mathbb{R}$ whose marginals are u and v on the first and second factors respectively. If U and V are the respective CDF's of u and v, this distance equals to:

$$l_1(u,v) = \int_{-\infty}^{+\infty} |U-V|$$

The input distributions can be empirical, therefore coming from samples whose values are effectively inputs of the function, or they can be seen as generalized functions, in which case they are weighted sums of Dirac delta functions located at the specified values.

The last sentence in Definition 1 means, that the distance between two distributions can be computed by using samples from these distributions.

(Please note, that Maël Forcier implemented the Earth Mover's Distance for his own. This code is usable in this script, too. But usually we use the Wasserstein Distance from scipy.stats. For more information about Maël Forciers code refer to his report "Research Internship Report".)

2.4 Diagonal

To get one dimensional samples (which are originally at least two dimensional in our case), we have to project them onto an one dimensional space. Regarding the dependence, the projection onto the marginal space would make no sense. Therefore we are projecting the data on one of the diagonals. Maël Forcier implemented some code for this purpose and explained it in his report "Research Internship Report". We are using the same code and provide his explanation in the next paragraphs.

2.4.1 Projection on the diagonal

First of all some definitions:

Definition 2. A **corner** of an hypercube $[0,1]^d$ is a point $\boldsymbol{a} = (a_1, \ldots, a_d) \in \{0,1\}^d$. So there are 2^d corners in a hypercube of dimension d.

[3] see [3]

Definition 3. A **diagonal** Δ is a segment which links to opposite corner $\boldsymbol{a}$ and $\boldsymbol{b}$:

$$\Delta = [\boldsymbol{a}, \boldsymbol{b}], \quad \text{where} \quad \forall i = 1, \dots, d : a_i = 0 \Leftrightarrow b_i = 1$$

Alternatively:

$$\Delta = \{(1-\lambda)\boldsymbol{a} + \lambda\boldsymbol{b}, \lambda \in [0,1]\} \quad \text{where} \quad \forall i = 1, \dots, d : a_i = b_i + 1 \mod 2$$

Because one diagonal can be written as $[\boldsymbol{a}, \boldsymbol{b}]$ or $[\boldsymbol{b}, \boldsymbol{a}]$, we will always consider $a_1 = 0$ so that each diagonal has a unique way notation.

Definition 4. We can also define the **direction** of a diagonal as the vector:

$$U_\Delta = \frac{1}{\sqrt{d}}(\boldsymbol{b} - \boldsymbol{a})$$

Definition 5. The **matrix of projection** on the linear space will be:

$$M_\Delta = U_\Delta U_\Delta^T$$

Finally, the **projection on the diagonal** which is an affine space is the function P_Δ such that:

$$P_\Delta(X) + M_\Delta(X - C) + C,$$

with $C = (\frac{1}{2}, \dots, \frac{1}{2})$.

Remark. So there are 2^{d-1} diagonals, directions and matrices of projection in an hypercube of dimension d. The division by $\sqrt{d}$ in the definition of direction permits to have a unit vector. M is indeed a matrix thanks to the order of the factors (and not a scalar product like $U^T U$). One should not confuse the matrix of projection on the linear space with the traditional projection on the diagonal. That is why we need to translate everything with the center C of the hypercube.

2.4.2 Distribution on the diagonal

We now want to study the distribution of the points projected on the diagonal to compare it to a uniform distribution. Since the diagonal is a segment, each point x of the diagonal can be described by only one scalar number λ: $x = (1-\lambda)a + \lambda b)$ (definition of the diagonal). λ can be understood as the normalized distance between a and x:

$$||x - a|| = ||(1-\lambda)a + \lambda b - a|| = \lambda||a - b|| = \lambda\sqrt{d}$$

Where $||\cdot||$ is a norm in our space.
But λ can be easily evaluated by taking the first coordinates of x:

$$x_1 = (1-\lambda)a_1 + \lambda b_1 = (1-\lambda)\cdot 0 + \lambda \cdot 1 = \lambda$$

This equality is valid thanks to our useful convention $(a_1, b_1) = (0, 1)$. Now we have a unique number that should be uniformly distributed on $[0, 1]$.

3 Algorithm and implementation

In this section we provide the implemented algorithm and explain it.

3.1 Algorithm

This algorithm was first introduced in Maël Forciers report "Research Internship Report". So we used primarily his existing code to implement this algorithm and to run some experiments.

I. Compute a sample S by doing for each day j:

1. Take the historic data and fit a copula to it by doing:
   (a) Fit marginal distributions for every hour of interest to the respective data (usually the data is first segmented).
   (b) Transform the data the copula will be fitted to into the copula space (i.e. $[0,1]^d$) using the cumulative distribution functions of the marginals. (for now all the historic data is used to fit the copula to, because we haven't found a good method to segment the data for fitting the copula yet)
   (c) Fit the copula to the transformed data.
2. Generate a sample U_j of n realization of the copula.
3. Project U_j onto one diagonal Δ of the copula space:
$$V_\Delta = (P_\Delta(U_{j,1}), \ldots, P_\Delta(U_{j,d}))$$
   P_Δ is the projection matrix defined in definition 5.
4. Take the observation of today O_j and transform it to the copula space using the marginals cumulative distribution functions:
$$Q_j = (F_{j,1}(O_{j,1}), \ldots, F_{j,d}(O_{j,d}))$$
   $F_{j,i}$ is the cumulative distribution function of day j and marginal i.
5. Project Q_j onto the same diagonal Δ:
$$R_\Delta = P_\Delta(Q_j)$$
6. Compute S_j:
$$S_j = F_\Delta(R_\Delta)$$
   $F_\Delta(x) = \frac{1}{n}\sum_{k=1}^{d} \mathbb{1}_{V_{\Delta,k} \leq x}$ is the empirical distribution of V_Δ.
7. Append the sample S with S_j.

II. Evaluate S:

- Compute the Wasserstein Distance between S and a uniform random or fixed sample X (which means you are computing the Wasserstein Distance between the empirical distribution of S and the uniform distribution, see Definition 1).
- Compute the Earth Mover's Distance between S and a uniform random or fixed sample X (which means you are computing the Earth Mover's Distance between the empirical distribution of S and the uniform distribution, see Definition 1).
- Plot the rank histogram of a uniform random or fixed sample X populated with S.

3.2 Explanation

The main theoretical result this algorithm is based on is the Probability Integral Transformation (PIT) [4]. If the copula, which was fitted to the historical data, describes the real distribution of the historical data and todays observations, the generated sample U_j and the computed Qj must have the same distribution for each day j. Thus, the projected samples V_Δ and R_Δ must have the same distribution. Now using the PIT implies that $S = F(R_\Delta)$, where F is the empirical distribution of V_Δ, must be uniformly distributed. That means, the distance from the empirical distribution of the computed (with the algorithm above) sample S and the uniform distribution is a good way to see how accurate the copula describes the real distribution. And of course you can compare different copulas or ways of computing them and evaluate, which one works better for some given data.

[4] see Theorem 1

4 Experiments and results

All these experiments were run with a fixed uniform sample X.

4.1 2012-2013 BPA data

Diagonal	**Copula**						
	Gaussian	Gumbel	Frank	Clayton	Empirical	Independence	Student
[0, 0], [1, 1]	0.0454	0.0459	0.0434	0.0419	0.0376	0.0856	
[0, 1], [1, 0]	0.0269	0.0174	0.0197	0.017	0.0395	0.1697	

Table 1: Wasserstein distances for 339 observations of pairs of BPA wind forecast errors beginning 2013-01-17:05:00, 2013-01-17:06:00. The copulas were fit using an average of 347 observations beginning 2012-06-02. The marginals were fit using an average of 139 observations of respective hours (i.e. the marginals for the two hours were computed separately) with a MW segmentation filter for the marginals of 0.4. The data file used was 2012-2013_BPA_forecasts_actuals.csv.

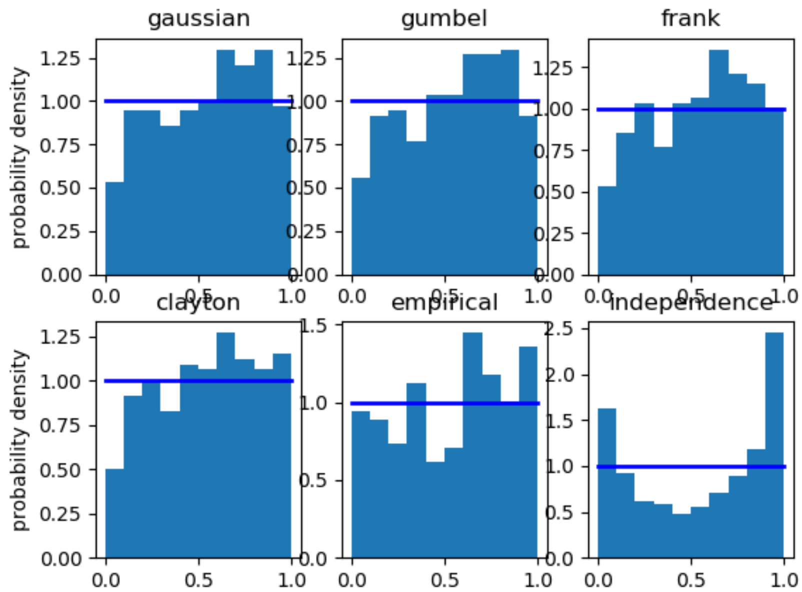

Figure 1: Diagonal [0,0], [1,1]

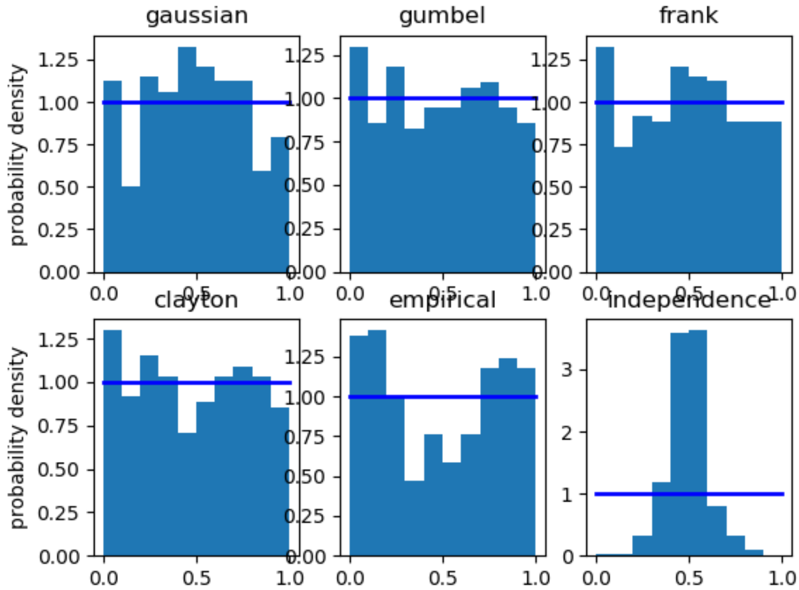

Figure 2: Diagonal [0,1], [1,0]

Diagonal	**Copula**						
	Gaussian	Gumbel	Frank	Clayton	Empirical	Independence	Student
$[0,0],[1,1]$	0.0553	0.0543	0.0543	0.0538	0.0392	0.0843	
$[0,1],[1,0]$	0.0466	0.0342	0.0346	0.0332	0.0369	0.1738	

Table 2: Wasserstein distances for 263 observations of pairs of BPA wind forecast errors beginning 2013-04-05:12:00, 2013-04-05:13:00. The copulas were fit using an average of 384 observations beginning 2012-06-02. The marginals were fit using an average of 154 observations of respective hours (i.e. the marginals for the two hours were computed separately) with a MW segmentation filter for the marginals of 0.4. The data file used was 2012-2013_BPA_forecasts_actuals.csv.

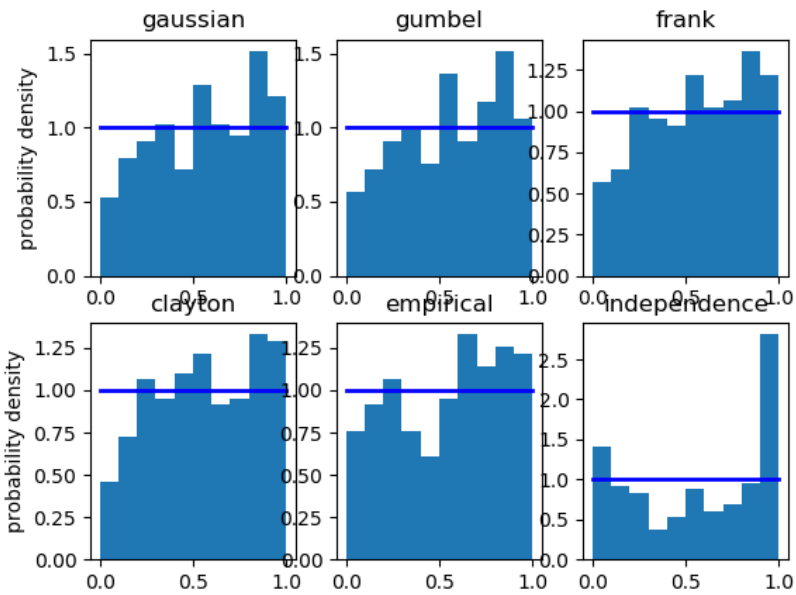

Figure 3: Diagonal [0,0], [1,1]

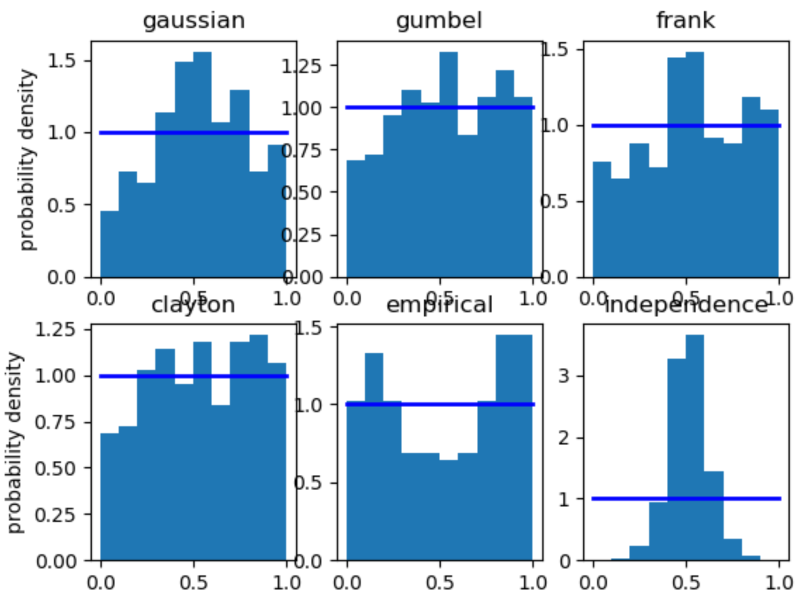

Figure 4: Diagonal [0,1], [1,0]

Diagonal	**Copula**						
	Gaussian	Gumbel	Frank	Clayton	Empirical	Independence	Student
$[0,0],[1,1]$	0.061	0.0622	0.0614	0.0594	0.0601	0.0845	
$[0,1],[1,0]$	0.0734	0.0477	0.0536	0.0484	0.0448	0.1855	

Table 3: Wasserstein distances for 66 observations of pairs of BPA wind forecast errors beginning 2013-10-23:17:00, 2013-10-23:18:00. The copulas were fit using an average of 484 observations beginning 2012-06-02. The marginals were fit using an average of 194 observations of respective hours (i.e. the marginals for the two hours were computed separately) with a MW segmentation filter for the marginals of 0.4. The data file used was 2012-2013_BPA_forecasts_actuals.csv.

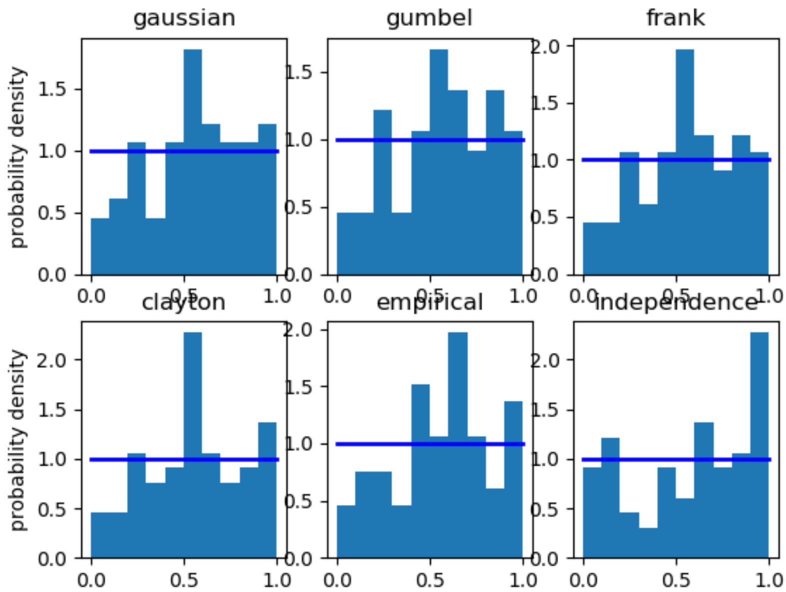

Figure 5: Diagonal [0,0], [1,1]

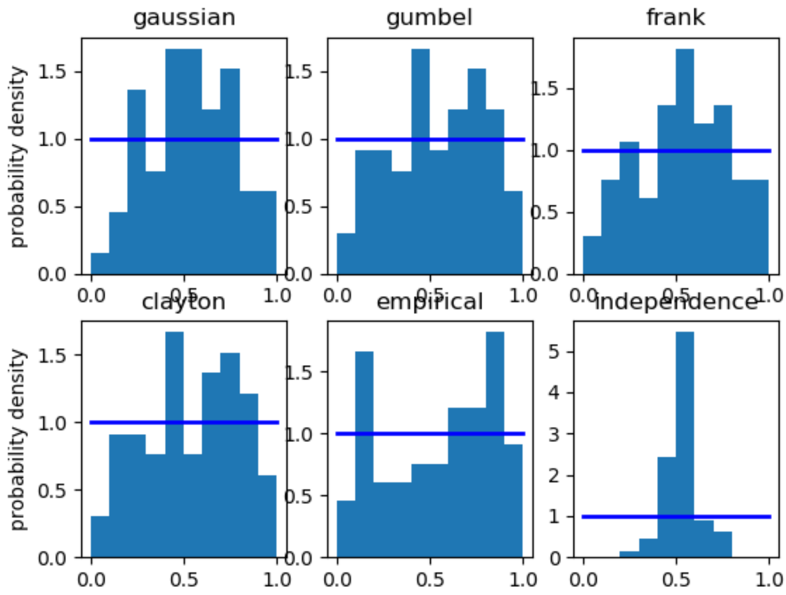

Figure 6: Diagonal [0,1], [1,0]

Diagonal	**Copula**						
	Gaussian	Gumbel	Frank	Clayton	Empirical	Independence	Student
$[0, 0], [1, 1]$	0.0843	0.0858	0.0824	0.0809	0.0698	0.1053	
$[0, 1], [1, 0]$	0.0586	0.0303	0.0409	0.0271	0.0248	0.1798	

Table 4: Wasserstein distances for 136 observations of pairs of BPA wind forecast errors beginning 2013-08-12:21:00, 2013-08-12:22:00. The copulas were fit using an average of 448 observations beginning 2012-06-02. The marginals were fit using an average of 180 observations of respective hours (i.e. the marginals for the two hours were computed separately) with a MW segmentation filter for the marginals of 0.4. The data file used was 2012-2013_BPA_forecasts_actuals.csv.

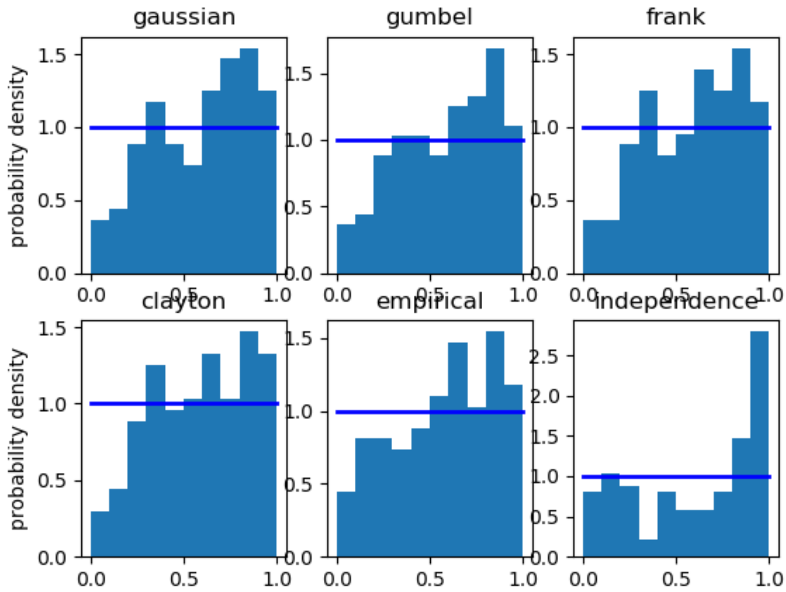

Figure 7: Diagonal [0,0], [1,1]

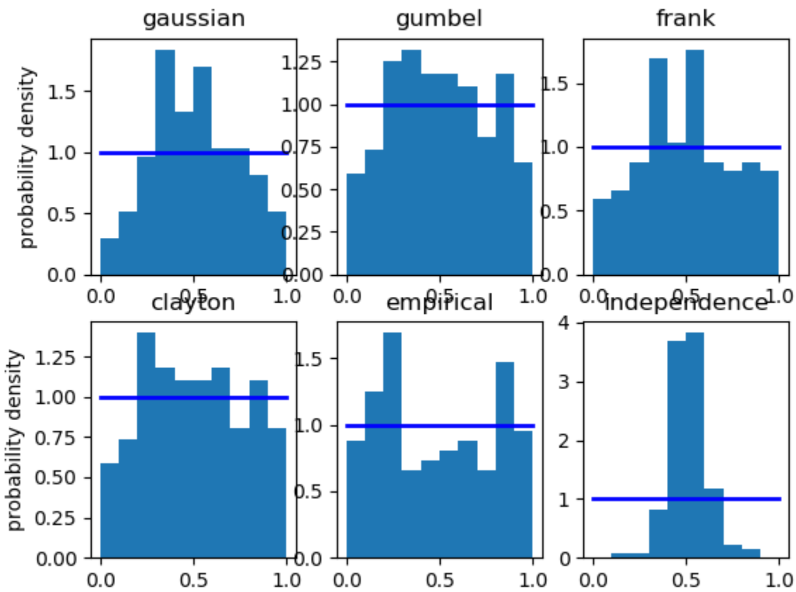

Figure 8: Diagonal [0,1], [1,0]

4.2 All BPA data

Diagonal	**Copula**						
	Gaussian	Gumbel	Frank	Clayton	Empirical	Independence	Student
$[0,0],[1,1]$	0.0215	0.021	0.021	0.0228	0.0306	0.0815	
$[0,1],[1,0]$	0.0585	0.035	0.0415	0.0348	0.0177	0.1723	

Table 5: Wasserstein distances for 400 observations of pairs of BPA wind forecast errors beginning 2016-04-03:15:00, 2016-04-03:16:00. The copulas were fit using an average of 294 observations beginning 2015-12-28. The marginals were fit using an average of 118 observations of respective hours (i.e. the marginals for the two hours were computed separately) with a MW segmentation filter for the marginals of 0.4. The data file used was all_bpa_data.csv.

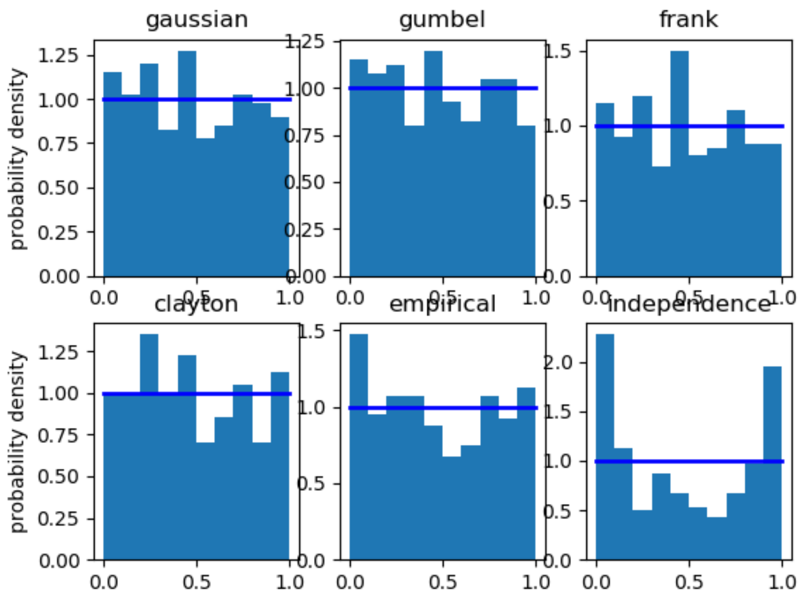

Figure 9: Diagonal [0,0], [1,1]

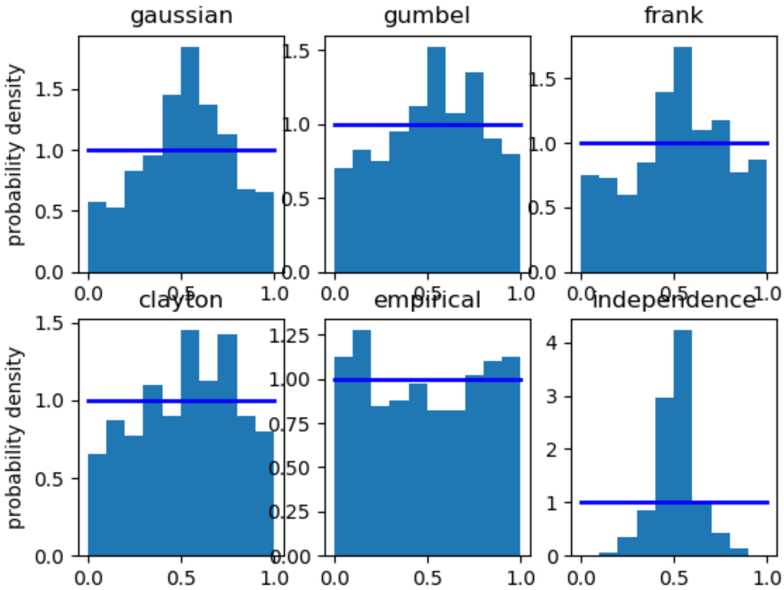

Figure 10: Diagonal [0,1], [1,0]

Diagonal	**Copula**						
	Gaussian	Gumbel	Frank	Clayton	Empirical	Independence	Student
$[0, 0], [1, 1]$	0.023	0.0218	0.0229	0.0271	0.0438	0.0605	
$[0, 1], [1, 0]$	0.0603	0.031	0.0408	0.0286	0.0219	0.1791	

Table 6: Wasserstein distances for 299 observations of pairs of BPA wind forecast errors beginning 2016-07-18:02:00, 2016-07-18:03:00. The copulas were fit using an average of 349 observations beginning 2015-12-28. The marginals were fit using an average of 140 observations of respective hours (i.e. the marginals for the two hours were computed separately) with a MW segmentation filter for the marginals of 0.4. The data file used was all_bpa_data.csv.

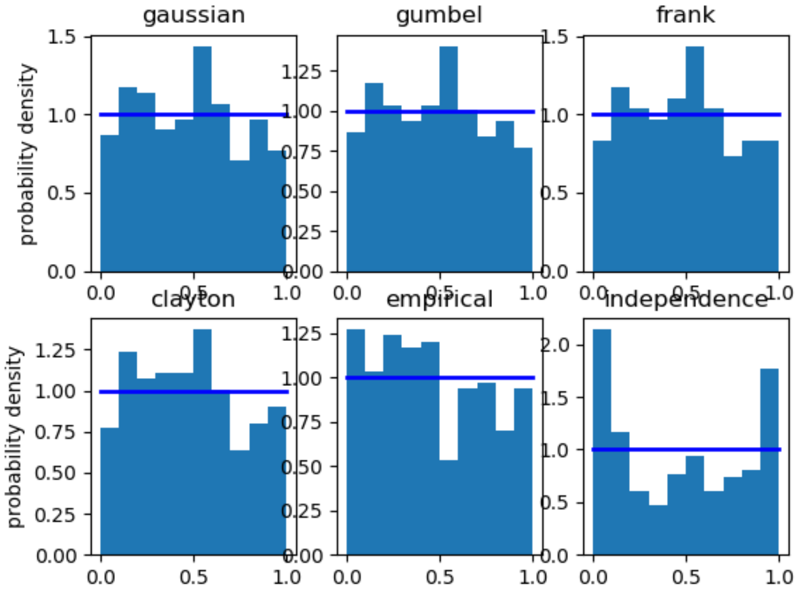

Figure 11: Diagonal [0,0], [1,1]

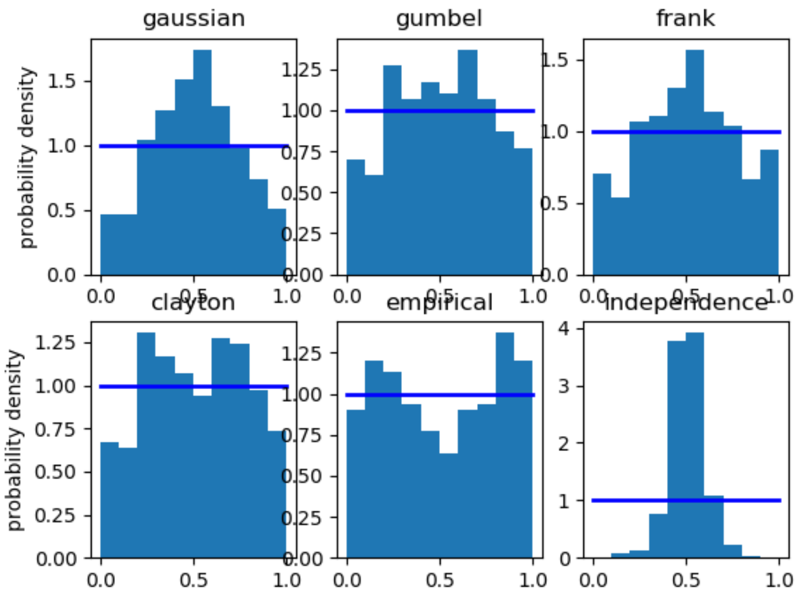

Figure 12: Diagonal [0,1], [1,0]

Diagonal	**Copula**						
	Gaussian	Gumbel	Frank	Clayton	Empirical	Independence	Student
$[0,0],[1,1]$	0.024	0.0244	0.0289	0.0305	0.0302	0.0614	
$[0,1],[1,0]$	0.036	0.0192	0.027	0.0176	0.0324	0.1639	

Table 7: Wasserstein distances for 211 observations of pairs of BPA wind forecast errors beginning 2016-10-11:20:00, 2016-10-11:21:00. The copulas were fit using an average of 390 observations beginning 2015-12-28. The marginals were fit using an average of 156 observations of respective hours (i.e. the marginals for the two hours were computed separately) with a MW segmentation filter for the marginals of 0.4. The data file used was all_bpa_data.csv.

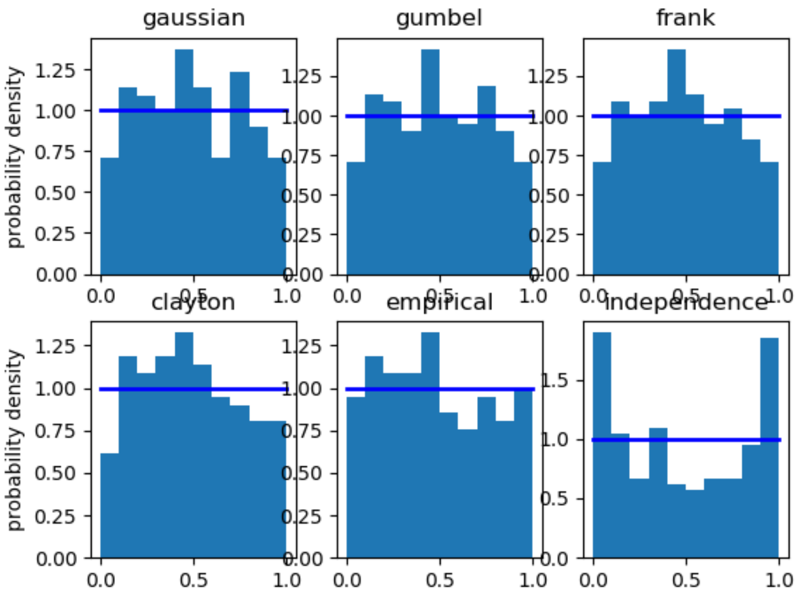

Figure 13: Diagonal [0,0], [1,1]

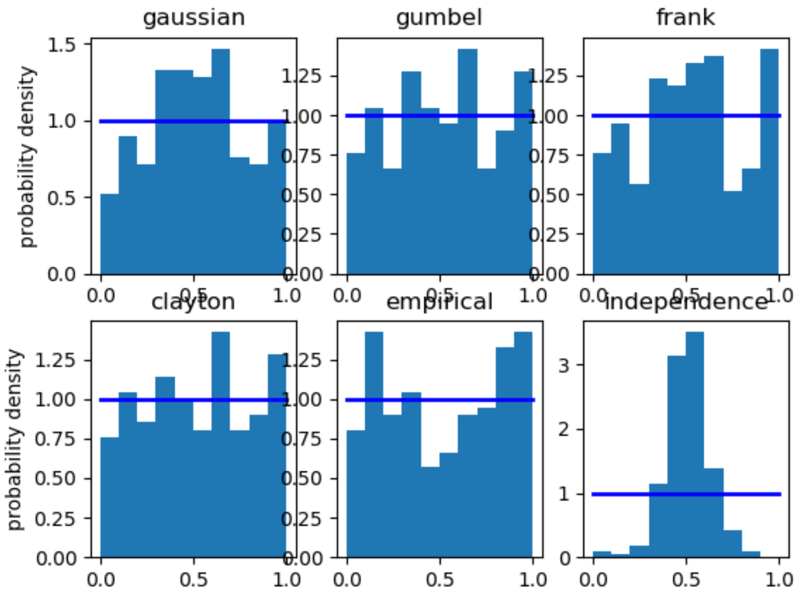

Figure 14: Diagonal [0,1], [1,0]

Diagonal	**Copula**						
	Gaussian	Gumbel	Frank	Clayton	Empirical	Independence	Student
$[0,0],[1,1]$	0.0145	0.0173	0.0191	0.0193	0.0244	0.0638	
$[0,1],[1,0]$	0.0396	0.0216	0.0264	0.0216	0.0329	0.1685	

Table 8: Wasserstein distances for 143 observations of pairs of BPA wind forecast errors beginning 2016-12-24:08:00, 2016-12-24:09:00. The copulas were fit using an average of 423 observations beginning 2015-12-28. The marginals were fit using an average of 170 observations of respective hours (i.e. the marginals for the two hours were computed separately) with a MW segmentation filter for the marginals of 0.4. The data file used was all_bpa_data.csv.

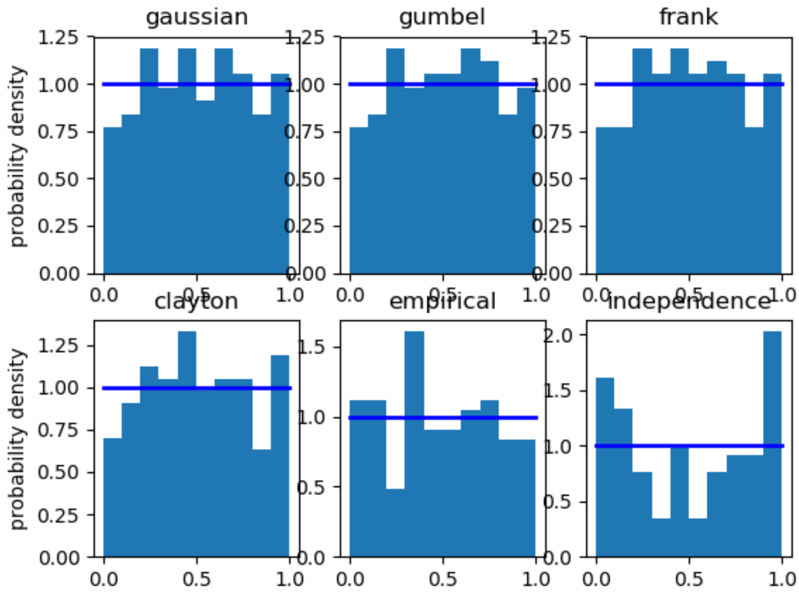

Figure 15: Diagonal [0,0], [1,1]

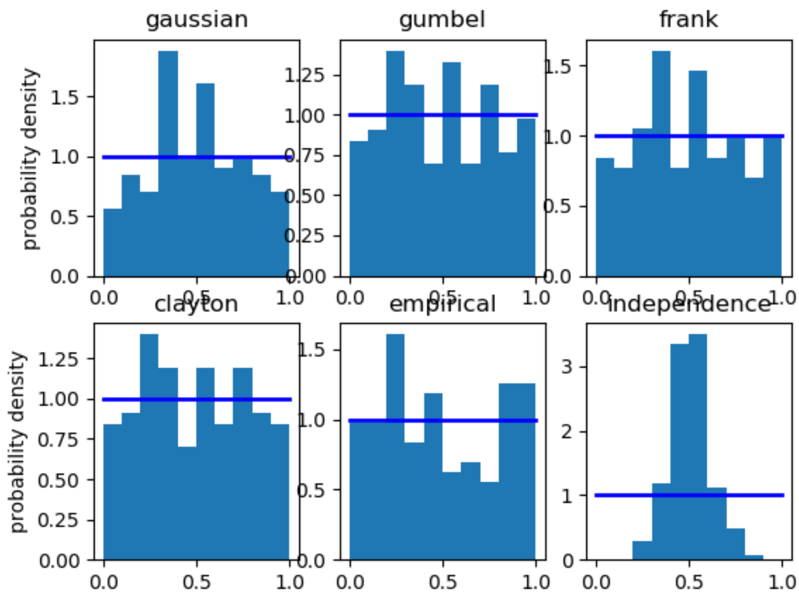

Figure 16: Diagonal [0,1], [1,0]

4.3 RTS data (area 3)

Diagonal	Copula						
	Gaussian	Gumbel	Frank	Clayton	Empirical	Independence	Student
$[0, 0], [1, 1]$	0.0213	0.0246	0.0257	0.0227	0.0091	0.0655	
$[0, 1], [1, 0]$	0.0563	0.0435	0.042	0.0446	0.0288	0.1779	

Table 9: Wasserstein distances for 318 observations of pairs of BPA wind forecast errors beginning 2020-02-18:06:00, 2020-02-18:07:00. The copulas were fit using an average of 206 observations beginning 2020-01-01. The marginals were fit using an average of 83 observations of respective hours (i.e. the marginals for the two hours were computed separately) with a MW segmentation filter for the marginals of 0.4. The data file used was WIND_forecasts_actuals.csv.

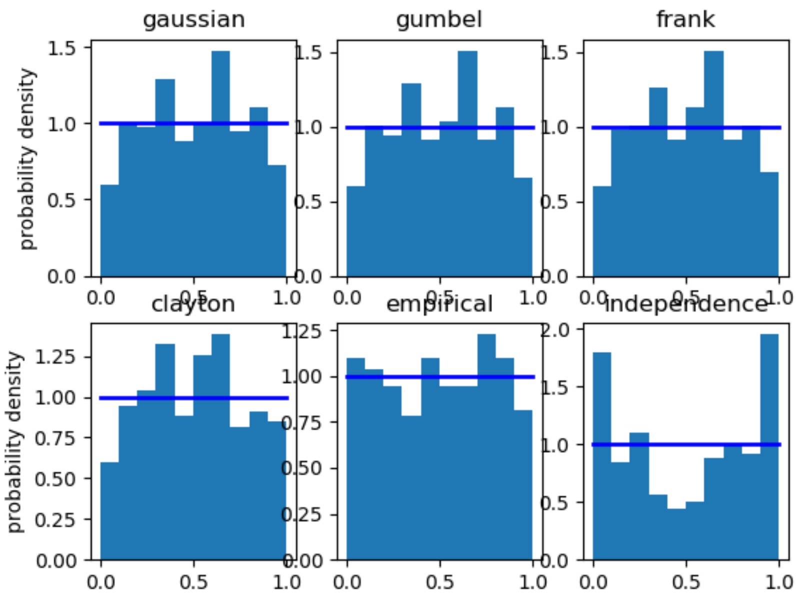

Figure 17: Diagonal [0,0], [1,1]

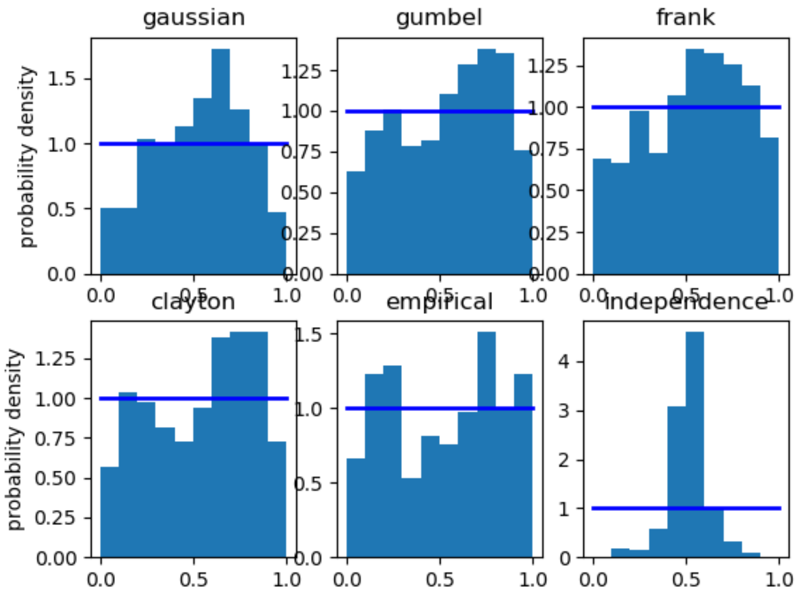

Figure 18: Diagonal [0,1], [1,0]

Diagonal	**Copula**						
	Gaussian	Gumbel	Frank	Clayton	Empirical	Independence	Student
$[0,0],[1,1]$	0.01	0.0111	0.0129	0.0159	0.0315	0.0709	
$[0,1],[1,0]$	0.0564	0.039	0.0394	0.0402	0.0321	0.156	

Table 10: Wasserstein distances for 273 observations of pairs of BPA wind forecast errors beginning 2020-04-03:15:00, 2020-04-03:16:00. The copulas were fit using an average of 229 observations beginning 2020-01-01. The marginals were fit using an average of 92 observations of respective hours (i.e. the marginals for the two hours were computed separately) with a MW segmentation filter for the marginals of 0.4. The data file used was WIND_forecasts_actuals.csv.

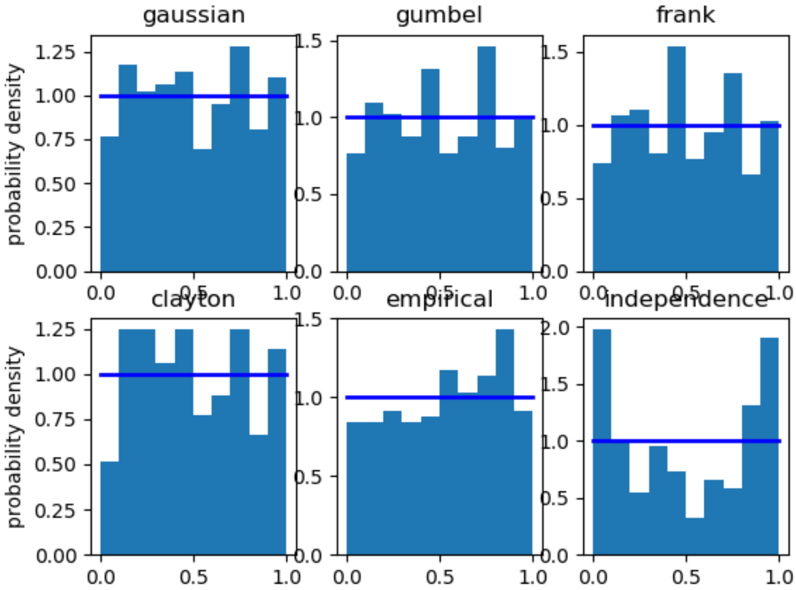

Figure 19: Diagonal [0,0], [1,1]

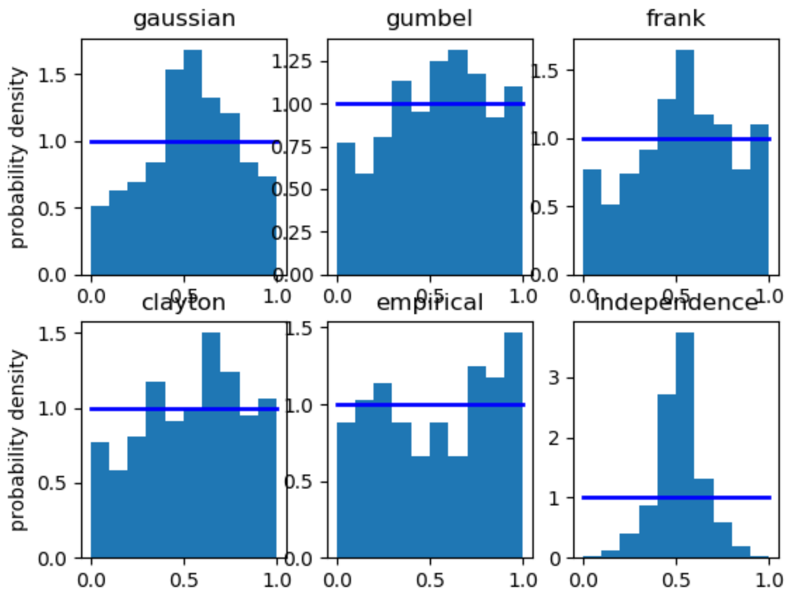

Figure 20: Diagonal [0,1], [1,0]

Diagonal	**Copula**						
	Gaussian	Gumbel	Frank	Clayton	Empirical	Independence	Student
$[0, 0], [1, 1]$	0.0288	0.0317	0.034	0.0328	0.0458	0.0544	
$[0, 1], [1, 0]$	0.0666	0.0526	0.0583	0.0514	0.0284	0.1635	

Table 11: Wasserstein distances for 162 observations of pairs of BPA wind forecast errors beginning 2020-07-23:21:00, 2020-07-23:22:00. The copulas were fit using an average of 284 observations beginning 2020-01-01. The marginals were fit using an average of 114 observations of respective hours (i.e. the marginals for the two hours were computed separately) with a MW segmentation filter for the marginals of 0.4. The data file used was WIND_forecasts_actuals.csv.

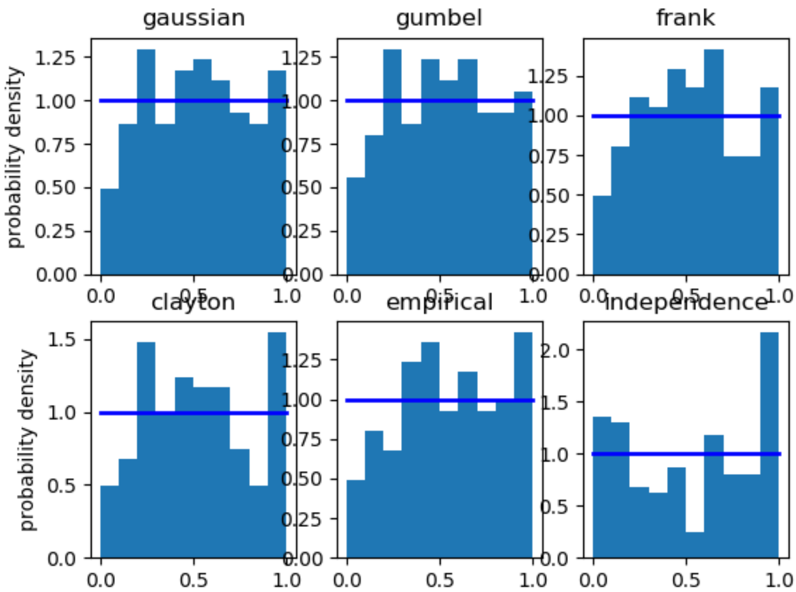

Figure 21: Diagonal [0,0], [1,1]

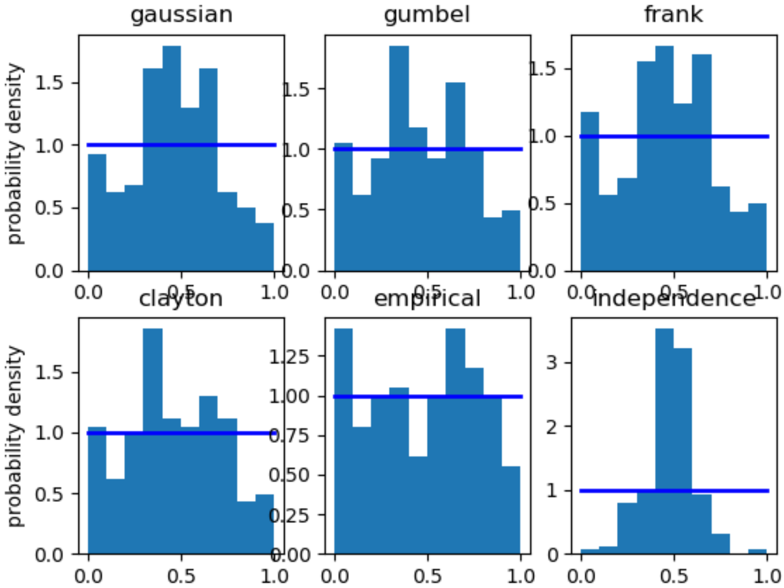

Figure 22: Diagonal [0,1], [1,0]

Diagonal	**Copula**						
	Gaussian	Gumbel	Frank	Clayton	Empirical	Independence	Student
$[0, 0], [1, 1]$	0.0281	0.0256	0.03	0.0341	0.0419	0.0758	
$[0, 1], [1, 0]$	0.0355	0.0192	0.0289	0.0173	0.042	0.1582	

Table 12: Wasserstein distances for 84 observations of pairs of BPA wind forecast errors beginning 2020-10-09:12:00, 2020-10-09:13:00. The copulas were fit using an average of 324 observations beginning 2020-01-01. The marginals were fit using an average of 130 observations of respective hours (i.e. the marginals for the two hours were computed separately) with a MW segmentation filter for the marginals of 0.4. The data file used was WIND_forecasts_actuals.csv.

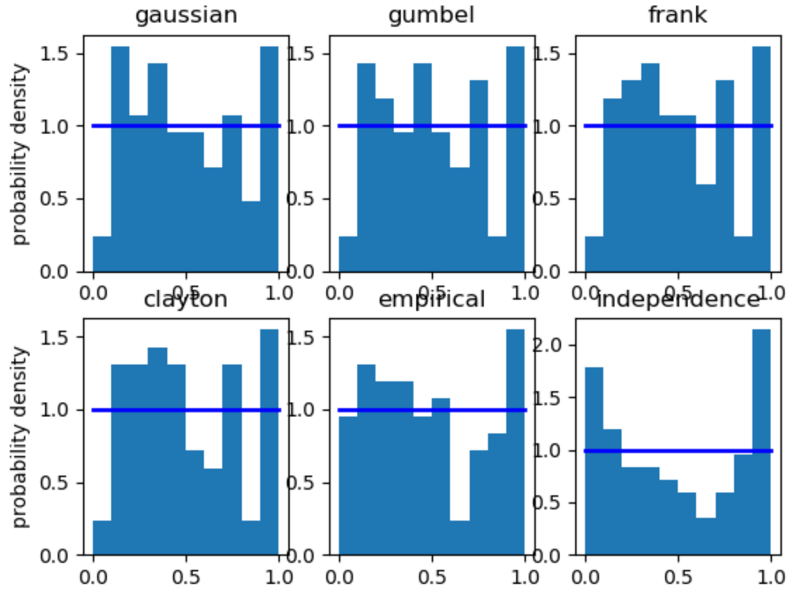

Figure 23: Diagonal [0,0], [1,1]

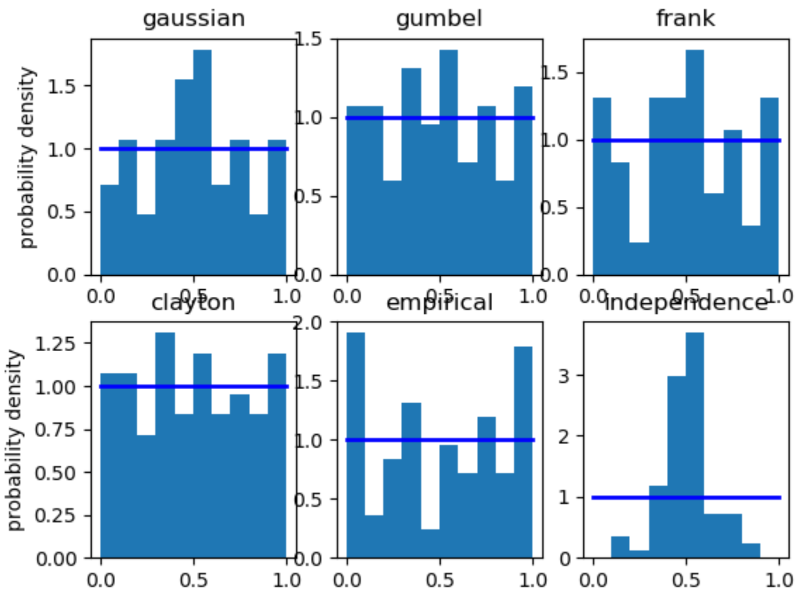

Figure 24: Diagonal [0,1], [1,0]

4.4 RTS data (area 1)

Diagonal	**Copula**						
	Gaussian	Gumbel	Frank	Clayton	Empirical	Independence	Student
$[0,0],[1,1]$	0.0234	0.0293	0.033	0.0308	0.0198	0.0457	
$[0,1],[1,0]$	0.0547	0.022	0.0308	0.0205	0.0376	0.1536	

Table 13: Wasserstein distances for 333 observations of pairs of BPA wind forecast errors beginning 2020-02-03:08:00, 2020-02-03:09:00. The copulas were fit using an average of 199 observations beginning 2020-01-01. The marginals were fit using an average of 82 observations of respective hours (i.e. the marginals for the two hours were computed separately) with a MW segmentation filter for the marginals of 0.4. The data file used was 122_WIND_1_forecasts_actuals.csv.

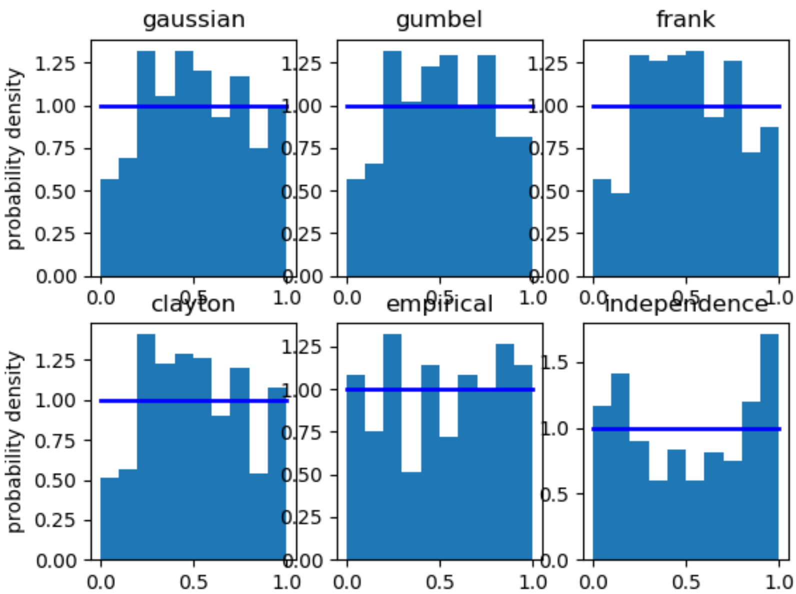

Figure 25: Diagonal [0,0], [1,1]

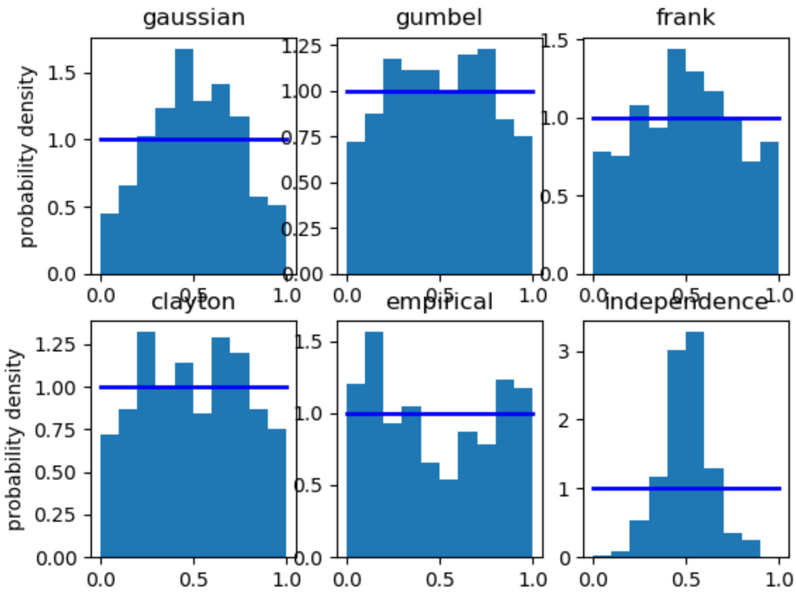

Figure 26: Diagonal [0,1], [1,0]

Diagonal	Copula						
	Gaussian	Gumbel	Frank	Clayton	Empirical	Independence	Student
$[0,0],[1,1]$	0.0285	0.0293	0.03	0.0294	0.0491	0.0705	
$[0,1],[1,0]$	0.0724	0.0378	0.0502	0.0349	0.0319	0.1627	

Table 14: Wasserstein distances for 230 observations of pairs of BPA wind forecast errors beginning 2020-05-16:13:00, 2020-05-16:14:00. The copulas were fit using an average of 250 observations beginning 2020-01-01. The marginals were fit using an average of 101 observations of respective hours (i.e. the marginals for the two hours were computed separately) with a MW segmentation filter for the marginals of 0.4. The data file used was 122_WIND_1_forecasts_actuals.csv.

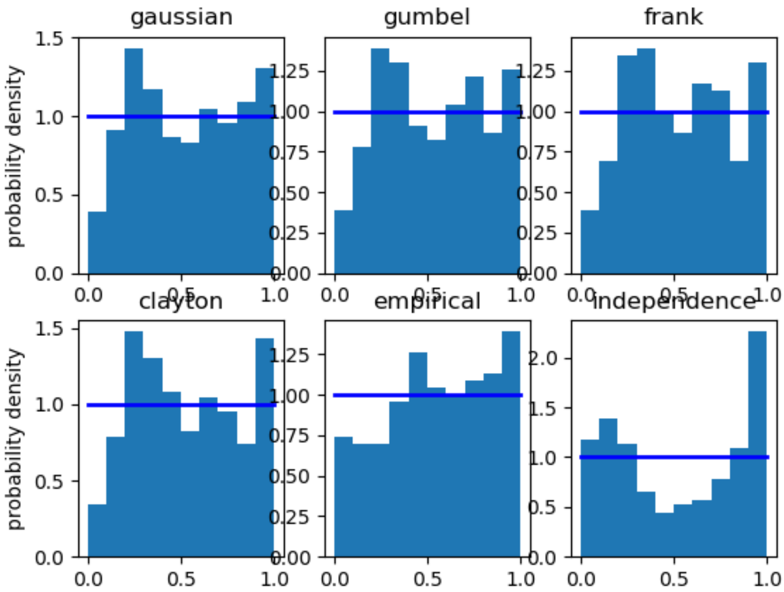

Figure 27: Diagonal [0,0], [1,1]

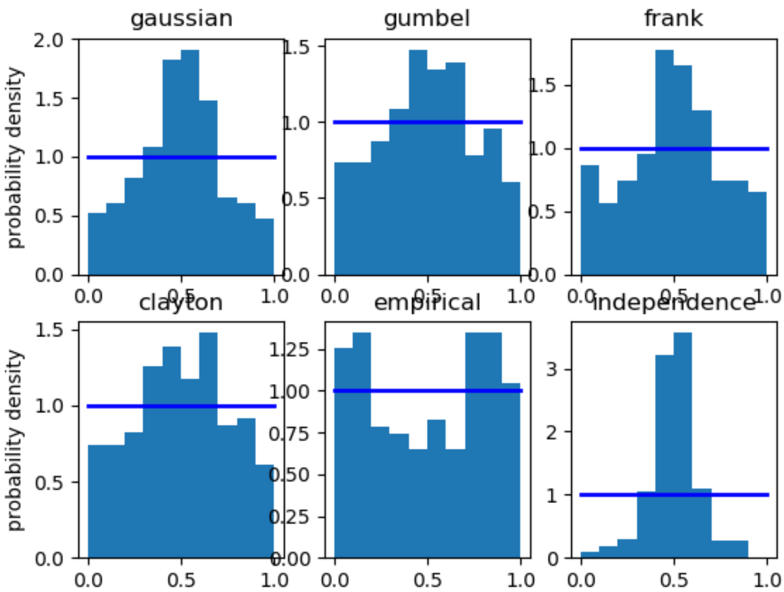

Figure 28: Diagonal [0,1], [1,0]

Diagonal	**Copula**						
	Gaussian	Gumbel	Frank	Clayton	Empirical	Independence	Student
$[0,0],[1,1]$	0.0483	0.0487	0.0485	0.0469	0.0426	0.0929	
$[0,1],[1,0]$	0.0577	0.0315	0.0403	0.0308	0.0325	0.1609	

Table 15: Wasserstein distances for 127 observations of pairs of BPA wind forecast errors beginning 2020-08-27:18:00, 2020-08-27:19:00. The copulas were fit using an average of 302 observations beginning 2020-01-01. The marginals were fit using an average of 122 observations of respective hours (i.e. the marginals for the two hours were computed separately) with a MW segmentation filter for the marginals of 0.4. The data file used was 122_WIND_1_forecasts_actuals.csv.

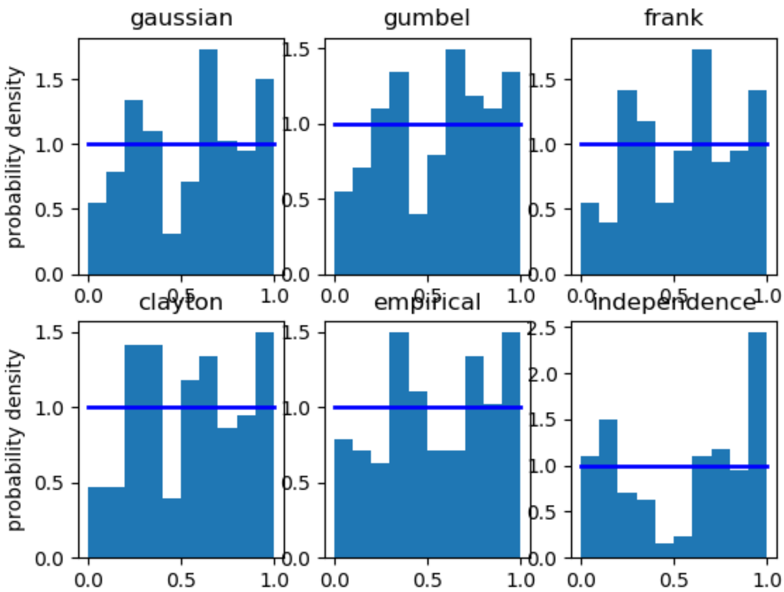

Figure 29: Diagonal [0,0], [1,1]

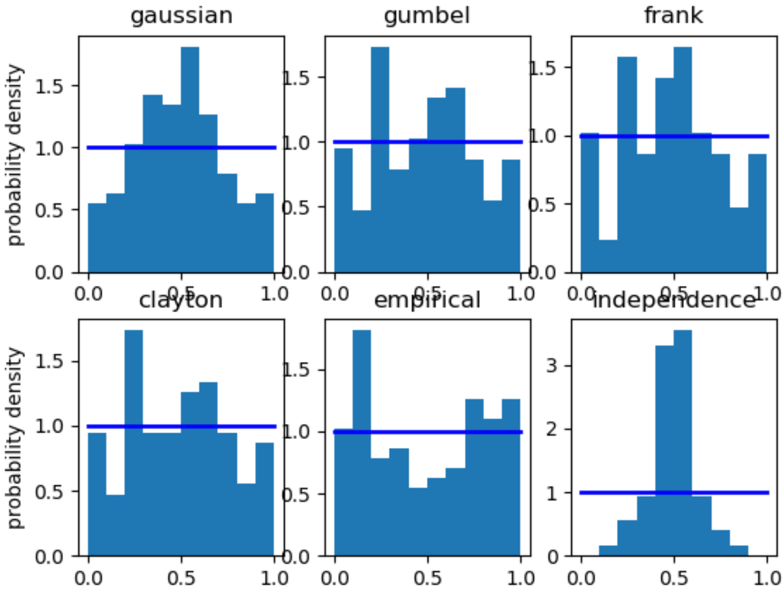

Figure 30: Diagonal [0,1], [1,0]

Diagonal	**Copula**						
	Gaussian	Gumbel	Frank	Clayton	Empirical	Independence	Student
$[0, 0], [1, 1]$	0.0573	0.0581	0.0591	0.0572	0.0294	0.0842	
$[0, 1], [1, 0]$	0.0825	0.0857	0.084	0.0861	0.0601	0.1709	

Table 16: Wasserstein distances for 51 observations of pairs of BPA wind forecast errors beginning 2020-11-11:02:00, 2020-11-11:03:00. The copulas were fit using an average of 340 observations beginning 2020-01-01. The marginals were fit using an average of 137 observations of respective hours (i.e. the marginals for the two hours were computed separately) with a MW segmentation filter for the marginals of 0.4. The data file used was 122_WIND_1_forecasts_actuals.csv.

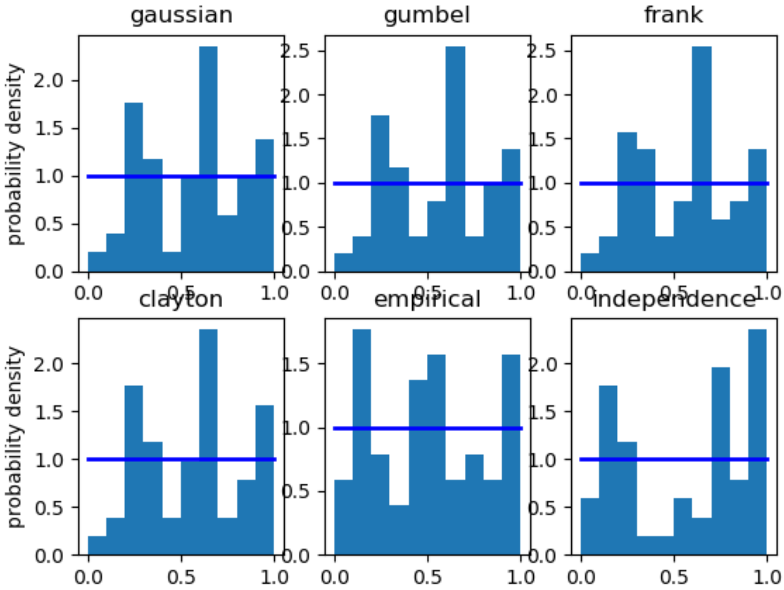

Figure 31: Diagonal [0,0], [1,1]

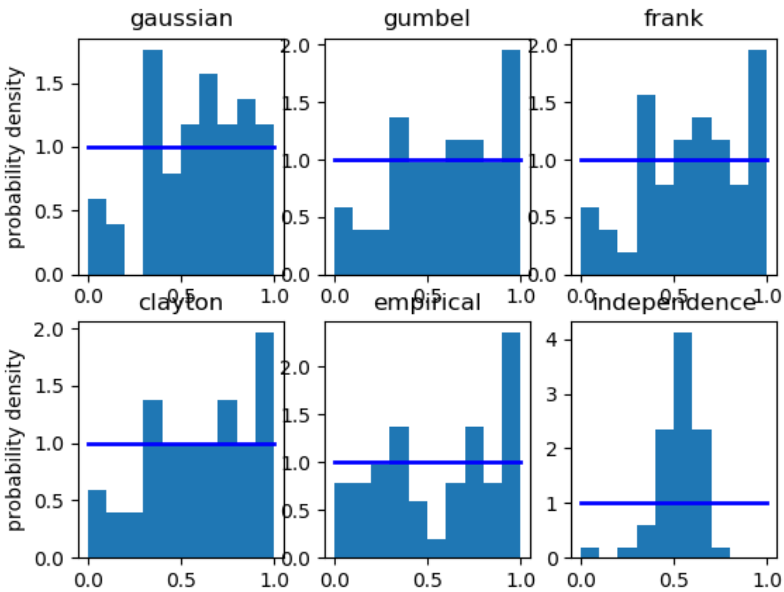

Figure 32: Diagonal [0,1], [1,0]

4.5 BPA 2012-2017

Diagonal	**Copula**						
	Gaussian	Gumbel	Frank	Clayton	Empirical	Independence	Student
$[0,0],[1,1]$	0.0249	0.0247	0.0267	0.0296	0.0235	0.0702	
$[0,1],[1,0]$	0.0482	0.0164	0.0305	0.0135	0.0411	0.1779	

Table 17: Wasserstein distances for 870 observations of pairs of BPA wind forecast errors beginning 2015-06-15:12:00, 2015-06-15:13:00. The copulas were fit using an average of 1504 observations beginning 2012-06-02. The marginals were fit using an average of 603 observations of respective hours (i.e. the marginals for the two hours were computed separately) with a MW segmentation filter for the marginals of 0.4. The data file used was bpa_wind_forecasts_actuals_2012_2017.csv.

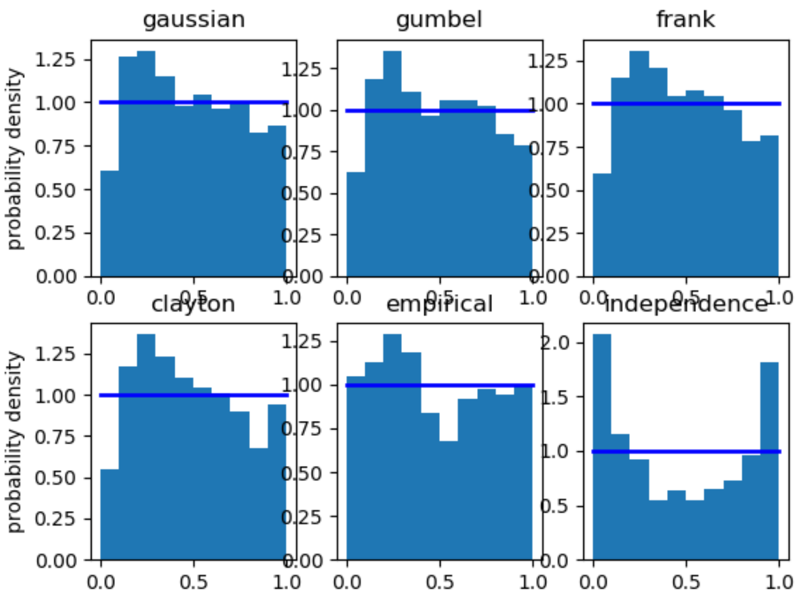

Figure 33: Diagonal [0,1], [1,0]

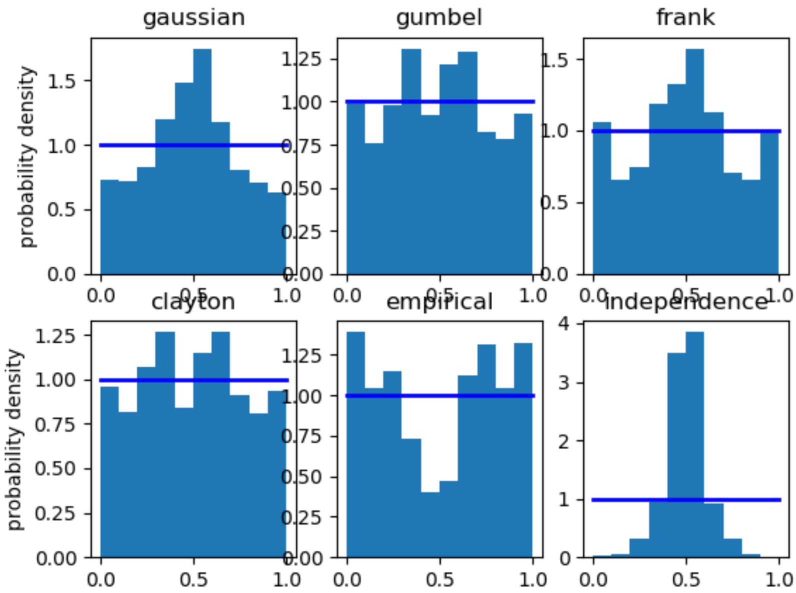

Figure 34: Diagonal [0,1], [1,0]

Diagonal	**Copula**						
	Gaussian	Gumbel	Frank	Clayton	Empirical	Independence	Student
$[0, 0], [1, 1]$	0.0175	0.0206	0.0208	0.0194	0.0101	0.0691	
$[0, 1], [1, 0]$	0.0462	0.0116	0.0258	0.0081	0.0407	0.1774	

Table 18: Wasserstein distances for 1727 observations of pairs of BPA wind forecast errors beginning 2013-01-01:12:00, 2013-01-01:13:00. The copulas were fit using an average of 1076 observations beginning 2012-06-02. The marginals were fit using an average of 431 observations of respective hours (i.e. the marginals for the two hours were computed separately) with a MW segmentation filter for the marginals of 0.4. The data file used was bpa_wind_forecasts_actuals_2012_2017.csv.

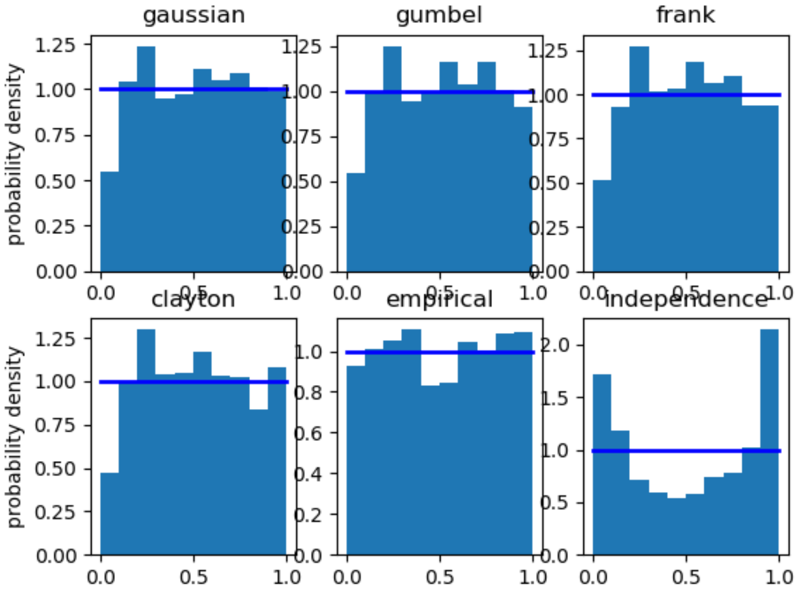

Figure 35: Diagonal [0,1], [1,0]

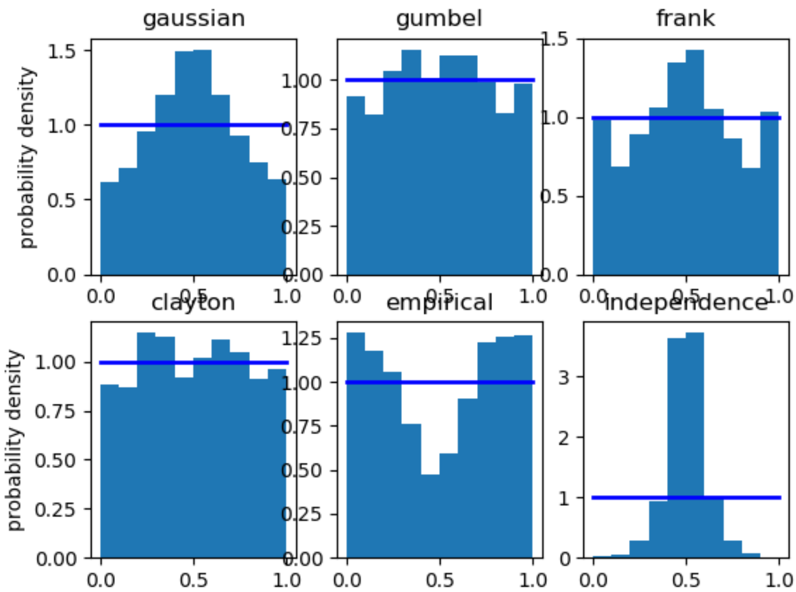

Figure 36: Diagonal [0,1], [1,0]

References

[1] *Encyclopedia of Statistical Sciences*, 2006, John Wiley & Sons, Inc.

[2] Thomas M. Hamill, *Interpretation of Rank Histograms for Verifying Ensemble Forecasts*, 2000

[3] *Scipy documentation about the Wasserstein Distance*, `https://docs.scipy.org/doc/scipy/reference/generated/scipy.stats.wasserstein_distance.html`