# Measuring Cultivation Parcels with GPS: a Statistical Evidence

Gabriele PALMEGIANI
*gabriele.palmegiani@gmail.com*

June 12, 2009

# Contents

# Studying the statistical relevance of measuring cultivation parcels with GPS: the data set

- African survey data (2005-2006) for Cameroon, Niger, Madagascar and Senegal.
- Cultivation parcel selection **was not random**.
- The surfaces of cultivation parcels were measured using:
  - the compass and meter method (or traditional method) as statistical point of reference.
  - more types of GPSs (GPS60, GPS72 and Magellan400).

# Studying the statistical relevance of measuring cultivation parcels with GPS: the data set

- African survey data (2005-2006) for Cameroon, Niger, Madagascar and Senegal.
- Cultivation parcel selection **was not random**.
- The surfaces of cultivation parcels were measured using:
  - the compass and meter method (or traditional method) as statistical point of reference.
  - more types of GPSs (GPS60, GPS72 and Magellan400).

# Studying the statistical relevance of measuring cultivation parcels with GPS: the data set

- African survey data (2005-2006) for Cameroon, Niger, Madagascar and Senegal.
- Cultivation parcel selection **was not random**.
- The surfaces of cultivation parcels were measured using:
  - the compass and meter method (or traditional method) as statistical point of reference.
  - more types of GPSs (GPS60, GPS72 and Magellan400).

# Studying the statistical relevance of measuring cultivation parcels with GPS: the covariates of interest

The covariates of interest (all expressed in squared meters) are:

- **S_1** = Compass and meter cultivation parcel surface.
- **S_21_1** = Garmin60 (GPS60) cultivation parcel surface.
- **S_22_1** = Garmin72 (GPS72) cultivation parcel surface.
- **S_24_1** = Magellan400 (MAG400) cultivation parcel surface.

We shall interested in the following differences:

1. **c_g60diff** = The difference between compass cultivation parcel surface and Garmin60 cultivation parcel surface.
2. **c_g72diff** = The difference between compass cultivation parcel surface and Garmin72 cultivation parcel surface.
3. **c_m400diff** = The difference between compass cultivation parcel surface and Magellan400 cultivation parcel surface.

# Studying the statistical relevance of measuring cultivation parcels with GPS: the covariates of interest

The covariates of interest (all expressed in squared meters) are:

- **S_1** = Compass and meter cultivation parcel surface.
- **S_21_1** = Garmin60 (GPS60) cultivation parcel surface.
- **S_22_1** = Garmin72 (GPS72) cultivation parcel surface.
- **S_24_1** = Magellan400 (MAG400) cultivation parcel surface.

We shall interested in the following differences:

1. **c_g60diff** = The difference between compass cultivation parcel surface and Garmin60 cultivation parcel surface.
2. **c_g72diff** = The difference between compass cultivation parcel surface and Garmin72 cultivation parcel surface.
3. **c_m400diff** = The difference between compass cultivation parcel surface and Magellan400 cultivation parcel surface.

# Studying the statistical relevance of measuring cultivation parcels with GPS: the covariates of interest

The covariates of interest (all expressed in squared meters) are:

- **S_1** = Compass and meter cultivation parcel surface.
- **S_21_1** = Garmin60 (GPS60) cultivation parcel surface.
- **S_22_1** = Garmin72 (GPS72) cultivation parcel surface.
- **S_24_1** = Magellan400 (MAG400) cultivation parcel surface.

We shall interested in the following differences:

1. **c_g60diff** = The difference between compass cultivation parcel surface and Garmin60 cultivation parcel surface.
2. **c_g72diff** = The difference between compass cultivation parcel surface and Garmin72 cultivation parcel surface.
3. **c_m400diff** = The difference between compass cultivation parcel surface and Magellan400 cultivation parcel surface.

# Studying the statistical relevance of measuring cultivation parcels with GPS: the covariates of interest

The covariates of interest (all expressed in squared meters) are:

- **S_1** = Compass and meter cultivation parcel surface.
- **S_21_1** = Garmin60 (GPS60) cultivation parcel surface.
- **S_22_1** = Garmin72 (GPS72) cultivation parcel surface.
- **S_24_1** = Magellan400 (MAG400) cultivation parcel surface.

We shall interested in the following differences:

1. **c_g60diff** = The difference between compass cultivation parcel surface and Garmin60 cultivation parcel surface.
2. **c_g72diff** = The difference between compass cultivation parcel surface and Garmin72 cultivation parcel surface.
3. **c_m400diff** = The difference between compass cultivation parcel surface and Magellan400 cultivation parcel surface.

# Studying the statistical relevance of measuring cultivation parcels with GPS: the covariates of interest

The covariates of interest (all expressed in squared meters) are:

- **S_1** = Compass and meter cultivation parcel surface.
- **S_21_1** = Garmin60 (GPS60) cultivation parcel surface.
- **S_22_1** = Garmin72 (GPS72) cultivation parcel surface.
- **S_24_1** = Magellan400 (MAG400) cultivation parcel surface.

We shall interested in the following differences:

1. **c_g60diff** = The difference between compass cultivation parcel surface and Garmin60 cultivation parcel surface.
2. **c_g72diff** = The difference between compass cultivation parcel surface and Garmin72 cultivation parcel surface.
3. **c_m400diff** = The difference between compass cultivation parcel surface and Magellan400 cultivation parcel surface.

# Studying the statistical relevance of measuring cultivation parcels with GPS: the main matter

The cultivation parcel selection **was not random**. Indeed, samples have not been generated by a random mechanism (intentional selection exists). It implies that:

- Common statistical tests, such as *t*-**tests**, **cannot work well** since they assume randomness.
- The **lose of accurancy** may become greater when samples are smaller.
- The **statistical inference** cannot have nothing to do with the sampling (we must take care of the nonrandomness of the samples).

# Studying the statistical relevance of measuring cultivation parcels with GPS: the main matter

The cultivation parcel selection **was not random**. Indeed, samples have not been generated by a random mechanism (intentional selection exists). It implies that:

- Common statistical tests, such as $t$-**tests**, **cannot work well** since they assume randomness.
- The **lose of accurancy** may become greater when samples are smaller.
- The **statistical inference** cannot have nothing to do with the sampling (we must take care of the nonrandomness of the samples).

# Studying the statistical relevance of measuring cultivation parcels with GPS: the main matter

The cultivation parcel selection **was not random**. Indeed, samples have not been generated by a random mechanism (intentional selection exists). It implies that:

- Common statistical tests, such as *t*-**tests**, **cannot work well** since they assume randomness.
- The **lose of accurancy** may become greater when samples are smaller.
- The **statistical inference** cannot have nothing to do with the sampling (we must take care of the nonrandomness of the samples).

# Studying the statistical relevance of measuring cultivation parcels with GPS: the main matter

The cultivation parcel selection **was not random**. Indeed, samples have not been generated by a random mechanism (intentional selection exists). It implies that:

- Common statistical tests, such as *t*-**tests**, **cannot work well** since they assume randomness.
- The **lose of accurancy** may become greater when samples are smaller.
- The **statistical inference** cannot have nothing to do with the sampling (we must take care of the nonrandomness of the samples).

# Statistical Inference: the theory

Basically, there are two ways to do statistical inference:

- Unconditional (or simple) Inference
  - *randomness* of parcel selection **is assumed**.
  - we may assume *normality* (parametric approach) or not (nonparametric approach) of the differences.
  - we may assume *paired samples* (pairing is effective when two measurements are strongly correlated) *unpared samples* (measurements are uncorrelated - they cannot influence each other).
- Conditional Inference
  - *randomness* of parcel selection **is not assumed**.
  - parametric statistical tables (such as $t$ or $F$ tables) are not valid, because they are based on *theoretical distributions* which assume randomness. *Empirical distributions* of the test statistic should be used instead.
  - the *resampling* is the tool to depart from theoretical

# Statistical Inference: the theory

Basically, there are two ways to do statistical inference:

- Unconditional (or simple) Inference
  - *randomness* of parcel selection **is assumed**.
  - we may assume *normality* (parametric approach) or not (nonparametric approach) of the differences.
  - we may assume *paired samples* (pairing is effective when two measurements are strongly correlated) *unpared samples* (measurements are uncorrelated - they cannot influence each other).
- Conditional Inference
  - *randomness* of parcel selection **is not assumed**.
  - parametric statistical tables (such as $t$ or $F$ tables) are not valid, because they are based on *theoretical distributions* which assume randomness. *Empirical distributions* of the test statistic should be used instead.
  - the *resampling* is the tool to depart from theoretical

# Statistical Inference: the theory

Basically, there are two ways to do statistical inference:

- Unconditional (or simple) Inference
  - *randomness* of parcel selection **is assumed**.
  - we may assume *normality* (parametric approach) or not (nonparametric approach) of the differences.
  - we may assume *paired samples* (pairing is effective when two measurements are strongly correlated) *unpared samples* (measurements are uncorrelated - they cannot influence each other).
- Conditional Inference
  - *randomness* of parcel selection **is not assumed**.
  - parametric statistical tables (such as $t$ or $F$ tables) are not valid, because they are based on *theoretical distributions* which assume randomness. *Empirical distributions* of the test statistic should be used instead.
  - the *resampling* is the tool to depart from theoretical distributions to empirical distributions

# Statistical Inference: the Unconditional Inference

The Unconditional inference approach can be done according to this table:

| Statistical aim | Data-set features | | | |
|---|---|---|---|---|
| | Continous Data | | Non-Continous Data | |
| | Measurements from a Gaussian distribution | Measurements, Ranks or Scores, from a non-Gaussian distribution | Binominal Data (two possible outcomes) | Survival Data |
| Describe one sample | Mean<br>Standard deviation | Median<br>Interquartile range | Proportion test | Kaplan Meier Survival Curve |
| Compare one sample to a hypothetical value | One sample *t*-test | One sample Wilcoxon test | Chi-square test<br>Binominal test | |
| Compare two unpaired samples | Unpaired *t*-test | Unpaired Wilcoxon test (Wilcoxon Rank Sum test) | Fisher's test (chi-square test for large samples) | Log-rank test |
| Compare two paired samples | Paired *t*-test | Paired Wilcoxon test (Wilcoxon Sign Rank test) | McNemar's | Hazards regression |
| Compare three or more unmatched samples | One-way ANOVA | Kruskal-Wallis test | Chi-square test | Cox-regression |
| Compare three or more matched samples | Repeated-measures ANOVA | Friedman test | Cochrane Q | Hazards regression |
| Quantify association between two variables | Pearson correlation | Sperman correlation | Contingency coefficients | |
| Predict value from another measured variable | Linear regression<br>Non-linear regression | Non-parametric regression | Logistic regression | Cox-regression |
| Predict value from several measured or binominal variables | Multiple linear regr<br>Multiple non-linear regression | | Multiple logistic regression | Cox-regression |

- We found strong correlation between couples of measurements, it implies that pairing is effective, then samples must be treated as **paired**.
- Then, to study the statistical relevence of measuring cultivation parcels with GPS, we may use both **parametric** and **nonparametric** tests. Since the data set is continuous, we shall have to deal with **paired *t*-test** and the **Wilcoxon Sign Rank test** respectively.

# Statistical Inference: the Unconditional Inference

The Unconditional inference approach can be done according to this table:

| Statistical aim | Data-set features | | | |
|---|---|---|---|---|
| | Continous Data | | Non-Continous Data | |
| | Measurements from a Gaussian distribution | Measurements, Ranks or Scores, from a non-Gaussian distribution | Binominal Data (two possible outcomes) | Survival Data |
| Describe one sample | Mean<br>Standard deviation | Median<br>Interquartile range | Proportion test | Kaplan Meier Survival Curve |
| Compare one sample to a hypothetical value | One sample $t$-test | One sample Wilcoxon test | Chi-square test<br>Binominal test | |
| Compare two unpaired samples | Unpaired $t$-test | Unpaired Wilcoxon test (Wilcoxon Rank Sum test) | Fisher's test (chi-square test for large samples) | Log-rank test |
| Compare two paired samples | Paired $t$-test | Paired Wilcoxon test (Wilcoxon Sign Rank test) | McNemar's | Hazards regression |
| Compare three or more unmatched samples | One-way ANOVA | Kruskal-Wallis test | Chi-square test | Cox-regression |
| Compare three or more matched samples | Repeated-measures ANOVA | Friedman test | Cochrane Q | Hazards regression |
| Quantify association between two variables | Pearson correlation | Sperman correlation | Contingency coefficients | |
| Predict value from another measured variable | Linear regression<br>Non-linear regression | Non-parametric regression | Logistic regression | Cox-regression |
| Predict value from several measured or binominal variables | Multiple linear regr<br>Multiple non-linear regression | | Multiple logistic regression | Cox-regression |

- We found strong correlation between couples of measurements, it implies that pairing is effective, then samples must be treated as **paired**.
- Then, to study the statistical relevence of measuring cultivation parcels with GPS, we may use both **parametric** and **nonparametric** tests. Since the data set is continuous, we shall have to deal with **paired $t$-test** and the **Wilcoxon Sign Rank test** respectively.

# Statistical Inference: the Unconditional Inference

For **paired samples**, unconditional inference can be done by using:

### The *paired t-test* statistic

It's a parametric test on means:

$$t_{paired} = \frac{\overline{X}_{1;i} - \overline{X}_{2;i}}{SE_{(\overline{X}_{1;i} - \overline{X}_{2;i})}} = \frac{\overline{d}}{SE_{\overline{d}}} = \frac{\overline{d}}{S_{\overline{d}}/\sqrt{n}} \xrightarrow{d} t_{n-1} \quad (1)$$

Where: $\overline{d} = n^{-1} \sum_{i=1}^{n} d_i$ ; $S_{\overline{d}} = \sqrt{\frac{\sum_{i=1}^{n}(d_i - \overline{d})^2}{n-1}}$

### The *Wilcoxon Sign Rank test* statistic

It's the nonparametric analogous test of the *paired t-test*:

$$WSR^{+} = \Sigma_{i=1}^{n} 1\{d_i > 0\} \times R_i \quad (2)$$

Where $n$ is the number of pairs, $1\{\cdot\}$ is an *indicator function* that permits just the summation of positive differences, and $R_i$ is the signed rank of the $i$-th obs.

# Statistical Inference: the Unconditional Inference

For **paired samples**, unconditional inference can be done by using:

## The *paired t-test* statistic

It's a parametric test on means:

$$t_{paired} = \frac{\overline{X}_{1;i} - \overline{X}_{2;i}}{SE_{(\overline{X}_{1;i} - \overline{X}_{2;i})}} = \frac{\overline{d}}{SE_{\overline{d}}} = \frac{\overline{d}}{S_{\overline{d}}/\sqrt{n}} \xrightarrow{d} t_{n-1} \quad (1)$$

Where: $\overline{d} = n^{-1} \sum_{i=1}^{n} d_i$ ; $S_{\overline{d}} = \sqrt{\frac{\sum_{i=1}^{n}(d_i - \overline{d})^2}{n-1}}$

## The *Wilcoxon Sign Rank test* statistic

It's the nonparametric analogous test of the *paired t-test*:

$$WSR^{+} = \Sigma_{i=1}^{n} 1\{d_i > 0\} \times R_i \quad (2)$$

Where $n$ is the number of pairs, $1\{\cdot\}$ is an *indicator function* that permits just the summation of positive differences, and $R_i$ is the signed rank of the $i$-th obs.

# Statistical Inference: the Unconditional Inference

For **paired samples**, unconditional inference can be done by using:

## The *paired t-test* statistic

It's a parametric test on means:

$$t_{paired} = \frac{\overline{X}_{1;i} - \overline{X}_{2;i}}{SE_{(\overline{X}_{1;i} - \overline{X}_{2;i})}} = \frac{\overline{d}}{SE_{\overline{d}}} = \frac{\overline{d}}{S_{\overline{d}}/\sqrt{n}} \xrightarrow{d} t_{n-1} \quad (1)$$

Where: $\overline{d} = n^{-1} \sum_{i=1}^{n} d_i \quad ; \quad S_{\overline{d}} = \sqrt{\frac{\sum_{i=1}^{n} (d_i - \overline{d})^2}{n-1}}$

## The *Wilcoxon Sign Rank test* statistic

It's the nonparametric analogous test of the *paired t-test*:

$$WSR^{+} = \Sigma_{i=1}^{n} 1\{d_i > 0\} \times R_i \quad (2)$$

Where $n$ is the number of pairs, $1\{\cdot\}$ is an *indicator function* that permits just the summation of positive differences, and $R_i$ is the signed rank of the $i$-th obs.

# Statistical Inference: the Conditional Inference

When the random sampling assumption **is not satisfied**, statistical inference can be done through *resampling methods*:

- **Randomization tests:** Also known as *permutation tests*. The aim is to make inference by estimating the empirical distribution of the test statistic (called the permutation distribution).
- **Jackknife:** The aim is to estimate the population distribution by deleting one observation at time (standard error and bias estimation).
- **Bootstrap:** The aim is to estimate the distribution of a population by resampling *with replacement* (standard error and bias estimation).

Because we are interested in inference and not in estimators' reliability, we shall have to dealing with permutations tests.

# Statistical Inference: the Conditional Inference

When the random sampling assumption **is not satisfied**, statistical inference can be done through *resampling methods*:

- **Randomization tests:** Also known as *permutation tests*. The aim is to make inference by estimating the empirical distribution of the test statistic (called the permutation distribution).
- **Jackknife:** The aim is to estimate the population distribution by deleting one observation at time (standard error and bias estimation).
- **Bootstrap:** The aim is to estimate the distribution of a population by resampling *with replacement* (standard error and bias estimation).

Because we are interested in inference and not in estimators' reliability, we shall have to dealing with permutations tests.

# Statistical Inference: the Conditional Inference

When the random sampling assumption **is not satisfied**, statistical inference can be done through *resampling methods*:

- **Randomization tests:** Also known as *permutation tests*. The aim is to make inference by estimating the empirical distribution of the test statistic (called the permutation distribution).
- **Jackknife:** The aim is to estimate the population distribution by deleting one observation at time (standard error and bias estimation).
- **Bootstrap:** The aim is to estimate the distribution of a population by resampling *with replacement* (standard error and bias estimation).

Because we are interested in inference and not in estimators' reliability, we shall have to dealing with permutations tests.

# Statistical Inference: the Conditional Inference

When the random sampling assumption **is not satisfied**, statistical inference can be done through *resampling methods*:

- **Randomization tests:** Also known as *permutation tests*. The aim is to make inference by estimating the empirical distribution of the test statistic (called the permutation distribution).
- **Jackknife:** The aim is to estimate the population distribution by deleting one observation at time (standard error and bias estimation).
- **Bootstrap:** The aim is to estimate the distribution of a population by resampling *with replacement* (standard error and bias estimation).

Because we are interested in inference and not in estimators' reliability, we shall have to dealing with permutations tests.

# The preliminary work

Working on differences, as preliminary work we have checked the satisfaction of all assumptions for applying correctly a *t*-test. We found:

| *t*-test assumptions | Features found |
|---|---|
| 1) Normality (measurements must be normal or near normal distributed) | 1a) c_g60diff can be considered near normal.<br>1b) c_g72diff can be considered near normal.<br>1c) c_m400diff cannot be considered normal. |
| 2) Independence (measurements cannot be influence each other) | 2a) It is satisfied, because measurements cannot influence each other. |
| 3) Random Sampling (measurements must be random variables) (random selection avoids conscious or unconscious bias) | We may assume that the selection was:<br>3a) Random ⟶ *t*-tests work well<br>3b) Not random ⟶ *t*-tests do not work well<br>⟶ use permutation tests |
| 4) Variance Homogeneity (measurements must have the same variance) | We must assume that samples are paired:<br>⟶ tests on variances have not any sense<br>⟶ the standard deviation of paired differences $S_{\overline{d}}$ is assumed as the only measure of variability |
| 5) No Outliers (all possible outliers must be deleted before the applied analysis) | 5a) Some evidence of outliers remains. |

So, the features found suggest the **conditional inference** approach.

# The preliminary work

Working on differences, as preliminary work we have checked the satisfaction of all assumptions for applying correctly a *t*-test. We found:

| ***t*-test assumptions** | **Features found** |
|---|---|
| 1) Normality (measurements must be normal or near normal distributed) | 1a) c_g60diff can be considered near normal.<br>1b) c_g72diff can be considered near normal.<br>1c) c_m400diff cannot be considered normal. |
| 2) Independence (measurements cannot be influence each other) | 2a) It is satisfied, because measurements cannot influence each other. |
| 3) Random Sampling (measurements must be random variables) (random selection avoids conscious or unconscious bias) | We may assume that the selection was:<br>3a) Random $\longrightarrow$ *t*-tests work well<br>3b) Not random $\longrightarrow$ *t*-tests do not work well<br>$\longrightarrow$ use permutation tests |
| 4) Variance Homogeneity (measurements must have the same variance) | We must assume that samples are paired:<br>$\longrightarrow$ tests on variances have not any sense<br>$\longrightarrow$ the standard deviation of paired differences $S_{\overline{d}}$ is assumed as the only measure of variability |
| 5) No Outliers (all possible outliers must be deleted before the applied analysis) | 5a) Some evidence of outliers remains. |

So, the features found suggest the **conditional inference** approach.

# The preliminary work

Working on differences, as preliminary work we have checked the satisfaction of all assumptions for applying correctly a *t*-test. We found:

| *t*-test assumptions | Features found |
|---|---|
| 1) Normality (measurements must be normal or near normal distributed) | 1a) c_g60diff can be considered near normal.<br>1b) c_g72diff can be considered near normal.<br>1c) c_m400diff cannot be considered normal. |
| 2) Independence (measurements cannot be influence each other) | 2a) It is satisfied, because measurements cannot influence each other. |
| 3) Random Sampling (measurements must be random variables) (random selection avoids conscious or unconscious bias) | We may assume that the selection was:<br>3a) Random $\longrightarrow$ *t*-tests work well<br>3b) Not random $\longrightarrow$ *t*-tests do not work well<br>$\longrightarrow$ use permutation tests |
| 4) Variance Homogeneity (measurements must have the same variance) | We must assume that samples are paired:<br>$\longrightarrow$ tests on variances have not any sense<br>$\longrightarrow$ the standard deviation of paired differences $S_{\overline{d}}$ is assumed as the only measure of variability |
| 5) No Outliers (all possible outliers must be deleted before the applied analysis) | 5a) Some evidence of outliers remains. |

So, the features found suggest the **conditional inference** approach.

# The results: the paired $t$-test

**Paired $t$-test**

**Panel: A**
data: S.1 and S.21.1
$t_{paired} = 0.7262$; $df = 87$; p-value = **0.4697**
alter. hypothesis: true difference in means is not equal to 0
95 percent confidence interval: [-30.61026 ; 65.85185]
mean of the differences: 17.62080

**Panel: B**
data: S.1 and S.22.1
$t_{paired} = 3.9084$; $df = 125$; p-value = **0.0001514**
alter. hypothesis: true difference in means is not equal to 0
95 percent confidence interval: [50.44549 ; 153.94531]
mean of the differences: 102.1954

**Panel: C**
data: S.1 and S.24.1
$t_{paired} = 5.8224$; $df = 125$; p-value = **4.601e-08**
alter. hypothesis: true difference in means is not equal to 0
95 percent confidence interval: [95.06151 ; 192.96658]
mean of the differences: 144.0140

- In Panel A, traditional method and Garmin60 method are compared. Since the p-value is greater than the significance level ($0.4697 > 0.05$), **we cannot reject** the null hypothesis that the true difference in **means** is equal to zero, then we can conclude stating that parcel estimates using Garmin60 **are not statistically different** from parcel estimates using traditional method.
- In Panel B, traditional method and Garmin72 method are faced toward. Since the p-value is now smaller than the significance level ($0.0001514 < 0.05$), **we can reject** the null hypothesis and we can conclude stating that parcel estimates using Garmin72 **are statistically different** from parcel estimates using traditional method.
- In Panel C, traditional method and Magellan400 method are compared. Since the p-value is smaller than the significance level ($4.601e^{-08} < 0.05$), **we can reject** the null hypothesis and we can conclude stating that parcel estimates using Magellan400 are **statistically different** from parcel estimates using traditional method.
- In panel B and C, the **mean of the differences** is positive, it implies that parcel estimates using traditional method are **larger** than parcel estimates using GPS72 and MAG400.

# The results: the paired $t$-test

| Paired $t$-test |
|---|
| **Panel: A** |
| data: S.1 and S.21.1 |
| $t_{paired} = 0.7262$; $df = 87$; p-value = **0.4697** |
| alter. hypothesis: true difference in means is not equal to 0 |
| 95 percent confidence interval: [-30.61026 ; 65.85185] |
| mean of the differences: 17.62080 |
| **Panel: B** |
| data: S.1 and S.22.1 |
| $t_{paired} = 3.9084$; $df = 125$; p-value = **0.0001514** |
| alter. hypothesis: true difference in means is not equal to 0 |
| 95 percent confidence interval: [50.44549 ; 153.94531] |
| mean of the differences: 102.1954 |
| **Panel: C** |
| data: S.1 and S.24.1 |
| $t_{paired} = 5.8224$; $df = 125$; p-value = **4.601e-08** |
| alter. hypothesis: true difference in means is not equal to 0 |
| 95 percent confidence interval: [95.06151 ; 192.96658] |
| mean of the differences: 144.0140 |

- In Panel A, traditional method and Garmin60 method are compared. Since the p-value is greater than the significance level ($0.4697 > 0.05$), **we cannot reject** the null hypothesis that the true difference in **means** is equal to zero, then we can conclude stating that parcel estimates using Garmin60 **are not statistically different** from parcel estimates using traditional method.
- In Panel B, traditional method and Garmin72 method are faced toward. Since the p-value is now smaller than the significance level ($0.0001514 < 0.05$), **we can reject** the null hypothesis and we can conclude stating that parcel estimates using Garmin72 **are statistically different** from parcel estimates using traditional method.
- In Panel C, traditional method and Magellan400 method are compared. Since the p-value is smaller than the significance level ($4.601e^{-08} < 0.05$), **we can reject** the null hypothesis and we can conclude stating that parcel estimates using Magellan400 are **statistically different** from parcel estimates using traditional method.
- In panel B and C, the **mean of the differences** is positive, it implies that parcel estimates using traditional method are **larger** than parcel estimates using GPS72 and MAG400.

# The results: the paired *t*-test

**Paired *t*-test**

**Panel: A**

data: S.1 and S.21.1
$t_{paired} = 0.7262$; $df = 87$; p-value = **0.4697**
alter. hypothesis: true difference in means is not equal to 0
95 percent confidence interval: [-30.61026 ; 65.85185]
mean of the differences: 17.62080

**Panel: B**

data: S.1 and S.22.1
$t_{paired} = 3.9084$; $df = 125$; p-value = **0.0001514**
alter. hypothesis: true difference in means is not equal to 0
95 percent confidence interval: [50.44549 ; 153.94531]
mean of the differences: 102.1954

**Panel: C**

data: S.1 and S.24.1
$t_{paired} = 5.8224$; $df = 125$; p-value = **4.601e-08**
alter. hypothesis: true difference in means is not equal to 0
95 percent confidence interval: [95.06151 ; 192.96658]
mean of the differences: 144.0140

- In Panel A, traditional method and Garmin60 method are compared. Since the p-value is greater than the significance level ($0.4697 > 0.05$), **we cannot reject** the null hypothesis that the true difference in **means** is equal to zero, then we can conclude stating that parcel estimates using Garmin60 **are not statistically different** from parcel estimates using traditional method.
- In Panel B, traditional method and Garmin72 method are faced toward. Since the p-value is now smaller than the significance level ($0.0001514 < 0.05$), **we can reject** the null hypothesis and we can conclude stating that parcel estimates using Garmin72 **are statistically different** from parcel estimates using traditional method.
- In Panel C, traditional method and Magellan400 method are compared. Since the p-value is smaller than the significance level ($4.601e^{-08} < 0.05$), **we can reject** the null hypothesis and we can conclude stating that parcel estimates using Magellan400 are **statistically different** from parcel estimates using traditional method.
- In panel B and C, the **mean of the differences** is positive, it implies that parcel estimates using traditional method are **larger** than parcel estimates using GPS72 and MAG400.

# The results: the paired $t$-test

**Paired $t$-test**

**Panel: A**
data: S.1 and S.21.1
$t_{paired} = 0.7262$; $df = 87$; p-value = **0.4697**
alter. hypothesis: true difference in means is not equal to 0
95 percent confidence interval: [-30.61026 ; 65.85185]
mean of the differences: 17.62080

**Panel: B**
data: S.1 and S.22.1
$t_{paired} = 3.9084$; $df = 125$; p-value = **0.0001514**
alter. hypothesis: true difference in means is not equal to 0
95 percent confidence interval: [50.44549 ; 153.94531]
mean of the differences: 102.1954

**Panel: C**
data: S.1 and S.24.1
$t_{paired} = 5.8224$; $df = 125$; p-value = **4.601e-08**
alter. hypothesis: true difference in means is not equal to 0
95 percent confidence interval: [95.06151 ; 192.96658]
mean of the differences: 144.0140

- In Panel A, traditional method and Garmin60 method are compared. Since the p-value is greater than the significance level ($0.4697 > 0.05$), **we cannot reject** the null hypothesis that the true difference in **means** is equal to zero, then we can conclude stating that parcel estimates using Garmin60 **are not statistically different** from parcel estimates using traditional method.
- In Panel B, traditional method and Garmin72 method are faced toward. Since the p-value is now smaller than the significance level ($0.0001514 < 0.05$), **we can reject** the null hypothesis and we can conclude stating that parcel estimates using Garmin72 **are statistically different** from parcel estimates using traditional method.
- In Panel C, traditional method and Magellan400 method are compared. Since the p-value is smaller than the significance level ($4.601e^{-08} < 0.05$), **we can reject** the null hypothesis and we can conclude stating that parcel estimates using Magellan400 are **statistically different** from parcel estimates using traditional method.
- In panel B and C, the **mean of the differences** is positive, it implies that parcel estimates using traditional method are **larger** than parcel estimates using GPS72 and MAG400.

# The results: the paired *t*-test

Summarizing, the paired *t*-test states that:

**Paired *t*-test: the results**

1) Statistical equivalency:
* traditional method is *statistically equivalent* to the Garmin60 method.
* traditional method is *statistically different* to the Garmin72 method.
* traditional method is *statistically different* to the Mag400 method.

2) Parcel estimates:
* using traditional method are *larger* than parcel estimates using Garmin72 method.
* using traditional method are *much more larger* than parcel estimates using Mag400 method.

Here, **randomness** and **normality** are supposed, but these assumptions are not satisfied by the data set. We need: first, to relax the assumption of normality mainteining randomness (Wilcoxon Sign Rank test (WSR)), second, to relax both the assumption of normality and randomness (paired *t*-permutation test).

# The results: the paired $t$-test

Summarizing, the paired $t$-test states that:

**Paired $t$-test: the results**

1) Statistical equivalency:
* traditional method is *statistically equivalent* to the Garmin60 method.
* traditional method is *statistically different* to the Garmin72 method.
* traditional method is *statistically different* to the Mag400 method.

2) Parcel estimates:
* using traditional method are *larger* than parcel estimates using Garmin72 method.
* using traditional method are *much more larger* than parcel estimates using Mag400 method.

Here, **randomness** and **normality** are supposed, but these assumptions are not satisfied by the data set. We need: **first**, to relax the assumption of normality mainteining randomness (Wilcoxon Sign Rank test (WSR)), **second**, to relax both the assumption of normality and randomness (paired $t$-permutation test).

# The results: the paired $t$-test

Summarizing, the paired $t$-test states that:

| **Paired $t$-test: the results** |
|---|
| 1) Statistical equivalency:<br>* traditional method is *statistically equivalent* to the Garmin60 method.<br>* traditional method is *statistically different* to the Garmin72 method.<br>* traditional method is *statistically different* to the Mag400 method.<br><br>2) Parcel estimates:<br>* using traditional method are *larger* than parcel estimates using Garmin72 method.<br>* using traditional method are *much more larger* than parcel estimates using Mag400 method. |

Here, **randomness** and **normality** are supposed, but these assumptions are not satisfied by the data set. We need: **first**, to relax the assumption of normality mainteining randomness (Wilcoxon Sign Rank test (WSR)), **second**, to relax both the assumption of normality and randomness (paired $t$-permutation test).

# The results: the paired *t*-test

Summarizing, the paired *t*-test states that:

**Paired *t*-test: the results**

1) Statistical equivalency:
* traditional method is *statistically equivalent* to the Garmin60 method.
* traditional method is *statistically different* to the Garmin72 method.
* traditional method is *statistically different* to the Mag400 method.

2) Parcel estimates:
* using traditional method are *larger* than parcel estimates using Garmin72 method.
* using traditional method are *much more larger* than parcel estimates using Mag400 method.

Here, **randomness** and **normality** are supposed, but these assumptions are not satisfied by the data set. We need: **first**, to relax the assumption of normality mainteining randomness (Wilcoxon Sign Rank test (WSR)), **second**, to relax both the assumption of normality and randomness (paired *t*-permutation test).

# The results: the WSR test

**Wilcoxon Sign Rank test**

**Panel: AW**
data: S.1 and S.21.1
$W^+ = 1886.5$; p-value = **0.796**
alternative hypothesis: true location shift is not equal to 0
pseudo-median of the differences: 5.99996

**Panel: BW**
data: S.1 and S.22.1
$W^+ = 5428$; p-value = **0.0001076**
alternative hypothesis: true location shift is not equal to 0
pseudo-median of the differences: 91

**Panel: CW**
data: S.1 and S.24.1
$W^+ = 6276$; p-value = **3.018e-08**
alternative hypothesis: true location shift is not equal to 0
pseudo-median of the differences: 109.565

- In Panel AW, since the p-value is greater than the significance level ($0.7977 > 0.05$), **we cannot reject** the null hypothesis that the true difference in **medians** is equal to zero, then we can conclude stating that parcel estimates using Garmin60 **are not statistically different** from parcel estimates using traditional method.
- In Panel BW and CW, since the p-value is smaller than the significance level, **we can reject** the null hypothesis concluding that both Garmin72 and Magellan400 parcel estimates should be considered **statistically different** from parcel estimates using the traditional method.
- In Panel BW and CW, the median of the differences is positive, it implies that parcel estimate using traditional method are **slightly larger** and **larger** than GPS72 and MAG400 respectively.

# The results: the WSR test

**Wilcoxon Sign Rank test**

**Panel: AW**
data: S.1 and S.21.1
$W^{+} = 1886.5$; p-value = **0.796**
alternative hypothesis: true location shift is not equal to 0
pseudo-median of the differences: 5.99996

**Panel: BW**
data: S.1 and S.22.1
$W^{+} = 5428$; p-value = **0.0001076**
alternative hypothesis: true location shift is not equal to 0
pseudo-median of the differences: 91

**Panel: CW**
data: S.1 and S.24.1
$W^{+} = 6276$; p-value = **3.018e-08**
alternative hypothesis: true location shift is not equal to 0
pseudo-median of the differences: 109.565

- In Panel AW, since the p-value is greater than the significance level ($0.7977 > 0.05$), **we cannot reject** the null hypothesis that the true difference in **medians** is equal to zero, then we can conclude stating that parcel estimates using Garmin60 **are not statistically different** from parcel estimates using traditional method.
- In Panel BW and CW, since the p-value is smaller than the significance level, **we can reject** the null hypothesis concluding that both Garmin72 and Magellan400 parcel estimates should be considered **statistically different** from parcel estimates using the traditional method.
- In Panel BW and CW, the median of the differences is positive, it implies that parcel estimate using traditional method are **slightly larger** and **larger** than GPS72 and MAG400 respectively.

# The results: the WSR test

**Wilcoxon Sign Rank test**

**Panel: AW**

data: S.1 and S.21.1
$W^+ = 1886.5$; p-value = **0.796**
alternative hypothesis: true location shift is not equal to 0
pseudo-median of the differences: 5.99996

**Panel: BW**

data: S.1 and S.22.1
$W^+ = 5428$; p-value = **0.0001076**
alternative hypothesis: true location shift is not equal to 0
pseudo-median of the differences: 91

**Panel: CW**

data: S.1 and S.24.1
$W^+ = 6276$; p-value = **3.018e-08**
alternative hypothesis: true location shift is not equal to 0
pseudo-median of the differences: 109.565

- In Panel AW, since the p-value is greater than the significance level ($0.7977 > 0.05$), **we cannot reject** the null hypothesis that the true difference in **medians** is equal to zero, then we can conclude stating that parcel estimates using Garmin60 **are not statistically different** from parcel estimates using traditional method.
- In Panel BW and CW, since the p-value is smaller than the significance level, **we can reject** the null hypothesis concluding that both Garmin72 and Magellan400 parcel estimates should be considered **statistically different** from parcel estimates using the traditional method.
- In Panel BW and CW, the median of the differences is positive, it implies that parcel estimate using traditional method are **slightly larger** and **larger** than GPS72 and MAG400 respectively.

# The results: the WSR test

Summarizing, we WSR test states that:

**Wilcoxon Sign Rank test: the results**

1) Statistical equivalency:
* traditional method is *statistically equivalent* to the Garmin60 method.
* traditional method is *statistically different* to the Garmin72 method.
* traditional method is *statistically different* to the Magellan400 method.

2) Parcel estimates:
* using traditional method are *slightly larger* than parcel estimates using Garmin72 method.
* using traditional method are *larger* than parcel estimates using Magellan400 method.

Here, normality is relaxed, but **randomness is mainteined**. The second step is to relax randomness too. We will provide that, in the conditional inference approch.

# The results: the WSR test

Summarizing, we WSR test states that:

| **Wilcoxon Sign Rank test: the results** |
|---|
| 1) Statistical equivalency:<br>* traditional method is *statistically equivalent* to the Garmin60 method.<br>* traditional method is *statistically different* to the Garmin72 method.<br>* traditional method is *statistically different* to the Magellan400 method.<br><br>2) Parcel estimates:<br>* using traditional method are *slightly larger* than parcel estimates using Garmin72 method.<br>* using traditional method are *larger* than parcel estimates using Magellan400 method. |

Here, normality is relaxed, but **randomness is mainteined**. The second step is to relax randomness too. We will provide that, in the conditional inference approch.

# Comparing the results

Bringing the unconditional inference results together:

**Paired $t$-test: the results**

1) Statistical equivalency:
* traditional method is *statistically equivalent* to the Garmin60 method.
* traditional method is *statistically different* to the Garmin72 method.
* traditional method is *statistically different* to the Magellan400 method.

2) Parcel estimates:
* using traditional method are *larger* than parcel estimates using Garmin72 method.
* using traditional method are *much more larger* than parcel estimates using Magellan400 method.

**Wilcoxon Sign Rank test: the results**

1) Statistical equivalency:
* traditional method is *statistically equivalent* to the Garmin60 method.
* traditional method is *statistically different* to the Garmin72 method.
* traditional method is *statistically different* to the Magellan400 method.

2) Parcel estimates:
* using traditional method are *slightly larger* than parcel estimates using Garmin72 method.
* using traditional method are *larger* than parcel estimates using Magellan400 method.

we can state that:

# Comparing the results

Bringing the unconditional inference results together:

**Paired *t*-test: the results**

1) Statistical equivalency:
* traditional method is *statistically equivalent* to the Garmin60 method.
* traditional method is *statistically different* to the Garmin72 method.
* traditional method is *statistically different* to the Magellan400 method.

2) Parcel estimates:
* using traditional method are *larger* than parcel estimates using Garmin72 method.
* using traditional method are *much more larger* than parcel estimates using Magellan400 method.

**Wilcoxon Sign Rank test: the results**

1) Statistical equivalency:
* traditional method is *statistically equivalent* to the Garmin60 method.
* traditional method is *statistically different* to the Garmin72 method.
* traditional method is *statistically different* to the Magellan400 method.

2) Parcel estimates:
* using traditional method are *slightly larger* than parcel estimates using Garmin72 method.
* using traditional method are *larger* than parcel estimates using Magellan400 method.

we can state that:

# The results

... we can state that:

## Unconditional Inference Result

The nonparametric unconditional inference approach based on WSR test supports the parametric conclusions from the paired *t*-test:

- On the statistical equivalency hand, **only** Garmin60 is found **statistically equivalent** to the traditional method.
- On the parcel estimates hand, the traditional method tends to produce **larger parcels** estimates respect to **all** GPSs measurements methods.

# The results

... we can state that:

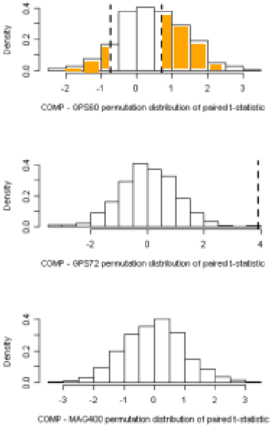
## Unconditional Inference Result

The nonparametric unconditional inference approach based on WSR test supports the parametric conclusions from the paired *t*-test:

- On the statistical equivalency hand, **only** Garmin60 is found **statistically equivalent** to the traditional method.
- On the parcel estimates hand, the traditional method tends to produce **larger parcels** estimates respect to **all** GPSs measurements methods.

# The permutation distribution of paired *t*-statistic

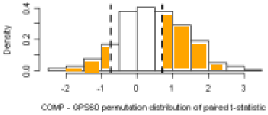
Generating sample replicates *without replacement* from the observed samples, we have estimated the **permutation distribution function** of the paired *t*-statistic:

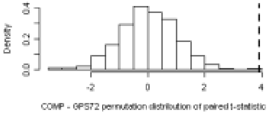


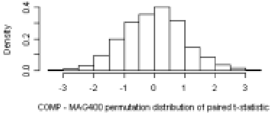

- These histogram estimates are the **empirical distribution functions** of paired *t*-test statistic worked out *conditionally* to the data at hand.
- Under the null hypothesis (that samples have the same distribution), the **empirical p-value** of the test will be the area of the histogram estimate outside the absolute value interval of the observed statistic (this area is emphasised in gold brown).
- A smooth estimate of these histogram densities can be obtained by using an appropriated **Kernel function**.

# The permutation distribution of paired $t$-statistic

Generating sample replicates *without replacement* from the observed samples, we have estimated the **permutation distribution function** of the paired $t$-statistic:



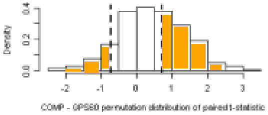



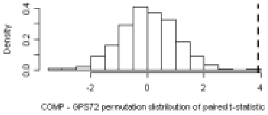



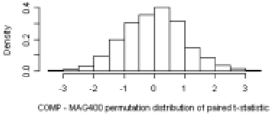

- These histogram estimates are the **empirical distribution functions** of paired $t$-test statistic worked out *conditionally* to the data at hand.
- Under the null hypothesis (that samples have the same distribution), the **empirical p-value** of the test will be the area of the histogram estimate outside the absolute value interval of the observed statistic (this area is emphasised in gold brown).
- A smooth estimate of these histogram densities can be obtained by using an appropriated **Kernel function**.

# The permutation distribution of paired *t*-statistic

Generating sample replicates *without replacement* from the observed samples, we have estimated the **permutation distribution function** of the paired *t*-statistic:




- These histogram estimates are the **empirical distribution functions** of paired *t*-test statistic worked out *conditionally* to the data at hand.
- Under the null hypothesis (that samples have the same distribution), the **empirical p-value** of the test will be the area of the histogram estimate outside the absolute value interval of the observed statistic (this area is emphasised in gold brown).
- A smooth estimate of these histogram densities can be obtained by using an appropriated **Kernel function**.

# The result: the permutation paired $t$-test

**Permutation paired $t$-test**

**Panel: AR**

data: S.1 and S.21.1
Observed test statistic: $\hat{\theta} = 0.7262$; Empirical p-value $=$ **0.472**
alternative hypothesis; $H_A : \; F \neq 0$

**Panel: BR**

data: S.1 and S.22.1
Observed test statistic: $\hat{\theta} = 3.9084$; Empirical p-value $=$ **0.001**
alternative hypothesis; $H_A : \; F \neq 0$

**Panel: CR**

data: S.1 and S.24.1
Observed test statistic $\hat{\theta} = 5.8224$; Empirical p-value $=$ **0.001**
alternative hypothesis; $H_A : \; F \neq 0$

- In Panel AR, the empirical p-value is worked out for the differences between traditional method and Garmin60 method. Since it is greater than the significance level ($0.472 > 0.05$) **we cannot reject** the null hypothesis, then we can conclude that parcel estimate using Garmin60 **are not "conditionally" statistical different** from parcel estimates using traditional method.
- In Panels BR and CR, the empirical p-values are smaller than the significance level ($0.001 < 0.05$), then **we can reject the null hypothesis**, stating that parcel estimate using Garmin72 or Magellan400 **are "conditionally" statistical different** from parcel estimates using traditional method.

# The result: the permutation paired $t$-test

**Permutation paired $t$-test**

**Panel: AR**

data: S.1 and S.21.1
Observed test statistic: $\hat{\theta} = 0.7262$; Empirical p-value = **0.472**
alternative hypothesis; $H_A: \ F \neq 0$

**Panel: BR**

data: S.1 and S.22.1
Observed test statistic: $\hat{\theta} = 3.9084$; Empirical p-value = **0.001**
alternative hypothesis; $H_A: \ F \neq 0$

**Panel: CR**

data: S.1 and S.24.1
Observed test statistic $\hat{\theta} = 5.8224$; Empirical p-value = **0.001**
alternative hypothesis; $H_A: \ F \neq 0$

- In Panel AR, the empirical p-value is worked out for the differences between traditional method and Garmin60 method. Since it is greater than the significance level ($0.472 > 0.05$) **we cannot reject** the null hypothesis, then we can conclude that parcel estimate using Garmin60 **are not "conditionally" statistical different** from parcel estimates using traditional method.
- In Panels BR and CR, the empirical p-values are smaller than the significance level ($0.001 < 0.05$), then **we can reject the null hypothesis**, stating that parcel estimate using Garmin72 or Magellan400 **are "conditionally" statistical different** from parcel estimates using traditional method.

# Comparing the results

By comparing the paired $t$-test with its resampling test analogous:

| **Paired $t$-test** | **Permutation paired $t$-test** |
|---|---|
| **Panel: A**: data S.1 and S.21.1 | |
| $t_{paired} = 0.7262$; $df = 87$; p-value = **0.4697**<br>$H_A: \mu \neq 0$<br>95 percent confidence interval: [-30.61 ; 65.851]<br>mean of the differences: 17.62080 | $\hat{\theta} = 0.7262$; Empirical p-value = **0.472**<br>$H_A: F \neq 0$ |
| **Panel: B**: data S.1 and S.22.1 | |
| $t_{paired} = 3.9084$; $df = 125$; p-value = **0.0001514**<br>$H_A: \mu \neq 0$<br>95 percent confidence interval: [50.445; 153.945]<br>mean of the differences: 102.1954 | $\hat{\theta} = 3.9084$; Empirical p-value: **0.001**<br>$H_A: F \neq 0$ |
| **Panel: C**: data S.1 and S.24.1 | |
| $t_{paired} = 5.8224$; $df = 125$; p-value = **4.601e-08**<br>$H_A: \mu \neq 0$<br>95 percent confidence interval: [95.061;192.966]<br>mean of the differences: 144.0140 | $\hat{\theta} = 5.8224$; Empirical p-value: **0.001**<br>$H_A: F \neq 0$ |

... taking into account these last two tables, we can state that:

... we can state that:

### Conditional Inference Result

- On the statistical equivalency hand, empirical p-values produce the same conclusion of the theoretical ones: **only** cultivation parcel estimates using Garmin60 are found **statistically equivalent** to the traditional method.
- On the parcel estimates hand, **nothing can be state** about the **largeness** of the cultivation parcels using the conditional approach, then the **unconditional inference results remain valid**.

... then, collecting the two main findings:

... we can state that:

**Conditional Inference Result**

- On the statistical equivalency hand, empirical p-values produce the same conclusion of the theoretical ones: **only** cultivation parcel estimates using Garmin60 are found **statistically equivalent** to the traditional method.
- On the parcel estimates hand, **nothing can be state** about the **largeness** of the cultivation parcels using the conditional approach, then the **unconditional inference results remain valid**.

... then, collecting the two main findings:

... collecting the two main findings:

### Unconditional Inference Result

- On the statistical equivalency hand, **only** Garmin60 is found **statistically equivalent** to the traditional method.
- On the parcel estimates hand, the traditional method tends to produce **larger parcels** estimates respect to **all** GPSs measurements methods.

### Conditional Inference Result

- On the statistical equivalency hand, empirical p-values produce the same conclusion of the theoretical ones: **only** cultivation parcel estimates using Garmin60 are found **statistically equivalent** to the traditional method.
- On the parcel estimates hand, **nothing can be state** about the **largeness** of the cultivation parcels using the conditional approach, then the **unconditional inference results remain valid**.

... we can conclude stating that:

... collecting the two main findings:

## Unconditional Inference Result

- On the statistical equivalency hand, **only** Garmin60 is found **statistically equivalent** to the traditional method.
- On the parcel estimates hand, the traditional method tends to produce **larger parcels** estimates respect to **all** GPSs measurements methods.

## Conditional Inference Result

- On the statistical equivalency hand, empirical p-values produce the same conclusion of the theoretical ones: **only** cultivation parcel estimates using Garmin60 are found **statistically equivalent** to the traditional method.
- On the parcel estimates hand, **nothing can be state** about the **largeness** of the cultivation parcels using the conditional approach, then the **unconditional inference results remain valid**.

... we can conclude stating that:

... collecting the two main findings:

### Unconditional Inference Result

- On the statistical equivalency hand, **only** Garmin60 is found **statistically equivalent** to the traditional method.
- On the parcel estimates hand, the traditional method tends to produce **larger parcels** estimates respect to **all** GPSs measurements methods.

### Conditional Inference Result

- On the statistical equivalency hand, empirical p-values produce the same conclusion of the theoretical ones: **only** cultivation parcel estimates using Garmin60 are found **statistically equivalent** to the traditional method.
- On the parcel estimates hand, **nothing can be state** about the **largeness** of the cultivation parcels using the conditional approach, then the **unconditional inference results remain valid**.

... we can conclude stating that:

# … the final conclusion

… because the GPSs methods are globally cheaper than traditional method using compass and meter, it's strongly recommended the use of GPS60 to reduce the costs of the agricultural surveys.

# Future research

For the upcoming papers:

1. We suggest to take care about the random mechanism which has generated the samples. When randomness is not satisfied, we recommend permutation tests.
2. We advise to develop this paper: changing the statistical aim or/and changing the permutation procedure.

## Changing the statistical aim

- We may study the **statistical equivalency** taking into account of **smaller and bigger parcels separately** (find out, if exists, the *equivalency surface threshold*).
- We may study the **statistical equivalency** for the **time requested** to do the measurements (find out, if exists, the *equivalency time threshold*).

## Changing the permutation procedure

- We may consider permutation tests which take care of **any differences** of the two samples, not only in means or in medians (**Kolmogorov-Smirnov test**).
- We may compare more than two samples at time (**multivariate permutation tests (MPT)**). The main type of MPT is based on *nearest neighbors*.

... thank you!!!