



# Forecasting at scale

Taylor, Sean J and Letham, Benjamin

# Abstract & Introduction

시계열에 대한 도메인 지식을 가진 분석가들이 직관적으로 조정할 수 있는 해석 가능한 매개변수를 가진 모듈형 회귀 모델 (Prophet) 을 제안

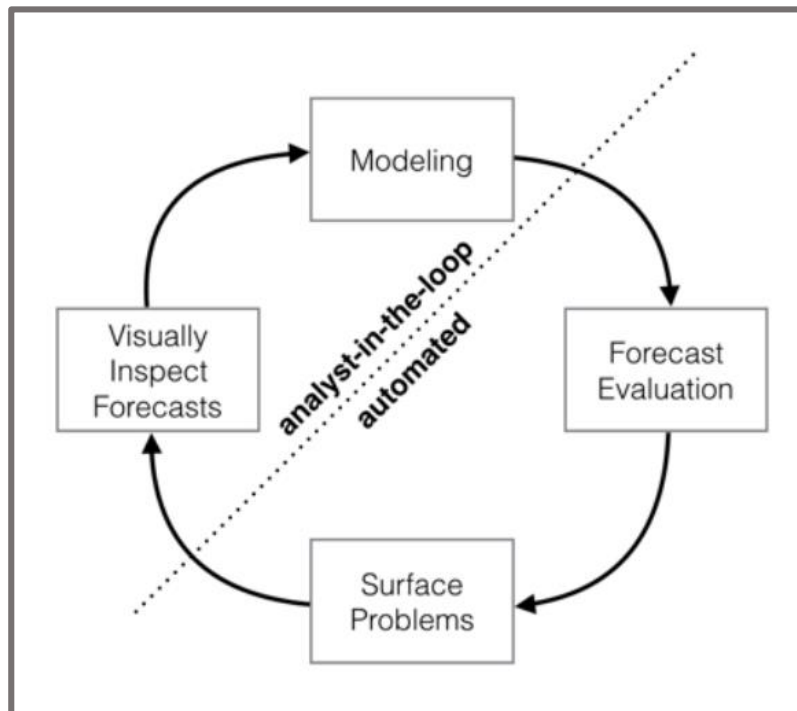
## **비즈니스 예측 생성 실무의 주요 주제:**

1. 완전 자동 예측 기술은 튜닝하기 어려울 수 있고 유용한 가정이나 heuristics를 통합하기에는 너무 유연성이 떨어지는 경우가 많음
2. 조직 전체에서 데이터 과학을 담당하는 분석가는 일반적으로 특정 제품 또는 서비스에 대한 심층 도메인 전문 지식을 보유하지만, 시계열 예측에 대한 교육을 받지 않는 경우가 많음

# Abstract & Introduction

## 규모(scale)에 따라 예측하기 위한 가이드:

1. 비즈니스 예측 방법은 시계열에 대한 지식없이 예측을 하는 사람들에게 적합해야함
2. 잠재적으로 특이한 feature을 가진 다양한 예측 문제
3. 효율적이고 자동화 된 평가 및 비교 수단이 필요하며, 성능이 저하 될 때를 감지할 수 있어야함

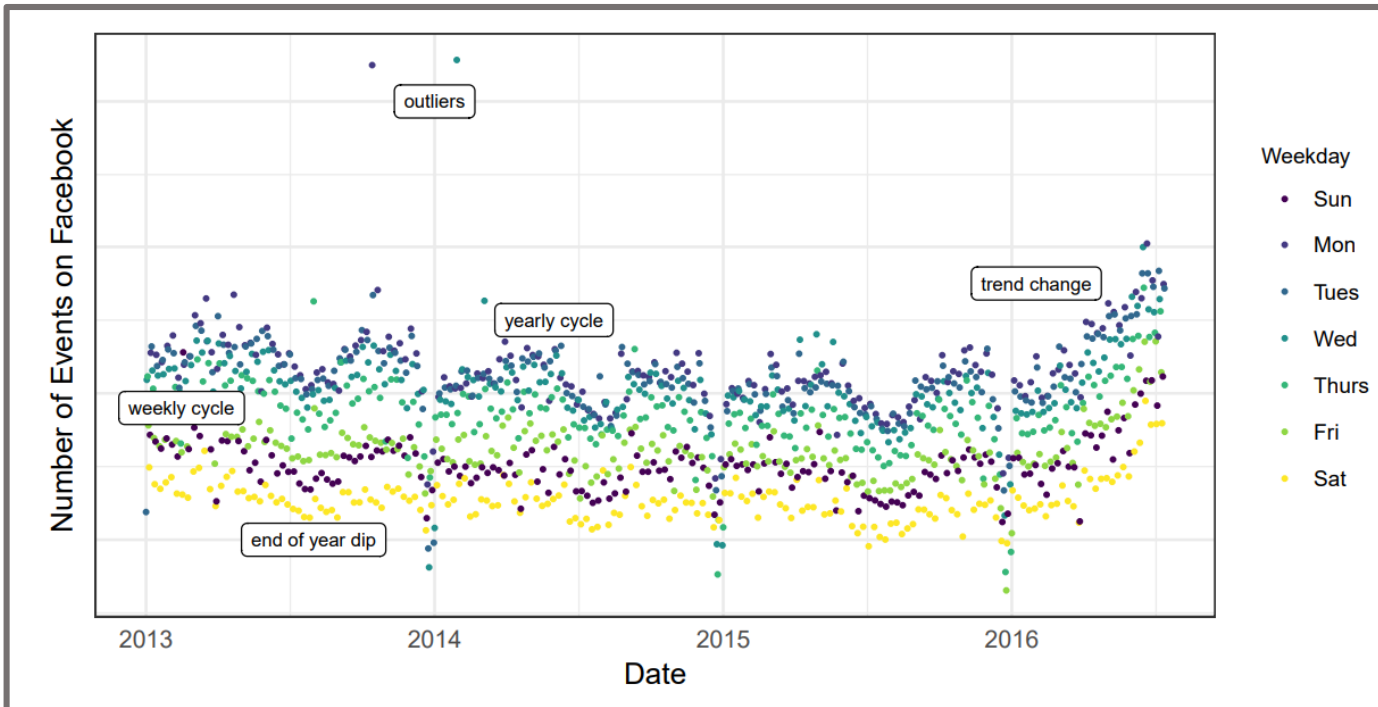


**analyst-in-the-loop approach**

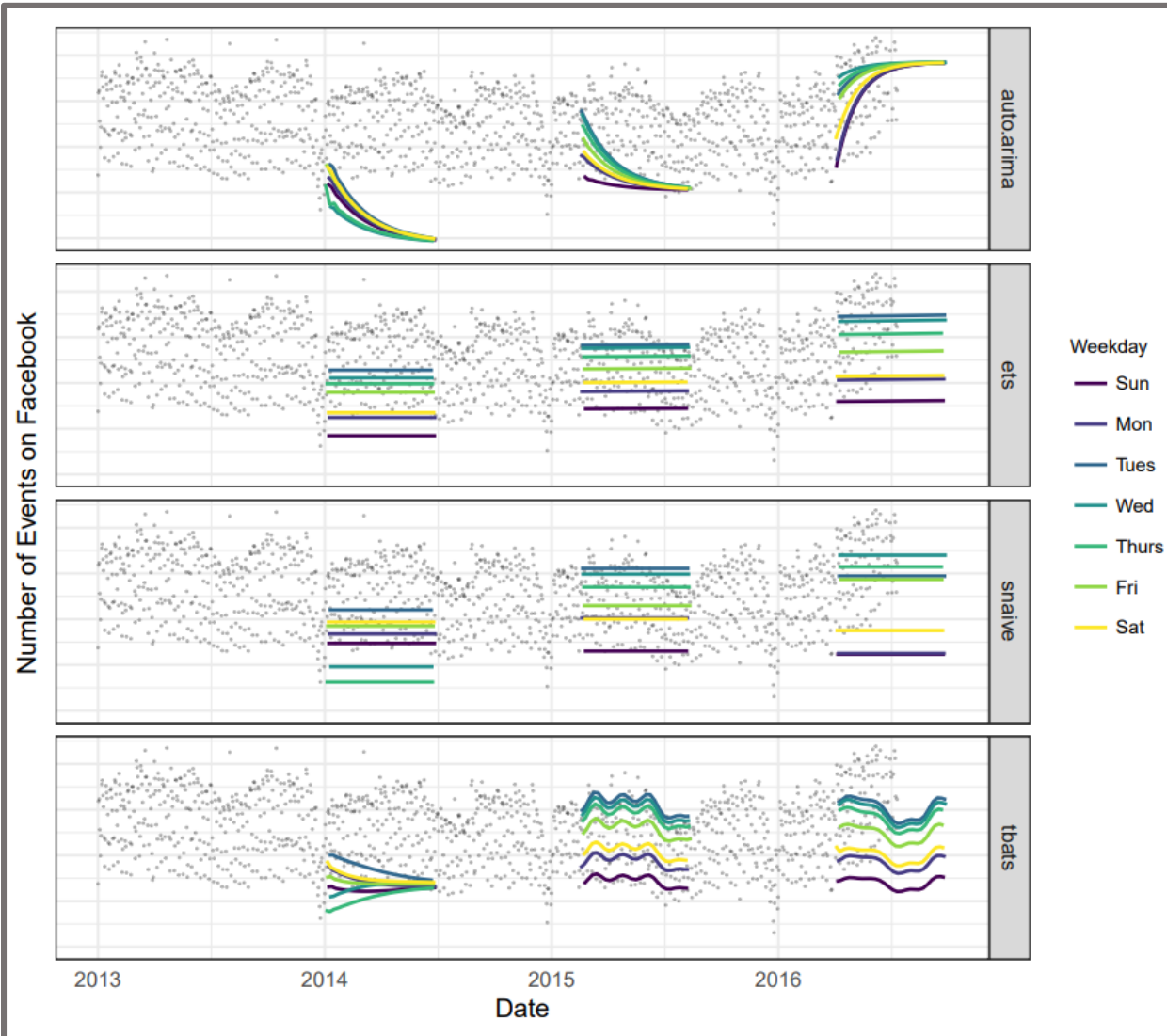
# Features of business time series

많은 비즈니스 예측 문제점에는 공통되는 feature들이 있음

- 이 시계열에서는 매주 및 매년 주기와, 크리스마스 및 신년을 전후로 확연히 볼 수 있는 몇 가지 계절 효과가 명확하게 표시됨
- 최근 6개월 동안의 추세 변화가 뚜렷하게 나타나는데, 이는 새로운 제품 또는 시장 변동에 의해 영향을 받는 시계열에서 발생
- outlier가 있을 수 있음



# Features of business time series



R 패키지를 사용한 시계열 예측 (Baseline)

## ARIMA 모델

cut-off 기간 근처에 추세가 변경되고,  
계절성을 포착하지 못하면서 추세 오류가 커짐

## Exponential smoothing

매주 계절성을 포착하지만 장기 계절성을 놓침

## Seasonal naïve

매주 계절성을 포착하지만 장기 계절성을 놓침

## TBATS 모델

모든 방법은 yearly seasonality를 적절하게 모델링  
하지 못하기 때문에 end-of-year dip에 과민 반응함

# The Prophet Forecasting Model

## Prophet

- 비즈니스 시계열의 일반적인 기능을 처리하도록 설계된 시계열 예측 모델.
- 모델의 세부 사항을 몰라도 조정이 가능하도록 직관적인 파라미터를 가짐
- 세 가지 주요 모델 구성 요소를 가진 분해 가능한 시계열 모델: trend, seasonality, and holidays

$$y(t) = g(t) + s(t) + h(t) + \epsilon_t.$$

$g(t)$  : 시계열 값의 비주기적 변화를 모델링하는 추세 함수

$s(t)$  : 주기 변화

$h(t)$  : 1일 이상에 걸쳐 잠재적으로 불규칙한 일정에 발생하는 휴일의 영향

$\epsilon_t$ : 모형에 의해 수용되지 않는 모든 특이적 변화 (정규 분포)

# The Prophet Forecasting Model

## Prophet

- generalized additive model (GAM) 과 유사함
  - 잠재적으로 non-linear smoother가 적용된 회귀모델
  - 우리는 시간을 regressor로 사용하지만 시간의 선형 및 비선형 함수를 구성 요소로 사용가능
  - 계절성의 새로운 원천이 확인되면 쉽게 분해되고 필요에 따라 새로운 구성 요소를 수용가능
- 우리는 예측 문제를 curve-fitting 문제로 framing하고 있는데, 이는 데이터에서 시간 의존 구조를 명시적으로 설명하는 시계열 모델과 본질적으로 다름

## Practical advantages of Prophet:

1. 유연성 : 계절성을 여러 기간으로 쉽게 수용할 수 있으며 분석가가 추세에 대해 다른 가정 가능
2. ARIMA 모델과 달리 측정 값은 규칙적으로 간격을 둘 필요가 없으며, 누락된 값을 보간할 필요가 없음 (아웃라이어 제거함)
3. 빠름
4. 분석가가 변경할 수 있는 쉽게 해석 가능한 매개 변수

# The Prophet Forecasting Model

## The trend model (saturating growth model and a piece wise linear model)

### Nonlinear, Saturating Growth

- 성장 예측의 경우, 데이터 생성 프로세스의 핵심 구성 요소는 인구가 어떻게 계속 증가 할 지를 예상하는 모델

→ 일반적으로 로지스틱 성장 모델을 사용하여 모델링

$$g(t) = \frac{C}{1 + \exp(-k(t - m))}$$

C the carrying capacity, k the growth rate, and m an offset parameter.

Carrying capacity 는 상수X.

→ fixed capacity C를 time-varying capacity C(t)로 변경

The growth rate k 는 상수X

→ 성장률이 변화할 수 있는 변화점을 명시적으로 정의하여 성장 모델에 추세 변화를 통합한다.



# The Prophet Forecasting Model

**The trend model (saturating growth model and a piece wise linear model)**  
**Nonlinear, Saturating Growth**

The piecewise logistic growth model:

$$g(t) = \frac{C(t)}{1 + \exp(-(k + \mathbf{a}(t)^\top \boldsymbol{\delta})(t - (m + \mathbf{a}(t)^\top \boldsymbol{\gamma})))}.$$

,where 
$$\gamma_j = \left( s_j - m - \sum_{l < j} \gamma_l \right) \left( 1 - \frac{k + \sum_{l < j} \delta_l}{k + \sum_{l \leq j} \delta_l} \right)$$

# The Prophet Forecasting Model

## The trend model (saturating growth model and a piece wise linear model) Linear Trend with Changepoints

- 포화 성장을 나타내지 않는 문제를 예측하기 위해, 부분별 일정한 성장률은 낙관적이고 유용한 모델을 제공

$$g(t) = (k + \mathbf{a}(t)^\top \boldsymbol{\delta})t + (m + \mathbf{a}(t)^\top \boldsymbol{\gamma})$$

## Automatic Changepoint Selection

- Changepoint  $s_j$ 는 알려진 제품 출시 날짜 및 기타 성장 변경 이벤트를 사용하여 분석가가 지정하거나, 자동으로 선택 될 수 있음
- 종종 많은 수의 changepoint를 지정하고,  $\delta_j \sim \text{Laplace}(0, \tau)$ 를 사용함
- 조정 전에 sparse한  $\delta$ 는 primary growth rate  $k$ 에 영향을 미치지 않으므로,  $\beta$ 가 0이 되면 적합치가 표준으로 감소한다.  
→ (not-piecewise) logistic or linear growth.

# The Prophet Forecasting Model

## The trend model (saturating growth model and a piece wise linear model)

### Trend Forecast Uncertainty

- 각각은 rate change  $\delta_j \sim \text{Laplace}(0, \tau)$  를 가짐
- 데이터에서 추론된 분산으로  $\tau$  를 대체함으로써, 과거의 것을 모방하는 미래 rate change를 시뮬레이션함
- The maximum likelihood estimate of the rate scale parameter:  $\lambda = \text{mean}(\text{abs}(\delta_j))$
- Changepoints의 평균 빈도는 history와 일치함

$$\forall j > T, \quad \begin{cases} \delta_j = 0 \text{ w.p. } \frac{T-S}{T}, \\ \delta_j \sim \text{Laplace}(0, \lambda) \text{ w.p. } \frac{S}{T}. \end{cases}$$

# The Prophet Forecasting Model

## The trend model (saturating growth model and a piece wise linear model) Seasonality

- t의 주기 함수인 계절성 모델을 지정
- 주기적인 효과에 유연한 모델을 만들기 위해 Fourier-series를 활용함
- 임의의 smooth seasonal effects를 근사가능

$$s(t) = \sum_{n=1}^N \left( a_n \cos \left( \frac{2\pi n t}{P} \right) + b_n \sin \left( \frac{2\pi n t}{P} \right) \right)$$

- 계절성을 맞추기 위해 2N 개의 매개 변수  $\beta = [a_1, b_1, \dots, a_N, b_N]^T$ 를 추정
- 계절성 요소 :  $s(t) = X(t)\beta$ .
- $\beta \sim \text{Normal}(0, \sigma^2)$  (계절성에 앞서 smoothing을 적용하기 위함)
- 연간 및 주간 계절성에서는 각각  $N = 10$ 과  $N = 3$ 이 대부분의 문제에서 잘 작동

# The Prophet Forecasting Model

## The trend model (saturating growth model and a piece wise linear model) Holidays and Events

휴일과 이벤트는 많은 비즈니스 시계열에 충격을 제공하며 주기적인 패턴을 따르지 않으므로 그 효과는 부드러운 사이클로 잘 모델링되지 않음  
→ 특정 휴가가 시계열에 미치는 영향은 종종 해마다 비슷하므로 이를 예측에 포함시키는 것이 중요

- 휴일 목록을 모델에 통합하는 것은 휴일의 영향이 독립적이라고 가정하여 간단하게 적용시킴  
→  $D_i$  : 각 휴가마다의 휴일의 과거와 미래 날짜 set  
→ 시간  $t$ 가 휴일  $i$ 인지 여부를 나타내는 지표 함수를 추가하고, 각 휴일에 예측의 해당 변화인 매개 변수  $\delta_i$ 를 할당

$$Z(t) = [1(t \in D_1), \dots, 1(t \in D_L)]$$

$$h(t) = Z(t)\kappa.$$

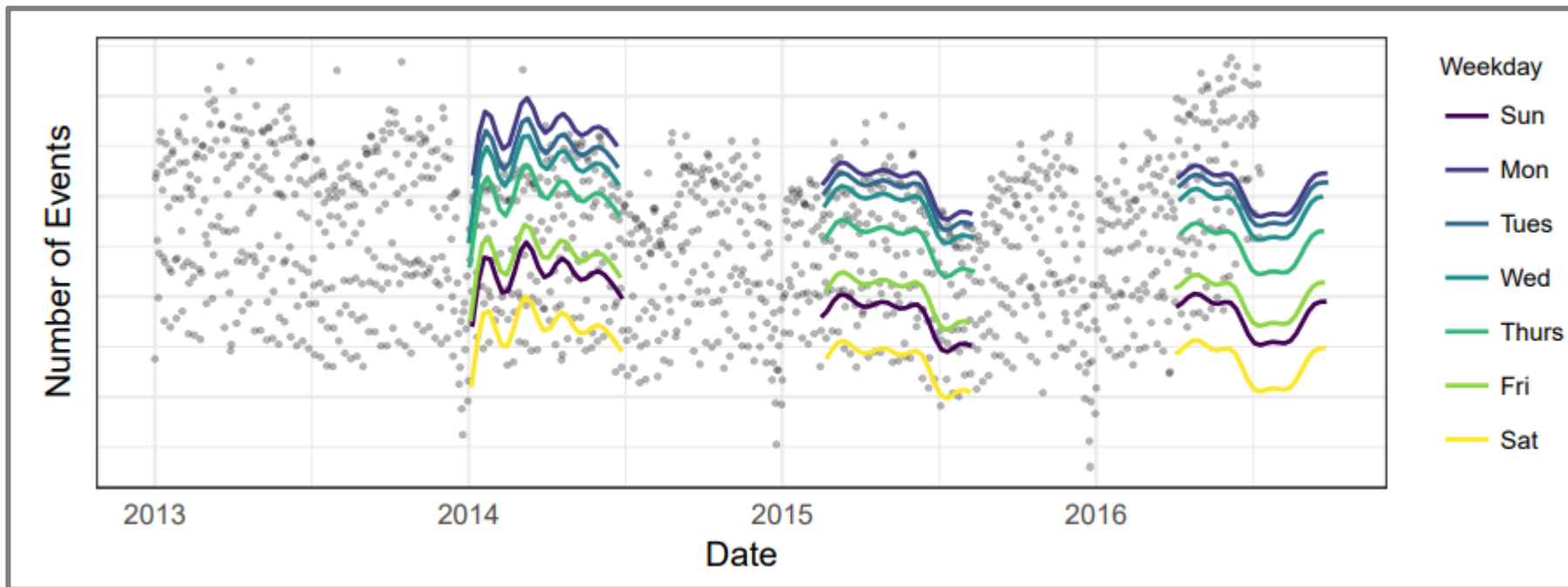
$$\kappa \sim \text{Normal}(0, \nu^2).$$

특정 휴일을 전후하여 며칠 동안 효과를 포함하는 것이 중요  
(휴일을 둘러싼 날에 대한 추가 매개변수를 포함)

# The Prophet Forecasting Model

**The trend model (saturating growth model and a piece wise linear model)  
Model Fitting**

- 주간 및 연간 계절성을 모두 예측할 수 있으며 baseline과 달리 첫 해의 holiday dip에 과민 반응X



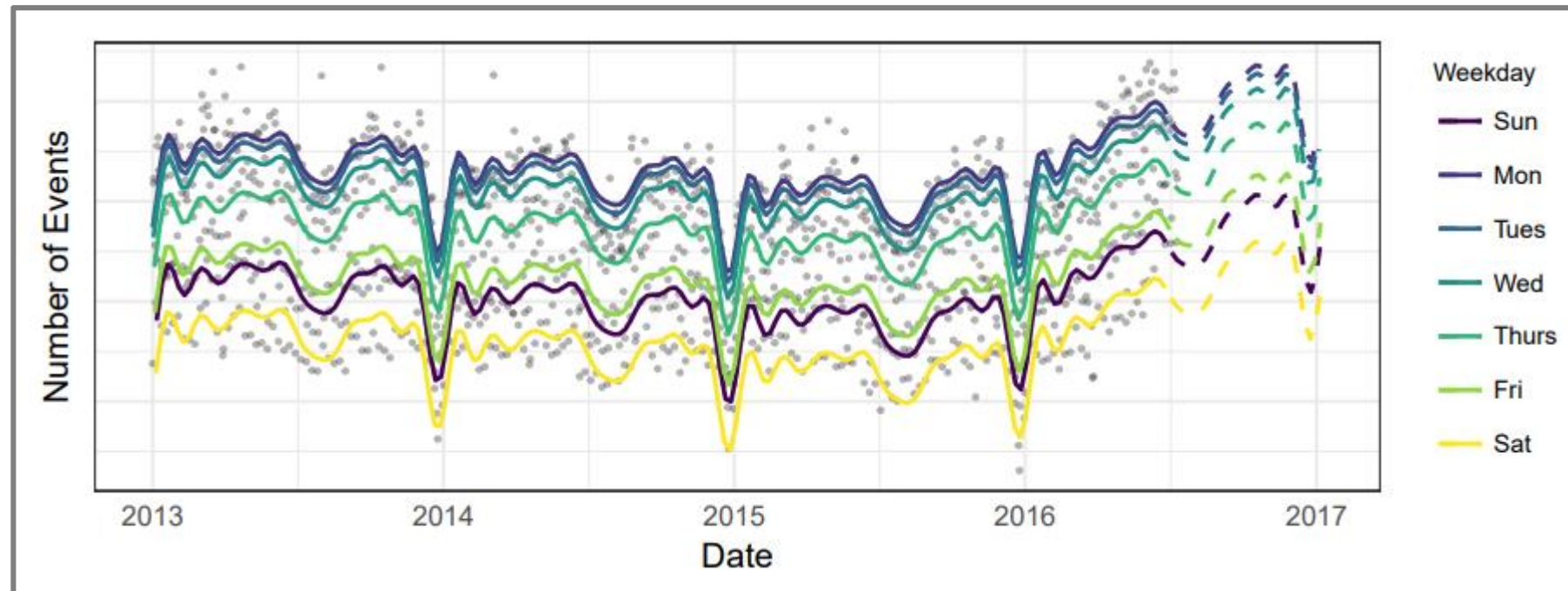
**Prophet model forecasts made on the same three dates as baseline:**

# The Prophet Forecasting Model

The trend model (saturating growth model and a piece wise linear model)

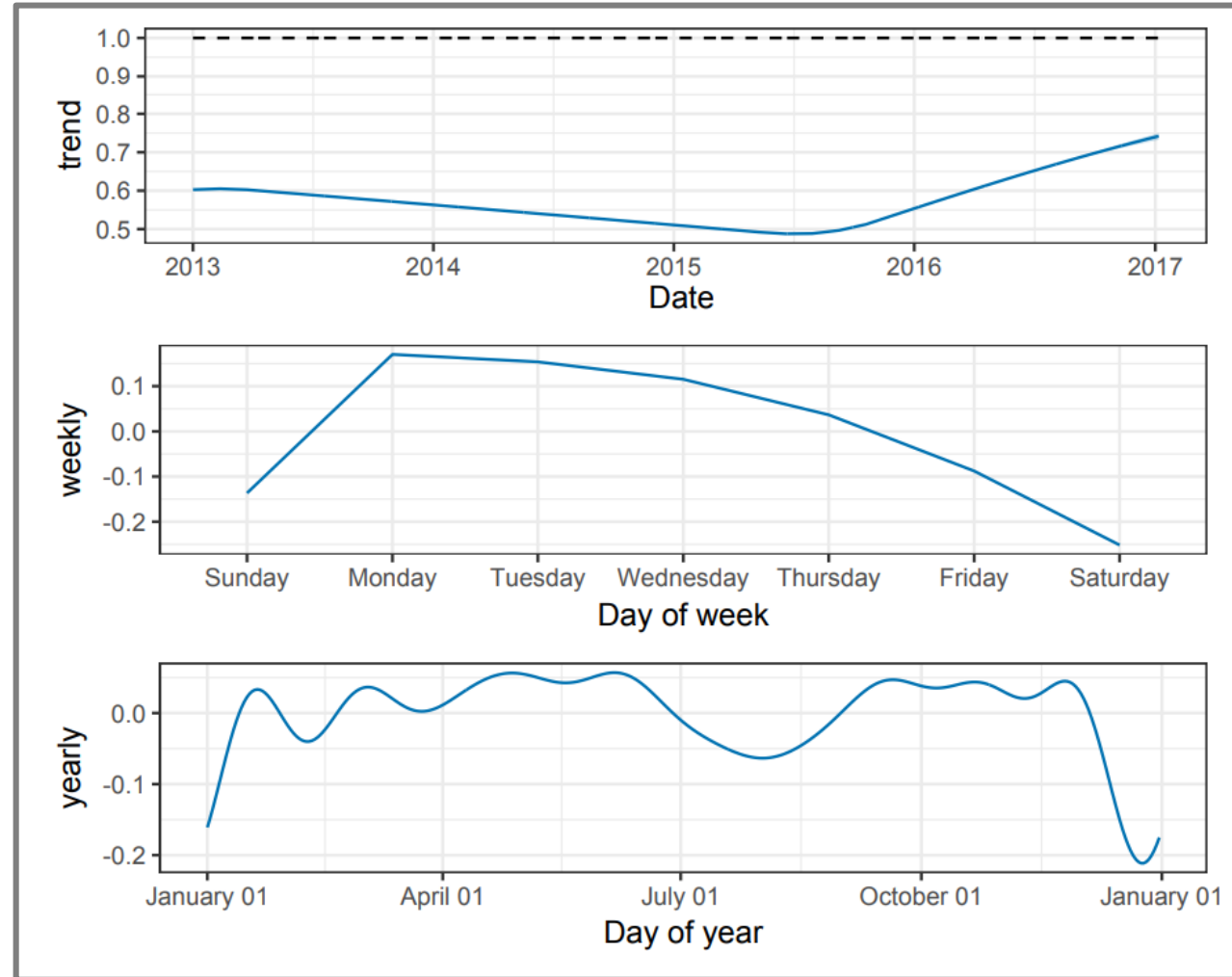
Model Fitting

- 추세 변화를 보여줌



Forecast incorporating the most recent three months of data

# The Prophet Forecasting Model



$\tau$  와  $\sigma$ 는 모델의 changepoint와 계절성에 대한 regularization을 제어함

예측의 각 요소를 따로 볼 수 있게 해줌

The trend, weekly seasonality, and yearly seasonality components



# The Prophet Forecasting Model

## The trend model (saturating growth model and a piece wise linear model)

### Analyst-in-the-Loop Modeling

- Prophet에는 분석가가 모델을 변경하여 해당 전문 지식 및 외부 지식을 적용가능:
  - 1) Capacities, 2) Changepoints, 3) Holidays and seasonality,
  - 4) Smoothing parameter
- $\tau$  매개 변수는 추세 유연성을 증가 시키거나 감소시키고  $\sigma$ 는 계절성 구성 요소의 강도를 증가 시키거나 감소시킬 수 있습니다.

# Automating Evaluation of Forecasts

다양한 방법을 비교하고 수동 개입이 보장될 수 있는 예측을 식별하여,  
예측 성능 평가를 자동화하기 위한 절차:

## 1. Baseline 예측 사용

→ 기본 과정에 대해 강력한 가정을 하지만 실제로 합리적인 예측을 산출할 수 있는 단순 예측을 사용하는 것을 선호함

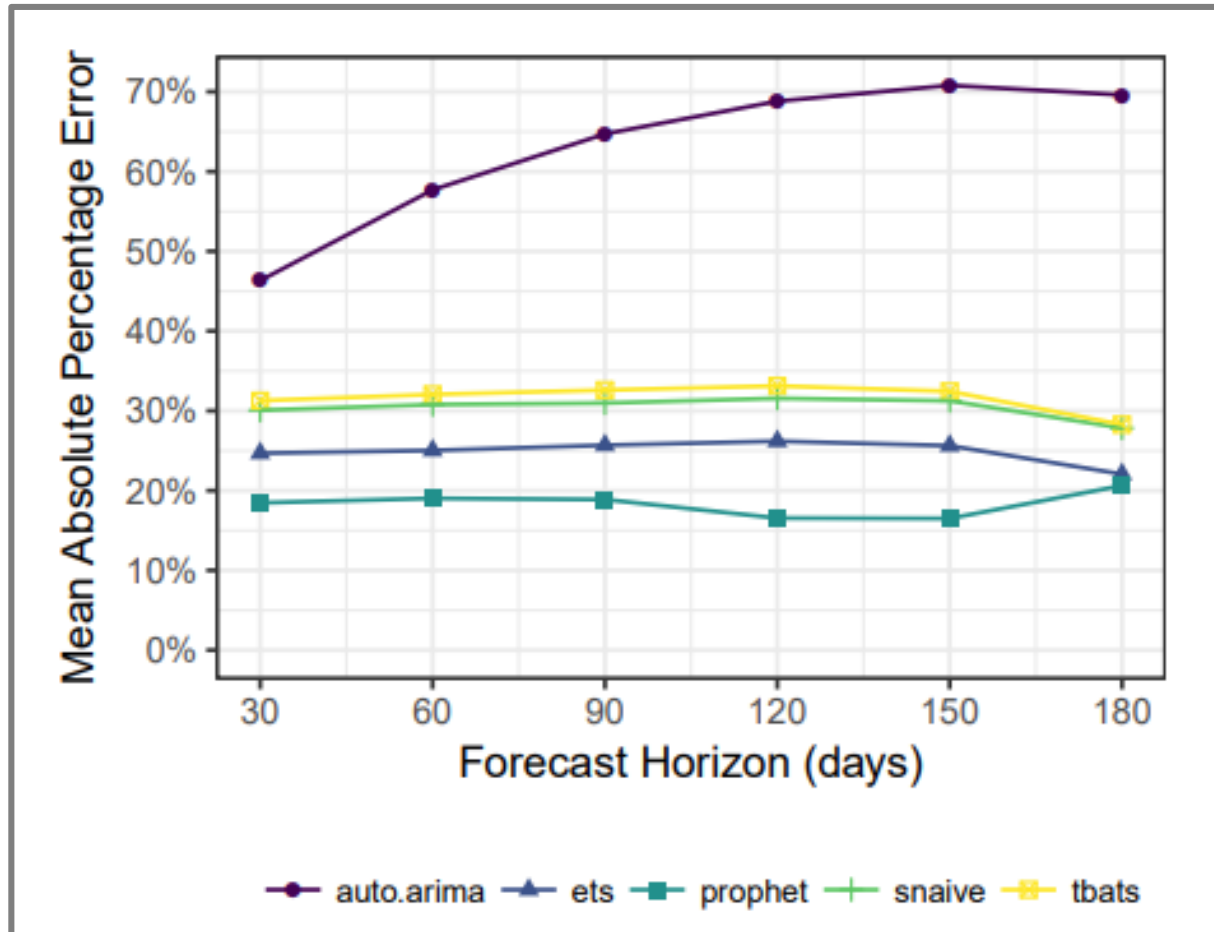
## 2. 예측 정확도 모델링

→ 일일 관측치가 있는 예측의 경우, 각각 오류와 관련된 미래 상태에 대한 최대  $H$ 개의 추정치를 산출  
→ 방법을 비교하고 성능을 추적하기 위해 예측 목표를 선언해야함

## 3. 시뮬레이션된 과거 예측

- 관측치를 교환 할 수 없기 때문에 cross-validation 과 같은 방법을 사용하기가 어려움
- 우리는 과거의 다양한 cut-off point에서  $K$ 개의 예측을 생성하여 simulated historical forecasts (SHF) 을 사용하며, 이는 수평선이 과거내에 놓이고 총 오류가 평가될 수 있도록 선택됨.
- SHF 방법론을 사용하여 예측을 평가하고 비교할 때 알아야 할 문제 :
  - 우리가 더 많은 시뮬레이션 예측을 할수록, 그들의 오차는 더 많은 상관관계를 갖는다.
  - 예측은 더 많은 데이터를 사용하면 더 잘 수행되거나 더 나쁘게 수행될 수 있습니다.

# Automating Evaluation of Forecasts



Prophet은 모든 예측 지평선에서 예측 오차가 더 낮음

Smoothed mean absolute percentage errors

# Automating Evaluation of Forecasts

다양한 방법을 비교하고 수동 개입이 보장될 수 있는 예측을 식별하여, 예측 성능 평가를 자동화하기 위한 절차:

## 4. 큰 예측 오류 식별

- SHF를 사용하여 예측과 관련된 문제를 식별
  - 예측이 baseline에 비해 큰 오차를 갖는 경우 모델을 잘못 지정됨
    - 추세 모델 및 계절성 조정
  - 특정 날짜의 모든 방법에 대한 큰 오류는 이상치를 암시
    - 이상치를 제거
  - 예측 방법에 대한 SHF 오류가 one cut-off에서 다음 cut-off로 급격히 증가하면, 데이터 생성 프로세스가 변경되었음을 나타냄
    - Changepoint를 추가하거나 다른 단계를 별도로 모델링

# Conclusion

## Prophet

- 페이스북에서 다양한 데이터를 예측하기 위해 여러 번 반복하여 개발한 새로운 모델
- 예측 정확도를 측정하고 추적하고 예측을 표시하는 시스템
- 분석가가 점진적으로 개선 할 수 있도록 시스템을 수동으로 점검해야함