Information theory and predictability

R. Kleeman

1 Center for Atmosphere Ocean Science
Courant Institute of Mathematical Sciences, New York University

Uncertainty Quantification Summer School, USC August 2016
Talk Outline

1. Predictability Definition
   - Basic Concepts
   - Universal information theoretic measures

2. Predictability and disequilibrium

3. Predictability variation
   - Introduction
   - Chaos
   - Stochastic Models
   - Turbulence
   - Mechanisms for predictability variation

4. References
1. Predictability Definition
   - Basic Concepts
     - Universal information theoretic measures

2. Predictability and disequilibrium

3. Predictability variation
   - Introduction
   - Chaos
   - Stochastic Models
   - Turbulence
   - Mechanisms for predictability variation

4. References
Perfect predictability

- Uncertainty in predictions occurs for two major reasons.
- An inaccurate specification of the initial or boundary conditions.
- An inaccurate dynamical model.
- We consider only the first case here as this is amenable to a comprehensive theoretical exploration.
Perfect predictability

- Uncertainty in predictions occurs for two major reasons.
- An inaccurate specification of the initial or boundary conditions.
- An inaccurate dynamical model.
- We consider only the first case here as this is amenable to a comprehensive theoretical exploration.
Perfect predictability

- Uncertainty in predictions occurs for two major reasons.
  - An inaccurate specification of the initial or boundary conditions.
  - An inaccurate dynamical model.
- We consider only the first case here as this is amenable to a comprehensive theoretical exploration.
Perfect predictability

- Uncertainty in predictions occurs for two major reasons.
- An inaccurate specification of the initial or boundary conditions.
- An inaccurate dynamical model.
- We consider only the first case here as this is amenable to a comprehensive theoretical exploration.
Simplest idea involves uncertainty. The smaller this is made the greater the predictability achieved.

A more useful idea takes into fundamental account that there is usually prior knowledge of a prediction random variable. This prior usually derives from historical or climatological observations of the system.

Prediction can be viewed then as a modification of a prior to a posterior random variable. This perspective underlies learning theory as well as Bayesian statistics.

This modification often involves more than just uncertainty reduction. Consider the following examples.....
Simplest idea involves uncertainty. The smaller this is made the greater the predictability achieved.

A more useful idea takes into fundamental account that there is usually prior knowledge of a prediction random variable. This prior usually derives from historical or climatological observations of the system.

Prediction can be viewed then as a modification of a prior to a posterior random variable. This perspective underlies learning theory as well as Bayesian statistics.

This modification often involves more than just uncertainty reduction. Consider the following examples.....
Predictability Definitions

- Simplest idea involves uncertainty. The smaller this is made the greater the predictability achieved.
- A more useful idea takes into fundamental account that there is usually prior knowledge of a prediction random variable. This prior usually derives from historical or climatological observations of the system.
- Prediction can be viewed then as a modification of a prior to a posterior random variable. This perspective underlies learning theory as well as Bayesian statistics.
- This modification often involves more than just uncertainty reduction. Consider the following examples.....
Predictability Definitions

- Simplest idea involves uncertainty. The smaller this is made the greater the predictability achieved.

- A more useful idea takes into fundamental account that there is usually prior knowledge of a prediction random variable. This prior usually derives from historical or climatological observations of the system.

- Prediction can be viewed then as a modification of a prior to a posterior random variable. This perspective underlies learning theory as well as Bayesian statistics.

- This modification often involves more than just uncertainty reduction. Consider the following examples.....
Suppose that climatologically the maximum temperature today is 34°C with a standard deviation of 4°C.

Imagine two different (hypothetical) statistical predictions for today. First is 34°C with standard deviation 0.5°C. Second is 20°C with a standard deviation of 4°C.

The usefulness of the first prediction derives from the reduction in uncertainty in the posterior from that of the prior. The usefulness of the second prediction derives not from any reduction in uncertainty but in the large difference of the prediction from “normal” or technically because the posterior mean differs markedly from the prior mean.
Simple Examples

- Suppose that climatologically the maximum temperature today is 34°C with a standard deviation of 4°C.

- Imagine two different (hypothetical) statistical predictions for today. First is 34°C with standard deviation 0.5°C. Second is 20°C with a standard deviation of 4°C.

- The usefulness of the first prediction derives from the reduction in uncertainty in the posterior from that of the prior. The usefulness of the second prediction derives not from any reduction in uncertainty but in the large difference of the prediction from “normal” or technically because the posterior mean differs markedly from the prior mean.
Simple Examples

- Suppose that climatologically the maximum temperature today is 34°C with a standard deviation of 4°C.

- Imagine two different (hypothetical) statistical predictions for today. First is 34°C with standard deviation 0.5°C. Second is 20°C with a standard deviation of 4°C.

- The usefulness of the first prediction derives from the reduction in uncertainty in the posterior from that of the prior. The usefulness of the second prediction derives not from any reduction in uncertainty but in the large difference of the prediction from “normal” or technically because the posterior mean differs markedly from the prior mean.
These different notions of predictability have traditionally been measured using a large variety of metrics (RMS error, anomaly correlation etc). They both derive however from a shift between a prior and posterior probability distribution.

In the first case the variances of the distributions differ while the means are identical. In the second case the means differ but the variances are identical. Is there a universal way of measuring the shift?
These different notions of predictability have traditionally been measured using a large variety of metrics (RMS error, anomaly correlation etc). They both derive however from a shift between a prior and posterior probability distribution.

In the first case the variances of the distributions differ while the means are identical. In the second case the means differ but the variances are identical. Is there a universal way of measuring the shift?
Predictability Definition

Basic Concepts

Universal information theoretic measures

Predictability and disequilibrium

Predictability variation

Introduction

Chaos

Stochastic Models

Turbulence

Mechanisms for predictability variation

References
In information theory and in statistical mechanics the uncertainty associated with a probability distribution $p$ is referred to as the (absolute) Entropy. This is the basic functional of both theories. There is a distinction between discrete and continuous random variables in the definition. We focus here on the latter.

**Absolute entropy**

$$H(p) \equiv - \int p \log(p) \, dx \geq 0$$

In the Gibbs approach to statistical mechanics this is the central definition of entropy which is the organizing principle for the entire field. Historically the statistical mechanics application came first and the information theoretical application showed the widespread applicability of the idea of uncertainty.
Information theoretic measures

- In information theory and in statistical mechanics the uncertainty associated with a probability distribution $p$ is referred to as the (absolute) Entropy. This is the basic functional of both theories. There is a distinction between discrete and continuous random variables in the definition. We focus here on the latter.

**Absolute entropy**

$$H(p) \equiv - \int p \log(p) \, dx \geq 0$$

- In the Gibbs approach to statistical mechanics this is the central definition of entropy which is the organizing principle for the entire field. Historically the statistical mechanics application came first and the information theoretical application showed the widespread applicability of the idea of uncertainty.
In information theory and in statistical mechanics the uncertainty associated with a probability distribution \( p \) is referred to as the (absolute) **Entropy**. This is the basic functional of both theories. There is a distinction between discrete and continuous random variables in the definition. We focus here on the latter.

**Absolute entropy**

\[
H(p) \equiv -\int p \log(p) \, dx \geq 0
\]

In the Gibbs approach to statistical mechanics this is the central definition of entropy which is the organizing principle for the entire field. Historically the statistical mechanics application came first and the information theoretical application showed the widespread applicability of the idea of uncertainty.
In learning theory the shift between prior and posterior is a measure of how much knowledge has been acquired. It is traditionally measured using the relative entropy of the two distributions. Let the posterior distribution be $p$ and the prior be $q$.

\[
D(p, q) \equiv \int p \log \left( \frac{p}{q} \right) \, dx \geq 0
\]

This is non-negative and measures effectively the “distance” between $p$ and $q$. It is invariant under general non-linear variable changes and is a non-increasing function of time for Markov processes. It is also not symmetric between prior and posterior as one might expect given that learning is directed.
In learning theory the shift between prior and posterior is a measure of how much knowledge has been acquired. It is traditionally measured using the relative entropy of the two distributions. Let the posterior distribution be $p$ and the prior be $q$.

Relative entropy

\[ D(p, q) \equiv \int p \log \left( \frac{p}{q} \right) \, dx \geq 0 \]

This is non-negative and measures effectively the “distance” between $p$ and $q$. It is invariant under general non-linear variable changes and is a non increasing function of time for Markov processes. It is also not symmetric between prior and posterior as one might expect given that learning is directed.
In learning theory the shift between prior and posterior is a measure of how much knowledge has been acquired. It is traditionally measured using the relative entropy of the two distributions. Let the posterior distribution be $p$ and the prior be $q$.

Relative entropy

$$D(p, q) \equiv \int p \log \left( \frac{p}{q} \right) dx \geq 0$$

This is non-negative and measures effectively the “distance” between $p$ and $q$. It is invariant under general non-linear variable changes and is a non increasing function of time for Markov processes. It is also not symmetric between prior and posterior as one might expect given that learning is directed.
Information theoretic measures

Note that uncertainty changes alone due to prediction can be measured using an entropy difference.

Entropy Difference

\[ H(p) - H(q) \equiv \int q \log(q) \, dx - \int p \log(p) \, dx \]

Note that this measure does not satisfy the list of nice mathematical properties noted above for the relative entropy. It is, for example, only invariant under linear transformations of dynamical variables. It is however useful in our discussion as it measures the simplest kind of predictability discussed earlier.
Information theoretic measures

- Note that uncertainty changes alone due to prediction can be measured using an entropy difference.

**Entropy Difference**

\[
H(p) - H(q) \equiv \int q \log(q) \, dx - \int p \log(p) \, dx
\]

- Note that this measure does not satisfy the list of nice mathematical properties noted above for the relative entropy. It is, for example, only invariant under linear transformations of dynamical variables. It is however useful in our discussion as it measures the simplest kind of predictability discussed earlier.
Information theoretic measures

- Note that uncertainty changes alone due to prediction can be measured using an entropy difference.

**Entropy Difference**

\[ H(p) - H(q) \equiv \int q \log(q) \, dx - \int p \log(p) \, dx \]

- Note that this measure does not satisfy the list of nice mathematical properties noted above for the relative entropy. It is, for example, only invariant under linear transformations of dynamical variables. It is however useful in our discussion as it measures the simplest kind of predictability discussed earlier.
We can also define a measure of independence of two random variables called mutual information.

**Mutual Information**

\[ I(X; Y) \equiv \int \int p(x, y) \log \left( \frac{p(x, y)}{p(x)p(y)} \right) dxdy = D(p(x, y) \| p(x)p(y)) \]

\[ = H(X) - H(X|Y) \]

\[ H(X|Y) \equiv \int \int p(x, y) \log p(x|y) dxdy \]

“Distance” between the joint distribution and that which would apply if the variables were independent or the uncertainty in X minus the uncertainty in X due to knowledge of Y.
We can also define a measure of independence of two random variables called mutual information.

**Mutual Information**

\[
I(X; Y) \equiv \int \int p(x, y) \log \left( \frac{p(x, y)}{p(x)p(y)} \right) dxdy = D(p(x, y)\|p(x)p(y))
\]

\[
= H(X) - H(X|Y)
\]

\[
H(X|Y) \equiv \int \int p(x, y) \log p(x|y) dxdy
\]

“Distance” between the joint distribution and that which would apply if the variables were independent or the uncertainty in \(X\) minus the uncertainty in \(X\) due to knowledge of \(Y\).
We can also define a measure of independence of two random variables called mutual information.

\[
I(X; Y) \equiv \int \int p(x, y) \log \left( \frac{p(x, y)}{p(x)p(y)} \right) \, dx\, dy = D(p(x, y) \| p(x)p(y))
\]

\[
= H(X) - H(X|Y)
\]

\[
H(X|Y) \equiv \int \int p(x, y) \log p(x|y) \, dx\, dy
\]

“Distance” between the joint distribution and that which would apply if the variables were independent or the uncertainty in \(X\) minus the uncertainty in \(X\) due to knowledge of \(Y\).
Gaussian distributions

- The relative entropy can be computed analytically in this instance. Let the prior random variable be denoted by $y$ and the posterior by $x$ and variances by $\sigma$.

\[
D(p \| q) = \frac{1}{2} \left[ \log \left( \frac{\det(\sigma_y)}{\det(\sigma_x)} \right) + tr \left( \sigma_x (\sigma_y)^{-1} \right) - n \right] + 
\frac{1}{2} (\bar{x} - \bar{y})^t \sigma_y^{-1} (\bar{x} - \bar{y})
\]

- The first term on the RHS here is the entropy difference and measures reduction in uncertainty. The third term is the shift in means. The second term is usually not important. We call the first line the dispersion since it depends only on variances while we call the second line the signal. The dispersion amounts to a multivariate ensemble spread measure. It is easily shown that the signal is the sum of the squares of EOF anomalies divided by their climatological variances.
The relative entropy can be computed analytically in this instance. Let the prior random variable be denoted by \( y \) and the posterior by \( x \) and variances by \( \sigma \).

The relative entropy is given by:

\[
D(p||q) = \frac{1}{2} \left[ \log \left( \frac{\det(\sigma_y)}{\det(\sigma_x)} \right) + tr \left( \sigma_x \sigma_y^{-1} \right) - n \right] + \frac{1}{2} (\bar{x} - \bar{y})^t \sigma_y^{-1} (\bar{x} - \bar{y})
\]

The first term on the RHS here is the entropy difference and measures reduction in uncertainty. The third term is the shift in means. The second term is usually not important. We call the first line the dispersion since it depends only on variances while we call the second line the signal. The dispersion amounts to a multivariate ensemble spread measure. It is easily shown that the signal is the sum of the squares of EOF anomalies divided by their climatological variances.

R. Kleeman

Abstract predictability
Gaussian distributions

- The relative entropy can be computed analytically in this instance. Let the prior random variable be denoted by $y$ and the posterior by $x$ and variances by $\sigma$

\[
D(p||q) = \frac{1}{2} \left[ \log \left( \frac{\det(\sigma_y)}{\det(\sigma_x)} \right) + tr \left( \sigma_x (\sigma_y)^{-1} \right) - n \right] \\
+ \frac{1}{2} (\bar{x} - \bar{y})^t \sigma_y^{-1} (\bar{x} - \bar{y})
\]

- The first term on the RHS here is the entropy difference and measures reduction in uncertainty. The third term is the shift in means. The second term is usually not important. We call the first line the dispersion since it depends only on variances while we call the second line the signal. The dispersion amounts to a multivariate ensemble spread measure. It is easily shown that the signal is the sum of the squares of EOF anomalies divided by their climatological variances.
In a Markov system convergence of distributions can be measured using the relative entropy of the transient and equilibrium distributions. Thus the utility of a prediction is also a measure of disequilibrium. The greater the utility the greater the disequilibrium.

A continuous time continuous outcome Markov process has distributions satisfying the Fokker Planck equation

\[
\partial_t p = -\partial_i [A_i(x, t)p] + \frac{1}{2} \partial_i \partial_j \{C_{ij}p\} \quad \text{(summation convention)}
\]
In a Markov system convergence of distributions can be measured using the relative entropy of the transient and equilibrium distributions. Thus the utility of a prediction is also a measure of disequilibrium. The greater the utility the greater the disequilibrium.

A continuous time continuous outcome Markov process has distributions satisfying the Fokker Planck equation

\[ \partial_t p = -\partial_i [A_i(x, t)p] + \frac{1}{2} \partial_i \partial_j \{C_{ij}p\} \quad \text{(summation convention)} \]
Equilibration

- The relative entropy evolution equation may be derived for two such processes satisfying this equation:

\[ \partial_t D(p||q) = - \int pC_{ij} \partial_i(\ln \frac{p}{q}) \partial_j(\ln \frac{p}{q}) < 0 \]

- Note that the monotonic decline to zero comes only from the noise i.e. from the \( C_{ij} \). Another way of viewing this is that the equilibration is purely a result of “coarse graining” the system i.e. representing the fast scales by noise. Coarse graining is fundamental to the second law of thermodynamics of which this equation is a generalized version.
The relative entropy evolution equation may be derived for two such processes satisfying this equation:

\[
\partial_t D(p||q) = -\int p C_{ij} \partial_i (\ln \frac{p}{q}) \partial_j (\ln \frac{p}{q}) < 0
\]

Note that the monotonic decline to zero comes only from the noise i.e. from the \( C_{ij} \). Another way of viewing this is that the equilibration is purely a result of “coarse graining” the system i.e. representing the fast scales by noise. Coarse graining is fundamental to the second law of thermodynamics of which this equation is a generalized version.
The relative entropy evolution equation may be derived for two such processes satisfying this equation:

\[
\partial_t D(p||q) = - \int p C_{ij} \partial_i (\ln \frac{p}{q}) \partial_j (\ln \frac{p}{q}) < 0
\]

Note that the monotonic decline to zero comes only from the noise i.e. from the \( C_{ij} \). Another way of viewing this is that the equilibration is purely a result of “coarse graining” the system i.e. representing the fast scales by noise. Coarse graining is fundamental to the second law of thermodynamics of which this equation is a generalized version.
Talk Outline

1. Predictability Definition
   - Basic Concepts
   - Universal information theoretic measures

2. Predictability and disequilibrium

3. Predictability variation
   - Introduction
   - Chaos
   - Stochastic Models
   - Turbulence
   - Mechanisms for predictability variation

4. References
Basic Motivation

Anyone who has done practical forecasting knows some predictions are more useful than others. We study here reasons for this variation from a theoretical perspective.

At any prediction time lag most dynamical systems exhibit large variations in the measure proposed above. We seek explanatory mechanisms for this variation since it appears to underly much utility variation.

Since mechanisms for variation in predictability are likely system dependent we study three widely known but very different dynamical systems.
Basic Motivation

- Anyone who has done practical forecasting knows some predictions are more useful than others. We study here reasons for this variation from a theoretical perspective.
- At any prediction time lag most dynamical systems exhibit large variations in the measure proposed above. We seek explanatory mechanisms for this variation since it appears to underly much utility variation.
- Since mechanisms for variation in predictability are likely system dependent we study three widely known but very different dynamical systems.
Basic Motivation

- Anyone who has done practical forecasting knows some predictions are more useful than others. We study here reasons for this variation from a theoretical perspective.
- At any prediction time lag most dynamical systems exhibit large variations in the measure proposed above. We seek explanatory mechanisms for this variation since it appears to underly much utility variation.
- Since mechanisms for variation in predictability are likely system dependent we study three widely known but very different dynamical systems.
Talk Outline

1. Predictability Definition
   - Basic Concepts
   - Universal information theoretic measures

2. Predictability and disequilibrium

3. Predictability variation
   - Introduction
   - Chaos
   - Stochastic Models
   - Turbulence
   - Mechanisms for predictability variation

4. References
Lorenz 1963 Model

- Characterized by three non-linearly interacting variables which exhibit similar times scales.
- The prior distribution is a strange attractor so is highly non-Gaussian. Posterior distributions become rapidly non-Gaussian as they relax toward the prior.
- Gaussian indicators such as signal and dispersion (which recall amounts to an ensemble spread measure) do not work in the non-Gaussian regime.
- Entropy difference i.e. uncertainty reduction, on the other hand, is a highly reliable determinant of predictability.
Lorenz 1963 Model

- Characterized by three non-linearly interacting variables which exhibit similar times scales.
- The prior distribution is a strange attractor so is highly non-Gaussian. Posterior distributions become rapidly non-Gaussian as they relax toward the prior.
- Gaussian indicators such as signal and dispersion (which recall amounts to an ensemble spread measure) do not work in the non-Gaussian regime.
- Entropy difference i.e. uncertainty reduction, on the other hand, is a highly reliable determinant of predictability.
Lorenz 1963 Model

- Characterized by three non-linearly interacting variables which exhibit similar times scales.
- The prior distribution is a strange attractor so is highly non-Gaussian. Posterior distributions become rapidly non-Gaussian as they relax toward the prior.
- Gaussian indicators such as signal and dispersion (which recall amounts to an ensemble spread measure) do not work in the non-Gaussian regime.
- Entropy difference i.e. uncertainty reduction, on the other hand, is a highly reliable determinant of predictability.
Lorenz 1963 Model

- Characterized by three non-linearly interacting variables which exhibit similar times scales.

- The prior distribution is a strange attractor so is highly non-Gaussian. Posterior distributions become rapidly non-Gaussian as they relax toward the prior.

- Gaussian indicators such as signal and dispersion (which recall amounts to an ensemble spread measure) do not work in the non-Gaussian regime.

- Entropy difference i.e. uncertainty reduction, on the other hand, is a highly reliable determinant of predictability.
Shown is the relationship between entropy difference and relative entropy approximately mid way through relaxation to the prior. Each dot is a different prediction whose initial condition means were chosen at random from the prior.
Predictions from this dynamical system exhibit considerable “durability” in that the level of predictability tends to persist for a considerable fraction of the relaxation (predictability limit) time. Shown is the predictability at early times (vertical) versus the predictability at later times (horizontal).
Talk Outline

1. Predictability Definition
   - Basic Concepts
   - Universal information theoretic measures

2. Predictability and disequilibrium

3. Predictability variation
   - Introduction
   - Chaos
   - Stochastic Models
   - Turbulence
   - Mechanisms for predictability variation

4. References
Stochastic Models

- Intended to represent dynamical systems with fast and slow modes. The simplest and most widely used models of this type are linear systems with additive stochastic forcing.

- Models of this type have been proposed to explain general SST climate variability; ENSO and aspects of mid-latitude atmospheric turbulence.

- With Gaussian initial conditions all posterior and prior distributions are Gaussian. In addition one may prove that the variances of posterior distributions depend only on time and not on the choice of initial condition.

- It follows that variations in the signal are the sole determinant of predictability variations. The mechanism for this variation is that some initial conditions have large slow mode anomalies and other initial conditions do not. These modal anomalies persist during a prediction until eroded by the fast modes (i.e. the noise forcing).
Stochastic Models

- Intended to represent dynamical systems with fast and slow modes. The simplest and most widely used models of this type are linear systems with additive stochastic forcing.

- Models of this type have been proposed to explain general SST climate variability; ENSO and aspects of mid-latitude atmospheric turbulence.

- With Gaussian initial conditions all posterior and prior distributions are Gaussian. In addition one may prove that the variances of posterior distributions depend only on time and not on the choice of initial condition.

- It follows that variations in the signal are the sole determinant of predictability variations. The mechanism for this variation is that some initial conditions have large slow mode anomalies and other initial conditions do not. These modal anomalies persist during a prediction until eroded by the fast modes (i.e. the noise forcing).
Stochastic Models

- Intended to represent dynamical systems with fast and slow modes. The simplest and most widely used models of this type are linear systems with additive stochastic forcing.

- Models of this type have been proposed to explain general SST climate variability; ENSO and aspects of mid-latitude atmospheric turbulence.

- With Gaussian initial conditions all posterior and prior distributions are Gaussian. In addition one may prove that the variances of posterior distributions depend only on time and not on the choice of initial condition.

- It follows that variations in the signal are the sole determinant of predictability variations. The mechanism for this variation is that some initial conditions have large slow mode anomalies and other initial conditions do not. These modal anomalies persist during a prediction until eroded by the fast modes (i.e. the noise forcing).
Stochastic Models

- Intended to represent dynamical systems with fast and slow modes. The simplest and most widely used models of this type are linear systems with additive stochastic forcing.

- Models of this type have been proposed to explain general SST climate variability; ENSO and aspects of mid-latitude atmospheric turbulence.

- With Gaussian initial conditions all posterior and prior distributions are Gaussian. In addition one may prove that the variances of posterior distributions depend only on time and not on the choice of initial condition.

- It follows that variations in the signal are the sole determinant of predictability variations. The mechanism for this variation is that some initial conditions have large slow mode anomalies and other initial conditions do not. These modal anomalies persist during a prediction until eroded by the fast modes (i.e. the noise forcing).
Stochastic Models

Like the chaotic model, stochastic model predictions also exhibit considerable durability. Shown is predictability for a random set of predictions (horizontal axis) for increasing prediction time (vertical axis). This result can be demonstrated analytically as well.
Talk Outline

1. Predictability Definition
   - Basic Concepts
   - Universal information theoretic measures

2. Predictability and disequilibrium

3. Predictability variation
   - Introduction
   - Chaos
   - Stochastic Models
   - Turbulence
   - Mechanisms for predictability variation

4. References
Mid-latitude atmospheric turbulence

- This is characterised both by modes of similar timescales interacting non-linearly as well as by the presence of fast and slow modes. One might expect predictability characteristics intermediate between the first two examples. This was studied in a T42 and 5 vertical level dry global primitive equation model with orography and a good simulation of storm tracks.

- In general distributions are quasi-Gaussian so we can use the analytical expression above as a good approximation. Numerical results show that signal actually dominates dispersion at all prediction lags. Shown is the correlation between signal and dispersion with relative entropy. This is calculated using 48 statistical predictions whose initial condition means were drawn from the model climatology at random.
Mid-latitude atmospheric turbulence

- This is characterised both by modes of similar times scales interacting non-linearly as well as by the presence of fast and slow modes. One might expect predictability characteristics intermediate between the first two examples. This was studied in a T42 and 5 vertical level dry global primitive equation model with orography and a good simulation of storm tracks.

- In general distributions are quasi-Gaussian so we can use the analytical expression above as a good approximation. Numerical results show that signal actually dominates dispersion at all prediction lags. Shown is the correlation between signal and dispersion with relative entropy. This is calculated using 48 statistical predictions whose initial condition means were drawn from the model climatology at random.
Mid-latitude atmospheric turbulence
Mid-latitude atmospheric turbulence

![Graph showing correlation coefficient over prediction time](image-url)
Mid-latitude atmospheric turbulence

Like the two earlier simpler models this more realistic model exhibits predictability durability although not quite as strongly. Shown is the signal for a random set of initial conditions (horizontal axis) and prediction times (vertical axis):
Talk Outline

1. Predictability Definition
   - Basic Concepts
   - Universal information theoretic measures

2. Predictability and disequilibrium

3. Predictability variation
   - Introduction
   - Chaos
   - Stochastic Models
   - Turbulence
   - Mechanisms for predictability variation

4. References
Different mechanisms for predictability variation

- High Signal Predictability
- Low Signal Predictability
- High Dispersion Predictability
- Low Dispersion Predictability

R. Kleeman
Abstract predictability
Different mechanisms for predictability variation

1. **Slow mode anomalies in the initial conditions.** Sometimes for purely random reasons, modes exhibiting slow temporal decorrelation are anomalously present in the initial conditions. This feature persists for a considerable time into a prediction thus increasing predictability. The reverse situation i.e. little anomalous slow mode presence in the initial condition results in reduced predictability since relaxation tends to be more rapid.

2. **Initial condition instability.** Sometimes again for purely random reasons the initial conditions of a prediction may be “more than normally” unstable to uncertainty growth. This could be linear or non-linear instability. Such a scenario results in decreased predictability. The reverse situation i.e. greater stability results in enhanced predictability.

3. **Which effect is more important?** Evidence gathered to date from realistic models of turbulence tends to favour the first mechanism but this conclusion needs deeper investigation.
For Further Reading I

R. Kleeman.

R. Kleeman.

R. Kleeman.