

# TREND FLAGGING TO AID RESOURCE ALLOCATION DECISIONS

*Dennis Egan, CCICADA Center, 96 Frelinghuysen Road, Piscataway, NJ 08854, 848-445-5928, [deegan@dimacs.rutgers.edu](mailto:deegan@dimacs.rutgers.edu)*

*Paul B. Kantor, CCICADA Center, 96 Frelinghuysen Road, Piscataway, NJ 08854, 848-445-5928, [kantorp@comminfo.rutgers.edu](mailto:kantorp@comminfo.rutgers.edu)*

*Joonhee Lee, CCICADA Center, 96 Frelinghuysen Road, Piscataway, NJ 08854, 848-445-5928, [joonhee@business.rutgers.edu](mailto:joonhee@business.rutgers.edu)*

*Christie Nelson, CCICADA Center, 96 Frelinghuysen Road, Piscataway, NJ 08854, 848-445-5928, [cgrewe@dimacs.rutgers.edu](mailto:cgrewe@dimacs.rutgers.edu)*

*Fred S. Roberts, CCICADA Center, 96 Frelinghuysen Road, Piscataway, NJ 08854, 848-445-5928, [froberts@dimacs.rutgers.edu](mailto:froberts@dimacs.rutgers.edu)*

*Helen M. Roberts, Department of Mathematical Sciences, Montclair State University, Montclair, NJ 07043, 973-655-7886, [robertsh@mail.montclair.edu](mailto:robertsh@mail.montclair.edu)*

## ABSTRACT

Decision makers in all kinds of organizations, and in particular those concerned with homeland security, need to be able to easily flag trends so that they can respond, for example with a reallocation of resources or a review of policies and procedures. This paper introduces a simple-to-use tool called the TrendFlagger that allows a decision maker to get evidence that a trend may be appearing without requiring statistical sophistication. The TrendFlagger will be illustrated using a source of data that organizations might use to study shipping trends and vessel behavior, the Automatic Identification System (AIS) now required by international agreement on all ships of 300 gross tons or more and all passenger ships.

**Keywords:** Decision Making, Trend Analysis, Resource Re-allocation, Automatic Identification System

## INTRODUCTION

Decision makers in all kinds of organizations, and in particular those concerned with homeland security, need to be able to easily flag trends so that they can respond to changes, by rethinking policies or procedures or reallocating resources. As a case in point, the US Coast Guard (USCG) may seek to reallocate resources to specific missions such as search and rescue (SAR), environmental protection, or migrant interdiction/law enforcement, based on increases in specific kinds of shipping or changing vessel behavior patterns. For example, increased shipping in the Arctic due to changing ice conditions will allow more eco-tourism and cruise ships, requiring preparation for more SAR events. Increased petrochemical shipping in the Houston Ship Channel increases the risk of an oil spill. Increasing frequency of vessels anchoring for unusually long (or short) periods of time at the same offshore location might suggest illegal activities such as smuggling of narcotics or weapons. Relevant data arrive monthly, weekly, daily, or even more frequently, and usually contain some random fluctuation, which should not drive reallocation decisions. This paper introduces a simple-to-use tool called the TrendFlagger that allows a decision maker to get evidence that a trend is appearing without requiring statistical sophistication. The TrendFlagger will be illustrated using a source of data that the USCG uses to study shipping trends and vessel behavior, the Automatic Identification System (AIS). AIS automatically transmits data identifying the ship, its location and course, its destination port and estimated time of arrival, etc., for over 1,000,000 ships worldwide. AIS data can be used to flag early

warning of changed shipping patterns or vessel behavior. In this paper, the TrendFlagger will be applied to AIS data using the Houston Ship Channel example as a case in point. Illustrative trends will be described, issues with unreliable AIS data mentioned, and other potential conclusions using TrendFlagger pointed out.

## AIS DATA

Since the year 2000, the International Maritime Organization (IMO) has required, via the International Convention for the Safety of Life at Sea [6], that all ships of 300 gross tons or more, and all passenger ships install an identification and location device on board that consistently and automatically transmits dynamic data (location, course, speed, etc.) and voyage-related information (destination, estimated time of arrival, etc.) to Vessel Traffic Services (VTS) stations as well as to other ships. Vessel identifiers such as vessel name and VHF call sign are programmed in the device and are also included in the transmittal. The global AIS system receives updates for each ship as frequently as every two seconds while in motion and every three minutes while at anchor. AIS tracks ships automatically by electronically linking data with other ships, AIS base stations, and satellites. This system enables ships to share positional data with other ships, and while its primary initial function was for traffic management, collision avoidance, and other safety applications, it has extensive other uses in national security. Overall, AIS offers awareness about vessels operating within the maritime transportation system [5][8].

The data is massive. For example, US coast and waterway related information accounts for 32 GB of AIS data each day. Some AIS data are available for download at <https://marinecadastre.gov/ais/>.

Users of AIS data should be aware of the fact that the data can have a variety of errors due to transmission problems, cyber attacks, or users forgetting to turn on the AIS transponder. Accuracy is an issue with some AIS data having 50% errors [10]. Since AIS data are compiled from participating vessels, AIS data can be intentionally inaccurate. An attacker could exploit weaknesses in AIS and intentionally falsify a vessel's identity or type, or its position, heading, and speed or make it invisible to authorities [2][3][4][5].

## THE TRENDFLAGGER

A tool called the TrendFlagger© was developed by one of the authors of this paper (Paul Kantor), with the goal of developing a systematic way to get an indication of a trend and distinguishing it from a random fluctuation.<sup>1</sup> In trend flagging, the decision maker must choose how long to wait to decide whether a trend has appeared and how strong an increase or decrease merits attention. How long to wait involves setting a “window” of number of reports that would indicate a trend. In the case of strength of increase or decrease that is considered significant, the decision maker must choose two thresholds, one for “uptrend” and one for “downtrend.” The two thresholds can be different, though in all of the examples in this paper they are the same. The three parameters can then be “tuned” by the decision maker.

The TrendFlagger software “watches” the data, using a moving window of the chosen size. It processes the data to find the (linear) rate of *increase or decrease in that window*. If the last point added confirms that the rate is above a target level (below a target level), then the latest number is flagged red (yellow) and an uptrend (downtrend) is indicated.

---

<sup>1</sup> The tool is Copyright © 2015 by the CCICADA Center at Rutgers University and is available for US government purpose use. For customization and further development, the developer *Paul Kantor* can be reached at [paul.kantor@rutgers.edu](mailto:paul.kantor@rutgers.edu).

In this way, the TrendFlagger lets the decision maker know whether the data are showing an up- or down-trend, at the level he or she has decided merits review of policies, resource allocations or procedures, and in an easy visualization. It provides the decision maker with data to back up a proposal, or to suggest that more sophisticated statistical analysis be performed.

To be precise, the choice of target level for uptrend or downtrend involves determining change per unit, or the slope of a line heading upward or downward that serves as minimum (maximum) for designating an upward (downward) trend. In practice, the user would be prompted to determine how big an increase per unit time would constitute an upward trend of concern, and similarly for decrease per unit time.

The TrendFlagger has a lot in common with the moving averages that are widely used to identify trends in stock prices, commodity prices, etc. A simple moving average chooses a window – a number of time periods – and reports the average value (average price) over that window. The average is “moving” in the sense that within each time period the oldest value is dropped and the newest one added. Moving averages are used to identify uptrends or downtrends. They can be used over the long term, the medium term or the short term, and there are things to be learned from all three or the combination of the three. Moving averages will be more sensitive to changes if they use smaller windows and less sensitive if they use larger windows [7]. The same is true for the TrendFlagger. Moreover, “There is no ‘right’ timeframe to use when setting up your moving average. The best way to figure out which one works best for you is to experiment with a number of different time periods until you find one that fits your strategy” [7]. We recommend doing this for the TrendFlagger as well, and indeed we have done a great many experiments with both the window size and the choice of upper and lower thresholds in applying the TrendFlagger. The conclusions reached are dependent upon choice of these parameters and in fact the same data can show dramatically different trend patterns with two different sets of parameters. This calls attention to the importance of real world experience in setting the window and the limits. “As with most technical analysis tools, moving averages should not be used on their own, but in conjunction with other complementary tools” [9]. This is again what we would recommend for the TrendFlagger.

It should be understood that the TrendFlagger is really only a decision support tool, and depends heavily on human intervention, and in particular experimentation with different parameters to see what kinds of interesting observations result. The choice of parameters will depend on familiarity with the application area and so on the expertise of the user in the subject matter to which this tool is applied. We show how this could be done when we give our first example.

The principles described here have been built into an Excel spreadsheet, using the conditional formatting power of the software, and a number of advanced built-in functions that support the statistical analyses that underlie the tool. Specifically, if the window size is  $n$ , the tool uses linear regression to find a straight line fit using the last  $n-1$  data points and the newest one. It then compares the slope of this line to the thresholds. If the slope is higher than the upper threshold, the  $n^{\text{th}}$  data point is colored red, and similarly if it is less than the negative of the lower threshold, the  $n^{\text{th}}$  data point is colored yellow. For instance, in the trivial case where the window size is 2, and the  $(n-1)^{\text{st}}$  and  $n^{\text{th}}$  points are 7 and 11, slope is 4 and if the upper threshold is less than 4, the  $n^{\text{th}}$  data point would be colored red. Points not colored red or yellow are colored white, except in the first  $n-1$  rows of the spreadsheet, where there are not enough data points to draw a conclusion, and the color pink is used. These datapoints shaded pink are, however, important because they are used in the first few trend decisions.

AIS data covers different UTM (Universal Transverse Mercator) zones, with 19 zones covering the United States. Figure 1 shows data arising from Zone 15, which includes the Houston Ship Channel. The column labeled Total shows the total number of vessels arriving in Zone 15 in a given month. The second column gives the proportion of vessels heading into Zone 15 in a given month that are headed to the Port of Houston. The first row gives the threshold for an Uptrend decision – it can of course be different for each column. Row 2 gives the threshold for the Downtrend and row 3 the window size. The user may change the number of time periods used to estimate change in the window cell in each column,

and also the uptrend and downtrend thresholds. It is recommended that different combinations of these parameter values be experimented with in order to develop some experience with useful values. Since the averaging window shown in the Total column has size 5, the first four rows of this column are colored pink and there is no trend indicated. (How we determined the Upper and Lower Thresholds is discussed below.) Consider the last five entries in this column. Linear regression yields a straight line with slope -116.4. Since 116.4 is larger than the DownTrend threshold 42.81, the downward slope indicates a downtrend and the entry for December 2014 is colored yellow. This may seem counter-intuitive if one just looks at the numbers, since the December 2014 entry of 1930 is higher than the previous four entries, but the reader should recall that it is the slope of the regression line that indicates a trend, not a particular number. The individual data point contributes to the slope, but does not itself indicate a trend. If the December 2014 entry were 2200, the slope of the regression line would be -62.4, and this would still indicate a downtrend. However, 2300 leads to a slope of -42.4, indicating no trend. If the December 2014 entry were 2700, the slope would be 37.6, and this is not yet high enough to indicate an upward trend. However, 2750 does, since it gives rise to a slope of 47.6.

**Figures 1 and 2: TrendFlagger Showing Vessel Arrivals in Zone 15**

		Total	Houston/Total
Year.Month	UpTrend	42.81	0.005
	DownTrend	42.81	0.005
	AveragingWindow	5	5
13.01	1	1801	0.198778456
13.02	2	1644	0.196
13.03	3	1874	0.191
13.04	4	1818	0.187
13.05	5	1983	0.184
13.06	6	1861	0.202
13.07	7	1931	0.190
13.08	8	1960	0.201
13.09	9	1893	0.192
13.10	10	2070	0.189
13.11	11	1947	0.216
13.12	12	2026	0.194
14.01	1	1915	0.207
14.02	2	1816	0.172
14.03	3	2026	0.187
14.04	4	2075	0.193
14.05	5	1969	0.198
14.06	6	1849	0.206
14.07	7	1918	0.209
14.08	8	1866	0.214
14.09	9	1714	0.235
14.10	10	1820	0.221
14.11	11	422	0.358
14.12	12	1930	0.216
alpha			0.3

		Total	Total
Year.Month	UpTrend	42.81	55.26
	DownTrend	42.81	55.26
	AveragingWindow	3	3
13.01	1	1801	1801
13.02	2	1644	1644
13.03	3	1874	1874
13.04	4	1818	1818
13.05	5	1983	1983
13.06	6	1861	1861
13.07	7	1931	1931
13.08	8	1960	1960
13.09	9	1893	1893
13.10	10	2070	2070
13.11	11	1947	1947
13.12	12	2026	2026
14.01	1	1915	1915
14.02	2	1816	1816
14.03	3	2026	2026
14.04	4	2075	2075
14.05	5	1969	1969
14.06	6	1849	1849
14.07	7	1918	1918
14.08	8	1866	1866
14.09	9	1714	1714
14.10	10	1820	1820
14.11	11	422	422
14.12	12	1930	1930
alpha			0.3

Figure 1. TrendFlagger displaying total number of vessels arriving in Zone 15 per month and the proportion of those vessels arriving in the Port of Houston in that month, under window size 5.	
---	--

Figure 2. TrendFlagger showing total number of vessels arriving in Zone 15 per month under different choices of upper and lower thresholds and window size 3.	
---	--

In this and other examples in this paper, we use thresholds and window sizes chosen to demonstrate the method and to fit the available data that could be presented in a reasonable-size table. How might we actually choose these parameters using knowledge of the application? Note that we have some leeway in how much we think our budget for oil spill response can handle. Suppose we figure we can handle up to a 10% increase in tanker traffic before we feel that we need to put more resources into oil spill response. Thus, we would use a 10% uptrend threshold. Similarly, because of pressures on our budget for other purposes such as counter-narcotics, we figure once we have a 15% decrease in traffic, we can reduce the oil spill response budget. Thus, we would use a 15% downtrend threshold. We should also not make such decisions based on small windows of time. In fact, there are annual fluctuations that can just be due to noise or weather or economic conditions. Thus, our window might be three years. The data we presented in this first example and in subsequent examples was by month, but we might want to divide the year into two halves, sum the traffic in each half, and then use a window of size 6, which would give us three years (e.g. looking at November-April as one half of a year and May-October as the other half). In this way, we would tie the choice of parameter values to some knowledge of or constraints on these values based on our experience and knowledge.

The first Total column of Figure 2 shows the same data of total Zone 15 vessel arrivals, the upper and lower thresholds are unchanged from those in Figure 1, but now the window size is changed to 3. With this change, the December 2014 entry of 1930 shown in the first Total column becomes red instead of yellow. That is because the straight line determined by the last three datapoints in the Total column has slope 55, which is larger than the upper threshold of 42.81. In the second Total column, we have changed the UpTrend/DownTrend thresholds. Now, the December 2014 entry of 1930 is white. This illustrates the sensitivity of conclusions to the three key parameters. As noted above, there is no hard and fast rule as to what values to choose for these parameters. Clearly the upper and lower thresholds should be chosen depending on the user's goals and experience. However, one rule of thumb we have found useful, all other things being equal, is to relate these numbers to the standard deviation  $s$  of column entries. Which column entries are used depends on how the TrendFlagger is used. If it is used to spot trends as data comes in, perhaps the standard deviation of the first  $n$  datapoints should be used. If all the data is in and then trends are identified, one could use the standard deviation of all datapoints. Because we study blocks of  $n$  data entries, it may also make sense to consider the standard deviation resulting from such a block, so to consider  $s/\sqrt{n}$ . Another thing that is occasionally useful is to multiply this number by a parameter  $\alpha$  to obtain the upper and lower thresholds. So, to summarize this heuristic, one takes the upper and lower thresholds to be  $\alpha s/\sqrt{n}$ . Note that in both columns of Figure 1, this heuristic is used to calculate the thresholds with  $\alpha$  taken to be 0.3 and all 24 datapoints used for the standard deviation. In the first column of Figure 2, we copy the thresholds from the first column of Figure 1 and so do not use this heuristic, whereas in the second column we do use the heuristic with  $\alpha = 0.3$  and  $s$  calculated from all 24 datapoints. The user of TrendFlagger is encouraged to vary  $\alpha$  as well as the window size in order to find interesting suggestions of trends, or of course to choose thresholds to match their experience or needs.

### **AIS DATA FOR THE HOUSTON SHIP CHANNEL: WHAT CAN WE LEARN FROM THE TRENDFLAGGER?**

To see what we might learn from the TrendFlagger, we gathered data for Zone 15 and for three ports in the Houston Ship Channel, Houston, Galveston, and Texas City. Figure 3 gives a portion of this data. The columns give total arrivals in each month during the years 2013 and 2014 at those ports. They also

specialize for specific types of vessels. Vessels listed in AIS are categorized by 2-digit codes, where the first digit gives the type of vessel and the second gives information about the cargo. An 8 in the first digit indicates a tanker, the type of vessel of interest in studying petrochemical movements. However, tankers can carry all kinds of liquids, including for example orange juice. The codes 80 to 89 indicate different types of tanker. Of particular interest are code 81 vessels, which carry hazardous cargo. Some of the data give total ship arrivals in a given month for ships of a given type.

It is not sufficient to simply download AIS data. It needs to be adjusted for purposes of analysis. For instance, AIS data gives a vessel's arrival port, but there is duplication since a given vessel will appear on multiple days with the same arrival port, so the AIS data needs to be modified to eliminate duplication if we are interested in how many vessels head to a given port in a given month. There are anomalies in the data. For example, there was a vessel of type 81 heading to Houston every day one month until it ended up in Texas City. Was this an error in the data? A reflection of a change of itinerary due to an emergency? There is no way to tell, so we counted it as an anomaly. Luckily, anomalies are rare.

In what follows, we discuss some of the observations we made using the TrendFlagger. It is important to emphasize that the conclusions are not definitive. Rather, they suggest things that should be investigated to get better understanding of seeming trends, possible anomalies, etc. The conclusions from the TrendFlagger should always be confirmed (or not) using other tools and approaches.

Returning to Figure 1 above, we see how the TrendFlagger can call attention to opposing trends in two sets of data positioned in adjacent columns of the spreadsheet. For both datasets, we have selected a window size of 5 to detect trends and we selected the size of the upper and lower thresholds to detect trends by using our heuristic with  $\alpha = 0.3$ . Looking at the column labeled "Total," we see that the trend for the months of 2013 after the first four was either red (increasing) or white (not enough deviation from zero slope to indicate a trend). However, this trend reverses briefly in February 2014 and then toward the latter half of 2014, where the total number of ships entering Zone 15 shows a fairly steady downward trend (yellow).

Now consider the column of Figure 1 showing the proportion of the total Zone 15 shipping that had the Port of Houston as the destination. During 2013, in months when the overall total was showing an upward trend, the proportion going into Houston was showing no trend. On the other hand, in the latter half of 2014, while the Zone 15 shipping showed a steady downward trend, the proportion going to Houston showed a steady upward trend. The TrendFlagger calls attention to these patterns by the contrasting colors in the last six months of 2014.

Figure 3 also leads to some interesting observations that need further investigation. For example, using the first three columns, we see that the total hazardous tanker traffic can have an opposite trend from the total traffic (see October and November 2013 and March and May 2014), while total tanker traffic never has an opposite trend from total traffic. However, there are no such opposite trends in hazardous traffic and total traffic heading to Houston.

## **Learning from Anomalies**

Looking for situations where two columns might be expected to have similar trend patterns can be useful. Anomalies can then lead to interesting and useful questions. Consider Figure 4, which shows total ship arrivals in the Port of Texas City in the years 2013 and 2014 in different columns. November 2013 shows an upward trend, but November 2014 shows a downward trend. A similar anomaly in the same months appears if we just consider all tankers (all vessels labeled 80 to 89) coming to Texas City in these two years and use a smaller window size – see Figure 5. The anomaly led us to observe that the November 2014 data values in all columns of Figure 3 are dramatically lower than those in other months.

The small values for data entries in November 2014 in Zone 15 led us to examine the magnitude of the overall AIS data sets available for downloading for different zones, and we found a similar low magnitude in many of the zones. Upon further research, we found that there was an AIS tracking system update in November 2014, and this probably makes the November 2014 data suspect. As noted in the section on AIS data, that data does have errors, and here the TrendFlagger has led us to identify a possible error in data. Once we discovered this issue, we decided to omit November and December 2014 data in our further analyses.

**Figure 3: Partial Set of Data for Zone 15**  
**Total Arrivals in Zone 15 and in Houston by Month and by Type of Vessel**

		Total	Total_80s	Total_81	Hous	Hous_80s	Hous_81
<i>Year.Month</i>	UpTrend	25.01	8.47	1.61	12.49	6.44	0.99
	DownTrend	25.01	8.47	1.61	12.49	6.44	0.99
	AveragingWindow	3	3	3	3	3	3
13.01	1	1801	528	46	358	163	12
13.02	2	1644	488	44	323	149	13
13.03	3	1874	538	53	358	150	15
13.04	4	1818	541	52	340	156	13
13.05	5	1983	567	57	365	176	17
13.06	6	1861	527	46	375	155	19
13.07	7	1931	561	49	366	162	17
13.08	8	1960	566	61	393	170	24
13.09	9	1893	553	51	364	163	16
13.10	10	2070	590	50	392	180	15
13.11	11	1947	586	45	421	217	19
13.12	12	2026	626	57	393	195	15
14.01	1	1915	566	57	396	201	20
14.02	2	1816	512	39	312	138	8
14.03	3	2026	575	50	378	175	16
14.04	4	2075	595	56	401	181	22
14.05	5	1969	568	57	390	194	20
14.06	6	1849	528	39	381	172	17
14.07	7	1918	561	47	401	174	13
14.08	8	1866	564	39	399	190	16
14.09	9	1714	483	39	402	173	14
14.10	10	1820	488	40	403	189	19
14.11	11	422	154	18	151	74	12
14.12	12	1930	580	50	417	198	28
alpha	0.4						

### Looking for Explanations

When TrendFlagger highlights an upward or downward trend, we want to look further. One thing to do is to perform other statistical tests on the data to confirm the suggestion from the informal TrendFlagger tool. One can look for example for one or multiple time series exhibiting trends [1]. Trend flagging will require development of new statistical methods to provide early warning of changed shipping patterns. The statistical methods needed include variants of sequential change detection methods and streaming algorithms suited to the particular spatio-temporal characteristics of shipping patterns.

Another approach is to look for explanations. Consider for example the data in Figure 6, which shows with a particular choice of up and down thresholds and window size that there was an uptrend in Zone 15 total in April 2014 and a downtrend in June 2014. In fact, in late March 2014, there was an oil spill in the region, so it is likely that some vessels might have been delayed in arriving, leading to an uptrend in April that didn't correspond to a seasonal uptrend. Similarly, in May 2014, there was significant flooding in the region, which likely impacted vessel traffic into June. The TrendFlagger has led us to understand the data.

**Figure 4: TrendFlagger for Texas City Vessel Arrivals, Comparing 2013 and 2014**

		Texas City 2013	Texas City 2014
	<b>UpTrend</b>	3.5	3.5
	<b>DownTrend</b>	3.5	3.5
<i>Month</i>	<b>AveragingWindow</b>	4	4
1		32	27
2		22	21
3		26	26
4		25	21
5		19	18
6		21	21
7		26	26
8		25	25
9		31	31
10		31	21
11		38	11
12		33	23

### FUTURE DIRECTIONS

The format presented here produces a good deal of information for users to utilize their expertise and judgment to investigate. There are, however, a few extensions that might make the TrendFlagger tool even more valuable:

1. We may want to consider the comparison of current data with a most appropriate point in the past. For example, high level decision makers may see only annual summary data. Because of that, together with annual variability related to market demand and weather factors, it might be valuable to flag meaningful changes from exactly a year earlier.
2. Allow for monthly variation in traffic patterns. For this, the current data would be automatically compared with data for the same month in an earlier year, in spotting trends up or down.
3. As with simple moving averages, the TrendFlagger weighs all data points equally. However, some argue that more recent data should have greater influence on conclusions, and that gives rise to weighted moving averages. A popular idea is the exponential moving average [7]. A similar modification in the TrendFlagger might be of interest to some users.
4. So far, our recommendation is that the choice of parameters in TrendFlagger should be primarily dependent on experience and goals of the user. There is need, however, to develop both theory and heuristics for parameter choice. The heuristic for choice of threshold discussed in the section called The TrendFlagger is a case in point.
5. As noted in the subsection called Looking for Explanations, trend flagging in general will require development of new statistical methods including variants of sequential change detection methods and streaming algorithms suited to particular spatio-temporal characteristics relevant to the application area in question.



**Figures 5 and 6: Investigating Trends**

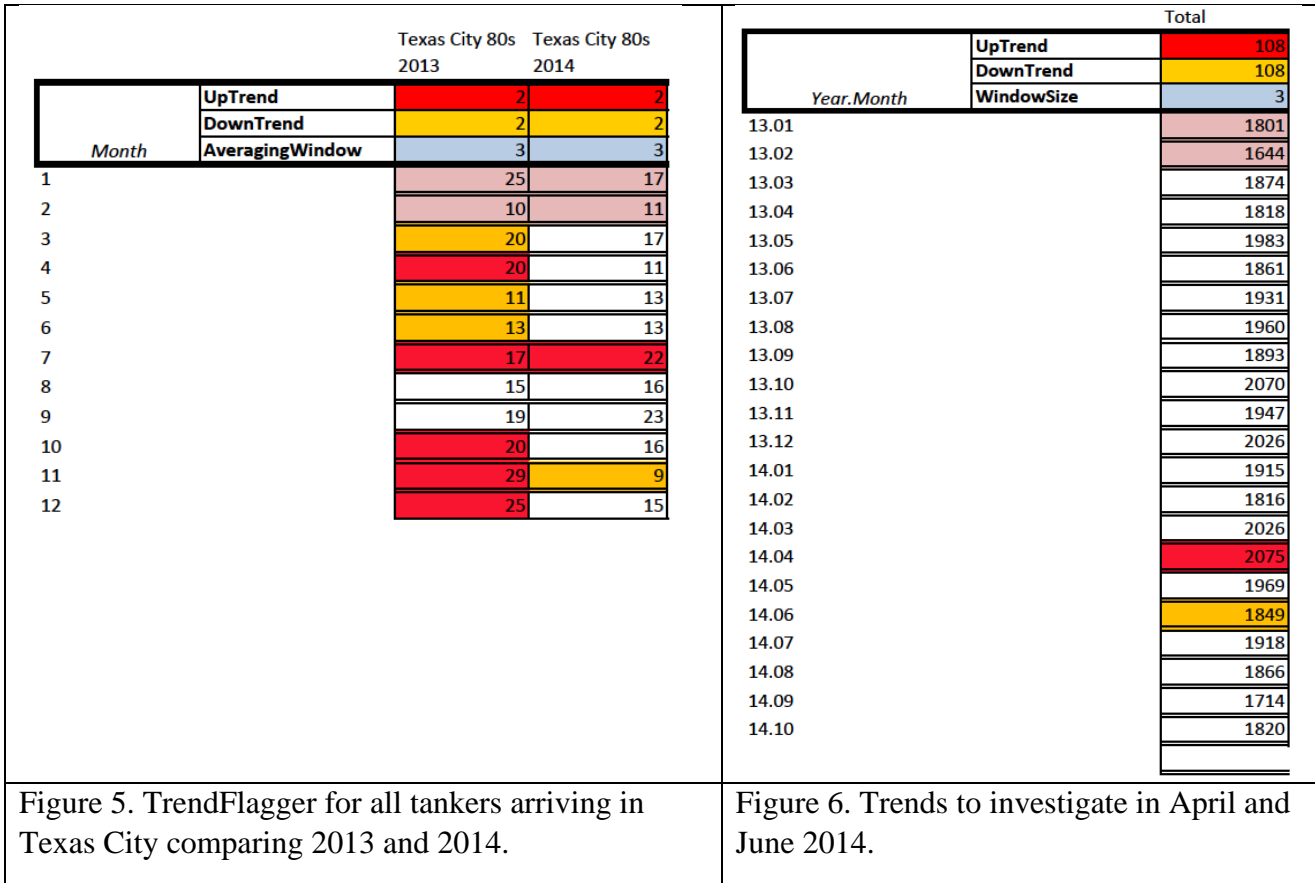


Figure 5. TrendFlagger for all tankers arriving in Texas City comparing 2013 and 2014.

Figure 6. Trends to investigate in April and June 2014.

**REFERENCES**

[1] Aarsaether, K.G., Moan, T. Estimating navigation patterns from AIS. *The Journal of Navigation*, 2009, 62, 587-607.

[2] Balduzzi, M., Wihoit, K., Pasta, A. Hey Captain, where’s your ship? Attacking vessel tracking systems for fun and profit. In *11<sup>th</sup> Annual HITB Security Conference in Asia*, <http://conference.hitb.org/hitbsecconf2013kul/materials/D1T1%20-%20Marco%20Balduzzi,%20Kyle%20Wilhoit%20Alessandro%20Pasta%20-%20Attacking%20Vessel%20Tracking%20Systems%20for%20Fun%20and%20Profit.pdf>, last accessed 2015/2/21.

[3] Cook, K. S., Nichols, D.L. Cyber seaworthiness: A call to action. In DiRenzo III, J., Drumhiller, N., & Roberts, F.S. (eds.), *Issues in Maritime Cyber Security*. PSO-Westphalia Press, 2017, pp. 81-85.

[4] CyberKeel. Maritime cyber-risks, Oct. 15, 2014, <http://www.sfm.org/support/amsc/cybersecurity/webdocs/Maritime%20Cyber%20Crime%2010-2014.pdf>, last accessed 2015/1/15.

[5] DiRenzo, J., Goward, D., Roberts, F.S. The little-known challenge of maritime cyber security. In *Proc. of the Int. Conf. on Information, Intelligence, Systems and Applications (IISA)*, pp. 1-5. IEEE, 2015. DOI: 10.1109/IISA.2015.7388071.

[6] International Maritime Organization. International convention for safety of life at sea (SOLAS), [http://www.imo.org/en/About/conventions/listofconventions/pages/international-convention-for-the-safety-of-life-at-sea-\(solas\)-1974.aspx](http://www.imo.org/en/About/conventions/listofconventions/pages/international-convention-for-the-safety-of-life-at-sea-(solas)-1974.aspx), last accessed 2017/4/17.

[7] Murphy, C. Moving averages: What are they? *Investopedia*: <http://www.investopedia.com/university/movingaverage/movingaverages1.asp>, last accessed 4/8/2017.

- [8] Net Help Security. Digital ship pirates: Researchers crack vessel tracking system, October 16, 2013. <http://www.net-security.org/secworld.php?id=15781>, last accessed 2015/2/21.
- [9] StockCharts.com. Moving averages – simple and exponential. [http://stockcharts.com/school/doku.php?id=chart\\_school:technical\\_indicators:moving\\_averages](http://stockcharts.com/school/doku.php?id=chart_school:technical_indicators:moving_averages) last accessed 2017/4/8.
- [10] Winkler, D. Enhancing the reliability of AIS through vessel identity verification. Presented at 2015 US DOT Datapalooza 16-17 June 2015, U.S. Department of Transportation, Washington DC. Technical Report, USCG Operations Systems Center, <https://www.fhwa.dot.gov/2015datapalooza/presentations/Safety.4-Winkler.pdf>

## ACKNOWLEDGEMENTS

The authors thank the United States Department of Homeland Security for its support under Grant 2009-ST-061-CCI002-08 to Rutgers University and *Fred Roberts* thanks the National Science Foundation for its support under Grant DMS1737857 to *Rutgers University*. The authors also thank their colleagues *Rong Chen, Minge Xie, and Linda Ness* who helped them explore the TrendFlagger and AIS data and its possible uses.