



# Early Warning with Calibrated and Sharper Probabilistic Forecasts

 [a. t. L](#) <sup>†</sup>

<sup>†</sup>Department of Mathematics and Statistics, University of Reading, United Kingdom  
[r.l.machete@reading.ac.uk](mailto:r.l.machete@reading.ac.uk)

## Abstract

Given a nonlinear deterministic model, a density forecast i

ar na ca brat on w. n , not a o at t sa t D ns t or asts ar  
 sa to b probab st a a brat an on t I say un of str but  
 ar na a brat on r rs to t as w n t av raf o a ns or asts  
 s qua to t at o a or asts r s no m p r a wa o ass snf x an  
 a brat on  
 Gn t nF a t n on tur t at w n a subs t o t s m o o a  
 brat on o s t n t pr t v str but ons ar at ast as spr out as t a  
 or asts w on tur t t a s arpn ss pr nc p t s arpn ss  
 s ms to av b n on b Bross an t sam asur o ow on ntrat  
 probab st or asts ar an a prop rt o o bof

to t s attr but s w n mato of s n u , m o s o a brat on as ss ar  
probab st an n ar f na brat on w t m p ass t at ntrop s ou b  
us to m asur s arpn ss n wpropost ont at a r ss s Gn tn f ta s  
on tur os xp ant n nfs r as Bro r mt g s rt s to  
not n u nfa n un tons nt inst smation was ts Fuarant n ras n  
varan ot pr tv str but on e par to t raw nsmb w monstrat  
t at t us o varay tom asur s arpn ss ou b m s a nf  
s pap r s orfan s as o ws c n xts ton s uss s or ast qua t st at  
ar h i at v m gasur b t of art m s or nfr u In part u ar a opo s  
ton o t s s or nfr u s pr s nt s arpn ss pr n p on tur Gn tn f  
ta g s s uss an ar vant propos tion pr s nt n s In s t m t o  
o of m po to pre u rst or asts s ou n su ts on rn nf inst  
or asts obtan vat of art m s or nfr u wt xsp t to a non n ar tron  
r ut a s uss n s ton Fv s on u nfr m ar s App n s A an B  
onta nt pico ot proposition on rn nft s arpn ss pr n p an app n x C  
ontans proos ort r sto t propositions

## 2 Probabilistic-Forecast Quality

I'm especially at on p's m'at ons on t' val' o' probab' st' or a' ts n  
 t' ot' r' an' onst' r' s o' r' a' ts m' a' m' an' pr' t' y' str but on t' at ar  
 bot' *ca' d' rat' an' s' urp' l' e' s' u' l' t' s* or a' ts ar' s' s' n' at on f' m' or on s t' n ar'  
 warn n' s' a' or' s' u' f' f' s' t' at t' s' q' u' a' n' t' s' an' h' t' u' a' t' quant' b  
 t' of ar t' m' s or n' f' r' u' p' p' o' s' b' Goo' r' ar' o' u' r' s' o' n' f' r' u' s  
 ava' ab' l' e' r' s' t' o' n' s' G' n' t' n' f' a' t' r' For instan' t' r' s' t' Br' r  
 s' o' n' Er' r' s' s' low' v' r' e' p' o' s' s' n' t' o' m' n' t' h' s' u' r' p' r' s' s' s' o' n'  
 o' z' w' ar' not' want to ou' r' s' u' s' s' o' n' an' t' s' s' a' b' l' e' or at f' e' r' a' v' n' t' s' A  
 f' n' r' a' s' a' t' o' n' o' t' Er' r' s' o' r' t' o' n' s' t' o' r' a' ts' s' t'



o two Gaussians with a low relative entropy Gaussian, but on the other hand, a Gaussian with a high relative entropy Gaussian. For a Gaussian with a low relative entropy, the relative entropy is also bounded by a constant. For a Gaussian with a high relative entropy, the relative entropy is also bounded by a constant. For a Gaussian with a low relative entropy, the relative entropy is also bounded by a constant. For a Gaussian with a high relative entropy, the relative entropy is also bounded by a constant.

## 2.3 Calibration

Calibration of a model's predictions is a well-known problem in machine learning. In this section, we discuss the concept of calibration and its relationship to the concept of proper scoring rules. We start by defining a proper scoring rule. A proper scoring rule is a function  $S(y, \hat{y})$  that maps a true outcome  $y$  and a predicted outcome  $\hat{y}$  to a real number. The function  $S$  is called proper if it satisfies the following property: for any two predicted outcomes  $\hat{y}_1$  and  $\hat{y}_2$ , the expected value of  $S(y, \hat{y}_1)$  is less than or equal to the expected value of  $S(y, \hat{y}_2)$  if and only if  $\hat{y}_1$  is more accurate than  $\hat{y}_2$ . This property ensures that the model's predictions are calibrated to the true outcomes.

The following theorem shows that a proper scoring rule is equivalent to a calibrated model.

$$\mathbb{E}[S(y, \hat{y}_1)] \leq \mathbb{E}[S(y, \hat{y}_2)] \iff \hat{y}_1 \text{ is more accurate than } \hat{y}_2$$

The following theorem shows that a calibrated model is equivalent to a proper scoring rule.

$$\mathbb{E}[S(y, \hat{y}_1)] \leq \mathbb{E}[S(y, \hat{y}_2)] \iff \hat{y}_1 \text{ is more accurate than } \hat{y}_2$$

For a proper scoring rule, the following property holds:

$$\mathbb{E}[S(y, \hat{y}_1)] \leq \mathbb{E}[S(y, \hat{y}_2)] \iff \hat{y}_1 \text{ is more accurate than } \hat{y}_2$$

I will have a theorem that says that if a model's predictions are calibrated, then the model's predictions are also proper. This is a well-known result in the literature. The proof of this theorem is straightforward. Suppose that a model's predictions are calibrated. Then, for any two predicted outcomes  $\hat{y}_1$  and  $\hat{y}_2$ , the expected value of  $S(y, \hat{y}_1)$  is less than or equal to the expected value of  $S(y, \hat{y}_2)$  if and only if  $\hat{y}_1$  is more accurate than  $\hat{y}_2$ . This property is exactly the definition of a proper scoring rule.

uppos w . av  $\theta$  s r s o , nst or asts  $\{ t x \}_{t=1}^T$  . n , n t  
 or ast r s m ato of as

$$T x \leftarrow \sum_{t=1}^T t x$$

, n a or ast r w. o ssu s

$$t x - T x \leftarrow \sum_{t=1}^T t x$$

or a  $\{$   $\}$  to b t . n t . c ato o ca or east . I  $t x$  m  $T x$   
 t. n w . av t. c ato o ca or east . A or ast r s n t . ar na y ca brat .

$T x$   $T x$   
 For a pra t a purpos s s n t an w . av no a ss to t .  $t x$  s H n  
 t s , ut to ass s n t m arf na a brat on I  $T$   $T$   $\approx$  w r s s  
 m tr t. n w an ta to b arf nouf to va uat m arf na a brat on o t s  
 n w an us t. H n f r , stan o ar an output

$$T c \leftarrow$$

Gndddd.6







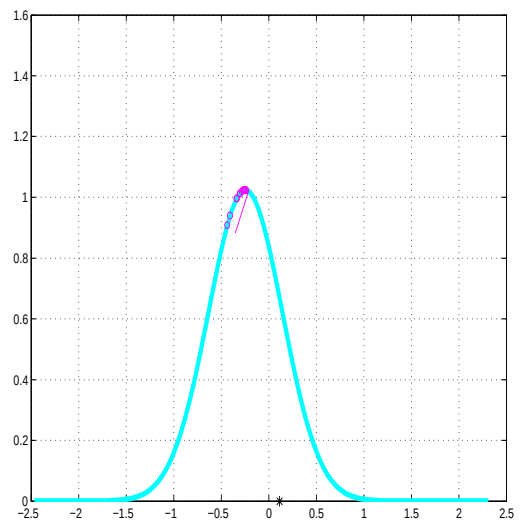
an<sup>1</sup> t<sup>1</sup> para<sup>1</sup> t<sup>1</sup> rs<sup>1</sup> c<sup>1</sup> an<sup>1</sup> c<sup>1</sup>

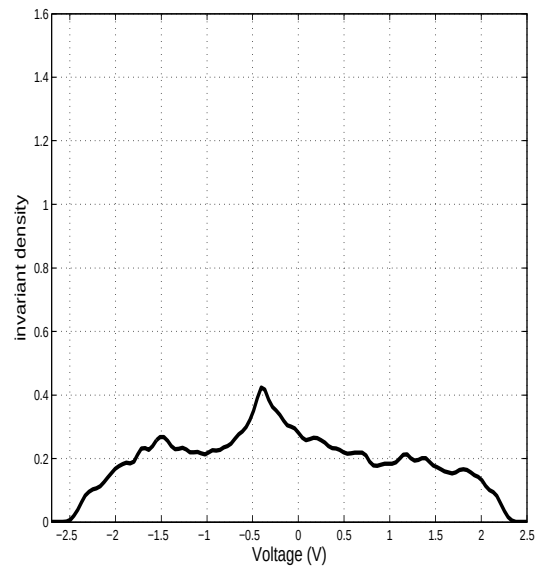
w. n. Dispt t. sr. uton se mxtur or asts m a st b ss s arp t. an  
 mato of nt s ns o ntrop. A sra F t orwar app at on o t. Ku ba  
 L b r an J ns ns n qua t s a s to t. r at ons

$$\{ (t) \} - c \leq \{ (t) \} \leq \{ (t) \} - \{ (t) \} c$$

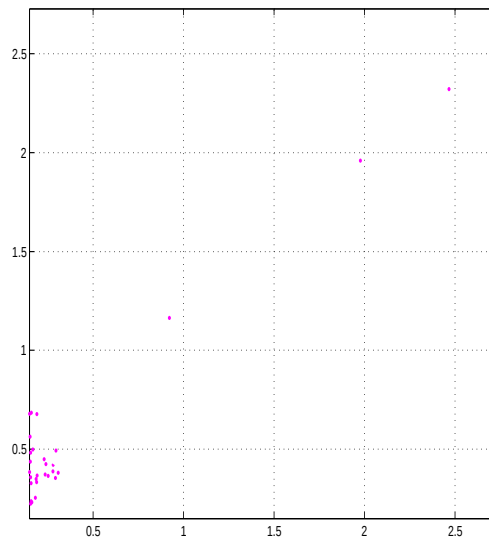
w. r  $-\int x$  of  $x$  an  $-\int x$  of  $x$  x ar t. ntrop  
 an ross ntrop r sp tv r or t. n ssar an su nt on t ons or  
 $\{ (t) \} c$  o. o. ar t. at  $\{ (t) \} c$  an  $\{ (t) \} c$  r sp  
 tv n v r t. mato of s s arp r t. an t. mxtur or ast t s ou b  
 ssu ast. or ast ns a o t. mxtur  
 It s not obv ous w at t. to t. mxtur s on a braton p t t. at n u n f  
 mato of m prov st. KI stan re m t. a or asts v r ss t. mxtur  
 para t r t at m s st. of ar t. s or s t. quat on

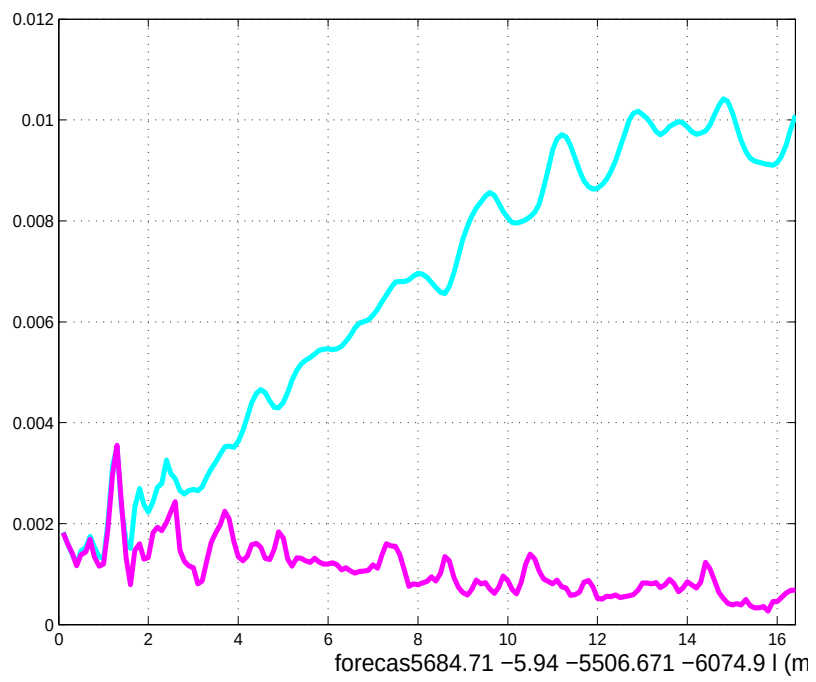
□

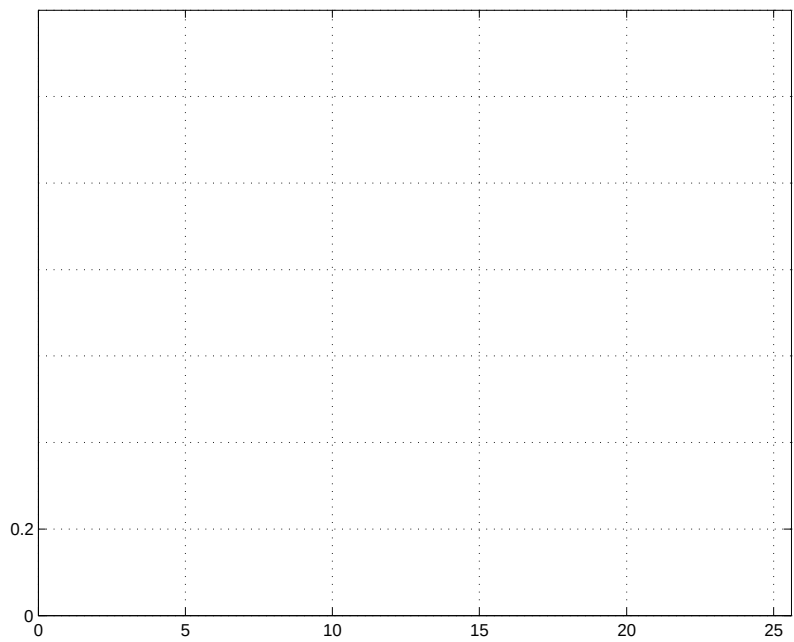


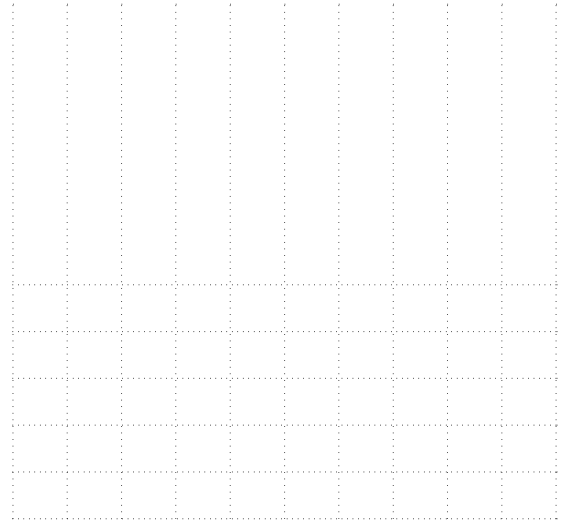
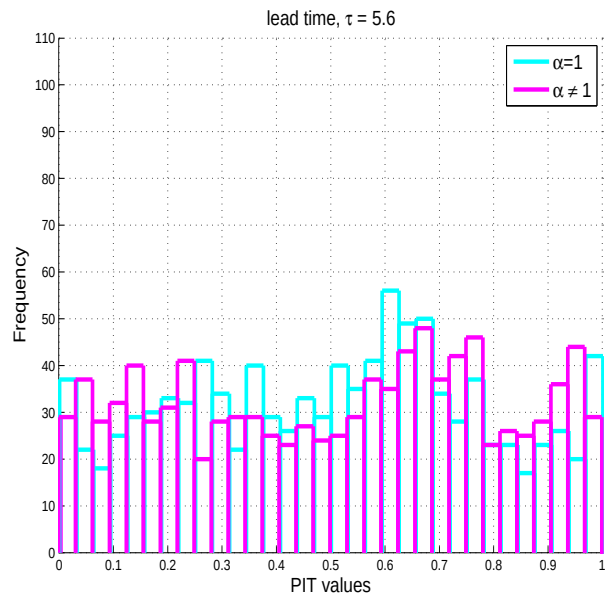


F  Fur *Graph of the climatology of the circuit estimated from data. Its entropy is 2.15, which is greater than the entropies shown in figure 1.*

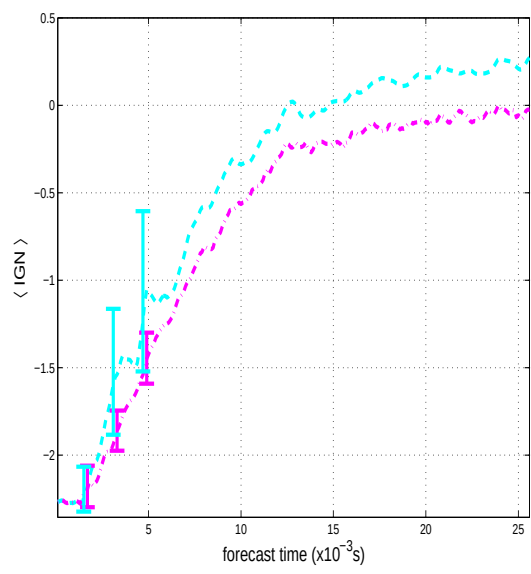


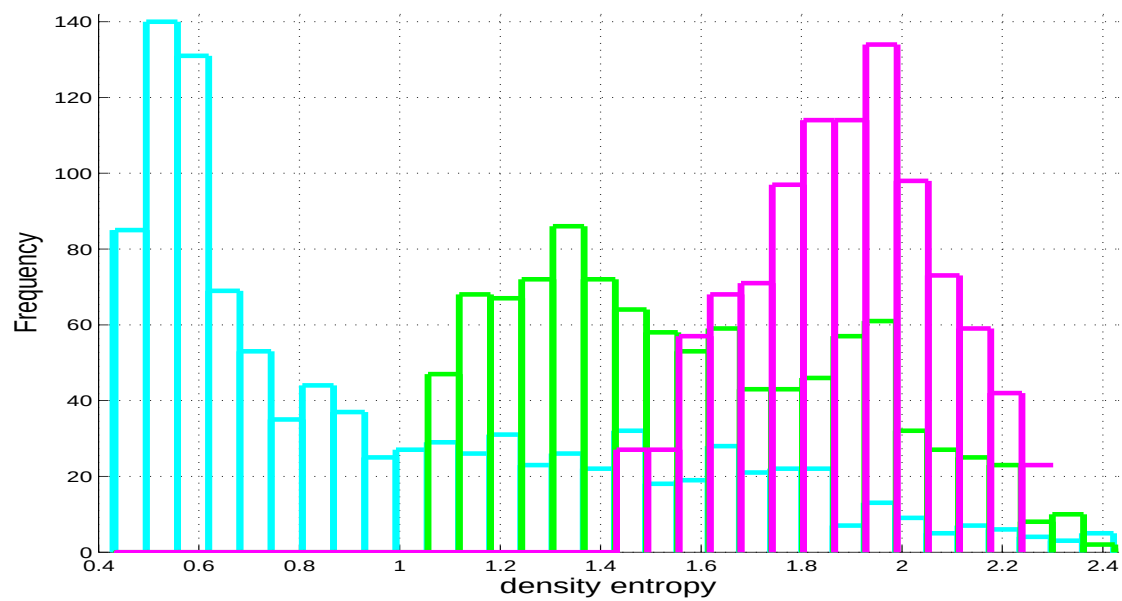














annon CE tor Co un cat on n t, pr s ac ro no s vo u ro

$t$   $x$   $t^{-1}$   $_1\{ \text{ }_1 x \}$  It  $t$  n o<sup>⤴</sup>ws  $t$  at

$$\subseteq \sum_{t=1}^T t \left[ t \left\{ \text{ }_s^{-1} x_s \right\} \right]$$



 t. r m a b at ast anot.  r p □ an' p su.   t at

\_\_\_\_\_