 doi:10.20944/preprints201708.0086.v1

*Article*

# Forecasting Algorithms for Recurrent Patterns in Consumer Demand

**Tetiana Boiko [1,†,‡,*], Oleg Karpenkov [2,‡] and Bulat Rakhimberdiev [3,‡]**

[1] University of Liverpool; t.boiko@liverpool.ac.uk
[2] University of Liverpool; karpenk@liverpool.ac.uk
[3] CheckMobile GmbH; rakhimberdiev@mail.ru
* Correspondence: t.boiko@liverpool.ac.uk; Tel.: +44-151-794-9972
† Current address: University of Liverpool, Mathematical Sciences Building, Liverpool L69 7ZL, United Kingdom
‡ These authors contributed equally to this work.

**Abstract:** In this paper we develop a forecasting algorithm for recurrent patterns in consumer demand. We study this problem in two different settings: pull and push models. We discuss several features of the algorithm concerning sampling, periodic approximation, denoising and forecasting.

**Keywords:** seasonality; forecasting; pull and push models; denoising

---

## 1. Introduction

In this paper we develop forecasting techniques for recurrent patterns in consumer demand. There are two different models of store operation which we deal with, they are as follows.

- *Seasonal pull model:* In this model the store immediately replenish a necessary amount of items for the next unit of time.
- *Seasonal push model:* In this model the store order items several units of time in advance. If one uses pull plan in such settings then storage loss expectation will be higher.

While developing forecasting models we take into account the following two types of losses.

- *Out-of-stock loss*: here the company looses customers due to shortage of items in the storage.
- *Storage loss*: storing exceeding number of items is also expensive. For instance this is essential for the market of electronics when the demand of old models strongly decreases when new models are offered for sale.

In both cases of seasonal pull and push models we aim to find the best forecast that minimizes the loss functional. In this paper we develop a method to find one of the minimizers of the total loss functional explicitly. We show that the set of values of such forecast is a subset of a certain finite set. Then this minimizer can be found by brute force case study (see Theorem 2). This method is faster than standard gradient algorithms for small values of $k$.

**Main stages of seasonal storage push and pull forecasting.** Both push and pull models follow the same general pattern.

Given an observed history of trading we complete the following stages.

 doi:10.20944/preprints201708.0086.v1



- **Sampling and normalization**: On that stage we synthesize a certain set of periodic functions from the input data, in addition we normalize every such function such that the sum over the period equals 1. We call such functions the *normalized periodic particles*. Here we use a non-probability stratified sampling, although on practice another types of sampling might be of used (e.g., see (R.L. Chambers, C.J. Skinner 2003)).
- **Periodic approximation**: First of all for both seasonal pull and push models we introduce a loss forecast functional on the spaces of all normalized periodic particles, constructed above. This functional essentially depends on the set of normalized periodic particles. Further, using the result of Theorem 2, we find best fitting periodic approximations that minimize such functionals.
- **Denoising**: Once a best fitting periodic approximation is computed, we can use it to denoise the set of normalized periodic particle, namely we introduce a weighted system on this set. A further normalized periodic particles from the best fitting periodic approximation is, a smaller weight it has. This helps to soften the noisy effects of sampling (for a general theory of denoising in finance and economics we refer to (R. Gençay, F. Selçuk, B.J. Whitcher 2001)).
- **Forecasting**: So we have constructed the set of normalized particles equipped with denoising weights. Now for both seasonal pull and push models we introduce a weighted loss forecast functional on the spaces of all normalized periodic particles. Again, using Theorem 2, we find a best forecast that minimize these functionals.
- **Rescaling**: The obtained best forecasts are computed for normalized periodic particles. So we should rescale back in order to get a true forecast. Here we should make a forecast for the sum of the values for one period. This can be done either iteratively applying the above techniques to averaging functions or by using some non-seasonal methods of forecasting (see Remark 9).

**A few words on relation to the existing forecasting models.** Recall several central forecasting methods that are in actual use and have proved to be successful when applied correctly.

- **Causal Methods:** *regression analysis, multiple regression, qualitative regression, logit and probit.* The classical approach to time series forecasting derives from regression analysis. The standard regression model involves specifying a linear parametric relationship between a set of explanatory variables and the dependent variables of the model. The parameters of the model can be estimated in a variety of ways, for example, by least squares method introduced by Gauss in 1794 or more "modern" approaches introduced by N. Wiener (see (N. Wiener 1949)) and A. Kolmogorov (see (A. Kolmogorov 1941)).
- **Time series methods:** *moving average, trend and seasonal decomposition, Box-Jenkins, ARIMA and exponential smoothing.* Methods such as seasonal decomposition, Box-Jenkins and ARIMA are designed to extract seasonal and other cyclical component signals from a series by means of an iterated finite moving average procedure (see (G.E.P. Box, G.M. Jenkins and G.C. Reinsel 1994), (D.F. Findley, B.C. Monsell, M.C. Otto, W.R. Bell, M. Pugh 1992), and for a general overview (D.F. Findley, B.C. Monsell, W.R. Bell, M.C. Otto 1996)). C.C. Holt (C.C. Holt 1957) and P.R. Winters (P.R. Winters 1960) further generalized this method to include a slope component in the forecast function. A few years later Brown (R.G. Brown 1963) reformulated the problem in terms of a discounted least squares regression (linear exponential smoothing).
- **Qualitative Methods:** *Delphi method, sales force composite and consumer market survey.* These methods are based on the judgement and opinions of other observers, they are mostly used when there is not enough data to support quantitative methods (for more details see (H.A. Linstone, M. Turoff 1975)).

According to this classification, the forecasting techniques we develop in this paper are *seasonal time series methods*. The main difference between the listed time series methods and our methods are

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 25 August 2017 doi:10.20944/preprints201708.0086.v1



as follows. First of all we do not assume the error term to be a random variable like in ARIMA, instead of that we deal with functional space of all samples optimizing a certain functional on it. Secondly we introduce a special weighted system (which is distinct to the mean absolute percentage error) in order to find the best fitting periodic approximation, which denoises the original data. The proposed techniques were successfully implemented in a chain of stores selling consumer electronics.

**This paper is organized as follows.** In Section 2 we describe the input data and describe sampling we employ. Further in Section 3 we define best seasonal pull plan. In Section 4 we introduce the push plan which generalizes the pull plan. We introduce a denoising weight system in order to reduce noisy nonseasonal effects on the space of normalized periodic particles in Section 5. In Section 6 we show how to find minimizers of periodic $\sigma$-discrepancy which is the key point to construct a best weighted seasonal pull plan. Further in this section we discuss criteria of uniqueness of best weighted seasonal pull plans. In Section 7 we show how to find explicitly the minimizers in the push model. Basing on the results of Sections 6 and 7 we introduce the algorithm for explicit computation of best weighted seasonal push plan in Section 8. We conclude this paper in Section 9 with a few examples.

## 2. Input data, sampling, and periodic seasonal approximation

In this section we give basic notions and definitions. In Subsection 2.1 we set the input data for the models we study in this articles. Further in Subsection 2.2 we show how to construct normalized periodic particles, this is a standard sampling procedure used for approximation. Finally in Subsection 2.3 we define the space of $P$-periodic seasonal approximations.

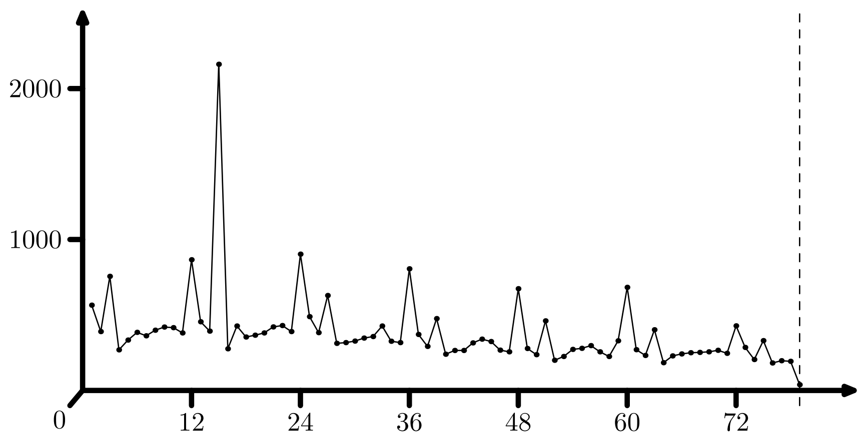
### *2.1. Input data*

The seasonal predictions described in this article are computed basing on the following *input data*:

- $T$: a *total number of observations* ($T$ is a positive integer);
- $P$: a *seasonal period* is a number of observations which we consider as a period ($P$ is a positive integer such that $P \leq T$);
- $p_I$: an *out-of-stock loss* value is the price that we pay for one out-of-stock loss ($p_I \geq 0$);
- $p_{II}$: a *storage loss* value is the price that we pay to store one item per one time unit ($p_{II} \geq 0$);
- $f : \{1, 2, \ldots, T\} \to \mathbb{R}_{\geq 0}$: an *observation data* is a function whose value $f(t)$ is the number of items sold between observations $t$ and $t+1$ (for $t = 1, \ldots, T$). Additionally we require that $f$ does not have $P$ consequent zero values.

Denote this input data by $(T, P, p_1, p_2, f)$.

**Example 1, part 1 of 2.** The input data of this example is a sample of real-life data from one company which is a chain of stores selling home electronics. The observation function $f$ is:



Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 25 August 2017 doi:10.20944/preprints201708.0086.v1

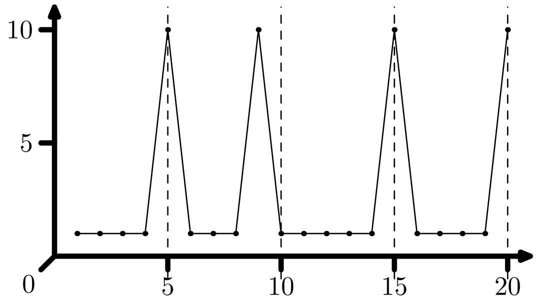


It contains 79 observations, and hence $T = 79$. There were recorded on a monthly basis, i.e. $P = 12$. The out-of-stock and storage loss values were set by the company as $p_I = 1$ and $p_{II} = 1$.

**Remark 1.** *While studying real-life examples of observation functions it is clear most of them have strong seasonality behaviour, like with one described in Example 1. Nevertheless the pikes might occur at the neighbouring time segments in different years (as discussed in Example 2 part 1 below), which makes them far from being periodic in the classical sense.*

Throughout this paper we will go through the following example.

**Example 2, part 1 of 6.** The observation function is




$$f = (1,1,1,1,10;1,1,1,10,1;1,1,1,1,10;1,1,1,1,10).$$

(Here and below we write functions as sequences of values, assuming that the first element of the sequence is the value at time 1, unless otherwise stated.) The observation function is defined for $t = 1,\ldots,20$ (hence $T = 20$). The seasonal period $P$ is considered to be 5. Finally set $p_I = 3$ and $p_{II} = 1$. So the input data is

$$\Big(20,5,3,1,(1,1,1,1,10;1,1,1,10,1;1,1,1,1,10;1,1,1,1,10)\Big).$$

Our task is to *compute a best weighted seasonal* $[20,26]$*-pull plan* (*i.e., a best weighted seasonal pull plan at time 20 till time 26*).

In this example we observe a seasonal behavior, although the observation function is not periodic in classical sense, we have

$$\max_{t\in\{1,\ldots,15\}} \big(|f(t+5) - f(t)|\big) = |f(10) - f(5)| = 9,$$

which is comparable to the maximum of the function itself.

*2.2. Sampling: normalized periodic particles*

There are several sampling strategies for construction of approximations (REFERENCES). In our models we use the following natural sampling.

**Definition 1.** *Given a triple* $(T,P,f)$*. For every integer* $i$ *such that* $1 \le i \le T-P+1$ *consider a periodic function* $f_i : \mathbb{Z} \to [0,1]$ *with period* $P$ *defined as follows:*

$$f_i(t) = \frac{f(t)}{S_i}, \quad \text{for } t = i,\ldots,i+P-1,$$

*where*

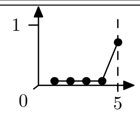
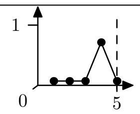
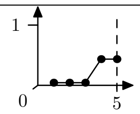
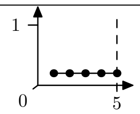
$$S_i = \sum_{k=i}^{i+T-1} f(k).$$

 doi:10.20944/preprints201708.0086.v1



*At all other values of the argument the function $f_i$ is defined by periodicity: $f_i(t+P) = f_i(t)$. The function $f_i$ is said to be a normalized periodic particle of a triple $(T, P, f)$. (It is important here that $f$ is not identical to the zero function at any consequent $P$ integers.)*

**Example 2, part 2 of 6.** In our testing example we will have the following normalized periodic particle types of normalized periodic particles:

| No. | Period | Graphs | Amount | Particles |
|---|---|---|---|---|
| I | $\left(\frac{1}{14}, \frac{1}{14}, \frac{1}{14}, \frac{1}{14}, \frac{5}{7}\right)$ |  | 10 | $f_1, \ldots, f_4, f_{11}, \ldots, f_{16}$ |
| II | $\left(\frac{1}{14}, \frac{1}{14}, \frac{1}{14}, \frac{5}{7}, \frac{1}{14}\right)$ |  | 4 | $f_6, f_7, f_8, f_9$ |
| III | $\left(\frac{1}{23}, \frac{1}{23}, \frac{1}{23}, \frac{10}{23}, \frac{10}{23}\right)$ |  | 1 | $f_5$ |
| IV | $\left(\frac{1}{5}, \frac{1}{5}, \frac{1}{5}, \frac{1}{5}, \frac{1}{5}\right)$ |  | 1 | $f_{10}$ |

### 2.3. *P-periodic seasonal approximations*

In this subsection we describe techniques of seasonal storage loss prediction. Later in Section 5 we improve it by introducing denoising weighted for normalized periodic particles. We start with some general notation.

For integers $a, b$ satisfying $a > b$ and $h : \mathbb{Z} \to \mathbb{R}$ we formally set (similar to integration)

$$\sum_{k=a}^{b} h(t) = -\sum_{k=b+1}^{a-1} h(t).$$

Note that as a consequence we have $\sum\limits_{k=a}^{a-1} h(t) = -\sum\limits_{k=a}^{a-1} h(t)$, and hence $\sum\limits_{k=a}^{a-1} h(t) = 0$.

**Definition 2.** *Given $h : \mathbb{Z} \to \mathbb{R}_{\geq 0}$ and an integer $t_0$. A discrete distribution function for $h$ with respect to $t_0$ is the function*

$$H^{t_0} : \mathbb{Z} \to \mathbb{R}, \quad \textit{where} \quad H^{t_0}(t) = \sum_{k=t_0}^{t} h(k).$$

Let us continue with the following formal definition.

**Definition 3.** *A periodic function $g(t) : \mathbb{Z} \to [0,1]$ with period $P$ is said to be a $P$-periodic seasonal approximation. The space of $P$-periodic seasonal approximation is naturally associated with $[0,1]^P$.*

## 3. Pull plan model

In this section we describe the pull plan model. First we start with the notion of Macaulay brackets that describe the ramp function:

$$\langle x \rangle = \begin{cases} x & \text{if } x \geq 0 \\ 0 & \text{if } x < 0 \end{cases}.$$

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 25 August 2017 doi:10.20944/preprints201708.0086.v1



Now we are ready to define the loss of a prediction $g$ with respect to an observation function $f$.

**Definition 4.** *Let $p_I$ and $p_{II}$ be the out-of-stock and storage loss values, and let $t_0$ be an integer. Consider a pair of periodic functions $f, g : \mathbb{Z} \to \mathbb{R}_{\geq 0}$ with period $P$ where $f$ is not zero at least at one point. Periodic $\delta$-storage loss of $g$ with respect to $f$ is*

$$\delta_f^{t_0}(g) = p_I \cdot \langle f(t_0) - g(t_0) \rangle + p_{II} \cdot \Big( \sum_{t=t_0}^{+\infty} \langle g(t_0) - F^{t_0}(t) \rangle \Big). \tag{1}$$

*where $F^{t_0}$ is the discrete distribution function for $f$ with respect to $t_0$ (recall that $F^{t_0}(t_0) = f(t_0)$).*

**Remark 2.** *Here we consider $\delta_f^{t_0}(g)$ as a single brick to construct best $P$-periodic seasonal approximations. Note that $\delta_f^{t_0}(g)$ depends entirely on $g(t_0)$, its behavior therefore is similar to the behavior of generalized Dirac $\delta$-functions in continuous settings.*

For a $P$-periodic function $h$ we set

$$E_P(h) = \sum_{t=1}^{P} h(t). \tag{2}$$

**Definition 5.** *Given an input data $(T, P, p_1, p_2, f)$ and a $P$-periodic seasonal approximation $g$.*

- *Periodic $\delta$-discrepancy of $g$ with respect to the input data is defined as follows:*

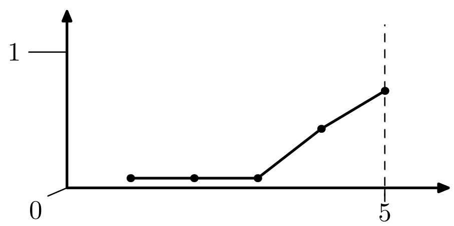
$$\Delta_f^{t_0}(g) = \frac{1}{T - P + 1} \sum_{i=1}^{T-P+1} \Big( \delta_{f_i}^{t_0}(g) \cdot E_P(f_i) \Big). \tag{3}$$

- *A periodic discrepancy of $g$ with respect to $f$ is*

$$\mathcal{D}_f(g) = \sum_{t=1}^{P} \Delta_f^t(g). \tag{4}$$

- *Global minimizers of $\mathcal{D}_f$ are said to be best $P$-periodic seasonal approximations.*
- *The discrepancy for a best $P$-periodic seasonal approximation is called the measure of periodicity/sesonality of the observation function $f$.*

**Example 2, part 3 of 6.** For our example the best 5-periodic seasonal approximation for the normalized periodic particles is the following function:




$$\Big( \frac{1}{14}, \frac{1}{14}, \frac{1}{14}, \frac{10}{23}, \frac{5}{7} \Big).$$

**Definition 6.** *Given an input data $(T, P, p_1, p_2, f)$, integers $t_0 \leq t_1$ and an interval (finite or infinite) $I$ such that $[t_0, t_1] \subset I$. Consider a function $g : I \cap \mathbb{N} \to [0, 1]$.*

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 25 August 2017 doi:10.20944/preprints201708.0086.v1



- *A pull plan loss at time $t_0$ till time $t_1$ is given by the following expression*

$$\Omega_f^{t_0,t_1}(g) = \sum_{t=t_0}^{t_1} \Delta_f^t(g). \tag{5}$$

- *Best seasonal pull plan at time $t_0$ till $t_1$ is a global minimizer of the functional $\Omega_f^{t_0,t_1}(g)$.*

In Example 2 we will go directly to best weighted seasonal pull plan (see further Example 2 part 4 below).

We conclude this section with two general remarks.

**Remark 3.** *Note that best P-periodic seasonal approximations do not necessarily sum up to 1 at the period. For instance this is the case in Example 2 (see Example 2 part 3). The reason for that is as follows: in order to catch the customers it is worthy to store items in some excess.*

**Remark 4.** *It is interesting to observe a continuous analog of a best P-periodic seasonal approximation. In this case all functions are defined on intervals. For continuous periodic $\delta$-discrepancy we have:*

$$\tilde{\Delta}_f^{t_0}(g) = \frac{1}{T-P} \int_0^{T-P} \left( \left( p_I \langle f_\lambda(t_0) - g(t_0) \rangle + p_{II} \Big( \int_{t_0}^{+\infty} \langle g(t_0) - f_\lambda^{t_0}(s) \rangle ds \Big) \right) \cdot \Big( \int_0^P f_\lambda(u) du \Big) \right) d\lambda,$$

*and a continuous periodic discrepancy is*

$$\int_0^P \tilde{\Delta}_f^t(g) dt.$$

*Here $f_\lambda$ is a periodic function obtained by means of periodicity starting with the function $f$ restricted to the segment $[\lambda, \lambda + P]$ and further normalized (i.e., divided by the value of the integral of $f$ over the segment $[\lambda, \lambda + P]$).*

## 4. Push plan model

In this section we briefly give main definitions for push plan model.

**Definition 7.** *Let $p_I$ and $p_{II}$ be the out-of stock and storage loss values. Let also $t_0$ and $t_1$ be two integers such that $t_1 \geq t_0$. Consider a function $f : \mathbb{Z} \to \mathbb{R}$ with an unbounded from above distribution function $F^{t_0}$ with respect to $t_0$. Let $g : \{t_0, \dots, t_1\} \to \mathbb{R}_{\geq 0}$. Then the $[t_0, t_1]$-push plan loss of $g$ with respect to $f$ is*

$$\begin{aligned} \Lambda_f^{t_0,t_1}(g) = \sum_{t=t_0}^{t_1} \Big( p_I \cdot \langle R_{f,g}^{t_0}(t) \rangle + p_{II} \cdot \langle -R_{f,g}^{t_0}(t) \rangle \Big) + \\ p_{II} \cdot \sum_{t=t_1+1}^{+\infty} \langle R_{f,g}^{t_0}(t_1) - F^{t_1}(t) \rangle. \end{aligned} \tag{6}$$

*Here the reminder function is defined iteratively*

$$\begin{aligned} &R_{f,g}^{t_0}(t_0) = g(t_0) - f(t_0); \\ &R_{f,g}^{t_0}(t+1) = \langle R_{f,g}^{t_0}(t) \rangle + g(t+1) - f(t+1), \quad \textit{for } t = t_0+1, \dots, t_1. \end{aligned}$$

*In other words*

$$R_{f,g}^{t_0}(t) = \langle \cdots \langle \langle g(t_0) - f(t_0) \rangle + g(t_0+1) - f(t_0+1) \rangle \cdots + g(t-1) - f(t-1) \rangle + g(t) - f(t)$$

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 25 August 2017 doi:10.20944/preprints201708.0086.v1



**Remark 5.** *The periodic $\delta$-storage loss in seasonal pull and push models are related by a simple formula:*

$$\delta_f^{t_0}(g) = \Lambda_f^{t_0,t_0}(g). \tag{7}$$

*In some sense the seasonal push model generalizes the pull plan model.*

Let us now give weighted analogs in seasonal push model.

**Definition 8.** *Let $(T, P, p_I, p_{II}, f)$ be an input data as above, let $t_0, t_1$ be nonnegative integers satisfying $t_1 \geq t_0$. Let also $f_i$ be normalized periodic particles for $i = 1, \ldots, T - P + 1$.*

- *Consider $g : \{t_0, \ldots, t_1\} \to \mathbb{R}_{\geq 0}$. Then the seasonal $[t_0, t_1]$-push plan loss of $g$ at time $t_0$ is as follows:*

$$\mathcal{L}_f^{t_0,t_1}(g) = \frac{1}{T-P+1} \sum_{i=1}^{T-P+1} \Big( \Lambda_{f_i}^{t_0,t_1}(g) \cdot E_P(f_i) \Big). \tag{8}$$

- *Best seasonal $[t_0, t_1]$-push plan are global minimizers of $\mathcal{L}_f^{t_0,t_1}$.*

## 5. Weighted seasonal predictions

In Example 2 we have spotted one serious problem with the methods described above. Some of the normalized periodic particles are noisy, they are rather far from an average normalized periodic particle. In particular a from common sense suggests that the constant normalized periodic particle of type **IV**: $(1/5, 1/5, 1/5, 1/5, 1/5)$ is noisy (see Example 2, part 2). The total sum here is 5, it is not 14 as expected. So it has a noisy contribution to every value of the period. In fact the noise of such normalized periodic particles can be reduced by the denoising techniques described in this section.

We introduce weights to seasonal push and pull plan models in Subsections 5.1 and 5.2 respectively. Further in Subsection 5.3 we discuss a particular normalized periodic particle denoising which we use in our weighted seasonal pull and push plan models.

### *5.1. Weighted seasonal pull plan*

Let us give the following general definition.

**Definition 9.** *Given an input data $(T, P, p_1, p_2, f)$, integers $t_0 \leq t_1$, and a finite or infinite interval $I$ such that $[t_0, t_1] \subset I$. Let $\mu = (\mu_1, \ldots, \mu_{T-P+1})$ be a collection of positive numbers. Consider a function $g : I \cap \mathbb{N} \to [0, 1]$.*

- *Weighted periodic $\delta$-discrepancy of $g$ with respect to the input data is defined as follows:*

$$\Delta_f^{t_0,\mu}(g) = \sum_{i=1}^{T-P+1} \Big( \mu_i(f) \delta_{f_i}^{t_0}(g) \cdot E_P(f_i) \Big). \tag{9}$$

- *A weighted $[t_0, t_1]$-pull plan loss at time $t_0$ till time $t_1$ is given by the following expression*

$$\Omega_f^{t_0,t_1,\mu}(g) = \sum_{t=t_0}^{t_1} \Delta_f^{t,\mu}(g). \tag{10}$$

- *A best weighted seasonal $[t_0, t_1]$-pull plan is a global minimizer of the functional $\Omega_f^{t_0,t_1,\mu}(g)$.*

**Remark 6.** *Note that the standard seasonal pull plan loss is the weighted seasonal pull plan loss with all weights being equal to $\frac{1}{T-P+1}$.*

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 25 August 2017 doi:10.20944/preprints201708.0086.v1



### 5.2. Weighted seasonal push plan

Consider a similar general definition for weighted seasonal push settings.

**Definition 10.** *Given an input data $(T, P, p_1, p_2, f)$, integers $t_0 \leq t_1$. Let also $f_i$ be normalized periodic particles with distributions functions $F_i$ where $i = 1, \ldots, T-P+1$. Consider a collection of positive numbers $\mu = (\mu_1, \ldots, \mu_{T-P+1})$. Fix a best P-periodic seasonal approximation $\tilde{f}$.*

- *Let $g : \{t_0, \ldots, t_1\} \to \mathbb{R}_{\geq 0}$. Then the weighted seasonal $[t_0, t_1]$-push plan loss of $g$ at time $t_0$ is as follows:*
$$\mathcal{L}_f^{t_0,t_1,\mu}(g) = \frac{1}{T-P+1} \sum_{i=1}^{T-P+1} \mu_i(f) \Lambda_{f_i}^{t_0,t_1}(g),$$
- *A best weighted seasonal $[t_0, t_1]$-push plan is a global minimizer of $\mathcal{L}_f^{\mu}$.*

### 5.3. Normalized periodic particle denoising

In our models we use the following natural noise function.

**Definition 11.** *Let $\tilde{f}$ be a best P-periodic seasonal approximation. Then the noise of a P-periodic seasonal approximation $g$ with respect to $\tilde{f}$ is defined as*

$$\Theta_{\tilde{f}}(g) = \frac{1}{1 + (\mathcal{D}_{\tilde{f}}(g))^2}.$$

**Remark 7.** *The set of all best P-periodic seasonal approximations is naturally ordered lexicographically with respect to their sequence of values. So one can always pick the smallest best P-periodic seasonal approximation with respect to lexicographical order.*

**Definition 12.** *Let $(T, P, p_I, p_{II}, f)$ be an input data as above, let $t_0, t_1$ be nonnegative integers satisfying $t_1 \geq t_0$. Let also $f_i$ be normalized periodic particles for $i = 1, \ldots, T-P+1$. Consider the smallest best P-periodic seasonal approximation $\tilde{f}$ with respect to the lexicographical order. The denoising weight of the normalized periodic particle $f_i$ is the following number*

$$\nu_i(f) = \frac{\Theta_{\tilde{f}}(f_i)}{\sum_{j=0}^{T-P+1} \Theta_{\tilde{f}}(f_j)}.$$

*for $i = 1, \ldots, T-P+1$.*

**Remark 8.** *For denoising we have used an approach similar to statistical bootstrapping. We have used the smallest best P-periodic seasonal approximation as the most probable prediction in order to generate weights for all normalized periodic particles.*

**Example 2, part 4 of 6.** In our testing example the denoising weights are as follows

| **Function No.** | **I** | **II** | **III** | **IV** |
|---|---|---|---|---|
| **Weights** $\nu_i$ | 0.076 | 0.038 | 0.057 | 0.027 |

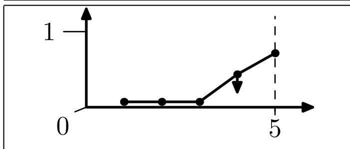
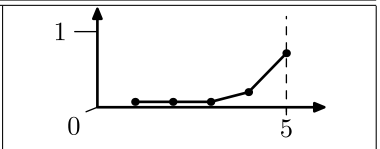
So the best weighted seasonal $[21, 25]$-pull plan for one period is

$$\left(\frac{1}{14}, \frac{1}{14}, \frac{1}{14}, \frac{1}{5}, \frac{5}{7}\right).$$

Here we have a correction at time 4:

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 25 August 2017 doi:10.20944/preprints201708.0086.v1



| **Standard best approximation** | **Best weighted** $[21,25]$**-pull plan** |
|---|---|
|  |  |
| $\left(\frac{1}{14},\frac{1}{14},\frac{1}{14},\frac{10}{23},\frac{5}{7}\right)$ | $\left(\frac{1}{14},\frac{1}{14},\frac{1}{14},\frac{1}{5},\frac{5}{7}\right)$ |

**Remark 9.** *Assume that we have computed a best weighted seasonal $[t_0,t_1]$-push plan $g$. Then one should consider*

$$E\cdot g:\{t_0,\ldots,t_1\}\to\mathbb{R}_{\geq 0},$$

*where $E$ is the expectation rate prediction of the total sum for all observations in the consequent $P$ steps.*

*The function $E$ can be computed by iteratively applying the above seasonality techniques to the averaging function formed by $E_P(f_i)$, i.e., to*

$$\big(E_P(f_1),\ldots,E_P(f_{T-P+1})\big).$$

*This function has a $T-P+1$ entry and a seasonal period $P$.*

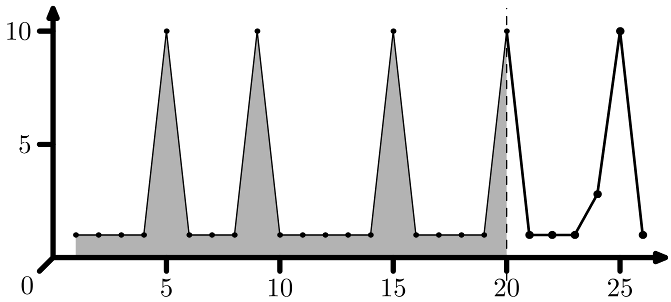
*Alternatively one can pick $E$ without referring to seasonality (we refer an interested reader to REFERENCES).*

**Example 2, part 5 of 6.** In our example **best weighted seasonal** $[21,26]$**-pull plan** is then

$$\left(\frac{1}{14},\frac{1}{14},\frac{1}{14},\frac{1}{5},\frac{5}{7},\frac{1}{14}\right).$$

In order to get a true prediction we multiply the obtained function by the expectation $E$ of a total amount of customers in during a period $P$ (as mentioned in Remark 9). In our case it is 14. Finally we have

$$\left(1,1,1,\frac{14}{5},10,1\right).$$



## 6. Properties of the weighted pull forecasting

In this section we prove main statements on weighted pull forecasting. In Subsection 6.1 we prove basic properties of weighted seasonal pull plan loss. Further in Subsection 6.2 we give a finite list which contains all the values for some best weighted seasonal pull plan. Finally in Subsection 6.3 we discuss the uniqueness of best seasonal pull plan.

### *6.1. Basic properties*

Let us collect basic properties of weighted seasonal $[t_0,t_1]$-pull plan loss in the following proposition.

**Proposition 1.** *For every integers $t_0<t_1$ we have*

*Preprints* (www.preprints.org) | NOT PEER-REVIEWED | Posted: 25 August 2017 doi:10.20944/preprints201708.0086.v1



*i) Let $g$ be a $P$-periodic seasonal approximation $g$. Then $\Omega_f^{t_0,t_0+P-1,\mu} = \mathcal{D}_f^{\mu}(g)$.*

*ii) Let $I$ be an interval containing $t_0$ and $t_1+P$ and let $g : I \cap \mathbb{N} \to [0,1]$. Then $\Omega_f^{t_0+P,t_1+P,\mu}(g) = \Omega_f^{t_0,t_1,\mu}(g)$;*

*iii) Let $I$ be an interval containing $t_0$ and $t_1+P$ and let $g : I \cap \mathbb{N} \to [0,1]$. Then $\Omega_f^{t_0,t_1+P,\mu}(g) = \Omega_f^{t_0,t_1,\mu} + \mathcal{D}_f^{\mu}(g)$.*

**Proof.** Item (i): We have

$$\Omega_f^{t_0,t_0+P,\mu}(g) = \sum_{t=t_0}^{t_0+P-1,\mu} \Delta_f^{t,\mu}(g) = \sum_{t=1}^{P,\mu} \Delta_f^{t,\mu}(g) = \mathcal{D}_f^{\mu}(g).$$

The second equality holds since both $f_i$ (for every admissible $i$) and $g$ are periodic with period $P$.

Item (ii) holds since both $f_i$ (for every admissible $i$) and $g$ are periodic with period $P$.

Item (iii): From the above two items we have

$$\Omega_f^{t_0,t_1+P,\mu}(g) = \Omega_f^{t_0,t_1,\mu}(g) + \Omega_f^{t_1+1,t_1+P,\mu}(g) = \Omega_f^{t_0,t_1,\mu}(g) + \Omega_f^{1,P,\mu}(g) = \Omega_f^{t_0,t_1,\mu} + \mathcal{D}_f^{\mu}(g).$$

This concludes the proof. □

*6.2. On values of periodic discrepancy*

The following theorem is one of the central theorems in this article.

**Theorem 1.** *Let $(T, P, p_I, p_{II}, f)$ be an input data, let $\mu$ be a collection of positive integers, and let $t_0$ be an integer. Then there exists a $P$-periodic seasonal approximation $g$ that fulfills the following two conditions:*

*— the approximation $g$ minimizes the weighted periodic $\delta$-discrepancy functional $\Delta_f^{t_0,\mu}$.*

*— the value $g(t_0)$ is contained in the union of all values of $F_i^{t_0}(t)$ for $i = 1, \dots, T-P+1$ and $t = t_0, \dots, t_0+P-1$.*

We start with the following simple statement.

**Lemma 1.** *The image of any best weighted $P$-periodic seasonal approximation is contained in* $[0,1]$.

**Proof.** It is clear that reducing the value to 1 or increasing negative value to 0 will reduce the value of the corresponding weighted periodic $\delta$-discrepancy. □

*Proof of Theorem 1.* The functional $\Delta_f^{t_0,\mu}$ is piecewise linear when we vary $g(t_0)$ and fix all the other values. In addition it is linear outside zeroes of Macaulay brackets involved in $\Delta_f^{t_0,\mu}$, see Equation (9). Zeroes of such Macaulay brackets are either at $f_i(t_0)$ or at $F_i^{t_0}(t)$ for some integer $t > 0$. Since be definition $f_i(t_0) = F_i^{t_0}(t_0)$, every non-linearity point is at $F_i^{t_0}(t)$ for some $t \geq 0$.

Since $\Delta_f^{t_0,\mu}$ is bounded from below by zero and piecewise linear, it has a global minimum on the real line. Since $\Delta_f^{t_0,\mu}$ is piecewise linear, one can choose the global minimum at the non-linearity point. As we have shown above, all such points are contained in the set of all values of $F_i^{t_0}$.

From the definition of normalized periodic particles we have

$$F_i^{t_0}(t_0 + P - 1) = 1$$

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 25 August 2017 doi:10.20944/preprints201708.0086.v1



for all admissible $i$. Hence all values of $F^{t_0}$ that are in the segment $[0,1]$ are attained at points $t_0,\ldots,t_0{+}P{-}1$ respectively. Now the statement of the theorem follows from Lemma 1. □

**Corollary 1.** *Let $(T,P,p_I,p_{II},f)$ be an input data, let $\mu$ be a collection of positive integers, and let $t_0 \le t_1$ be a pair of integers. Then there exists a best weighted seasonal $[t_0,t_1]$-pull plan $g$ such that for every admissible $\hat{t}$ the value $g(\hat{t})$ is contained in the union of all values of $F_i^{\hat{t}}(t)$ for $i=1,\ldots,T{-}P{+}1$ and for $t=\hat{t},\ldots,\hat{t}{+}P{-}1$.*

**Proof.** By definition we have

$$\Omega_f^{t_0,t_1,\mu} = \sum_{t=t_0}^{t_1} \Delta_f^{t,\mu}.$$

For every integer $t_2$ in the segment $[t_0,t_1]$ the value $\Delta_f^{t,\mu}(g)$ at $t_2$ depends only on $g(t_2)$ and does not depend on the other values of $g$ in the period. Hence $g$ minimizes $\Omega_f^{t_0,t_1,\mu}$ if and only if it minimizes every summand $\Delta_f^{t,\mu}$ in the sum. This reduces Corollary 1 to Theorem 1. □

*6.3. On uniqueness of a weighted pull plan*

We conclude this section with the following observation.

**Remark 10.** *A best weighted seasonal $[t_0,t_1]$-pull plan is uniquely defined if*

$$\left(\sum_{i=1}^{T-P+1} a_i\mu_i E_P(f_i)\right)p_I + \left(\sum_{i=1}^{T-P+1} b_i\mu_i E_P(f_i)\right)p_{II} \neq 0$$

*for all choices $a_i = 0,1$ and $b_i = -P,\ldots,P$. This directly follows from Equation 1.*

*In the case of unit weights and unit $E_P$'s, the condition of uniqueness is that $ap_I + bp_{II} \neq 0$ for all non-negative integers $a,b$, such that $|a| \le T-P+1$ and $|b| \le (T-P+1)\cdot P$. In particular if $p_I/p_{II}$ is irrational then the best weighted seasonal $[t_0,t_1]$-pull plan is uniquely defined in such settings.*

## 7. Detection of best weighted seasonal $[t_0,t_1]$-push plan

Note that the push plan loss functional $\Lambda_f^{t_0,t_1}$ is a piecewise linear function where $g(t_0),\ldots,g(t_1)$ are considered as variables. Global minima of such functions are obtained at points, where at least $t_1{-}t_0{+}2$ different linear domains come together. This gives $t_1{-}t_0{+}1$ linear equations on the variable values $g(t_0),\ldots,g(t_1)$.

**Theorem 2.** *There exists a minimizer (i.e., a best weighted seasonal $[t_0,t_1]$-push plan) of $\mathcal{L}_f^{t_0,t_1,\mu}$ which is described as an intersection point of $t_1{-}t_0{+}1$ planes of the following family:*

$$\sum_{t\in I} g(t) = \sum_{t\in I} f_i(t) + F_i^{t_1}(\tilde{t}) \tag{11}$$

*where $I$ is an arbitrary subset of the set $\{t_0,\ldots,t_1\}$, for the choice of normalized periodic particles we have $1 \le i \le T{-}P{+}1$, and the integer value $\tilde{t}$ satisfies $t_1-1 \le \tilde{t} \le t_1{+}P{-}1$. (Recall that $F_i^{t_1}(t_1{-}1)=0$.)*

**Proof.** Since $\mathcal{L}_f^{t_0,t_1,\mu}$ is piecewise linear, its minimizer is at one if the vertices of $\mathcal{L}_f^{t_0,t_1,\mu}$, i.e., at intersection of $t_1{-}t_0{+}2$ hyperplanes of nonlinearity of $\mathcal{L}_f^{t_0,t_1,\mu}$. Each of such hyperplanes is in fact a hyperplane of non-linearity for $\Lambda_{f_i}^{t_0,t_1}$ for some $i$ and it is defined by one of Equations (11).

There are two type of hyperplanes where $\Lambda_f^{t_0,t_1}$ is nonlinear. The hyperplanes of the first type are defined by

$$R_{f,g}^{t_0}(t) = 0.$$

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 25 August 2017 doi:10.20944/preprints201708.0086.v1



The hyperplanes of the second type are defined by

$$R^{t_0}_{f,g}(t) - F^{t_1}_i(\tilde{t}) = 0.$$

Similarly to the proof of Theorem 1, the time $\tilde{t}$ can be chosen from the set

$$\{t_1 - 1, \ldots, t_1 + P - 1\}.$$

Now both types are of the form of Equation (11). This concludes the proof. □

## 8. Algorithm to compute best weighted seasonal pull and push plans

Recall that we have the following stages of best weighted seasonal pull and push plan computation.

- **Data**: We start with an input data $(T, P, p_1, p_2, f)$ and integers $t_1 \geq t_0 \geq 0$.
- **Sampling and normalization:** This stage is described in Subsection 2.2 above.
- **Periodic approximation:** To compute a value at $\hat{t}$ of a best periodic seasonal approximations we use the below algorithm with equal weights and $t_0 = t_1 = \hat{t}$.
- **Denoising:** We generate denoising weights as in Subsection 5.3.
- **Forecasting:** Here we use the below algorithm with denoising weights $\nu_i$ constructed at previous stage.
- **Rescaling:** Finally we multiply the chosen best weighted seasonal $[t_0, t_1]$-push plan by the function $E$ as discussed in Remark 9.

Two of the above stages use the following algorithm.

**Algorithm of best weighted seasonal $[t_0, t_1]$-push plan computation**.

- **Data**: We start with an input data $(T, P, p_1, p_2, f)$, integers $t_1 \geq t_0 \geq 0$, and a weight system, i.e. a collection of positive integers $(\mu_1, \ldots, \mu_{T-P+1})$.
- **Task of the algorithm:** Compute a best weighted seasonal $[t_0, t_1]$-push plan.
- **Description of the algorithm:** According to Theorem 2 it is sufficient to find all intersection points of $k+1$ hyperplanes defined by some of Equations (11). In order to pick a plane from the family of Equations (11) one should choose a subset $I$, the index $i$ of a normalized periodic particle and the time $\tilde{t}$.

  For each choice of $(k+1)$-tuples of triples $(I, i, \tilde{t})$ we check if the corresponding hyperplanes are linearly independent (i.e., their intersection is a single point). Here one might consider only distinct $I$ in the triples, since otherwise the equations are linearly dependent.

  In case if they are linearly independent we compute the value of $\mathcal{L}^{t_0,t_1,\mu}_f$ at the intersection point. After that we should find the minimum of such values, it corresponds to a minimizer (i.e., to a best weighted seasonal $[t_0, t_1]$-push plan).
- **A few words on complexity of this algorithm.** A simple upper bound on the number of choices of $(k+1)$-tuples of triples $(I, i, \tilde{t})$ with distinct $I$ is

$$\big((T - P + 1)(P + 1)\big)^{t_1 - t_0 + 1} \cdot \binom{2^{t_1 - t_0 + 1}}{t_1 - t_0 + 1}.$$

  For each choice we solve a linear system on $t_1 - t_0 + 1$ variables with a $(0, 1)$-matrix (i.e., a matrix whose entries are either 0 or 1).

**Remark 11.** *Recall that a weighted seasonal $[t_0, t_1]$-pull plan $g$ is a combination of weighted seasonal $[\hat{t}, \hat{t}]$-push plans for each value $g(\hat{t})$ where $\hat{t} = t_0, t_0+1, \ldots, t_1$.*

*Preprints* (www.preprints.org) | NOT PEER-REVIEWED | Posted: 25 August 2017 doi:10.20944/preprints201708.0086.v1



**Remark 12.** *The algorithm of this section is rather effective if the number of steps is relatively small, e.g., for the pull model. The complexity grows exponentially with respect to the growth of number of steps (i.e., with respect to the growth of $t_1 - t_0$). For large number of steps a better method would be useful. In particular one can consider some greedy methods that finds the function $g$ at which $\mathcal{L}_f^{t_0,t_1,\mu}(g)$ is small but probably it is not a global minimum there.*

## 9. Examples

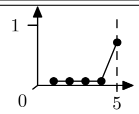
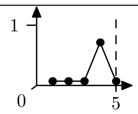
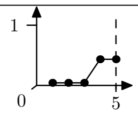
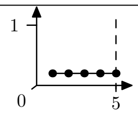
We conclude this paper with the summary of Examples 2 and the last part of Example 1s.

**Example 2, part 6 of 6: summary.**

- **Input data:**

$$\Big(T = 20, P = 5, p_I = 3, p_{II} = 1, f = (1,1,1,1,10;1,1,1,10,1;1,1,1,1,10;1,1,1,1,10)\Big).$$

- **Task:** to compute a best weighted seasonal $[21,26]$-pull plan.
- **Sampling and denoising:**

| No. | Period | Graphs | Amount | Particles | Weights $\nu_i$ |
|---|---|---|---|---|---|
| I | $\left(\frac{1}{14}, \frac{1}{14}, \frac{1}{14}, \frac{1}{14}, \frac{5}{7}\right)$ |  | 10 | $f_1, \ldots, f_4$, $f_{11}, \ldots, f_{16}$ | 0.076 |
| II | $\left(\frac{1}{14}, \frac{1}{14}, \frac{1}{14}, \frac{5}{7}, \frac{1}{14}\right)$ |  | 4 | $f_6, f_7, f_8, f_9$ | 0.038 |
| III | $\left(\frac{1}{23}, \frac{1}{23}, \frac{1}{23}, \frac{10}{23}, \frac{10}{23}\right)$ |  | 1 | $f_5$ | 0.057 |
| IV | $\left(\frac{1}{5}, \frac{1}{5}, \frac{1}{5}, \frac{1}{5}, \frac{1}{5}\right)$ |  | 1 | $f_{10}$ | 0.027 |

- **Best 5-periodic seasonal approximation** for such particles is the following function:

$$\left(\frac{1}{14}, \frac{1}{14}, \frac{1}{14}, \frac{10}{23}, \frac{5}{7}\right).$$

- **Forecasting:** Best weighted seasonal $[21,26]$-pull plan is then

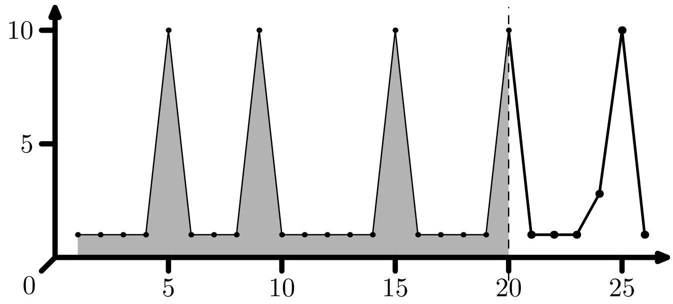
$$\left(\frac{1}{14}, \frac{1}{14}, \frac{1}{14}, \frac{1}{5}, \frac{5}{7}, \frac{1}{14}\right).$$

- **Rescaling:** finally we multiply the obtained function by the yearly expectation $E$ of a total amount of customers in during a period $P$, which is 14 in our case (see Remark 9).

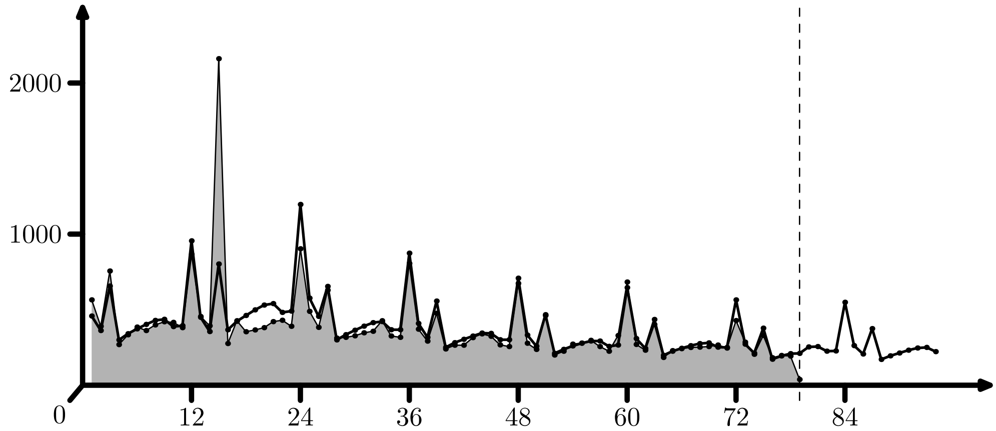
$$\left(1, 1, 1, \frac{14}{5}, 10, 1\right).$$

 doi:10.20944/preprints201708.0086.v1






**Example 1, part 2 of 2.** Let us finally conclude Example 1. On the following picture the observed function is filled with grey. The best weighted seasonal pull plan is shown with thick line. We extend the best weighted seasonal pull plan for the past history in order to examine the quality of prediction.




## References


G.E.P. Box, G.M. Jenkins and G.C. Reinsel, *Time series analysis : Forecasting and control*, Pearson Education, Delhi, 1994.

R.G. Brown, *Smoothing, Forecasting and Prediction*, Englewood Cliffs: Prentice Hall, 1963.

R.L. Chambers, C.J. Skinner, (editors), *Analysis of Survey Data*, Wiley Series in Survey Methodology, Wiley, 2003.

D.F. Findley, B.C. Monsell, M.C. Otto, W.R. Bell, M. Pugh *Towards X-12 ARIMA*, Technical report, Bureau of the Census, 1992.

D.F. Findley, B.C. Monsell, W.R. Bell, M.C. Otto, B.C. Chen *New capabilities and methods of the X-12 ARIMA seasonal adjustment program*, Technical report, U.S. Bureau of the Census, 1996.

R. Gençay, F. Selçuk, B.J. Whitcher, *An Introduction to Wavelets and Other Filtering Methods in Finance and Economics*, Academic Press, 2001.

C.C. Holt, *Forecasting seasonals and trends by exponentially weighted moving averages*, ONR Research Memorandum 52, Carnagie Institute of Technology, Pittsburgh, Pennsylvania, 1957.

A. Kolmogorov, *Interpolation and extrapolation von stationaren Zufalligen Folgen*, Bulletin of the Academy of Sciences, USSR, Ser. Math., 5, 3–14, 1941.

H.A. Linstone, M. Turoff, *The Delphi method*, Reading, MA: Addison-Wesley, 1975.

N. Wiener, *The extrapolation, interpolation and smoothing of stationary time series with engineering applications*, Wiley: New York, 1949.

P.R. Winters, Forecasting sales by exponentially weighted moving averages, *Management Science*, **1960**, *6*, 324-342.