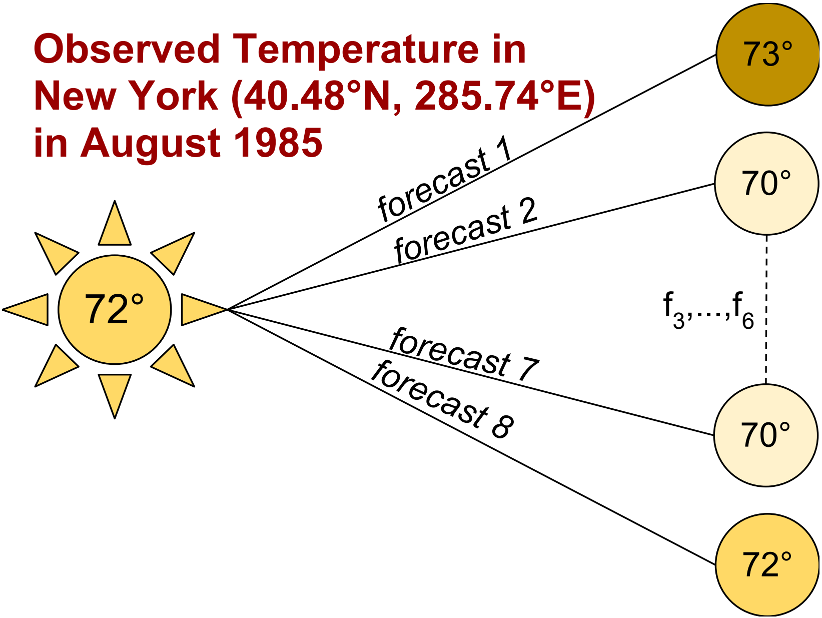
Evaluating Ensemble Seasonal Forecasts Using Information Metrics

Hannah Aizenman, Nir Krakauer,
Michael Grossberg, Irina Gladkova

NOAA-CREST at CCNY

haizenm00@ccny.cuny.edu

How do we evaluate the accuracy of ensembles?

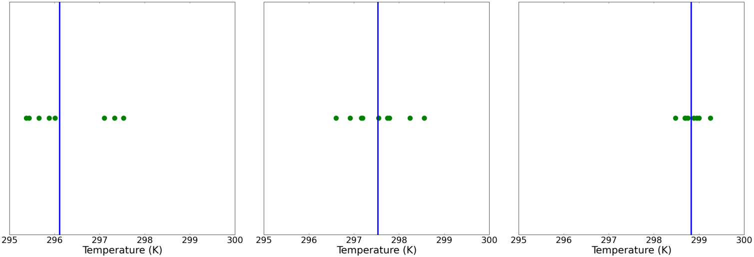


Case Study

NCEP Forecast System Version 2 (CFSv2)

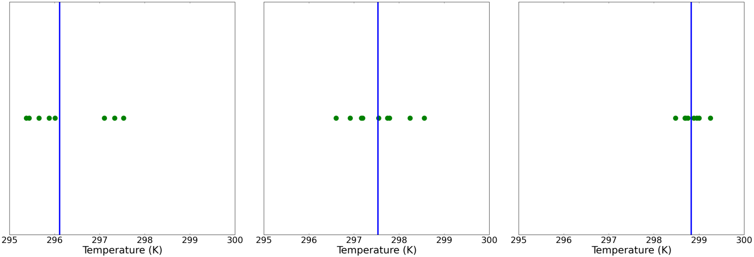
- Monthly mean 2-m temperature (K)
- Reanalysis observation: x_o
- Ensemble members 1-9, 0.5 month lag forecasts: $x_1, x_2, ..., x_k$
- 01/1979- 12/2009
- $1^\circ x 1^\circ$ spatial grid

Mean Error: $\mu\left(x_{1\ldots k}\right) - x_o$ at (t,lat,lon)



All spreads have an error score of 0

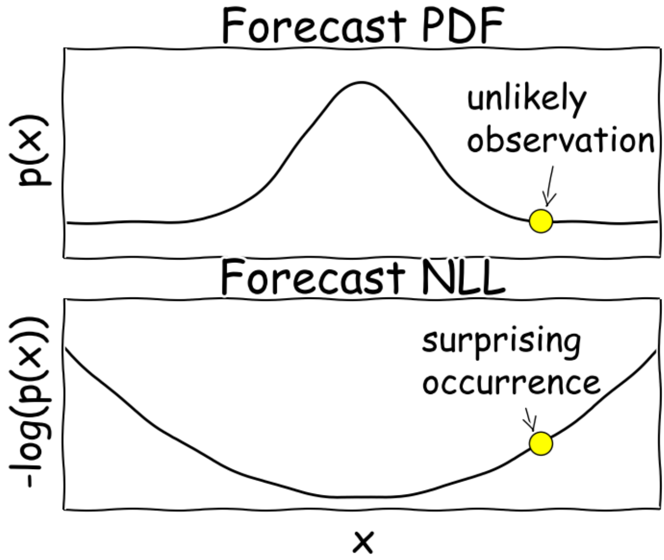
Mean Error: $\mu\left(x_{1\ldots k}\right) - x_o$ at (t,lat,lon)



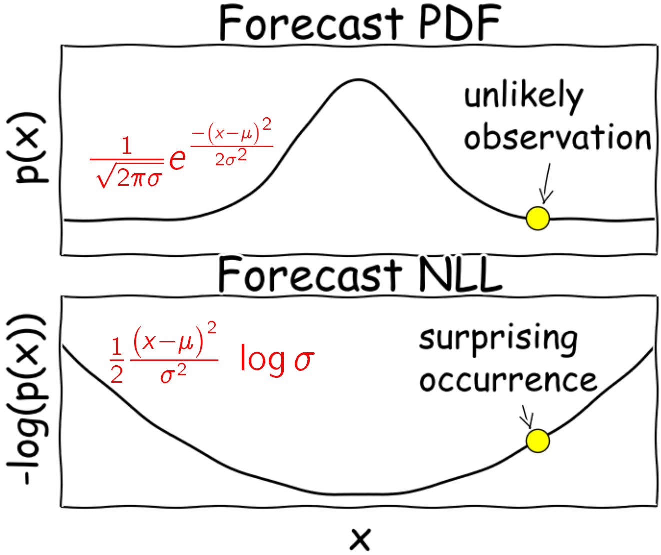
All spreads have an error score of 0

- skill is based on the error of the ensemble mean
- MSE $= \frac{\sum_{t=1} E_t^2}{N}$
- SS $= 1 - \frac{MSE_{forecast}}{MSE_{ref}}$

Negative Log Likelihood

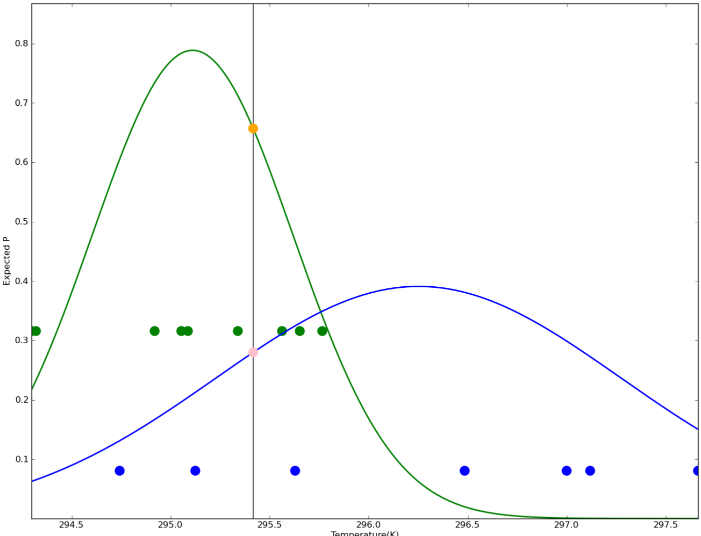


Negative Log Likelihood

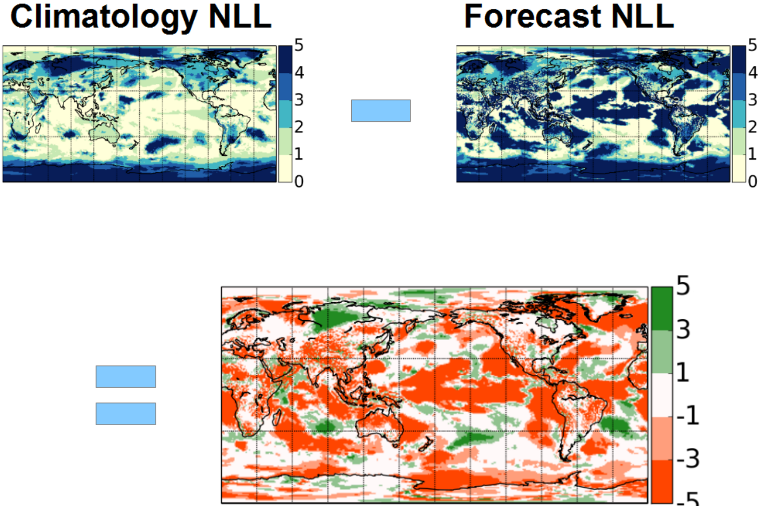


Converting from Observation to PDF

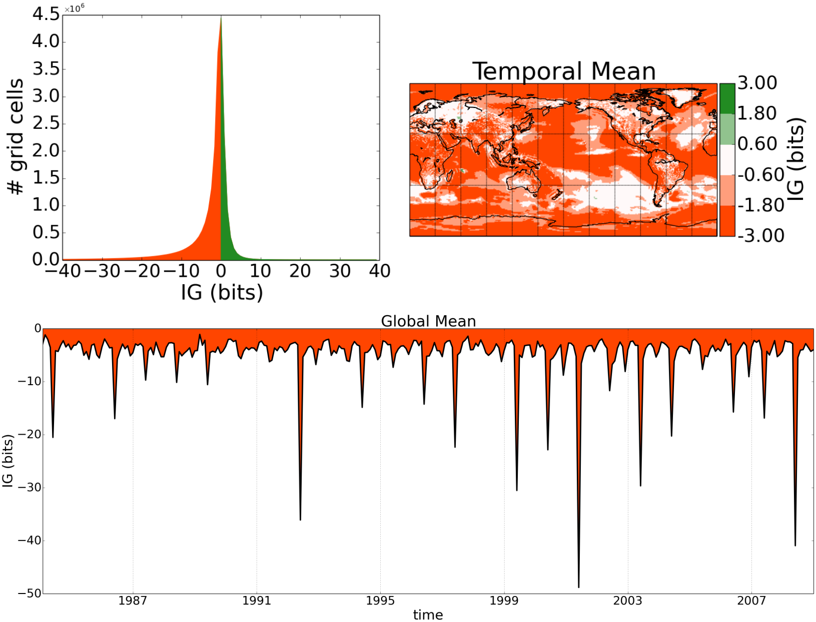
- p^f estimates $f_{i,t,lat,lon}$, i $1, 2, \ldots 9$
- p^c estimates $x_{j,lat,lon}, j$ $0, 1, \ldots, t-1,$



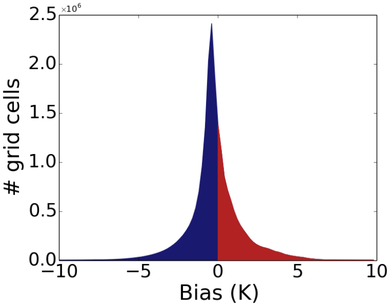
Information Gain: $-\log_2\left(p^f\left(x_o\right)\right) - -\log_2\left(p^c\left(x_o\right)\right)$



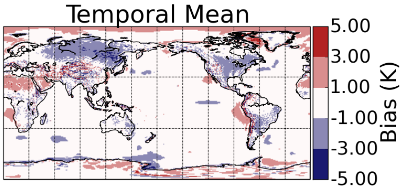
Aggregate IG: Normal PDF from Ensemble Members



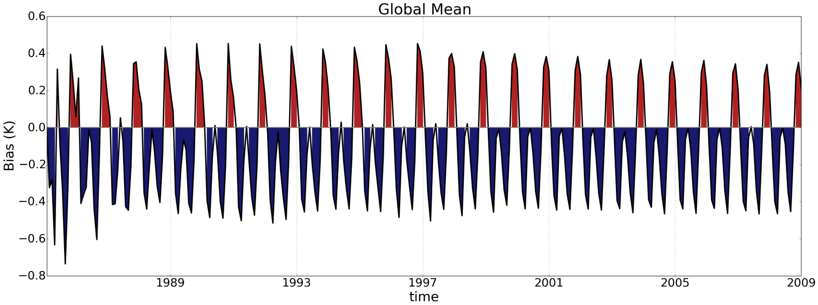
Bias



×10^6
2.5
2.0
1.5
1.0
0.5
0.0
grid cells
−10
−5
0
5
10
Bias (K)



Temporal Mean
5.00
3.00
1.00
-1.00
-3.00
-5.00
Bias (K)



Global Mean
0.6
0.4
0.2
0.0
−0.2
−0.4
−0.6
−0.8
Bias (K)
1989
1993
1997
2001
2005
2009
time

Bias Correction Estimated from Previous Hindcasts

$$bias_{i,t_k,la,ln} = \sum_{\mathrm{k=m}/y_0}^{k<S} \frac{f_{i,t_k,la,ln} - obs_{t_k,la,ln}}{S-1}$$

s.t. k increments in steps of 12

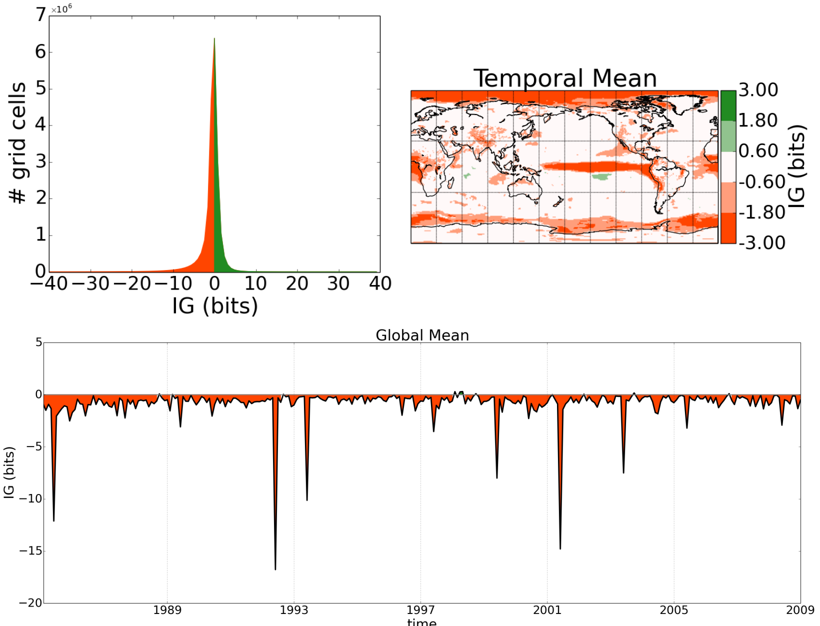
Bias Correction Estimated from Previous Hindcasts

$$bias_{i,t_k,la,ln} = \sum_{\mathrm{k=m}/y_0}^{k<S} \frac{f_{i,t_k,la,ln} - obs_{t_k,la,ln}}{S-1}$$

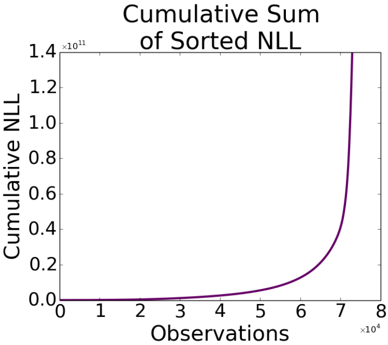
s.t. k increments in steps of 12

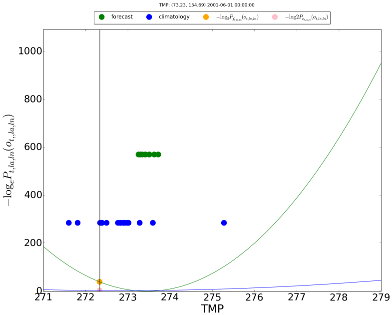
- i $= f_i \in forecasts \left[1, 9\right]$
- k $= month \in \left[1, 12\right], year \in \left[y_0, ..., y_f\right]$
- t, la, ln = time, latitude, longitude
- S = current observation's time-step

Improve Forecast by Subtracting the Bias

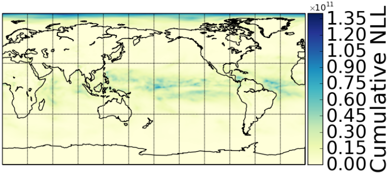


A Few Grid Points (e.g. in Arctic) have Extreme NLLs

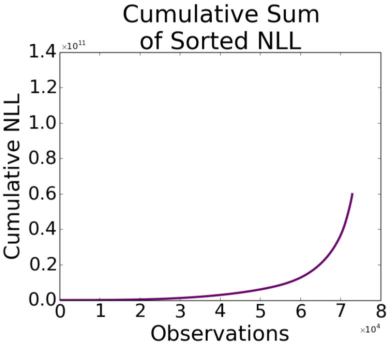


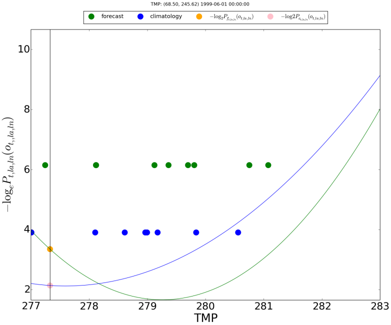


High certainty in wrong prediction = spike in negative IG



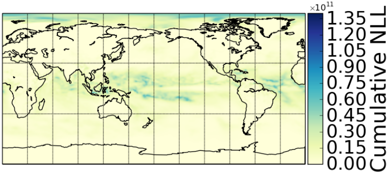
Most Grid Points have Normal NLLs



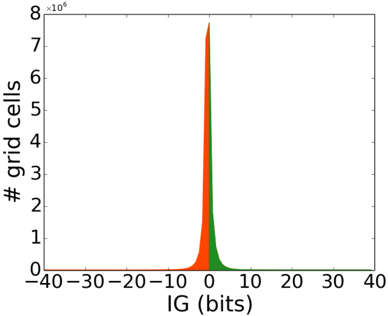


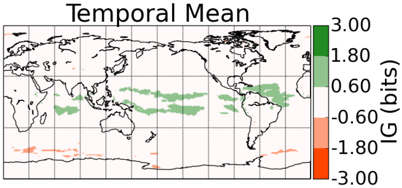
Low certainty

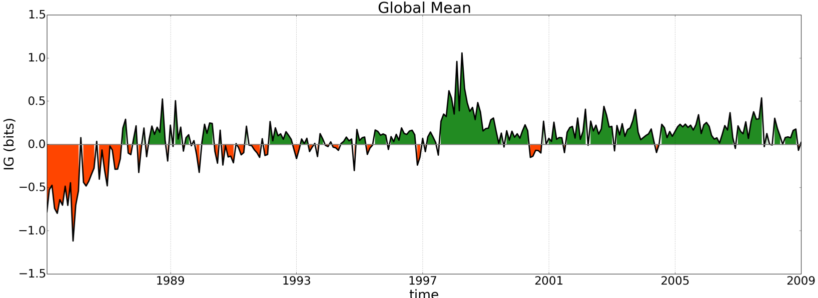
= no spike



IG Improves by Using Climatology σ Instead of Forecast





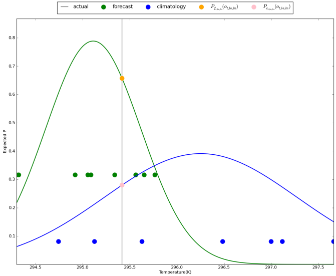


Global Mean IG

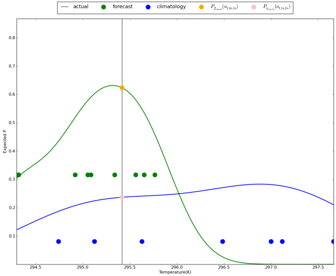
	TMP		PRATE	
	mean	**median**	**mean**	**median**
Forecast	-4.9134	-0.4207	-58.0248	-0.1401
Bias Corrected	-0.8877	0.0964	-21.8516	-0.1342
Climatology SD	0.0558	0.0147	-2.3272	-0.0276

- IG improves after bias correction and substituting climatology SD for ensemble SD
- T reforecasts now average slightly better than climatology, P still worse

How does a normal PDF fit to an ensemble differ from a kernel density estimator?

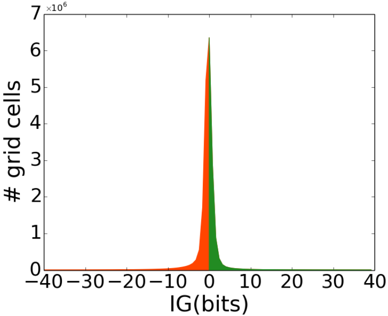


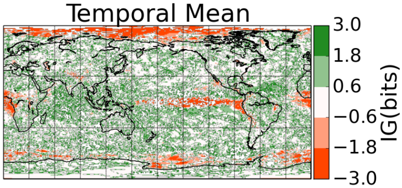
(a) Normal Density Estimation

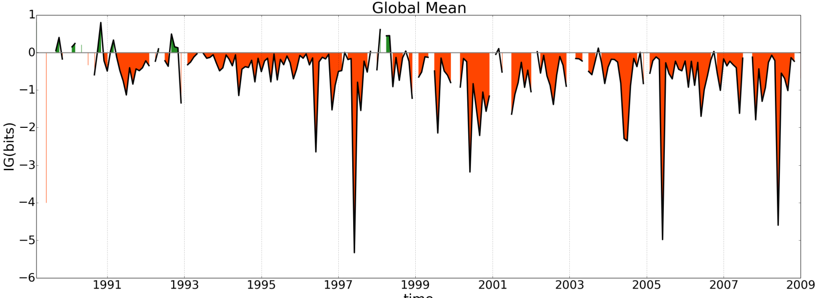


(b) Kernel Density Estimation

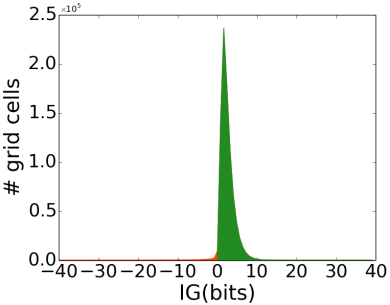
Aggregate IG: KDE Approximation

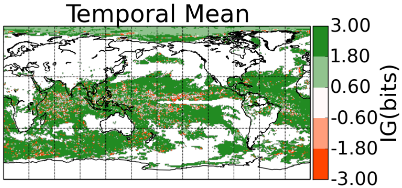


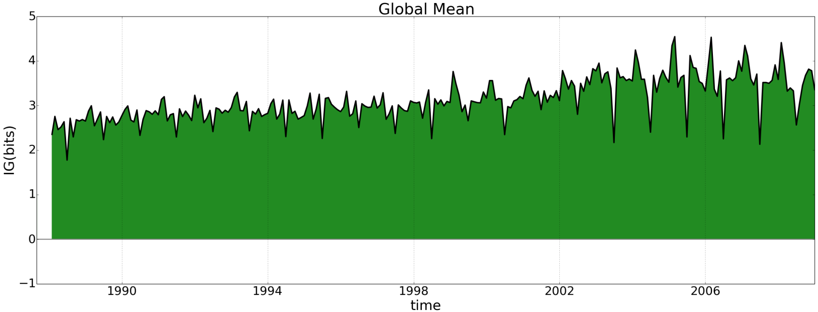




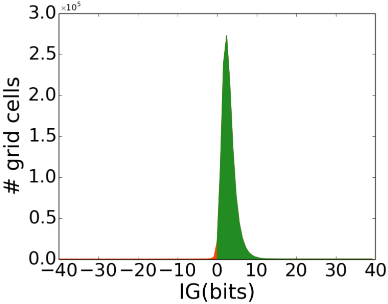
Aggregate IG: Climatology KDE vs. Normal

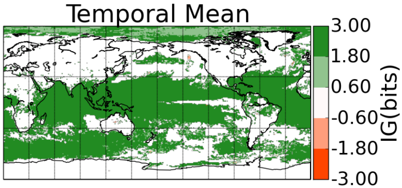


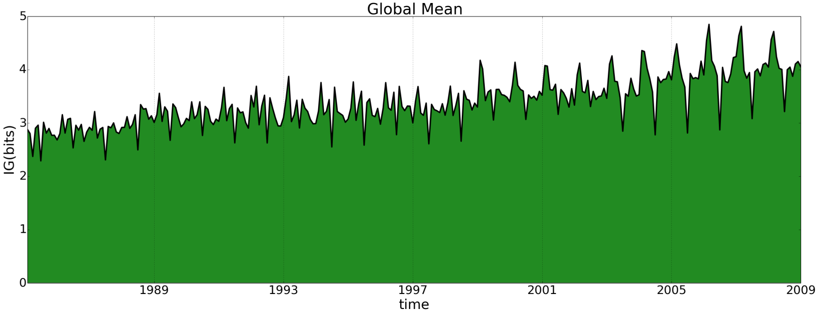




Aggregate IG: Forecast KDE vs. Normal







Mean Global IG for KDE vs. Normal Fit

	TMP	
	mean	**median**
Forecast	-4.9134	-0.4207
Bias Corrected	-0.8877	0.0964
Climatology SD	0.0558	0.0147
KDE	0.3587	0.1326

- IG improves even more using KDE
- KDE is sensitive to spread-spikes

Conclusion

- Gaussian estimation vs. kernel density estimation for converting ensemble forecasts to a PDF
- Treat ensemble as probabilistic predictions
- IG measures probabilistic predictions well
- Forecast can be improved by using IG to diagnose problems (e.g. mean and SD offsets)

Future Work

- Fit other distributions to climatology and ensemble forecasts
- Include trend estimates in probabilistic forecasts
- Develop predictive model that incorporates climatology and multiple model forecasts based on skill at a given region and season

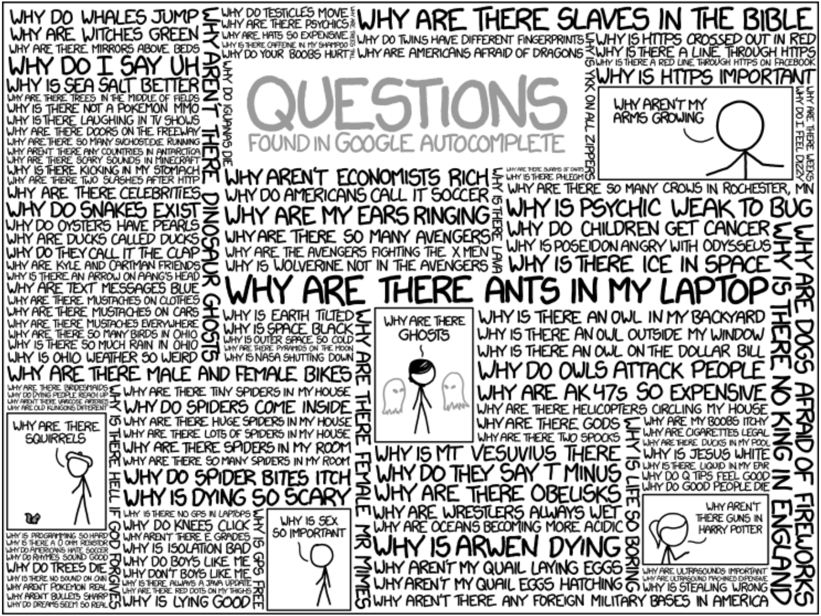
Acknowledgement

The authors gratefully acknowledge support from NOAA under Grants NA11SEC4810004 and NA12OAR4310084. All statements made are the views of the authors and not the opinions of the funding agency or the US government.

References

- Nir Y. Krakauer, Nir, Michael D. Grossberg, Irina Gladkova, and Hannah Aizenman, Information Content of Seasonal Forecasts in a Changing Climate, Advances in Meteorology, vol. 2013, Article ID 480210, 12 pages, 2013. doi: 10.1155/2013/480210
- Suranjana Saha, and Coauthors, "The NCEP Climate Forecast System Version 2", Journal of Climate (early online release.),2013: doi: 10.1175/JCLI-D-12-00823.1

Questions?


WHY DO WHALES JUMP
WHY ARE WITCHES GREEN
WHY ARE THERE MIRRORS ABOVE BEDS
WHY DO I SAY UH
WHY IS SEA SALT BETTER
WHY ARE THERE TREES IN THE MIDDLE OF FIELDS
WHY IS THERE NOT A POKEMON MMO
WHY IS THERE LAUGHING IN TV SHOWS
WHY ARE THERE DOORS ON THE FREEWAY
WHY ARE THERE SO MANY SVCHOST.EXE RUNNING
WHY AREN'T THERE ANY COUNTRIES IN ANTARCTICA
WHY ARE THERE SCARY SOUNDS IN MINECRAFT
WHY IS THERE KICKING IN MY STOMACH
WHY ARE THERE TWO SLASHES AFTER HTTP
WHY ARE THERE CELEBRITIES
WHY DO SNAKES EXIST
WHY DO OYSTERS HAVE PEARLS
WHY ARE DUCKS CALLED DUCKS
WHY DO THEY CALL IT THE CLAP
WHY ARE KYLE AND CARTMAN FRIENDS
WHY IS THERE AN ARROW ON AANG'S HEAD
WHY ARE TEXT MESSAGES BLUE
WHY ARE THERE MUSTACHES ON CLOTHES
WHY ARE THERE MUSTACHES ON CARS
WHY ARE THERE MUSTACHES EVERYWHERE
WHY ARE THERE SO MANY BIRDS IN OHIO
WHY IS THERE SO MUCH RAIN IN OHIO
WHY IS OHIO WEATHER SO WEIRD
WHY AREN'T THERE DINOSAUR GHOSTS
WHY DO IGUANAS DIE
WHY DO TESTICLES MOVE
WHY ARE THERE PSYCHICS
WHY ARE HATS SO EXPENSIVE
WHY IS THERE CAFFEINE IN MY SHAMPOO
WHY DO YOUR BOOBS HURT
WHY ARE THERE SLAVES IN THE BIBLE
WHY DO TWINS HAVE DIFFERENT FINGERPRINTS
WHY ARE AMERICANS AFRAID OF DRAGONS
QUESTIONS
FOUND IN GOOGLE AUTOCOMPLETE
WHY IS YKK ON ALL ZIPPERS
WHY IS HTTPS CROSSED OUT IN RED
WHY IS THERE A LINE THROUGH HTTPS
WHY IS THERE A RED LINE THROUGH HTTPS ON FACEBOOK
WHY IS HTTPS IMPORTANT
WHY AREN'T MY ARMS GROWING
WHY ARE THERE WEEKS
WHY DO I FEEL DIZZY
WHY AREN'T ECONOMISTS RICH
WHY DO AMERICANS CALL IT SOCCER
WHY ARE MY EARS RINGING
WHY ARE THERE SO MANY AVENGERS
WHY ARE THE AVENGERS FIGHTING THE X MEN
WHY IS WOLVERINE NOT IN THE AVENGERS
WHY IS THERE LAMA
WHY IS THERE PHLEGM
WHY ARE THERE SO MANY CROWS IN ROCHESTER, MN
WHY IS PSYCHIC WEAK TO BUG
WHY DO CHILDREN GET CANCER
WHY IS POSEIDON ANGRY WITH ODYSSEUS
WHY IS THERE ICE IN SPACE
WHY ARE THERE ANTS IN MY LAPTOP
WHY IS THERE NO KING IN ENGLAND
WHY ARE DOGS AFRAID OF FIREWORKS
WHY IS EARTH TILTED
WHY IS SPACE BLACK
WHY IS OUTER SPACE SO COLD
WHY ARE THERE PYRAMIDS ON THE MOON
WHY IS NASA SHUTTING DOWN
WHY ARE THERE FEMALE MR MIMES
WHY ARE THERE GHOSTS
WHY IS THERE AN OWL IN MY BACKYARD
WHY IS THERE AN OWL OUTSIDE MY WINDOW
WHY IS THERE AN OWL ON THE DOLLAR BILL
WHY DO OWLS ATTACK PEOPLE
WHY ARE AK 47s SO EXPENSIVE
WHY ARE THERE HELICOPTERS CIRCLING MY HOUSE
WHY ARE THERE GODS
WHY ARE THERE TWO SPOCKS
WHY ARE THERE MALE AND FEMALE BIKES
WHY ARE THERE BRIDESMAIDS
WHY DO DYING PEOPLE REACH UP
WHY AREN'T THERE VARICOSE ARTERIES
WHY ARE OLD KLINGONS DIFFERENT
WHY ARE THERE SQUIRRELS
WHY IS THERE HELL IF GOD FORGIVES
WHY ARE THERE TINY SPIDERS IN MY HOUSE
WHY DO SPIDERS COME INSIDE
WHY ARE THERE HUGE SPIDERS IN MY HOUSE
WHY ARE THERE LOTS OF SPIDERS IN MY HOUSE
WHY ARE THERE SPIDERS IN MY ROOM
WHY ARE THERE SO MANY SPIDERS IN MY ROOM
WHY DO SPIDER BITES ITCH
WHY IS DYING SO SCARY
WHY IS THERE NO GPS IN LAPTOPS
WHY DO KNEES CLICK
WHY AREN'T THERE E GRADES
WHY IS ISOLATION BAD
WHY DO BOYS LIKE ME
WHY DON'T BOYS LIKE ME
WHY IS THERE ALWAYS A JAVA UPDATE
WHY ARE THERE RED DOTS ON MY THIGHS
WHY IS LYING GOOD
WHY IS GPS FREE
WHY IS SEX SO IMPORTANT
WHY IS PROGRAMMING SO HARD
WHY IS THERE A 0 OHM RESISTOR
WHY DO AMERICANS HATE SOCCER
WHY DO RHYMES SOUND GOOD
WHY DO TREES DIE
WHY IS THERE NO SOUND ON CNN
WHY AREN'T POKEMON REAL
WHY AREN'T BULLETS SHARP
WHY DO DREAMS SEEM SO REAL
WHY IS MT VESUVIUS THERE
WHY DO THEY SAY T MINUS
WHY ARE THERE OBELISKS
WHY ARE WRESTLERS ALWAYS WET
WHY ARE OCEANS BECOMING MORE ACIDIC
WHY IS ARWEN DYING
WHY AREN'T MY QUAIL LAYING EGGS
WHY AREN'T MY QUAIL EGGS HATCHING
WHY AREN'T THERE ANY FOREIGN MILITARY BASES IN AMERICA
WHY IS LIFE SO BORING
WHY ARE MY BOOBS ITCHY
WHY ARE CIGARETTES LEGAL
WHY ARE THERE DUCKS IN MY POOL
WHY IS JESUS WHITE
WHY IS THERE LIQUID IN MY EAR
WHY DO Q TIPS FEEL GOOD
WHY DO GOOD PEOPLE DIE
WHY AREN'T THERE GUNS IN HARRY POTTER
WHY ARE ULTRASOUNDS IMPORTANT
WHY ARE ULTRASOUND MACHINES EXPENSIVE
WHY IS STEALING WRONG