Machine Learning for Semi-Strong Efficiency Test of Inter-Market Wheat Futures

Phélippé-Guinvarc'h, Martial and Cordier, Jean

Actuary and assistant professor at the Risk and Insurance Institute of Maine University and at the GAINS (Le Mans, France)

1 June 2015

Machine Learning for Semi-Strong Efficiency Test of Inter-Market Wheat Futures

by

Martial Phelippe- Guinvarc'h and Jean E. Cordier

Suggested citation format:

Phelippe-Guinvarc'h, M., and J. E. Cordier. 2015. "Machine Learning for Semi-Strong Efficiency Test of Inter-Market Wheat Futures." Proceedings of the NCCC-134 Conference on Applied Commodity Price Analysis, Forecasting, and Market Risk Management. St. Louis, MO. [http://www.farmdoc.illinois.edu/nccc134].

Machine Learning for Semi-Strong Efficiency Test of Inter-Market Wheat Futures

Martial Phélippé-Guinvarc'h and Jean E. Cordier

June 1, 2015

PROSPECTUS FOR:

2015 NCCC-134 Conference on Applied Commodity Price Analysis, Forecasting, and Market Risk Management

Actuary and assistant professor at the Risk and Insurance Institute of Maine University and at the GAINS (Le Mans, France)

Martial.Phelippe-GuinvarcH@univ-lemans.fr

Professor at the UMR1302 SMART and the Agrocampus Ouest (Rennes, France)

jean.cordier@agrocampus-ouest.fr

Arbitrage is one of the central tenets of financial economics, enforcing the law of one price and keeping markets efficient. Theoretically, when spreads are observed, the arbitrage gives a positive return, requires no capital and is riskless. But, in reality, important impediments, due to market frictions and imperfect information, can limit arbitrage.

The semi-strong efficiency form tests, defined by Fama (1970, 1991), are concerned whether current prices "fully reflect" all obviously publicly available information. This article establishes a large literature on the efficiency of markets, including agricultural commodities futures, although **?** note the relatively few semi-strong studies with respect to weak tests.

Machine learning methods, detailed for example by Hastie et al. (2008), are used in finance literature to test market efficiency. Recently, Hsu and Chen (2014) used growing hierarchical self-organizing maps about hedging. Under Fama's framework, authors used neural networks on agricultural commodities, such as Hamm and Brorsen (2000) or Hamm et al. (1993) for weak tests. In the case of Taiwan stocks indexes, Hsu et al. (2011) used machine learning to model the inter-market opportunities of arbitrage. First, they compute arbitrage conditions on a training period, then they apply a machine learning approach (an extended classifier system method) to estimate the model and finally they test the model's quality over a testing period. Implicitly, such a method is used as a weak form test of *inter-market* efficiency.

Commodity futures prices are more complex processes than those of stocks, bonds and other financial assets. They strongly depend on the cash market, with supply and demand seasonality, quality and storage issues. Moreover, inter-market arbitrage cannot be computed directly on one unique futures price but only on their spread. In fact, wheat futures specifications vary by Exchanges with respect to quality specifications, delivery point(s), maturity dates and currency of quotation. As a consequence, Garcia and Leuthold (2004), in their review of the literature, indicate the pertinence of using projected balance sheets and transport index-based values for inter-market studies.

This paper aims to provide a semi-strong test of the efficiency of inter-market wheat futures, using an original machine learning method. Unlike Hsu et al. who use extended the classifier system method, we choose the classification and regression tree (CART) algorithm first proposed by Breiman et al. (1984). CART allows a shift from a weak to a semi-strong test by introducing public information into the model.

1 Data and empirical procedures

The semi-strong form of market efficiency hypothesis requires that prices reflect all publicly available information. An investor cannot benefit over and above the market by trading on new information. Then, to test this hypothesis, we compose data base with wheat quotes and with public information.

1.1 Wheat futures prices

The data include US and EU Wheat futures exchange-listed on the Chicago Mercantile Exchange (Chicago Soft Red Winter Wheat) and on the EuroNext (Milling Wheat). [1]

US Exchange quote in dollar per bushel and EU Exchanges quote in euros per metric ton. Clearly, we have to convert EU prices in dollar per bushel or convert US prices in euro per metric ton to compute the spread. Convert EU prices is the best solution because:

1. CME Wheat begins in 1960 and EuroNext Wheat begins in 1998. Because of Euro FX Futures (EC) begins in 1999, we would lose the 1969-1999 historical data by converting US price in euro per ton.

2. Transportation index published by the Baltic Exchange (see subsection 1.4 for details) is computed from prices in dollar per ton and per days. Explain the value of arbitrage in euro with transportation data in dollar seems inappropriate.

European Futures Wheat prices are converted in dollar cents per bushel since 1999 with a technical rate of 36.7437 bushels per metric ton.

To obtain the price in dollar, we use the Euro Dollar Futures quotation of Chicago Mercantile Exchange (CME Euro FX Futures (EC)) with the nearest maturity.[2]

1.2 Inter-market arbitrage computation

Arbitrage is taking advantage of a price spread of same financial product (or equivalent product). Triantafyllopoulos and Montana (2011) explain that 'possibly the simplest of such strategies consists of a portfolio of only two assets, as in pairs trading.'

In this paper, this trading approach consists in going long a wheat futures A while shorting another wheat futures B. *This portfolio has only exposure to changes of spreads but*

[1] Kansas City Hard Red Winter Wheat and Black Sea Wheat (CME) and UK Feed Wheat (ICE) could be include in the next stage of research.

[2] The CME Euro FX Futures (EC) offers four maturities per year. For example, the January futures prices in dollar per bushel are obtained from the January futures prices in euros per ton and from the mars CME Euro FX Futures prices.

not to wheat price trends. The underlying assumption that futures wheat must be priced with a explainable spread and then, pairs trading is a mean-reverting strategy.

Even if quotations concern the same commodity, such as wheat, each futures contract presents different specifications (maturity date, delivery point(s), wheat quality specifications...). We compute spread arbitrage on daily historical futures data and calculate the return and time to close trading. For each settlement day d, for each futures combination A_B, the ${}_{A\_B}ARB(t)$ arbitrage return is, if positive, the max of spread in N next days minus initial spread $(B(t) - A(t))$ and transaction costs.

We note the name of pair trading as A_B. Using the settlement price, we take a short position on A and a long position on B. When spread $B - A$ increases, this both positions give a positive return. Arbitrage is computed to retain the best return during the next 6 weeks. When spread $B - A$ deacreases, the best return is realized at time t, ie 0.

$${}_{A\_B}ARB(t) = \max_{d \leq 42} \{A(t) - A(t+d) - (B(t) - B(t+d))\}$$

where t is the present time and d represents duration in days, the first $t + d$ where the max is obtained is the cloture date of arbitrage. The close time is the nearest day where max is reached.

When A is a European wheat quote in euros per metric ton, and B is an US wheat quote, we adjust the previous formula:

$${}_{A\_B}ARB(t) = \max_{d \leq 42} \{\gamma\,(A(t) - A(t+d)) \times EC(t+d) - (B(t) - B(t+d))\}$$

where γ is the conversion from metric ton to bushel cent (equal to 36.7437/100) and $EC(t+d)$ is the euro dollar spot price at $t + d$.[3]

The figure 1.2 shows respectively (1) EBMH2015 and WH2005 quotations (maturity Mars 2015), (2) the synchronized spread, (3) the result of arbitrage EBMH2015_WH2015, (4) the result of arbitrage WH2015_EBMH2015 and (5) the euro dollar parity ECH2015. All values are in US dollar cents per bushel except euro dollar.

A naive reading of figure 1.2 led to think that EBMH2015 and WH2005 track closely and, consequently, that the world wheat market is working. Nevertheless, the spread moves

[3] Just by reversing the roles of A and B, when A is an US wheat contract, and B is an european wheat, the previous formula becomes:

$${}_{A\_B}ARB(t) = \max_{d \leq 42} \{(A(t) - A(t+d)) - \gamma\,(B(t) - B(t+d)) \times EC(t+d)\}$$

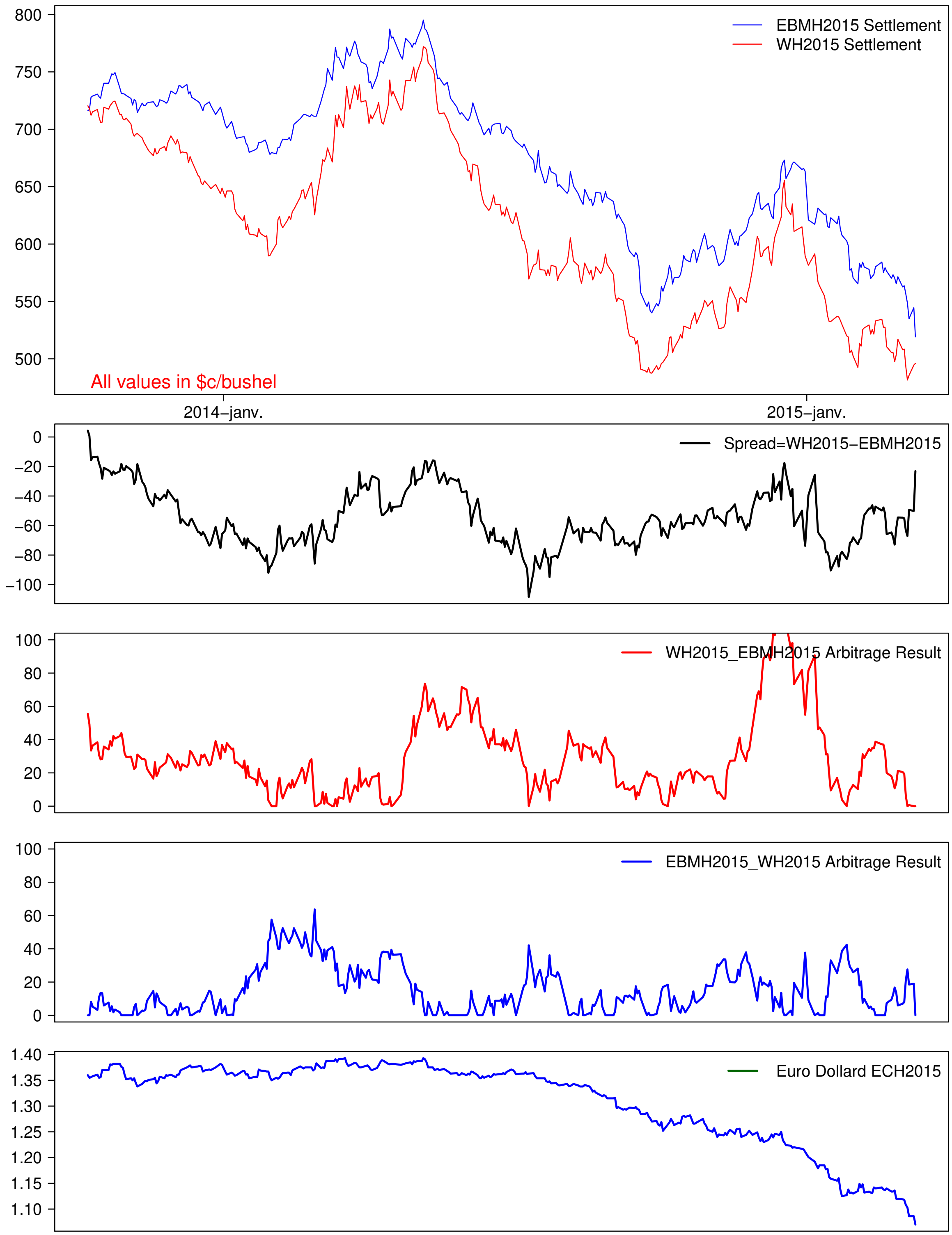

Figure 1: Synchronized prices and arbitrage computation (maturity Mars 2015)

between -108 to 4\$c/bu. The volatility of spread creates pairs trading arbitrage opportunities up 120\$c/bu (18.45% of the WH2015 price). The mean of EBMH2015_WH2015 arbitrage is 29.03 and represents 4.6% of the WH2015 price.

We test too the presence of trend in response variable. In figure 1.2 (on the left), we compute the mean by harvest of arbitrage return calculated in \$c/bu. We observe a significative (5%) positive trend, 0.6825\$ per year. However, when the arbitrage is divided by the price of the long position, the trend is no more significative.

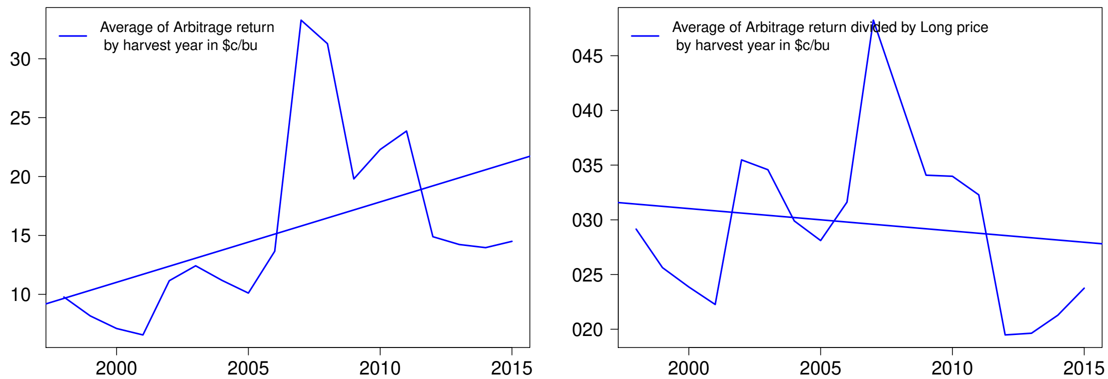

Figure 2: Trend of arbitrage returns

The figure 3 presents the density of arbitrage returns in cents per bushel divided by the price of the long position. As understood, arbitrage returns exhibits a non-normal density. Because arbitrage returns are calculated as a maximum of random variables, the general extreme theorem suggests that arbitrage distribution converges to an extremum generalized distribution. A simple fit of the generalized extrem value density on data gives a significatif tail index ξ estimated to 0.71. The $\xi > 0$ suggests the presence of heavy tail density and that arbitrage returns follows Fréchet law.[4]

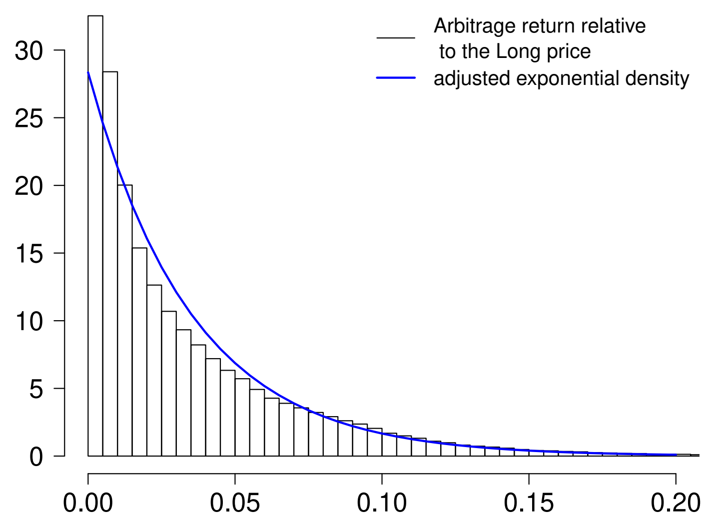

Figure 3: L'arbitrage

[4]The Fréchet distribution is the generalized extreme value distribution in the case $\xi > 0$. It has the

1.3 World Agricultural Supply and Demand Estimates

We add historical World Agricultural Supply and Demand Estimates (WASDE) data as provided by the USDA (www.usda.gov/oce/commodity/wasde). Monthly, WASDE reports a full balance sheet for each commodity. It includes forecasts for US, EU and world wheat. Separate estimates are made for the components of supply (beginning stocks, imports, and production) and demand (domestic use, exports, and ending stocks). Forecast balances could explain calendar spread and zone details could explain spatial spreads.

We would add in quotes and arbitrage database new variables as beginning stocks of production, imports, and production, domestic use, exports, and ending stocks for the long contract and for the short contract. This merge presents two difficulties:

1. we have to join the daily calculated arbitrage database with the monthly WASDE database. For each cotation date, we take on the last published WASDE data in the merge. For example, we consider the arbitrage at April 20^{th} 2015 between the CME wheat futures (long WK2015), maturity September 2015 and the European Wheat futures, maturity December 2015 (short EBMZ2015). At the April 20^{th} 2015, the last WASDE publication date sets for April 9^{th} 2015. Then, for the calculated arbitrage starting the April 20^{th} 2015, all merged WASDE data come from the report published the April 9^{th} 2015.

2. Spread arbitrage combines two futures contracts (WK2015 and EBMZ2015). Both contracts can be on two different crop years or on two different crop regions. Then, we have to merge WASDE data twice, one merge per contract. Each merge is performed for wheat on regions (respectively US and EU), and harvest year (production for 2014/2015)..

Then, we calculate some variables as the ratio of production on domestic total use ($ratioPU$), the ratio of export on sum of international exchange ($ratioEI$), the ratio of ending stocks on beginning stocks ($ratioS$). We calculate the harvest year spread ($DiffC$) and some ratios of ratio between contracts: $DiffratioS$ refers to the ratio of $ratioS$, $DiffPU$ the ratio of $ratioPU$ and $DiffEI$ the ratio of two $ratioEI$. First, ratio aims to synthetize information from several variables to perfom the regression tree. Second, ratio is less sensitive to trend.

cumulative distribution function

$$\Pr(X \le x) = e^{-x^{-1/\xi}} \text{ if } x > 0.$$

where $1/\xi > 0$ is a shape parameter.

1.4 Transportation data

To explain spatial spread, we have to include the value of transportation in the model. At this stage of research, we propose to use the Baltic Exchange Dry Index (BDI). BDI is a measure of the price of shipping major raw materials such as metals, grains, and fossil fuels by sea. It is created by the London Baltic Exchange based on daily assessments from a panel of shipbrokers and is quote-listed on the Baltic Exchange.[5]

The graph illustrates the BDI freight index levels and volatilities. From 2000 to mid-2008, freight rates increased by to rose to 11 793, the record high of BDI, on 20^{th} May 2008. In few month, it dropped to 663 on 5^{th} December 2008.

Empirical results of Chang et al. (2014) confirm that the density of the BDI exhibits the significant asymmetric long-memory property of volatility and the fat-tail phenomenon. As the consequence, the estimation of the trend largely depends on the used period for estimation. Therefore, detrending the BDI time-series does not appear to be relevant. The ratio BDI on begin long price is computed also.

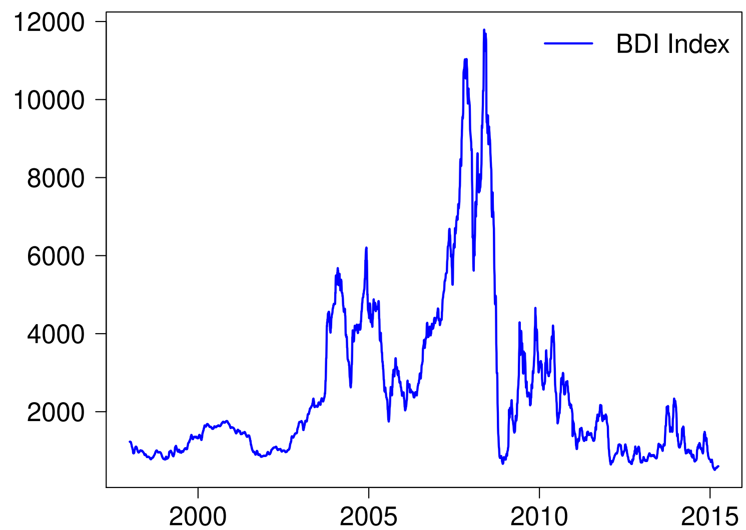

Figure 4: Historical data of BDI freight index

2 Methodology

2.1 CART formulation

Classification And Regression Tree (CART), first proposed by Breiman et al. (1984), could be used for classification or regression even if the method differs to determine where to split. In Classification, CART compute the Gini impurity coefficient to split. [6] In Regression,

[5] The BDI is a composite of 3 sub-indices, each covering a different carrier size: Capesize, Panamax, and Supramax. Capesize carriers are the largest ships with a capacity greater than 150,000 DWT. Panamax refers to the maximum size allowed for ships travelling through the Panama Canal, typically 65,000 - 80,000 DWT. The Supramax Index covers carriers with a capacity of 50,000 - 60,000 DWT.

[6] It is a measure of how often a randomly chosen element from the set would be incorrectly labeled if it were randomly labeled according to the distribution of labels in the subset.

variance reduction is often employed to determine where to split Hastie et al. (2008).

CART algorithm fits the response value Y, the arbitrage results, is a function as $\hat{f}_w(x)$ that depends on X (i.e. $E[Y|X]$) and a complexity parameter w, the number of final node in CART algorithm. After the data were split into independent samples, the training and the testing periods. It is carried out in two stages:

First, the algorithm is chosing the splitting variable (sv) and the splitting point (sp) using training sample. CART algorithm estimates on the training period the function $\hat{f}_w(x)$, the regression tree of arbitrage return.

Secondly, the resultant model is sometimes too complex and overfits data. This stage of procedure consists to trim back the full tree and aims to establish the predictability of the model. Cross-validation is a model validation technique for assessing how the results of a statistical analysis will generalize to an independent data set. Then, on the testing periods, we estimate w that minimize the error of function $\hat{f}_w(x)$ and prune the tree as a consequence.

Where splitting node ? In the general case, CART estimates the function $\hat{f}_w(x)$ could be writed as:

$$\hat{f}_w(x) = \sum_{j=1}^{w} \bar{Y}_{j,w} \times \mathbb{1}_{\{x \in R_{j,w}\}}$$

where w is the number of final node (the complexity), $\mathbb{1}_{\{x \in R_{j,w}\}}$ is the indicator function of j-node $R_{j,w}$ and $\bar{Y_{j,w}}$ is the average in subset j. If we split a node j into two sons (left and right sons), we will have

$$D_{parent} = \sum_{x_i \in R_j} (y_i - \bar{y_i})^2$$

For each $j-$node split, the deviance reduction R could be computed as $R = D_{parent} - (D_{left\ son} + D_{right\ son})$. The splitting process must be stop only when some minimum node size is reached. *This general case assumes the normality of errors.*

Davis and Anderson (1989) grew trees by assuming survival times to be exponential within a given node. The method is similar. Instead estimated y, the response variable, the model estimates the hazard rate of y, $h(y)$ and instead mean square criteria to split, the model uses the log-likelihood.

The models specifies the *hazard rate function* as following:

$$\hat{h}_w(y_i) = \sum_{j=1}^{w} \lambda_j \mathbb{1}_{\{y_i \in R_{j,w}\}} \quad j = 1, 2, \cdots, w$$

for j^{th} terminal node y_j is estimated by $\hat{f}_{j,w} = 1/\lambda_j$, because of expected value of exponential density.

Algorithm splits node on the basis of exponential log-likelihood. The split selected is the partition that minimizes for a node j, the proposed loss function is

$$D_j = \hat{l}_j = n_j - n_j \ln\left(\frac{n_j}{\sum_{x_i \in R_j} y_i}\right)$$

where n_j is the number of complete observations at the node j.

How pruning the tree ? Cross-validation involves partitioning a sample of data into complementary subsets, performing the analysis on the training set, and validates the analysis on the other subset, the testing set. Steps could be repeated using different partitions to reduce de variability of estimator. This step gives the predictive capacity on model. CART uses the cost complexity criterion to the validation and overfits nodes are prune on the tree:

$$C_\alpha(T) = \sum_{j=1}^{|T|} D_j + \alpha|T|$$

when terminal nodes of tree T are indexed by m (Region R_m).In usual regression case, D_j is the sum of square errors but with exponential errors:

$$C_\alpha(T) = \sum_{j=1}^{|T|} n_j - n_j \ln\left(\frac{n_j}{\sum_{x_i \in R_j} y_i}\right) + \alpha|T|$$

2.2 Computing specifications

Because of volum, we choose an SQL database manager (mysql[7]) to build and manage databases. We compute with the useful rpart packages in R [8] to build CART regression tree (Therneau et al. (2009)). Please note that in CART, the first subgroup is set to be the reference level, ie $\hat{h}(y_i) = \frac{\lambda_j}{\lambda_0} \mathbb{1}_{\{x \in R_{j,w}\}}$. Then, in the root node, hazard rate estimated is always equal to 1 and the predicted value is:

$$\hat{h}(y_i) = \frac{\lambda_0}{\lambda_j} \mathbb{1}_{\{x_i \in R_{j,w}\}}$$

[7] https://www.mysql.fr/

[8] http://cran.r-project.org/

'exponential' method is selected in CART to compute the hazard rate estimation.

rpart function needs 6 parameters : *xval*, *minbucket*, *maxcompete*, *maxsurrogate*, *cp* et *maxdepth*.

The *xval* parameter refers to the number of cross-validation to prune the tree. This method has proved very reliable for screening out 'white noise'variables in the data set. The cross-validation has to divide the data set into *xval* groups $G_1, G_2 \ldots G_{xval}$, the size of each group is reduced into $n/xval$. Because of the high variance of response variable, this method is not appropriate.

CART algorithm is a stepwise procedure. The question of when to stop algorithm is important. *minbucket* parameter gives one of stop criteria and refers to the minimum size of a final node. A *minbucket* too high is loosing the singularity estimation of the data. A value too small lets the algorithm cost time computation and build a large number of nodes that will be pruned in second stage of CART.

maxcompete is useful for printing or ploting results.

CART performs the fit of tree in presence of missing values. One approach is to estimate the missing datum using the other independent variables; rpart uses a variation of this to define surrogate variables. When an explained variable is missing the split use the first surrogate variable, or if missing that, the second surrogate is us, etc... *maxsurrogate* parameter defines the maximal variables useful to surrogate a missing value. A missing value in explain data do not implies that rows is loosed in regression. Our data sources give rarely missing values, the *maxsurrogate* is fixed to 2.

cp is a second algorithm stop parameter that uses the complexity w and that aims to optimise time calculation. This cost parameter is fixed to 0.001.

maxdepth is the last stop parameters and refers to the maximal depth of tree between 1 to 30. We fixe this value to 12.

2.3 Model

Our model estimates the arbitrage returns from publicly available information with the following hazard equation:

$$\hat{h}_j(y_i) = \lambda_j \mathbb{1}_{\{x_i \in R_{j,w}\}} \quad j = 1, 2, \cdots, w$$

If the predictibility of this model is prooved, we deduce that the Fama's semi-strong efficiency form hypothesis is not verified in world wheat futures markets and deduce pair trading strategies.

3 Results and practical implications

3.1 Tree analysis

Major result is the tree proposed in figure 5 page 11 and figures 6, 7, 8 pages 15–17. The first figures presents the high part of tree and the followings present the tree sub-trees. Tree is fit on 1999-2009 harvest years data and is cross validated since 2010 harvest year. In *split nodes*, the tree indicates the split variable (sv) and the (number of rows) in node. Over theedge from a parent to a child node, the tree indicates the split point (sp). In *final nodes*, the tree indicates the estimated response value, the return arbitrage divided by the long futures price (Y) in percent (%) and the (number of rows) in final node.

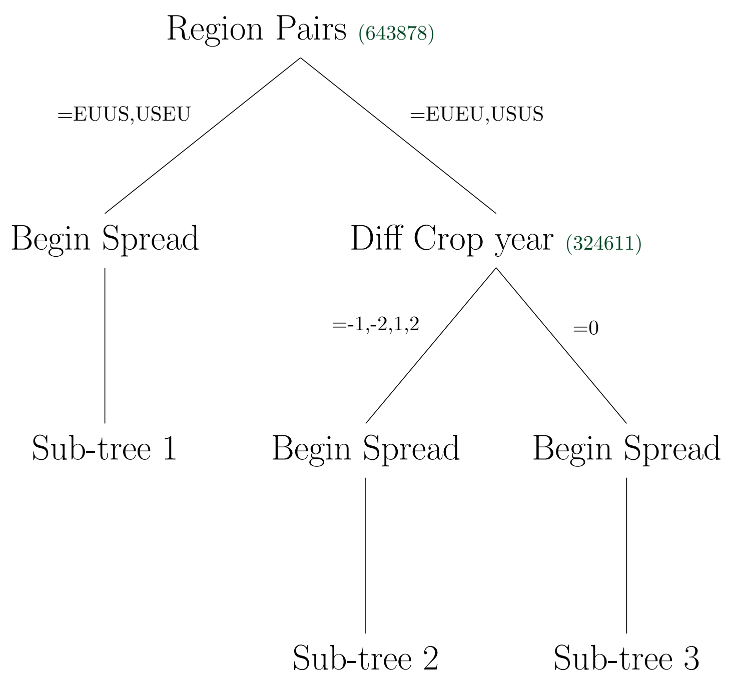

Figure 5: The high part of tree

The tree tells us an hierarchy of variables in model. First variable split intra-regional arbitrage than inter-regional arbitrage. Second variable cut off the arbitrage that have the same crop years from others. Les us note that these both major variables concern the proximity or homogenity of the two futures in the pair trading. Next, the Begin Spread is a recurrent pertinent separator variable. This result is logical if we considere that pair trading is a mean-reverting strategy.

The month of quote is an interesting variable pertinent in the tree. SymbPM refers to the contract and the maturity month of contract, independly of the year. This both previous variables are intrinsic to market. Because not exogeneous public information is needed to predict arbitrage return (in first nodes), we deduce that weak test hyptothesis is not completly verified.

Public information variables perform the CART regression tree. We found wasde data ratios of begining stocks ond ending stocks and ratios of production on uses. The transportation index on begin long futures price split two nodes. This variables in tree prooves that public information help to predict arbitrage return.

3.2 Some statistics

Table on the next page gives the cost complexity calculated on the testing period for each created node:

$$C_\alpha(T) = \sum_{j=1}^{|T|} n_j - n_j \ln\left(\frac{n_j}{\sum_{x_i \in R_j} y_i}\right) + \alpha|T|$$

with α fixed to 0.001. $\alpha > 0$ assure that equal deviance, the valided tree will be as small as possible. When α is too high, nodes are not splitting even if the split reduces significatively the deviance. The min is reached for the 19^{th} node split and the validated tree has 23 final nodes.

The tail statisitics of errors is really interesting with $\xi = -0.0170$. The tail index of errors is negative and near of 0. This result is really different than that of the arbitrage returns ($\xi = 0.707$) and confirms that the choice of exponential errors is reasonable.

We want first to validate the use of CART for a semi-strong efficiency test, then to conclude on the inter-market inefficiency of wheat futures and finally, to give a public and useful arbitrage filter on wheat spread.

Summary and conclusions

This paper proposes an original work on world wheat futures market efficiency test to conclude on the semi-strong inefficiency of wheat futures.

Our model uses american and european data together to estimate pair trading arbitrage returns on the wheat futures market. Some variables like transportation and balance sheet of USDA are significative in CART regression. Then, pair trading arbitrage is predictible with public information and we deduce of the semi-strong inefficiency of inter-market wheat futures.

Splits	Cost complexity
1	1.0010
2	0.8397
3	0.8224
4	0.8227
5	0.8066
6	0.8051
7	0.8055
8	0.7867
9	0.7730
10	0.7636
11	0.7597
12	0.7607
13	0.7615
14	0.7601
15	0.7609
16	0.7614
17	0.7619
18	0.7610
19	0.7593
20	0.7597
21	0.7607
22	0.7617
23	0.7626
24	0.7636
25	0.7646
26	0.7655
27	0.7665
28	0.7673
29	0.7679
30	0.7689
31	0.7698
32	0.7707
33	0.7716
34	0.7720
35	0.7714
36	0.7722
37	0.7732
38	0.7740
39	0.7630

The recent paper of Garcia et al. (2014) deals with spread between spot and futures at maturity and explains also the failure of the futures market. They show that the magnitude of the non convergence could be explain by the wedge, the difference between the price of carrying physical grain and the cost of carrying delivery instruments.

In contrast, some recent papers conlude in favor of efficiency of wheat futures contracts. For example, Hamm et al. (1993; 2000) use neural network on wheat futures contracts and Kristoufek and Vosvrda (2014) use econometric approach to support this result. In fact, the random-walk-based efficiency tests (or martingale efficiency tests) imply that price returns of one contract integrate new informations independantly of past information. But these tests do not imply that two prices returns integrate new information with consistency.

Our research confirms that efficiency tests based on spread movement could provide the opposite result that efficiency tests based on prices returns. To go further, we could include new wheat futures contracts in data set, experiment new public informations and test new machine learning algorithms in this context.

References

Breiman, L., Friedman, J. H., Olshen, R. A., and Stone, C. J. (1984). *Classification and regression trees*. Chapman & Hall/CRC.

Chang, C.-C., Chih Chou, H., and Chou Wu, C. (2014). Value-at-risk analysis of the asymmetric long-memory volatility process of dry bulk freight rates. *Marit Econ Logist*, 16(3):298–320.

Davis, R. B. and Anderson, J. R. (1989). Exponential survival trees. *Statistics in Medicine*, 8(8):947–961.

Fama, E. F. (1970). Efficient capital markets: A review of theory and empirical work*. *The journal of Finance*, 25(2):383–417.

Fama, E. F. (1991). Efficient capital markets: Ii. *The Journal of Finance*, 46(5):1575–1617.

Garcia, P., Irwin, S. H., and Smith, A. (2014). Futures market failure? *American Journal of Agricultural Economics*, 97(1):40–64.

Garcia, P. and Leuthold, R. M. (2004). A selected review of agricultural commodity futures and options markets. *European Review of Agricultural Economics*, 31(3):235–272.

Hamm, L. and Brorsen, W. B. (2000). Trading futures markets based on signals from a neural network. *Applied Economics Letters*, 7(2):137–140.

Hamm, L., Brorsen, W. B., and Sharda, R. (1993). Futures trading with a neural network. In *NCR-134 Conference on Applied Commodity Analysis, Price Forecasting and Market Risk Management Proceedings*, pages 286–296.

Hastie, T. M., Tibshirani, R., and Friedman, J. (2008). *The Elements of Statistical Learning*. Springer Series in Statistics.

Hsu, Y.-C. and Chen, A.-P. (2014). A clustering time series model for the optimal hedge ratio decision making. *Neurocomputing*, 138:358–370.

Hsu, Y.-C., Chen, A.-P., and Chang, J.-H. (2011). An inter-market arbitrage trading system based on extended classifier systems. *Expert Systems with Applications*, 38(4):37–84.

Kristoufek, L. and Vosvrda, M. (2014). Commodity futures and market efficiency. *Energy Economics*, 42:50–57.

Therneau, T. M., Atkinson, B., and Ripley, B. (2009). *Rpart: Recursive Partitioning*. R package version 3.1-45.

Triantafyllopoulos, K. and Montana, G. (2011). Dynamic modeling of mean-reverting spreads for statistical arbitrage. *Computational Management Science*, 8(1-2):23.

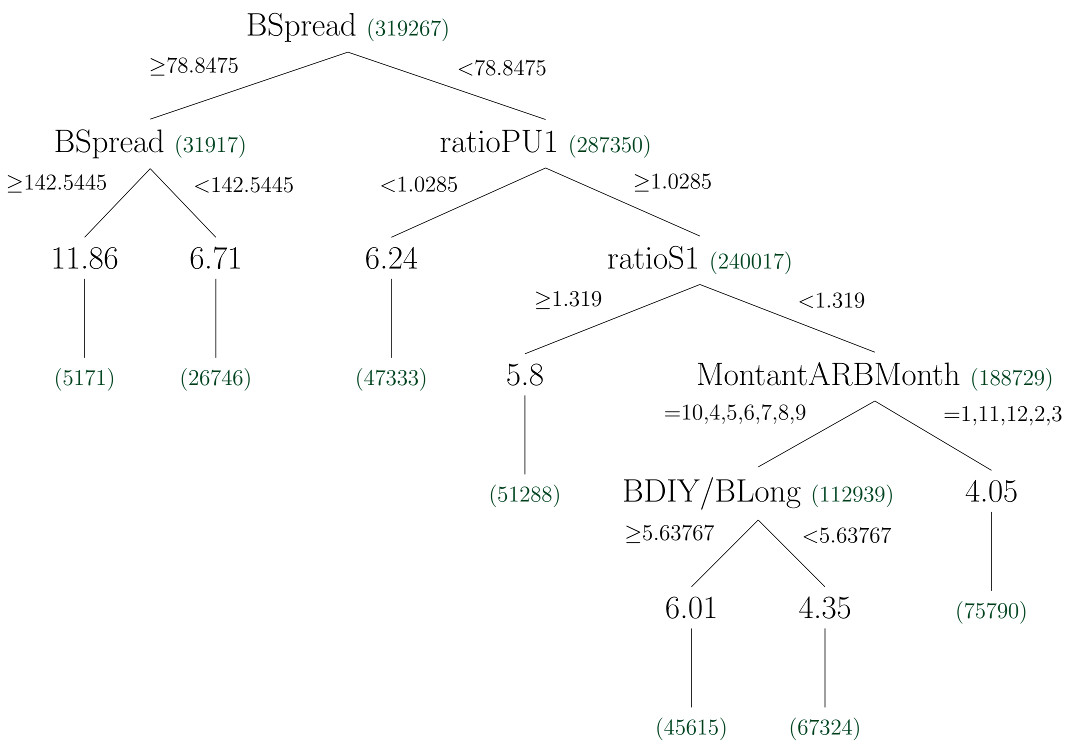

Figure 6: First sub-tree

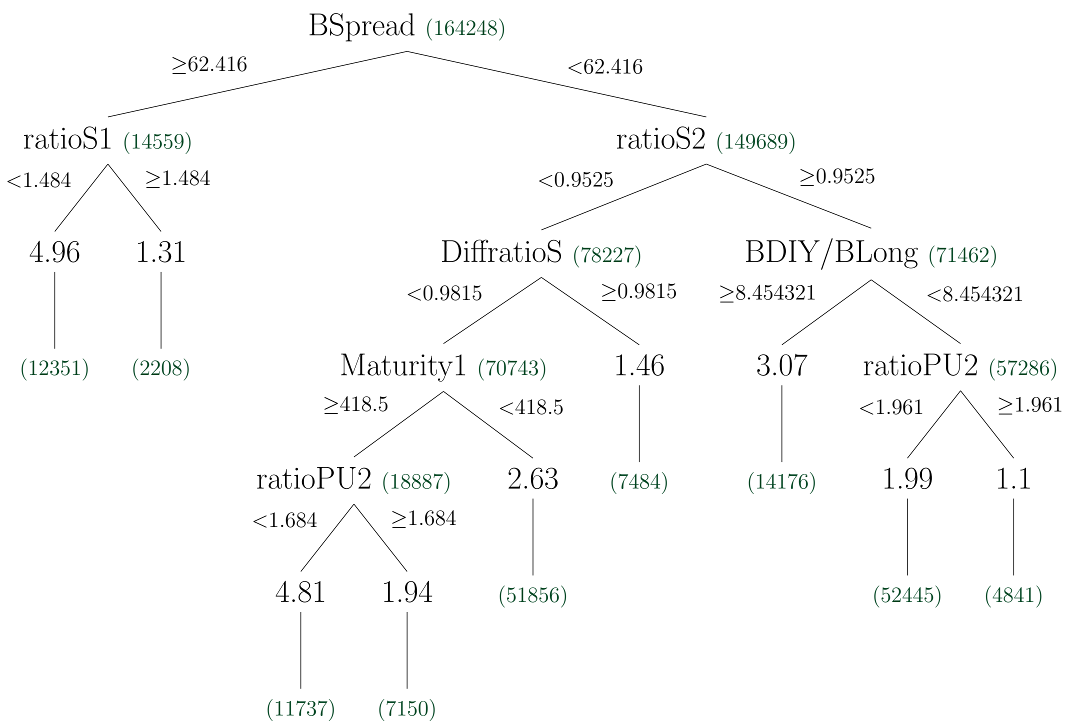

Figure 7: Second sub-tree

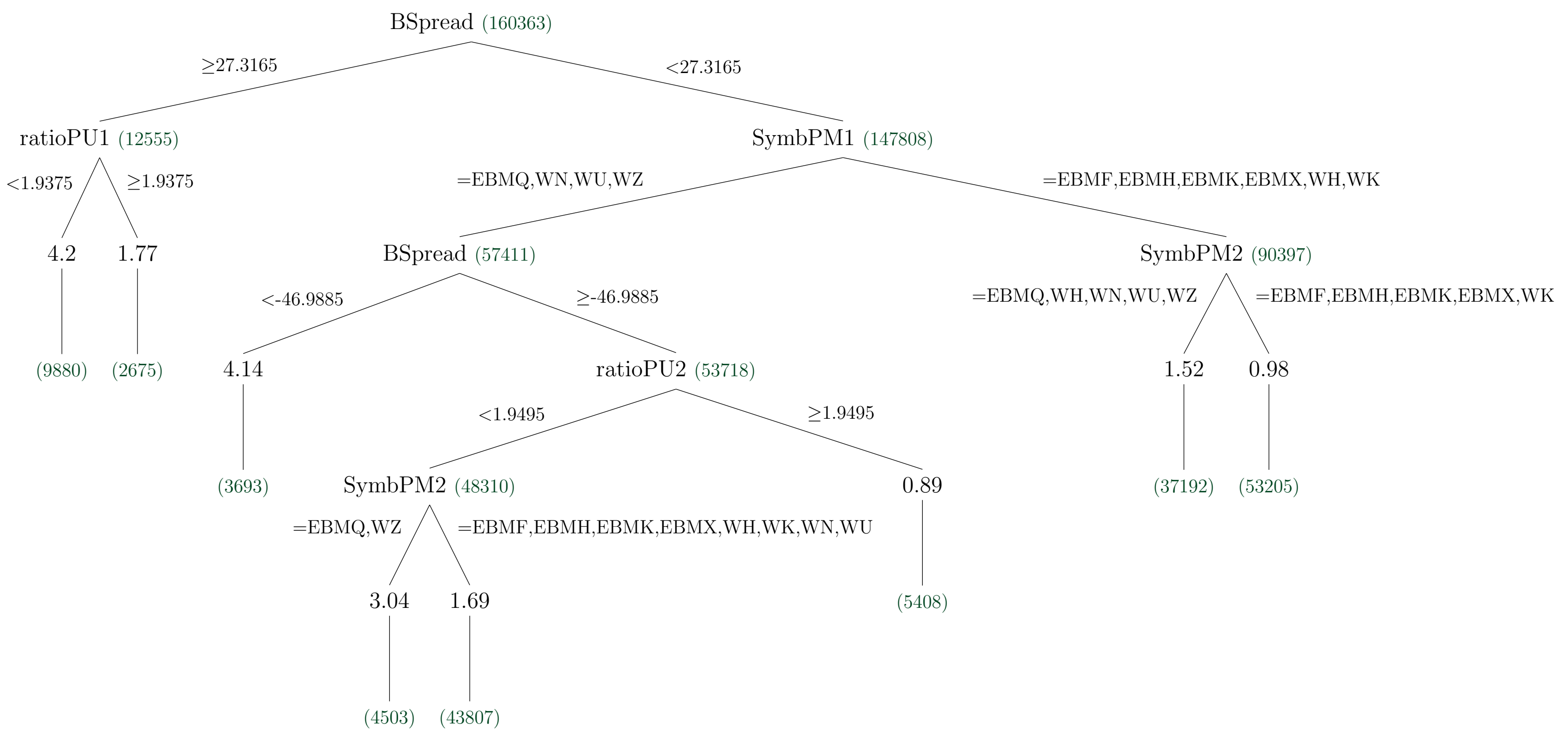

Figure 8: Third sub-tree