RFID White paper

A new definition for stock accuracy in the retail sector

1 Introduction

Customers are more demanding than ever, both in brick and mortar stores and online, and want to buy what they want to buy – and not something else. Instantly, and not later. Therefore, making sure that a product in the right size and the right colour is available for customers is crucial these days. However, merchandise availability should not result in overstocked stores and the associated high capital cost. That is why accurate stock information at all times is key. Consequently, inaccurate stock information is disastrous for retail management systems and eventually leads to delayed replenishment and out-of-stocks [1,3,4].

In this white paper, we will first analyze what is meant with 'stock accuracy' and then take a close look at the existing methods to calculate stock accuracy. Finally, we will introduce a new definition to calculate the stock accuracy that suits the retail sector.

1.1 What is accurate stock?

To answer this question, we briefly address the concept of *accuracy* and subsequently the definition of *stock*. Several sources describe accuracy as being the degree to which a measurement's quantity corresponds to the quantity that is being measured [2].

From this, we already see that accuracy involves a degree of correspondence, a measurement and a measured quantity. In this paper, the quantity we measure is stock, this activity is also known as stock counting or stock-taking. Stock is defined as a collection of stock keeping units (SKUs) associated with integer quantities (see Table 1).

SKU	Quantity
SKU_1	1
SKU_2	0
SKU_3	4
SKU_4	2

Table 1: A certain stock

Accuracy is thus about the degree of correspondence of a (measured) stock with another (ideally the actual or real) stock. Without defining what that degree is, an example situation of comparing two stocks is described in Table 2.

SKU	Measured quantity	Actual quantity
SKU_1	1	1
SKU_2	0	1
SKU_3	4	3

Table 2: Stock comparison (measured versus real quantity)

Table 2 shows a measured stock on the left and the real stock on the right for three different SKUs. The measured quantity of SKU_1 corresponds with the real quantity. SKU_2 is not measured but is present in reality, and for SKU_3 the measured quantity exceeds the real quantity. A later section will elaborate on how to derive a quantitative accuracy (as a percentage) from these individual differences.

1.2 Stock as measurement

In retail stores, there is a multifold of techniques to count stock, such as:

- Human / manual
- Barcode scanner

- ERP system (based on barcode transactions)
- RFID handheld count
- Etc.

Of course, each individual technique again has a plethora of sub-techniques to count stock. We emphasize that a human/ manual count is considered to be a measurement. If a retailer maintains three SKUs where each SKU has a maximum quantity of 4 (as in the above example), a manual count is feasible and probably corresponds to reality.

However, measuring stock by manually counting SKUs and associated quantities quickly becomes infeasible and diverges from the real stock if the number of SKUs exceeds multiples of thousand (which is not unrealistic for a retail store).

An ERP system often diverges from reality over time due to all sorts of events, like incorrect processing of deliveries or stock movements or (internal & external) theft [1]. Also, RFID technology can sporadically fail to detect a label due to programming errors or shielding of labels by densely packed surroundings. Taken together, this leads to an important and paradoxical insight in the discussion about stock accuracy: in general, we do not know the real stock, we can only do measurements.

Fortunately, we are at a point in time where RFID technology is so advanced, that doing a stock take of thousands of SKUs is feasible and close to reality. Most environments can, without noticeable changes, be prepared in such a way that the probability of not reading labels is negligible.

Updating the ERP system with RFID stock information subsequently results in the ERP system of a retailer being more close to reality. Hence, when talking about accurate stock data, we mean that the ERP system contains stock data, which is close to the stock data derived from the RFID technology, meaning the closeness of ERP to RFID is used to express stock accuracy. In this context, accuracy is used as performance indicator for the closeness of the ERP stock to the RFID stock.

2 Expressing stock accuracy as percentage

In the following examples, the stocks of three different SKU's of an RFID and an ERP system are compared:

SKU	RFID	ERP
SKU_1	1	1
SKU_2	2	2
SKU_3	4	4

Table 3a: Comparison between RFID and ERP system, 100% accurate

SKU	RFID	ERP
SKU_1	1	1
SKU_2	2	2
SKU_3	0	4

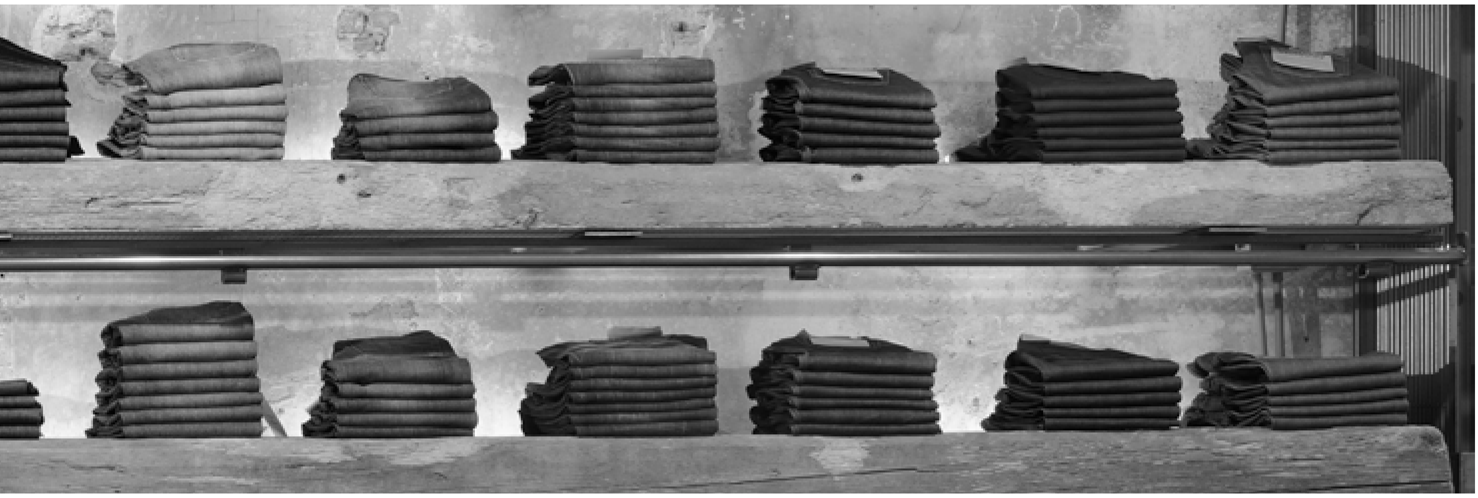
Table 3b: Comparison between RFID and ERP system, SKU_3 differs

Without explicit calculation, the reader will probably agree with the example in Table 3a being 100% accurate. But how to express the (in)accurateness of Table 3b?

2.1 Currently used methods to calculate accuracy

In general, retailers currently use two methods to calculate their stock accuracy:

1. *Summed absolute differences*, relative to the total items in the ERP stock
2. *Percentage of (in)correct SKU's*, relative to the SKU quantity in the ERP stock



The *summed absolute difference (SAD)* method sums the absolute differences between RFID and ERP, and divides those by the total quantity of items in the ERP system.

SKU	RFID	ERP	Difference
SKU$_1$	1	1	0
SKU$_2$	3	3	0
SKU$_3$	0	4	-4
SKU$_4$	2	1	1
SUM	**6**	**9**	**5**
Difference: 5 / 9 = 56% Accuracy: 100% - 56% = 44%			

Table 4: SAD calculation for stock accuracy

The *percentage of (in)correct SKU's (PCS)* method checks the total number of SKU's that do not match between RFID and ERP, and divides that by the total number of SKU's in the stock file.

SKU	RFID	ERP	Correct?
SKU$_1$	1	1	YES
SKU$_2$	3	3	YES
SKU$_3$	0	4	NO
SKU$_4$	2	1	NO
SUM	**6**	**9**	**2**
Difference: 2 / 4 = 50% Accuracy: 100% - 50% = 50%			

Table 5: PCS calculation for stock accuracy

While it seems like both methods provide a reasonable indication for stock accuracy, there are actually some fundamental problems that arise when investigating deeper.

Example 1: more than 100% difference (SAD)

With the SAD method, a difference of more than 100% can occur in certain cases. This leads to a negative accuracy, which does not make sense.

SKU	RFID	ERP	Difference
SKU$_1$	1	1	0
SKU$_2$	2	2	0
SKU$_3$	4	0	4
SUM	**7**	**3**	**4**
Difference: 4 / 3 = 133% Accuracy: 100% - 133% = ?			

Example 1a: SAD calculation resulting in more than 100% difference

However, if the difference between the RFID and ERP system on the same SKU (in this case SKU$_3$) is the other way around, the difference is less than 100%.

SKU	RFID	ERP	Difference
SKU$_1$	1	1	0
SKU$_2$	2	2	0
SKU$_3$	0	4	-4
SUM	**3**	**7**	**4**
Difference: 4 / 7 = 57% Accuracy: 100% - 57% = 43%			

Example 1b: SAD calculation; resulting in less than 100% difference

Example 2: minor difference, huge impact (PCS)

With the PCS method, differences with a minor impact on product availability (because there is still enough available for customers to buy) might have a huge impact on the difference and stock accuracy indicators.

SKU	RFID	ERP	Correct?
SKU_1	47	48	NO
SKU_2	33	32	NO
SKU_3	27	26	NO
SUM			**3**
Difference: 3 / 3 = 100%			
Accuracy: 100% - 100% = 0%			

Example 2: PCS calculation; minor difference, huge impact

The above example illustrates that PCS can overreact for differences with a minor impact on the business.

Example 3: high unavailability of product, low difference (SAD)

In cases where there is a significant difference in the stock levels of SKU's, unavailability with a big impact might get unnoticed in the difference/accuracy figures with the SAD method. In this case, two SKU's (out of three) are out of stock, but still the stock accuracy is 98%.

SKU	RFID	ERP	Difference
SKU_1	0	1	1
SKU_2	0	1	1
SKU_3	100	100	0
SUM	**100**	**102**	**2**
Difference: 2 / 102 = 2%			
Accuracy: 100% - 2% = 98%			

Example 3a: SAD calculation; high unavailability of products, low difference

The next example is a lot less troubling, because all SKU's are still available to potential customers, but the outcome is exactly the same.

SKU	RFID	ERP	Difference
SKU_1	1	1	0
SKU_2	1	1	0
SKU_3	98	100	2
SUM	**100**	**102**	**2**
Difference: 2 / 102 = 2%			
Accuracy: 100% - 2% = 98%			

Example 3b: SAD calculation; low unavailability of products, low difference

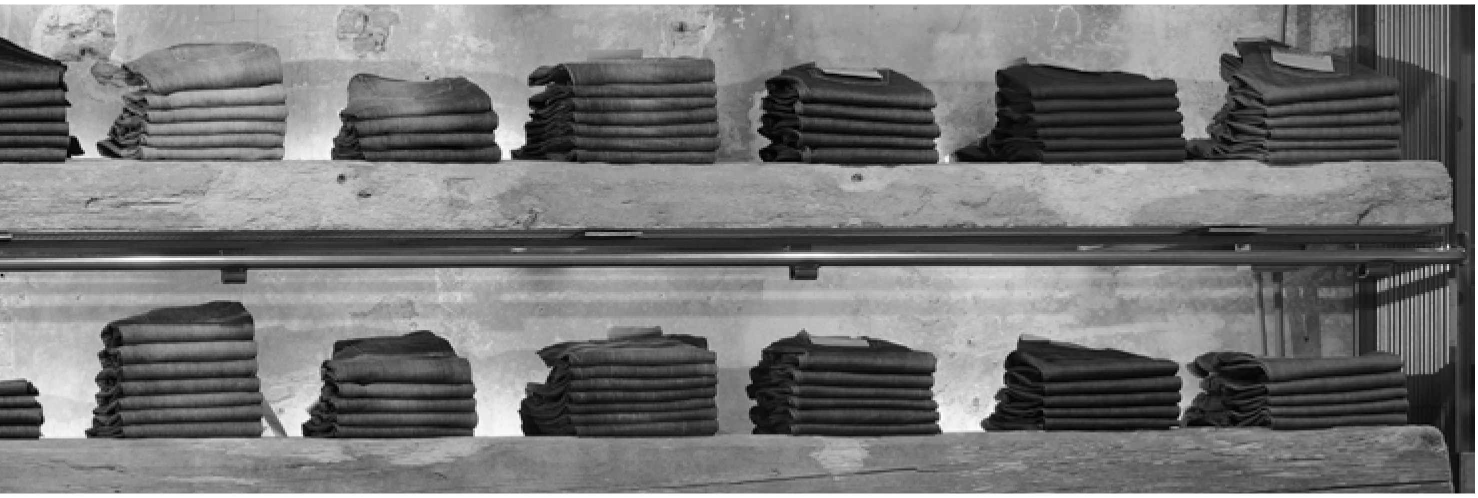
Main problems of SAD & PCS calculations

The above examples illustrate that there are three main problems of the existing SAD and PCS stock accuracy indicators used in retail:

1. Lack of symmetry – based on what you take as a reference, the outcome can differ significantly.
2. Stock differences that have a minor impact on business might have a huge impact on stock accuracy figures, and the other way around.
3. In some cases, the calculated accuracy is meaningless, while accuracy is the positive minded indicator that we prefer to use.

2.2 SMAPE method

We propose using the Symmetric Mean Absolute Percentage Error (SMAPE) to solve the problems listed before. SMAPE is used to express accuracy of forecasting models [5]. It also has very desirable characteristics for retail purposes, as we will see in this section.



SMAPE is defined as:

d_{SMAPE}(RFID, ERP) := average sum of

$$\frac{\text{absolute difference between SKU quantities}}{\text{sum of both individual absolute SKU quantities}}$$

Note: The absolute value of a number is the non-negative value of x, without regard to its sign. So -6 becomes 6, and 5 will remain 5. It is denoted by using two vertical bars: |-6|=6.

Accuracy is defined as $100\% - d_{SMAPE}$

SKU	RFID	ERP	SMAPE Difference
SKU_1	1	1	|(1 - 1)|/(|1|+|1|) = 0
SKU_2	2	2	|(2 - 2)|/(|2|+|2|) = 0
SKU_3	0	4	|(0 - 4)|/(|0|+|4|) = 1
SKU_4	2	1	|(2 - 1)|/(|2|+|1|) = 0,33
SUM	5	9	1,33
Difference: 1,33 / 4 (SKU's) = 33% Accuracy: 100% - 33% = 67%			

Table 6: SMAPE calculation for stock accuracy

Note: If for one SKU, the RFID and ERP quantity are both 0, the record can be neglected.

In the next section, more examples will be given to demonstrate the properties of SMAPE.

2.3 Difference in importance of a difference

Here is an overview of the earlier shown examples (on page 3 & 4) that were used to indicate the flaws of the SAD and PCS methods, but now compared to the SMAPE method. When comparing the results for example 1, we see that by using SMAPE, it is possible to calculate an accuracy (which was not possible with the SAD method). SMAPE also proves to give the same outcome when the numbers for SKU_3 are interchanged between ERP and RFID, thus the method is symmetric.

Example 1a	SAD	PCS	SMAPE
Difference	133%	33%	33%
Accuracy	?	67%	67%

Example 1a: comparison of SAD, PCS and SMAPE

Example 1b	SAD	PCS	SMAPE
Difference	57%	33%	33%
Accuracy	43%	67%	67%

Example 1b: comparison of SAD, PCS and SMAPE

In example 2, the impact of differences with a minor impact on business had a huge impact on the difference and accuracy with the PCS method, whereas with SMAPE this is corrected for. The accuracy is in line with what you intuitively would expect.

Example 2	SAD	PCS	SMAPE
Difference	3%	100%	1%
Accuracy	97%	0%	99%

Example 2: comparison of SAD, PCS and SMAPE

In example 3, two cases gave the same output with the SAD method, but had a different impact on the business. The first case had a major impact on business (two out of three products were out of stock), while the second case had a minor impact on business (no products out of stock). This example shows that PCS handles the first case correctly, but the second one completely overstates the impact of the minor difference. SMAPE handles both cases correctly with an outcome that reflects the actual business impact.

Example 3a	SAD	PCS	SMAPE
Difference	2%	67%	67%
Accuracy	98%	33%	33%

Example 3a: comparison of SAD, PCS and SMAPE

Example 3b	SAD	PCS	SMAPE
Difference	2%	33%	0%
Accuracy	98%	67%	100%

Example 3b: comparison of SAD, PCS and SMAPE

In general, we see that SMAPE smoothly weighs the importance of differences. The closer the differences are to zero, the closer the difference is to product (un)availability. This means that an individual difference between 1 and 0 is more important than a difference between 2 and 1, and similarly a difference between 100 and 99 is considered even less important.

Furthermore, it can be proven that the difference d_{SMAPE} always gives an outcome between 0% and 100%, which means that the accuracy 100% – d_{SMAPE} will also always give an outcome between 0% and 100%. This property gives the retailer the ability to use SMAPE accuracy with RFID as a safe performance indicator for its ERP system, where 100% is the target performance.

3 Conclusion

First a context for stock accuracy in retail business was sketched. This was necessary for the understanding of stock and comparisons between measured stocks. Also the need for product availability in retail business was indicated.

Subsequently, we defined SMAPE, a way to express stock accuracy quantitatively. This definition inherits some desirable properties, which strongly take into account product availability in a smooth weighing way.

The closer the difference is to product unavailability, the stronger the negative effect on accuracy. In the same way, a difference between large quantities has far less impact on the accuracy than a difference between 0 and 1.

Together with the other discussed characteristics, this makes SMAPE a quantitative accuracy definition that suits the retail sector and can be used as a performance indicator for stock accuracy in a safe way.

References

[1] Nicole DeHoratius and Ananth Raman, *Inventory record inaccuracy: an empirical analysis*, Management Science 54 (2008), no. 4, 627–641.

[2] Oxford Dictionaries. *https://en.oxforddictionaries.com/definition/accuracy* [Online; accessed 24 November 2016].

[3] H. Sebastian Heese, *Inventory record inaccuracy, double marginalization, and rfid adoption*, Production and Operations Management 16 (2007), no. 5, 542.

[4] Yun Kang and Stanley B Gershwin, *Information inaccuracy in inventory systems: stock loss and stockout*, IIE transactions 37 (2005), no. 9, 843–859.

[5] Chris Tofallis, *A better measure of relative prediction accuracy for model selection and model estimation*, Journal of the Operational Research Society 66 (2015), no. 8, 1352–1362.