Approaches to Modeling, Types of uncertainty

Ecological Forecasting, Jan 25th

First, checking in on lab from Tuesday.

Any questions, or issues?

Revisiting logistic growth

$$N_{t+1} = N_t + rN_t(1 - N_t/K)$$

Low density
growth rate

Carrying
Capacity

Forward Modeling

$$N_{t+1} = N_t + rN_t(1 - N_t/K)$$

Estimated

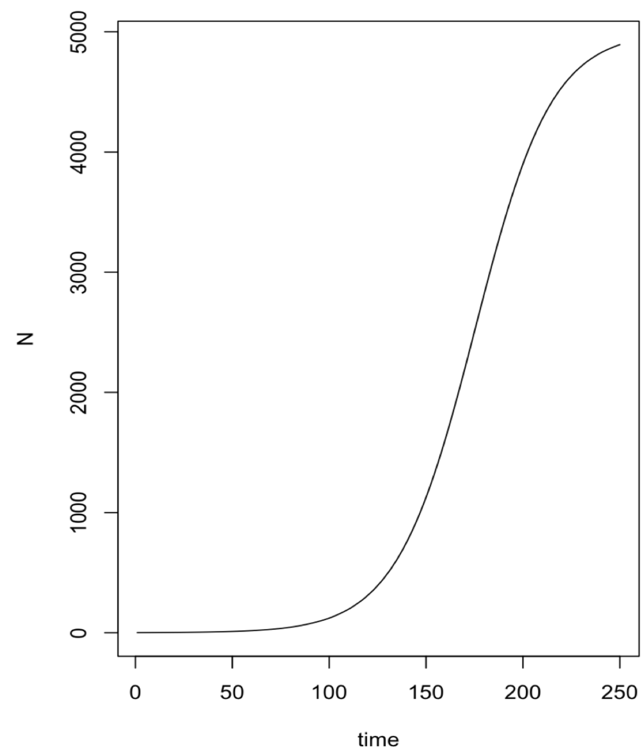
Data

Data

Forward Modeling

$$N_{t+1} = N_t + rN_t(1 - N_t/K)$$

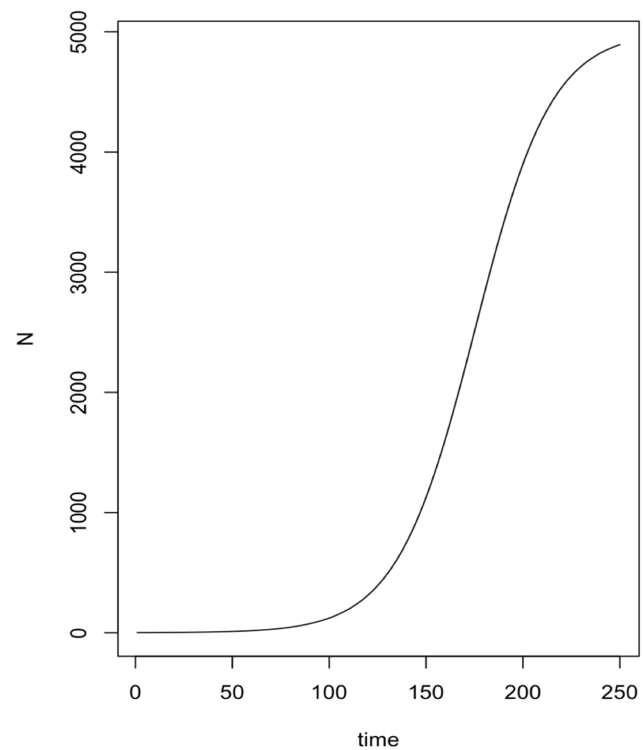
Model is purely deterministic,
stochasticity can be added in some forms



Forward Modeling

$$N_{t+1} = N_t + rN_t(1 - N_t/K)$$

Model is purely deterministic,
stochasticity can be added in some forms



Forward Modeling

Anyone know of any applications of forward modeling?

Forward Modeling

Anyone know of any applications of forward modeling?

Parameter	Quantity	Units
Aboveground state variables		
$n(\mathbf{z}, \mathbf{x}, a, t)$	density† of size $\mathbf{z}$ type $\mathbf{x}$ plants in gaps of age a at time t	m^{-2}
$p(a, t)$	distribution of gap ages a at time t	dimensionless
State dimensions		
$\mathbf{z}$	plant size $\mathbf{z} = [z_s, z_a] = [B_s, B_a]$	kg C, kg C
$\mathbf{x}$	plant type $[x_1, x_2] \cdot x_1 = 0$ if C_3, 1 if C_4, x_2 = leaf longevity	dimensionless, yr
a	gap age	yr
t	time	yr
Plant resource environment		
$\mathbf{r}$	resource vector $\mathbf{r} = [\phi, W, N]$	$J \cdot m^{-2} \cdot s^{-1}$, $m^3\ H_2O/m^2$, kg N/m^2
ϕ	photosynthetically active radiation (PAR)	$J \cdot m^{-2} \cdot s^{-1}$
W	soil water content	$m^3\ H_2O/m^2$
N	plant available soil nitrogen content	kg N/m^2
Transition rates		
g_s	structural biomass growth rate	kg C/yr
g_a	living biomass tissue growth rate	kg C/yr
μ	mortality rate	yr^{-1}
f	fecundity	yr^{-1}
λ	total disturbance rate $\lambda = \lambda_F + \lambda_{DI}$	yr^{-1}
λ_F	fire frequency	yr^{-1}
λ_{DI}	rate of canopy gap formation	yr^{-1}
s	survivorship of plants following disturbance	dimensionless
Plant size characteristics		
h	height	m
B_a	living biomass $(B_l + B_r + B_{sw})$	kg C
B_s	structural stem biomass	kg C
B_l	leaf biomass	kg C
B_r	root biomass	kg C
B_{sw}	sapwood biomass	kg C
Leaf-level carbon and water fluxes		
A_n	net rate of carbon gain per unit leaf area	$\mu mol\ C \cdot m^{-2} \cdot s^{-1}$
Ψ	evapotranspiration rate per unit leaf area	$\mu mol\ H_2O \cdot m^{-2} \cdot yr^{-1}$
Decomposition model state variables and coefficients		
C_1	fast soil carbon pool	kg C/m^2
C_2	structural soil carbon pool	kg C/m^2
N_1	fast soil nitrogen pool	kg N/m^2
N_2	structural soil nitrogen pool	kg N/m^2
Miscellaneous		
$n_0(\mathbf{z}, \mathbf{x}, a)$	initial plant density $n(\mathbf{z}, \mathbf{x}, a, 0)$†	m^{-2}
$p_0(a)$	initial gap age distribution	dimensionless
$\mathbf{z}_0$	seedling size	kg C, kg C
y	integer gap number $(1 \ldots Q)$	dimensionless
h^*	height above which mortality is treated as disturbance	m
c_s	stomatal conductance	$\mu mol\ H_2O \cdot m^{-2} \cdot s^{-1}$
T_L	leaf temperature	°C
T_A	atmospheric air temperature	°C
C_i	interstitial concentration of CO_2	mol/mol
C_a	atmospheric concentration of CO_2	mol/mol

† $n(\mathbf{z}, \mathbf{x}, a, t)$ is technically a density distribution where $n(\mathbf{z}, \mathbf{x}, a, t)dz_s dz_a da$ is the per m^{-2} density of type $\mathbf{x}$ plants between size z_s and $z_s + dz_s$, and size z_a and $z_a + dz_a$ in gaps aged between a and $a + da$ at time t.

Paul Moorecroft, Ecosystem Demography model

Forward Modeling

Anyone know of any applications of forward modeling?

Forward modeling, historically, has been of the most common modeling approaches. Big push to make more and more complex forward models came in the 1970s, 1980s with increases in computing power.

Forward Modeling

Anyone know of any applications of forward modeling?

Forward modeling historically has been of the most common modeling approaches. Big push to make more and more complex forward models came in the 1970s, 1980s with increases in computing power.

While computers are powerful and important tools, adding additional processes and complexity to models does not always yield better predictions.

Inverse modeling

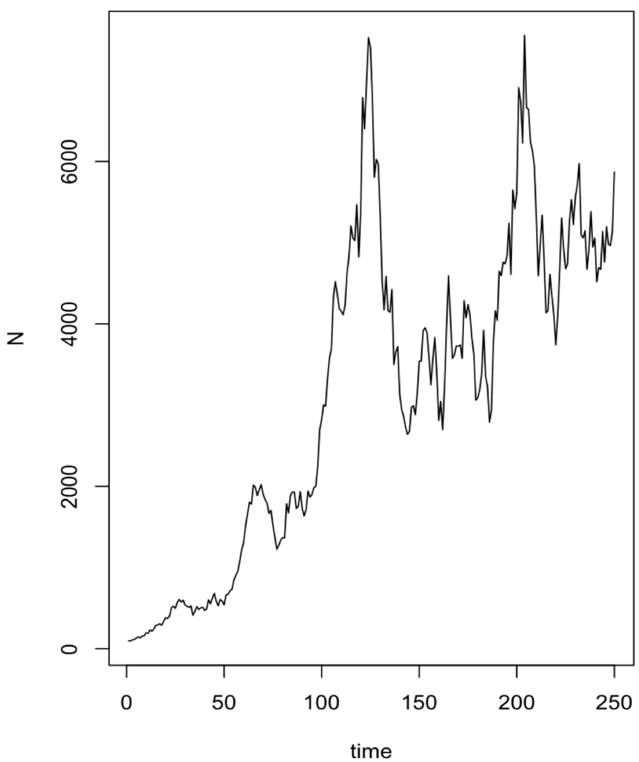
$$N_{t+1} = N_t + rN_t(1 - N_t/K)$$

Data

Estimated

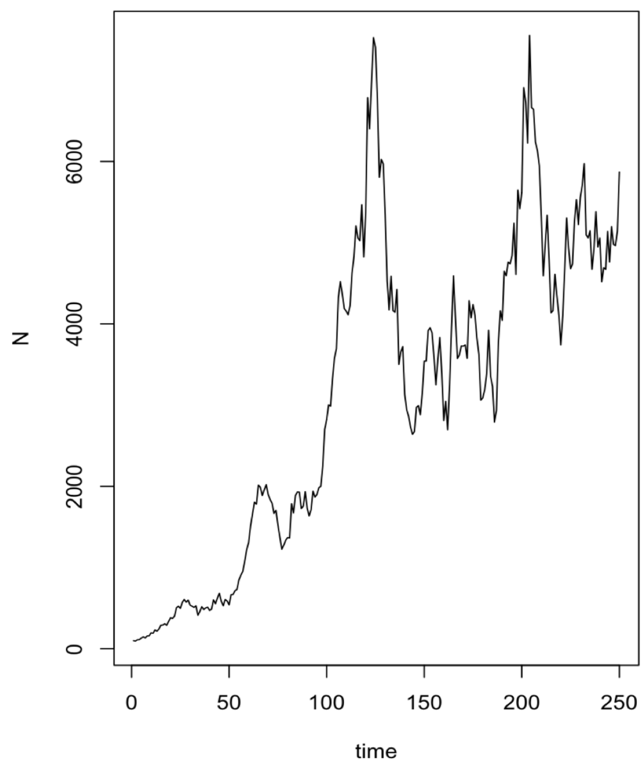
Estimated

Inverse Modeling



$$N_{t+1} = N_t + rN_t(1 - N_t/K)$$

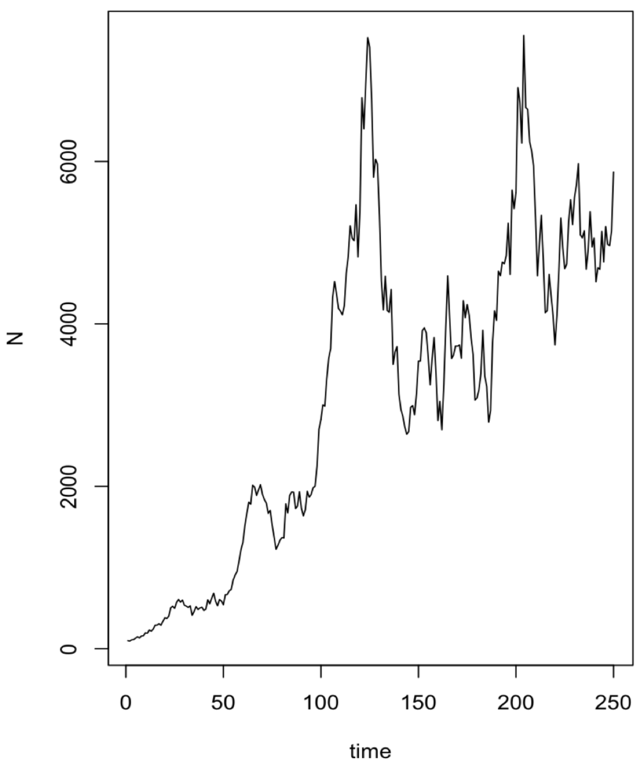
Inverse Modeling



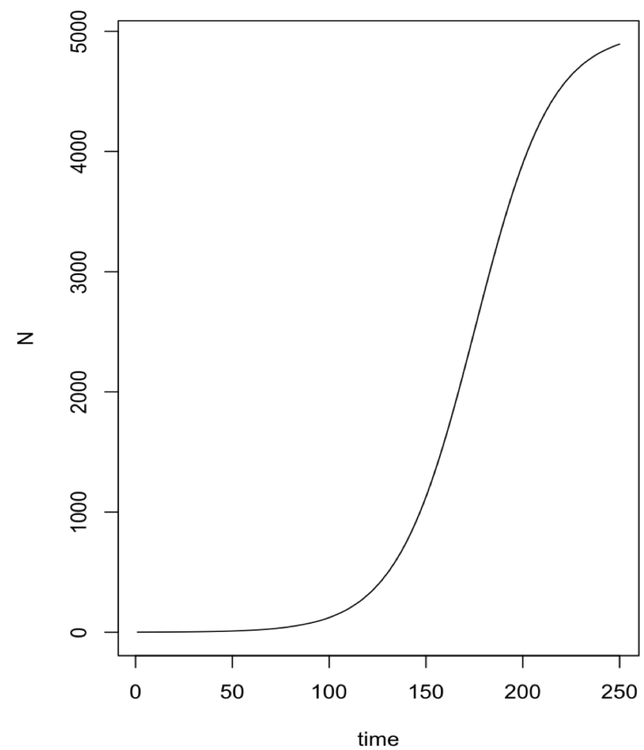
Probability distributions describes connection between data and model

$$N_{t+1} = N_t + rN_t(1 - N_t/K)$$

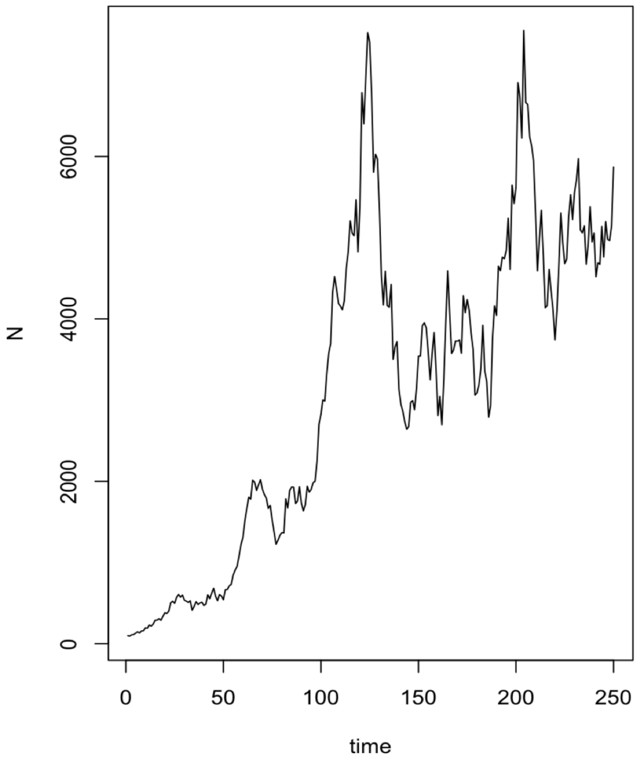
Forward and Inverse don’t match



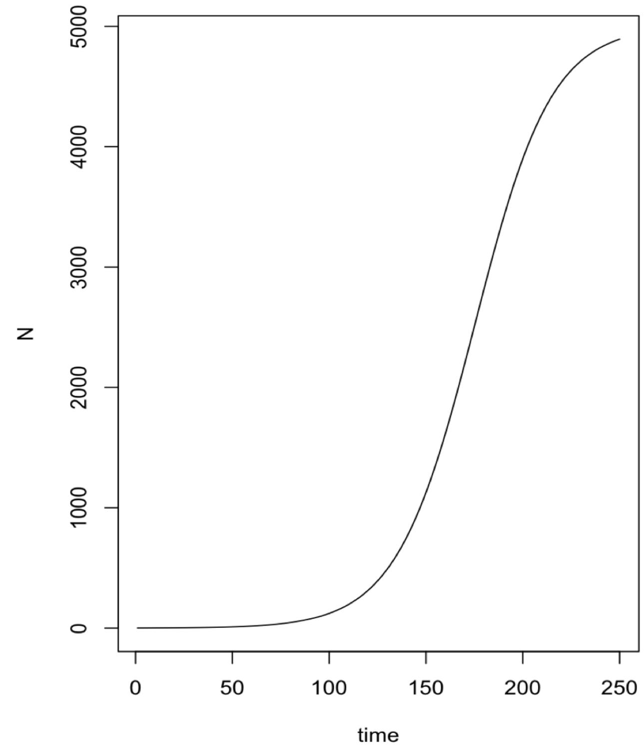
$\neq$



What accounts for this?



≠



What accounts for this?

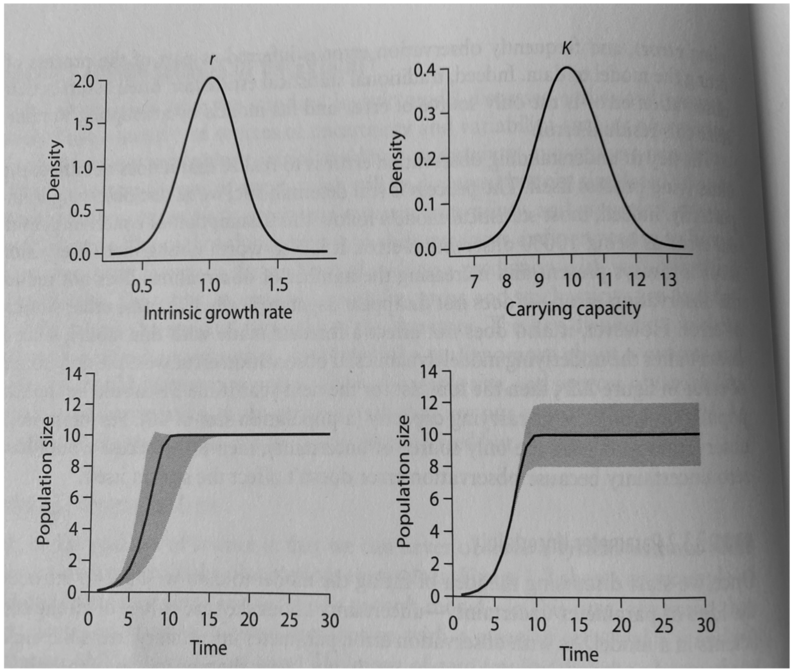
Variability and uncertainty

Types of variability and uncertainty

Parameter uncertainty: The true values of the parameter in the models are unknown

This type of uncertainty can be accounted for by both forward and inverse modeling approaches

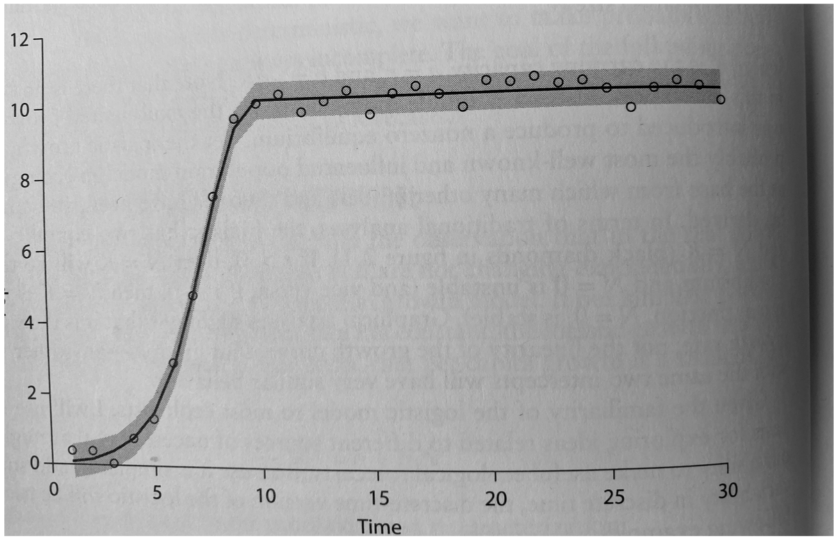
Parameter uncertainty will be reduced with more data collection.



Types of variability and uncertainty

Observation uncertainty: The true state of the system can not be observed without error.

If observation error is the only source of error then the underlying forecast is deterministic.

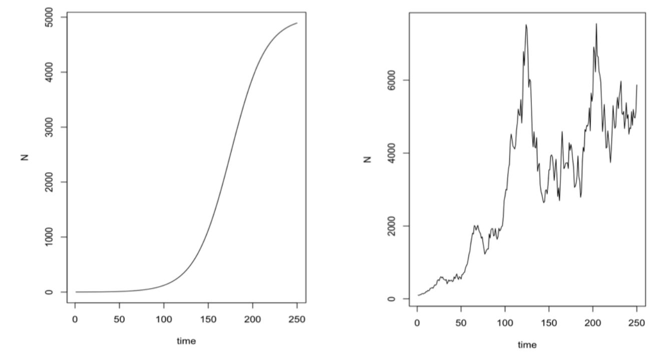


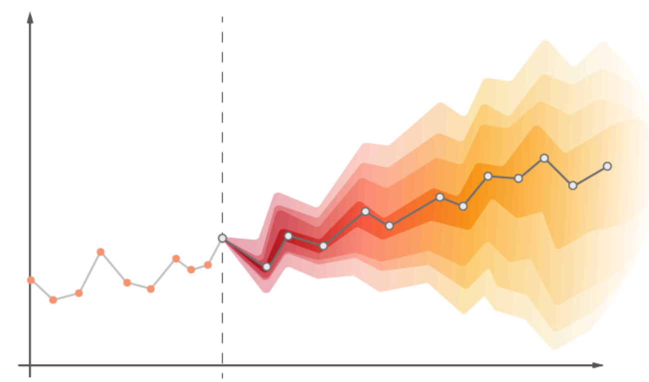
Types of variability and uncertainty

Process error/variability: Our models do not include all of the relevant processes.

This source of error is very common in ecology and environmental sciences, but hard to include/describe with forward modeling approaches. One way is using a model ensembling.

Because process error is uncertainty about the true state of a system, it propagates out in forecasts, often leading to greater uncertainty as time/space goes on.

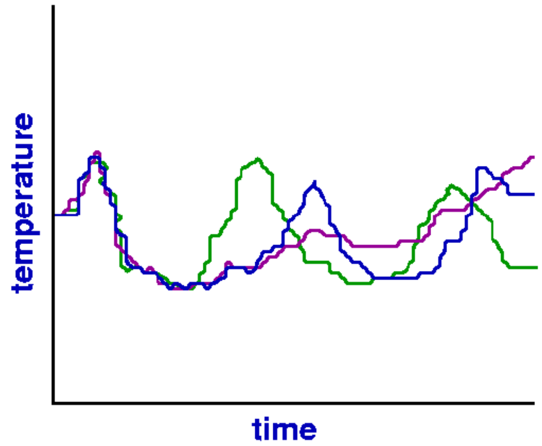




Types of variability and uncertainty

Initial condition uncertainty: Not knowing the current state of the system.

This is most important in chaotic systems like weather.



Grab bag of other uncertainties

Driver uncertainty: If forecast includes “drivers” like temperature, than uncertainty in the forecast of that driver propagate through to forecasts. One of the reasons why the most ecologically realistic model may not be the best for forecasting.

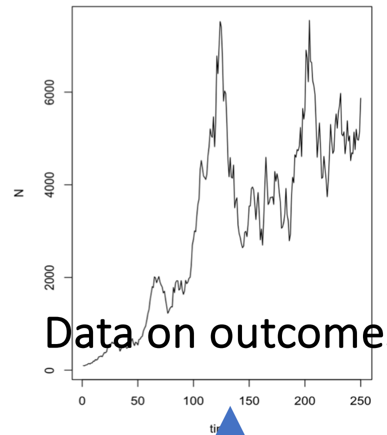
Numerical approximation: Computers make approximation errors. This rarely matters for most of our applications, but good to be aware of.

So how do we build models and forecasts in the face of all of these sources of uncertainty?

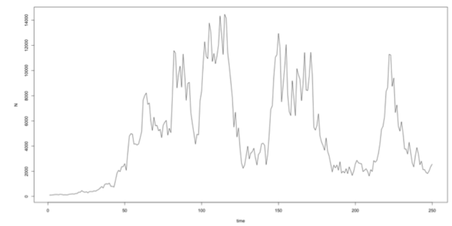
Models as scaffolds

We use models to describe the relationship (or covariance) between different sources of data, and data yet to be observed.

Models as scaffolds



Data on outcomes



Forecast

$$N_{t+1} = N_t + rN_t(1 - N_t/K)$$

Data on parameters

Soapbox

Models are sometimes described as process-based, mechanistic, statistical, phenomenological.

What does that mean?

Soapbox

Any (or at least most) useful models contain some element of all of these.

Readings- Clark et al.

What does "inherent uncertainty" mean and what are some examples?

What are some of the problems if a forecast doesn't handle uncertainty properly?

What is the "bet-hedging" described in the paper and can you think of any examples?

What is your impression of how the "next steps" discussed here have been implemented in the last 20 years?

Readings-Houlihan et al.

What is model transferability?

What is bias-variance tradeoff? How does it impact prediction?

Can we truly falsify forecast (in the NHT sense)?

Next week!

Beginning some analysis of simple timeseries data.

Fitting and forecasting timeseries using R functions.