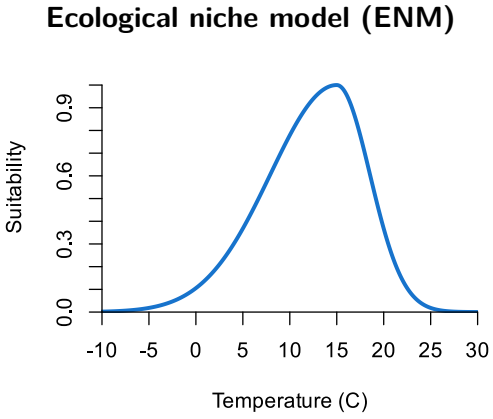
Fitting ENMs

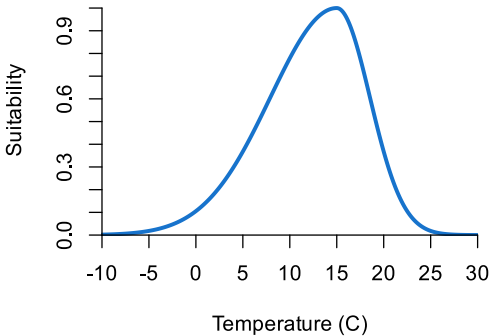
Generalized Additive Models (GAMs)

Duality of SDMs



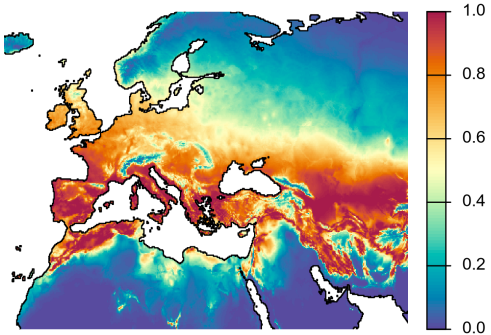
Duality of SDMs

Ecological niche model (ENM)



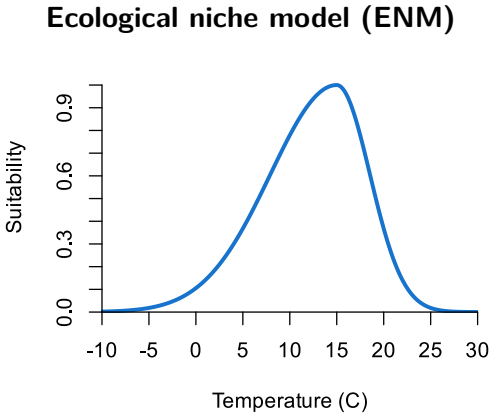
Ecological space

Species distribution model (SDM)



Geographic space

Ecological Niche Model (ENM)



Ecological space

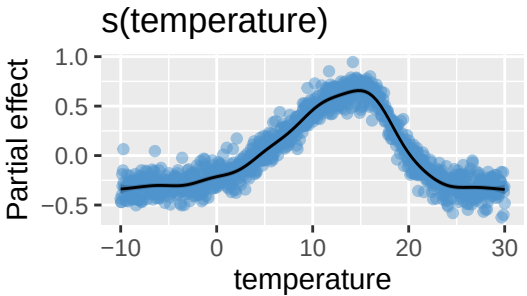
Introduction to GAMs

Generalized additive models (GAMs)

- Machine learning statistical framework
- Highly flexible and extremely powerful
- Easy to interpret
- Easy to implement in R (`mgcv` & `gratia`)

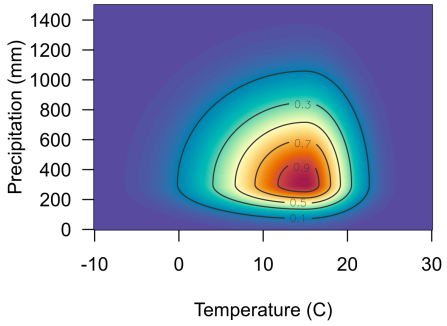
Example – Fitting a 1D niche

```
library(mgcv)
library(gratia)
d <- read.csv("../data/niche-1d.csv")
enm <- gam(suitability ~ s(temperature, k = 20), data = d)
draw(enm, residuals = TRUE, rug = FALSE)
```



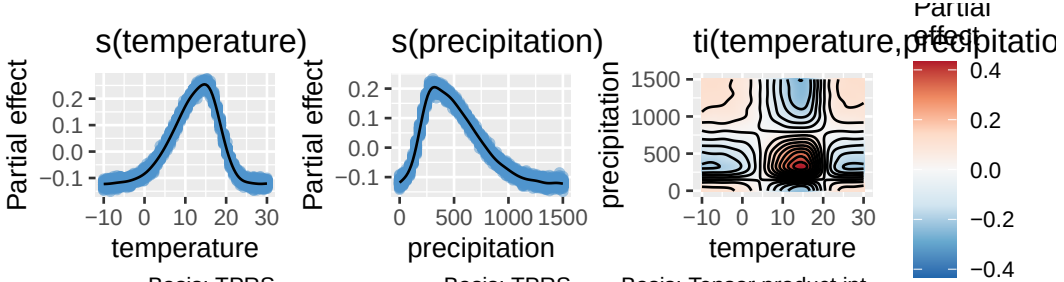
Basis: TPRS

Example – Fitting a 2D niche

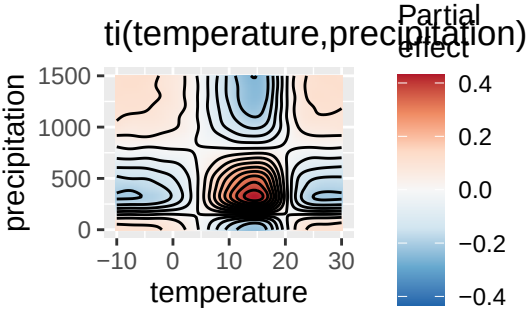


Example – Fitting a 2D niche

```r
d <- read.csv("../data/niche-2d.csv")
enm <- gam(
  suitability ~ s(temperature, k = 20) + s(precipitation, k = 20) +
                ti(temperature, precipitation, k = 16),
  data = d
)
draw(enm, residuals = TRUE, rug = FALSE, nc = 3)
```

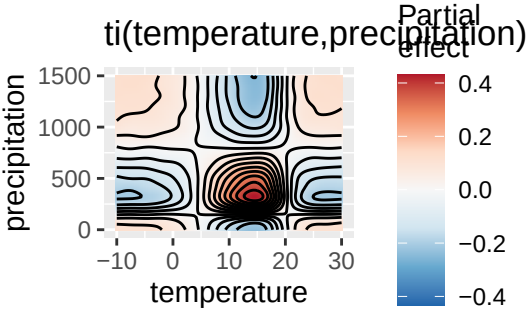


Example – Fitting a 2D niche

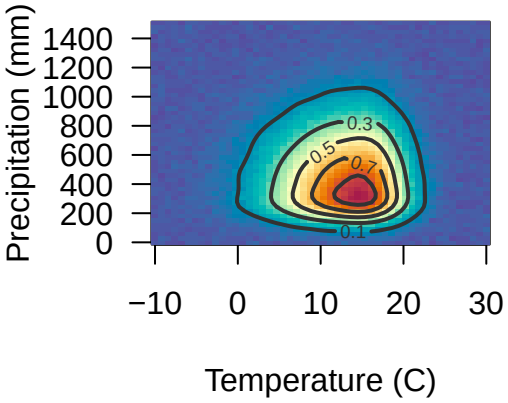


Basis: Tensor product int.

Example – Fitting a 2D niche



Basis: Tensor product int.



GAM workflow

1. Fit GAM model: `gam(<y> ~ <x>)`

GAM workflow

1. Fit GAM model: `gam(<y> ~ <x>)`
2. Simple covariate: `s(<x>, k = ...)`

GAM workflow

1. Fit GAM model: `gam(<y> ~ <x>)`
2. Simple covariate: `s(<x>, k = ...)`
3. Interaction term: `ti(x1, x2, k = ...)`

GAM workflow

1. Fit GAM model: `gam(<y> ~ <x>)`
2. Simple covariate: `s(<x>, k = ...)`
3. Interaction term: `ti(x1, x2, k = ...)`
4. Inspect: `draw(...)`

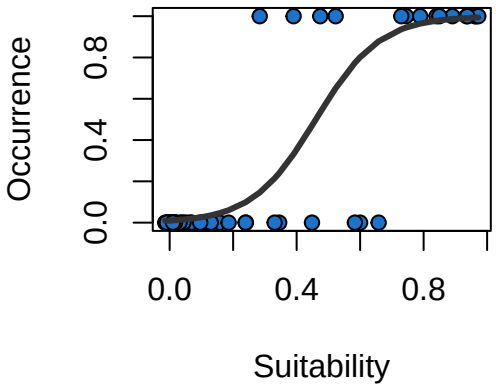
GAM workflow

1. Fit GAM model: `gam(<y> ~ <x>)`
2. Simple covariate: `s(<x>, k = ...)`
3. Interaction term: `ti(x1, x2, k = ...)`
4. Inspect: `draw(...)`
5. Appraise: `k.check(...)`
   - $k - index \ll 1 \rightarrow$ Increase$k$
   - $p - value < 0.05 \rightarrow$ Increase$k$

Data for ENM/SDM

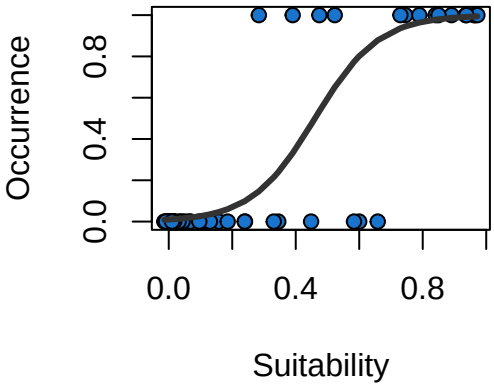
GAM – SDM data

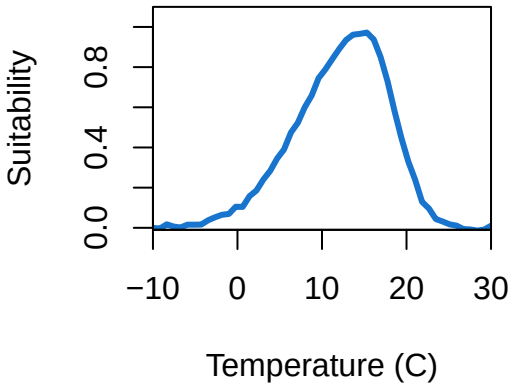
- Rarely (if ever) you will get suitability as data
- More likely (almost always) you will get presence (1) and absence (0) data



GAM – SDM data

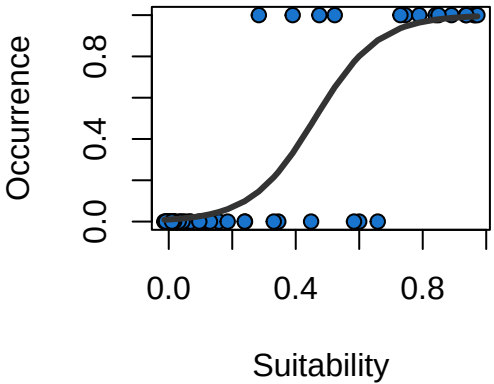
- Rarely (if ever) you will get suitability as data
- More likely (almost always) you will get presence (1) and absence (0) data

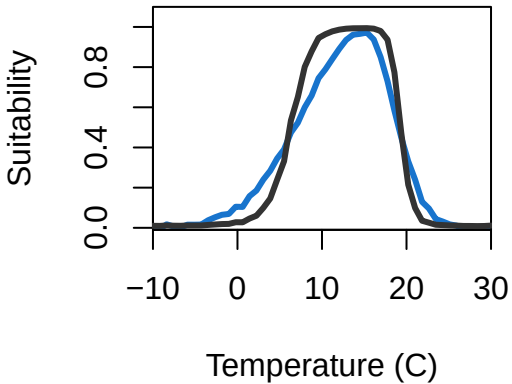




GAM – SDM data

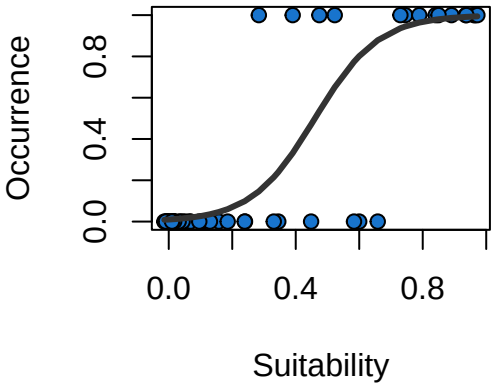
- Rarely (if ever) you will get suitability as data
- More likely (almost always) you will get presence (1) and absence (0) data

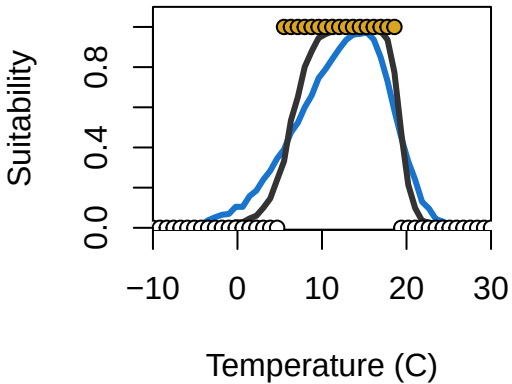




GAM – SDM data

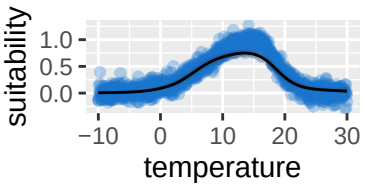
- Rarely (if ever) you will get suitability as data
- More likely (almost always) you will get presence (1) and absence (0) data





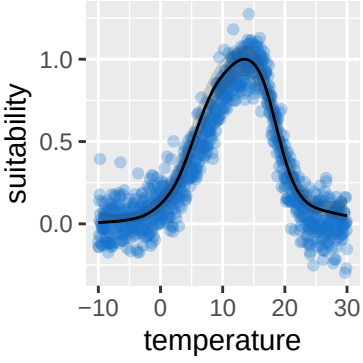
GAM – Example with SDM data

```r
d <- read.csv("../data/niche-1d.csv")
enm <- gam(occurrence ~ s(temperature, k = 30), data = d, family = binomial
# draw(enm, rug = FALSE, residuals = TRUE)
library(ggplot2)
fitted_values(enm) |>
  ggplot() +
  geom_point(data = d, aes(temperature, suitability), alpha = .3, col = "d
  geom_ribbon(aes(x = temperature, ymin = .lower_ci, ymax = .upper_ci), al
  geom_line(aes(temperature, .fitted))
```



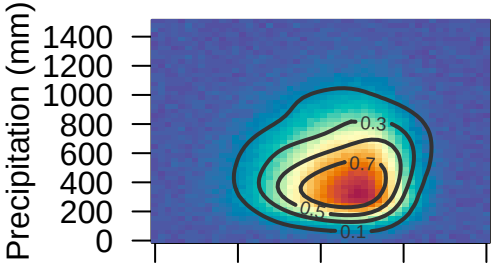
GAM – Fitted niche is scaled

```
fitted_values(enm) |>
  ggplot() +
  geom_point(data = d, aes(temperature, suitability), alpha = .3, col = "d
  geom_ribbon(aes(x = temperature, ymin = .lower_ci / max(.fitted), ymax =
  geom_line(aes(temperature, .fitted / max(.fitted)))
```



GAM – 2D example with SDM data

```
d <- read.csv("../data/niche-2d.csv")
enm <- gam(
  occurrence ~ s(temperature, k = 10) + s(precipitation, k = 10) +
               ti(temperature, precipitation, k = 10),
  data = d,
  family = binomial
)
niche <- predict(enm, type = "response")
```



Model selection

AIC

You can perform model selection on ENMs just as any other statistical models.

```
enm_simple <- gam(
  occurrence ~ s(temperature, k = 10) + s(precipitation, k = 10),
  data = d, family = binomial, method = "ML"
)
enm_int <- gam(
  occurrence ~ s(temperature, k = 10) + s(precipitation, k = 10) +
               ti(temperature, precipitation, k = 10),
  data = d, family = binomial, method = "ML"
)
AIC(enm_simple, enm_int)
```

```
                df      AIC
enm_simple 12.66876 1126.237
enm_int    28.21966 1113.738
```

Maximum Likelihood

When you perform model selection using AIC, remember to specify `gam(..., method = "ML")`.

Once you select the model, you can update it and remove `method = "ML"`.