Field Crops Research 298 (2023) 108967

Contents lists available at ScienceDirect

Field Crops Research

journal homepage: www.elsevier.com/locate/fcr

Conservation tillage effects on European crop yields: A meta-analysis

Etiendem Achankeng *, Wim Cornelis

Department of Environment, Ghent University, Belgium

ARTICLE INFO

Keywords:
Crop Yield
No-Tillage
Ridge Tillage
Strip Tillage
Conservation Tillage

ABSTRACT

Across Europe, many studies have investigated conservation tillage techniques, which are generally implemented as means of erosion control. Farmers are reluctant in shifting from their traditional tillage/conventional tillage (CT) methods to conservation tillage, their motive being driven by economic reasons stemming from higher crop yields under CT. While, no-till (NT), reduced tillage and minimum tillage have been widely covered by studies across Europe, little is done on ridge (RT) and strip tillage (ST) which might be a compromise between CT and NT. Thus, this study was set out to assess the impact of RT and ST on crop yields within Europe in comparison with NT and CT. The hypothesis tested was that unlike NT, RT and ST significantly increases European crop yields. This was tested by conducting a meta-analysis of 128 studies, based on 624 crop yield observations from 21 European countries. We analysed the influence of crop rotation, crop type, texture, climate, tillage depth, residue retention, and duration of experimentation on the relative yield (response ratio), i.e., yield under NT/RT/ST over yield under CT. In assessing RT and ST, ridge height and strip tilled depth were added. The results show that, on average, NT within Europe resulted in 5.1% reduction in crop yields while RT and ST each led to a 5% increase in crop yield over CT. The major moderator responsible for the increase in yields are ridge heights of at least 20 cm and, loam soil texture on ridges, and strips with tillage depth of 8 cm and a loamy sand texture on strips. Grain maize was the most negatively affected crop type under all conservation tillage techniques with yield reductions of 8% under NT and 18% under RT, except under ST where it showed a 7% increase.

1. Introduction

Conservation Tillage (CS) refers to practices utilizing a reduced number of tillage operations without any soil inversion and leaving at least 30% residues on the soil surface, which increases water infiltration and reduces erosion. It can be roughly divided into no-tillage (NT), mulch tillage (MT), strip tillage (ST), ridge tillage (RT) and reduced/minimum tillage (RMT) (FAO, 1993). Here, organic residues are not inverted into the soil but remain on the surface and protect the soil against erosion and soil moisture losses by evaporation (Baker et al., 2002). Intensive research into various aspects of NT and MT as dominant CS methods has been done in Europe (Soane and Ball, 1998, Rasmussen, 1999; Holland, 2004; Van den Putte et al., 2010; Soane et al., 2012; Kertesz and Madarász, 2014). However, when compared to other regions of the world, their adoption in Europe has been rather low (Holland, 2004; Soane et al., 2012).

Implementation of conservation tillage in Europe, has been majorly in response to protecting soils from erosion and compaction, to conserve soil moisture and reduce production costs (Holland, 2004). According to Panagos et al. (2020), more than 6% of European agricultural lands suffer from severe erosion (11 t ha^{-1} yr^{-1}) and a more incisive set of measures of soil conservation are needed to mitigate soil erosion across the EU. As it is well documented that conservation tillage is an effective measure at reducing soil erosion and improving water infiltration (Meyer et al., 1999; Strauss et al., 2003), the Common Agricultural Policy (CAP) of Europe has sought to ensure that farmers utilize these sustainable techniques so as to properly manage soils and hence maintain ecosystem services (Holland, 2004). A Europe-wide study showed that no and minimum tillage resulted overall in better soil quality, particularly on less fertile soils (Alaoui et al., 2020). However, farmers are economically driven and have thus shown less interest in adopting the new strategies proposed (Van den Putte et al., 2010) especially when techniques like NT are leading to lesser yields. In order to shun the use of traditional systems, other methods like ST and RT are being proposed but their adoption is also influenced by their impact on crop yield.

In brief, NT is a technique without intensive cultivation and involving the direct sowing of crops on the soil, hence minimal soil disturbance (Soane et al., 2012). Reduced tillage or non-inversion tillage

* Correspondence to: Coupure Links 653, Gent 9000, Belgium.
E-mail address: etiendem.achankeng@gmail.com (E. Achankeng).

https://doi.org/10.1016/j.fcr.2023.108967
Received 18 April 2022; Received in revised form 24 April 2023; Accepted 3 May 2023
Available online 18 May 2023

on the other hand loosens the soil to various depths but does not turn it like mouldboard ploughing in CT (Langhans et al., 2019). Reduced tillage is part of conservation tillage as it retains crop residues on the soil surface and preserves soil structure. NT typically leaves the soil covered with 30–100% crop residue. By not ploughing the soil, NT tends to result in soil compaction at the top (0–10 cm) soil layer as opposed to lower (30–40 cm) layers (Bogunovic et al., 2018). Compared with CT, the bulk density at top soil layers is often higher for NT and subsequently this leads to lower crop root density (FAO et al., 2016). However, a hard soil layer (plough pan) at the 25–30 cm depth under CT as a result of repeated ploughing with in-furrow tractor driving and smearing by ploughshares is also reported and below this layer, root density is lower compared to NT. NT improves the soil physical resilience through modifying the soil structural characteristics such as the size and arrangement of peds, stability, and size distribution of aggregates, and continuity, orientation, and size distribution of pores and voids (Lal, 1991). Just by minimal soil disturbance, Sapkota et al. (2012) showed that NT was associated with an increase in microbial diversity and activity. Mineral N content is higher with less nitrate leaching under NT thus resulting in greater N efficiency, and practicing NT for a long period may need less N fertilization than CT (Regina and Alakukku, 2010).

Ridge tillage is a method involving preparation of seedbeds that are elevated (Hatfield et al., 1998; Mitchell et al., 2016). These are usually permanent beds and are seeded in subsequent growing seasons with minimal soil disturbance between the times of harvest to planting (Mitchell et al., 2016). So, these beds are established either before the growing season or during the growing season. Between the beds are furrows that can be ploughed into if necessary. The nature of these permanent beds brings about changes in the physical characteristics of the soil like increased penetration resistance and bulk density during the early growth stages (Katsvairo et al., 2002; Shi et al., 2012) compared to CT. However, compared to NT, the study of Shi et al. (2012) showed that lesser penetration resistance and bulk density existed for RT. According to Bargar et al. (1999), infiltration on ridges is less. However, because infiltration is higher in adjacent furrows, there is usually lateral redistribution of water and solute transport into the ridges at < 45 cm depth. Also, the higher bulk densities of ridges and resulting lower hydraulic conductivity leads to less loss of nutrients like nitrates and phosphates through leaching (Müller et al., 2009; Bargar et al., 1999). Krause et al. (2009) reported that the enlarged surface of ridged soil increases net radiation absorption thus increasing soil temperature and enhancing germination. This was confirmed by Taivalmaa and Talvitie (1997) who found that low soil temperatures and wet soil in spring often delayed the germination and early growth of vegetables. Ridge tillage has been shown to be conducive to microbial reproduction and growth in surface soil, improve soil activity, enhance soil water-holding capacity and water use efficiency, and increase soil aggregate content (Araya et al., 2021), in addition to the mentioned increase in soil temperature, resulting in an effectively coordinate relation among soil, water, fertilizer, gas, heat, light, and temperature (Hatfield et al., 1998).

Strip tillage is a system where row crops are grown in narrow tilled strips with the area between the rows left undisturbed (Overstreet, 2009). Under ST, the soil is divided into sowing zones and soil management zones. The sowing zones are usually 5–15 cm wide and are worked mechanically to depths of 25 cm in order to optimize the soil and microclimate conditions for crop germination and growth. On the other hand, the soil management zones are left untilled (Lal, 1993) and can be combined with the use of mulching (Mitchell et al., 2016). It has been shown that ST accounts for higher penetration resistance and consequently lower root depth on non-tilled strips as opposed to tilled strips (Laufer and Koch, 2017; Ren et al., 2019). The reason behind these observations were the higher penetration resistance for plant roots and

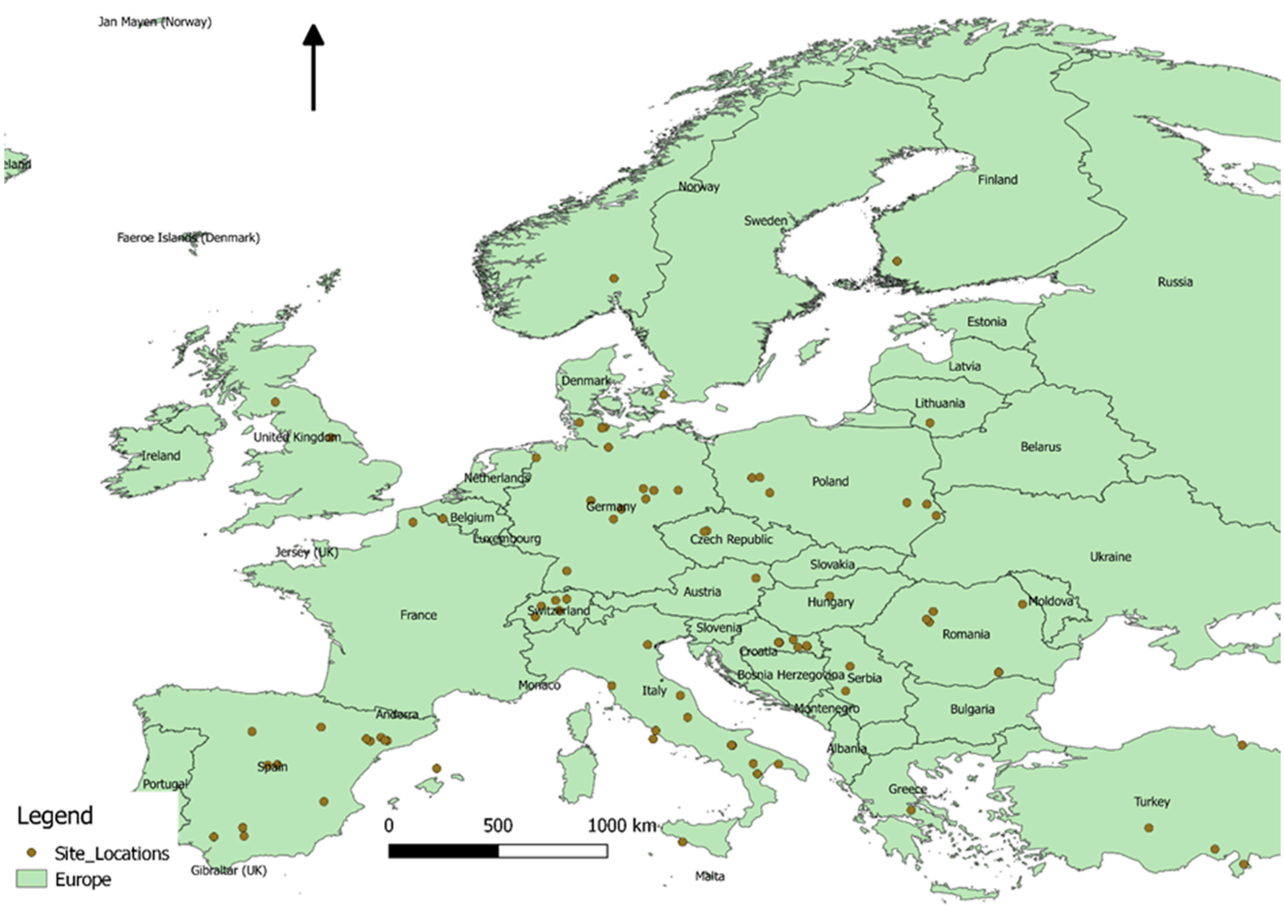

Fig. 1. Locations of selected study sites from which data for the meta-analysis was derived.

Table 1
Other major moderators considered with number of observations.

Texture		Observations	Rotation		Observations
	Clay	20		Yes	96
	Clay Loam	25		No	32
	Clay_Slt_loam	3	**Residue**		
	loam	12		L	89
	Loamy_Sand	6		R	39
	Sand	2	**Climate**		
	Sandy_Clay	3		Dry	39
	Sandy_Clay_Loam	2		Humid	89
	Sandy_ loam	17	**Duration**		
	Silt	2		st	86
	Silt_Clay_loam	11		lt	42
	Silt_Loam	23			
	Silty_Clay	2			

With rotation (Yes), without rotation (No), removed (R), left (L), st is short term duration (< 5years) and, lt is long term (>5years) duration.

the reduced availability of N in the non-tilled strips (Laufer and Koch, 2017). Also, higher bulk densities were observed on the non-tilled strips while those on tilled strips were lower and comparable to those under CT (Jaskulska and Jaskulski, 2020). In general, Jabro et al. (2011) found that soil is less compacted with ST as compared to CT. The tilled zone enhances evaporation of water from soil and warming of the seedbed while total soil disturbance is minimized (Licht and Al-Kaisi, 2005). Thus, strip-till has the potential to increase soil temperature in-row while using inter-row residue cover to conserve soil moisture for plant growth and development. And the crop residue left on the untilled strips account for the increased soil organic matter under ST (Fernández et al., 2015). Also, Debska et al. (2020) observed higher levels of C and N within the top layer of soil under ST and NT compared to CT. Compared to the conventional method with moldboard ploughing, ST requires smaller fuel consumption and shorter work time, thereby reducing cultivation costs and farm input (Overstreet, 2009). Trevini et al. (2013) noted that scarce research had been delivered to the subject of strip-till in Europe. The available scientific literature reports on application of ST focusses primarily on erosion control and not necessarily on crop production/improvements.

Most research on conservation tillage in Europe focused on NT and RMT which have been introduced there decades ago, as was reviewed by Van de Putte et al. (2010) in their meta-analysis. Since their review more than ten years ago, alternatives like ST and RT have emerged in Europe, but no overarching analysis of their effect on crop yield vis-à-vis that of NT, reduced tillage and conventional tillage has been done so far. Likewise, in their global meta-analysis, Pittelkow et al. (2015) considered no-till and/or reduced tillage only. Thus, global evaluations of other conservation tillage methods such as strip-till and ridge-till remain absent, especially in Europe. In addition, the previous studies did not included RT and ST in their analysis of CS. Also, the statistical analysis procedure employed by Van de Putte et al. (2010) was based on t-test but no information was given to show whether their data obeyed the rules for parametric tests whereas this study employs a detailed parametric approach. Thus, it is at this point that this study comes in and seeks to deepen the knowledge concerning the effects of tillage practices in Europe by hypothesizing that ridge-till and strip-till can improve crop yields under European agroecosystems in contrast with the often observed reduction in yields under NT and RMT. To test this hypothesis, this study applies a meta-analysis to answer specifically (i) what impacts do conservation tillage techniques have on crop yields in Europe, (ii) which are major moderators explaining the yield discrepancies if any, among these tillage techniques.

2. Methodology

2.1. Data Compilation

The meta-analysis conducted in this study followed a procedure similar to Nunes et al. (2020). Web of Science was employed for the literature search. In preparing the database, certain studies were selected and others rejected based on four criteria: i) the studies needed to include field experimentation with side by side comparisons of one or more conservation tillage practices of NT, ST and RT relative to CT (as control) with the latter using mouldboard or disc ploughing and based on maximum soil disturbance; ii) the studies must have been performed on experimental sites in Europe (using the United Nations definition of geographical Europe) with locations mentioned; iii) the studies must have compared crop yields as one major component under investigation; iv) the studies must have been documented in English. These criteria allowed for some uniformity across the studies finally included in the analysis. Since a lot of studies addressing tillage practices vary in the tillage methods used, location and subsequent climate, experimental setup, duration of the experiment, agronomic management implemented and other variables, applying these criteria allowed to compare the finally selected studies and to assess to what extent these variables affect their collective results and conclusions. Search terms used were: ridge* OR strip* OR mulch* OR no* OR direct drill* OR conservation OR till* in combination with the different European states (see Fig. 1). The literature search ended on 14/05/2020 and because as many studies as possible were desirable, the search date was from 1955 to 2020.

A total of 107 studies where retained after filtering based on the study inclusion criteria mentioned above including 86 NT, 13 RT and 8 ST studies. However, some studies reported yields of more than one crop type. So, in order to account for this in the analysis, they had to be re-entered as different studies for every crop type thus, leading to 106 NT, 14 RT and 8 ST studies. In effect, the total number of studies used in the analysis thus amounted to 128.

Data gathering was based on extracting mean crop yields under conventional tillage (CT) versus the respective conservation tillage (CS) practice. The following main crops were considered: silage/fodder maize (SM) with 38 observations, grain maize (GM) with 67 observations, potato (P) with 14 observations, sugar beet (SB) with 38 observations, spring cereal (SC) with 55 observations and winter cereals (WC) with 399 observations; these crops constitute some of the main crops grown in Europe (Van de Putte et al., 2010). Grain yield was used as a measure for cereals and grain maize, while dry matter yield was used as a measure for silage maize and fresh tuber yield used as a measure for sugar beet and potato. In order to have a larger number of studies for the ST and RT practices, other common crops were added like cabbage with two, cotton with six, and carrots with three observations, and with their fresh harvested mass being used as measure. All these measures were reported in tons per hectare (t/ha).

The tillage practices were compared by calculating the response ratio (RR) of the annual mean yields of no-till (NT), strip-till (ST), and reduced-till (RT) to the mean yield of conventional tillage (CT). Each RR corresponds to a single observation.

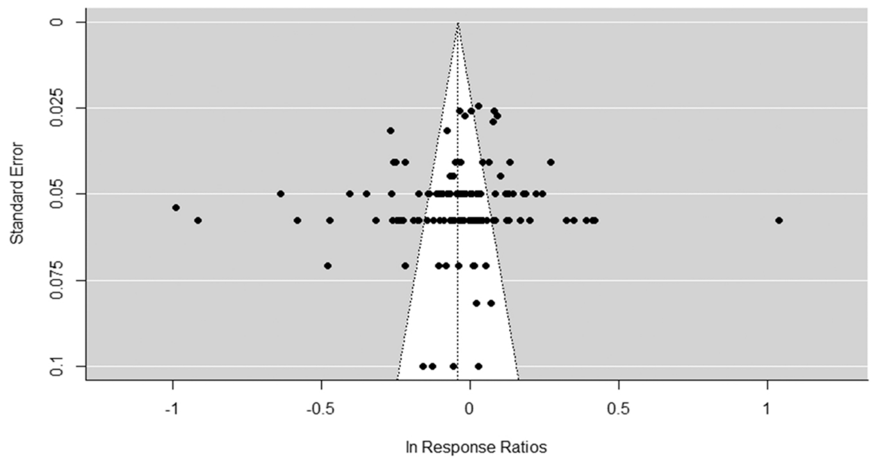

Fig. 2. Funnel plots *In* Response Ratio against standard error. Each data point represents the effect size and the standard error of one experimental condition. The white area represents a pseudo confidence-interval region around the effect-size estimate with bounds equal to ± 1.96standard error.

$$\mathrm{RR} = \frac{\boldsymbol{Yield}_{NT/ST/RT}}{\boldsymbol{Yield}_{CT}} \tag{1}$$

These observations were weighted by the number of plots and replicates. If a long term study had only reported one mean yield for the entire study period, then, observations were considered based on how long the yields had been recorded. The RR was naturally log-transformed ($In\ RR = In\ Yield_{NT/RT/ST} - InYield_{CT}$) to account for the effect size on which the meta-analysis was based.

Other moderators/variables (Table 1) considered were crop rotation, crop type, soil texture total precipitation and average temperature during the growing seasons, year of the study, whether or not the study was long or short term, and tillage depth (for CT). According to FAO (1993), "conservation tillage procedures must be related to the particular site. Their successful application and use over a wide range of soil conditions depends on matching the procedure to soil type, crop cultivar, climatic factors, and other aspects of the environment. Appropriate recommendations to farmers should be based on scientific data from well-designed and adequately equipped long-term experiments" and also, "long-term studies (>10 yrs) are required because following transition from a conventional system to a NT system, takes some years for the soil to reach a new equilibrium" (e.g. Alvarez and Steinbach, 2009).

Here, studies carried out for a period greater than or equal to five years were considered as long term and less than five years as short term, because Madarasz et al. (2016) showed that the soil already starts showing changes about 6 years into the adoption of conservation tillage. When extracting data from studies where the main objective was assessing fertilizer inputs on tillage treatments, the mean yields considered were those in which plots received 0 kg/ha of fertilizer. Reason being that these plots will provide crop yield differentials closer to being influenced by the tillage type implemented. In cases where all treatments received some level of fertilizer input, the averaged mean yield for all treatments was considered provided the controls received an equal amount of fertilizer. Climate was assessed with the aridity index. Rainfall was extracted from the studies as reported and in situations where this data was not reported, latitude and longitude coordinates of the study location were used together with the New_LocClim database to assess the climatic data based on the nearest meteorological station. Thus, rainfall and potential evapotranspiration were used from experiments when available or obtained from the New_LocClim (precipitation) or United Nations Environment Programme (potential evapotranspiration) databases. Here, we followed the procedure according to Pittelkow et al. (2015) and classified all the areas as either dry or humid using the UNEP's upper threshold value of 0.65. So, areas with values less than 0.65 were classified as dry and those equal or greater than 0.65 were considered humid.

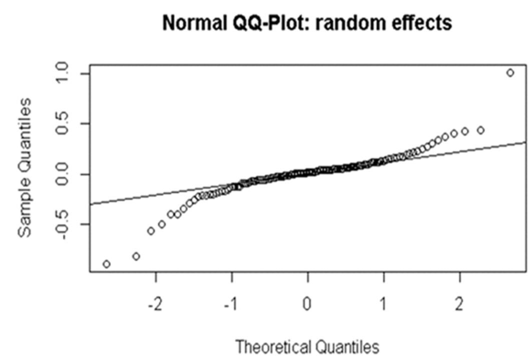

Fig. 3. QQ-plot test of normality assumption.

All the criteria for conservation agriculture were considered for data extraction. So, residue retention and crop rotation were entered as categorical variables. Studies reporting residue retention received the letter L (where L denotes leaving) and those reporting no residue retention received the letter R (R for removing). Studies with crop rotation practiced were labelled YES and those without crop rotation NO. Long term studies were entered as lt and short term as st. Tillage depth was considered for the conventional tillage method so as to check its effects on the variability in yield. Studies usually reported tillage depth in two ways, some reporting a defined depth while others reporting an interval, e.g., 30–35 cm. So, in case of tillage depth reported as intervals, the maximum depth was chosen meaning that, in case 30–35 cm was reported, 35 cm was chosen. However, some studies did not report the specific tillage depth, and in this case, no value was entered as an assumption of a tillage depth becomes erroneous.

2.2. Data Analysis and Model Building

Details about the data analysis and model building are given in the Supplementary Materials S1. In brief, to test the quality of the selected studies, diagnostics run on the selected data involved examination for publication bias evaluation, potential outliers and normality testing. Similar as in Viechtbauer (2010), checking for bias in the study selection was done by a funnel plot (Light and Pillemer, 1984; Sterne and Egger, 2001) whereby the observed *ln RR* on the horizontal axis were plotted against their corresponding standard errors (i.e., the square root of the sampling variances) on the vertical axis (Viechtbauer, 2010). Testing for funnel plot asymmetry, revealed that asymmetry was not significant

($p = 0.65$). A visual inspection already showed three outlying studies with two points to the left (- 1, 0.07) and one point to the extreme right (1, 0.07) on the random effect funnel plot (see Fig. 1). (Fig. 2 and Fig. 3).

A random-effect model was fitted and the residuals of the best linear unbiased prediction (BLUP) were used to check for normality by a normal probability/quantile-quantile plot (Q-Q).

To compute the effect size, the sample sizes and the standard deviations of the studies were supplied as arguments accounting for the measure (log response ratio) and its variance. The sample size (n) for each study was calculated (n = *Number of Plots * Replicate*) based on the number of plots per replicate (Dekemati et al., 2019). However, more than 70% of the studies did not report any measured form of variability. Thus, imputation as suggested by Luo et al. (2006) and used by others (Meurer et al., 2018; Nunes et al., 2020) was employed whereby, 1/10 of each mean yield was taken as the standard deviation. Analyses were made with and without moderators using the rma functions within the Metafor package of the R software (3.5.1). First, the data were plotted without moderators (a random model) and tested for heterogeneity. When significant heterogeneity existed ($p < 0.05$) various moderators (i.e., crop type, rotation, residue, texture, climate and study duration) were added before testing again for residual heterogeneity. All models used the restricted maximum likelihood (REML) which has been shown to be appropriate for comparing models (Zuur et al., 2009).

2.2.1. Meta Analytic Models Fitted

This study employed random and mixed effect models for the analysis. The fixed effect model is the third type often used in meta-analysis when the goal is to make a conditional inference only about the k studies included in the meta-analysis (Hedges and Vevea, 1998). Veichtbauer (2010) noted that meta-analyses are based on sets of studies that are not exactly identical in their methods and/or the characteristics of the included samples. He reiterated that most of these types of analysis starts off with $i = 1,., k$ independent effect size estimates, with each estimating a corresponding (true) effect size. It is assumed that.

$$y_i = \theta_i + e_i \tag{2}$$

where, y_i denotes the observed effect in the *i*-th study, θ_i the corresponding (unknown) true effect, e_i is the sampling error, and $e_i \sim N(0, v_i)$. Thus, the y_i's are assumed to be unbiased and normally distributed estimates of their corresponding true effects with the sampling variances (v_i values) are assumed to be known. Based on the outcome measure used, some form of normalizing, and/or variance stabilizing transformation is usually deemed necessary to ensure that these assumptions are (approximately) true. In the case of this study, a natural logarithm of response ratio was employed.

He further stated that differences in the methods and sample characteristics may introduce variability ("heterogeneity") among the true effects. One way to model the heterogeneity is to treat it as purely random. Thus, the random-effects model used was,

$$\theta_i = \mu + u_i \tag{3}$$

where, $u_i \sim N\ (0,\ \tau^2)$. Therefore, the true effects are assumed to be normally distributed.

with mean μ and τ^2, the estimated amount of total heterogeneity among the true effects. The objective was to estimate μ, the average true effect and τ^2 the amount of total heterogeneity among the true effects as a result of study variability. If $\tau^2 = 0$, it implies homogeneity among the true effects (i.e., $\theta_1 = \ldots = \theta_k \equiv \theta$), so that $\mu = \theta$ then denotes the true effect.

In order to account for the heterogeneity among the true effects, a mixed effect model was fitted. This is basically a random effect model with moderators included (Viechtbauer and Cheung, 2010). It is given as,

$$\theta_i = \beta_0 + \beta_1 x_{i1} + \ldots + \beta_j x_{ij} + u_i \tag{4}$$

where, x_{ij} denotes the value of the *j*-th moderator variable for the *i*-th study and we assume.

again that $ui \sim N(0, \tau^2)$. β_0, β_1, and β_j are regression coefficients. It should be noted that, when talking about a mixed effect model, τ^2 then denotes the amount of estimated residual heterogeneity among the true effects, that is, variability among the true effects that is not accounted for by the moderators included in the model. This model allowed for the examination to what extent the moderators included in the model influenced the size of the average true effect. τ^2 is an estimate of the between-study variance and reflects the extent to which the true effects of the studies in the meta-analysis vary from each other, while I^2 represents the proportion of the total variation in effect sizes that is due to heterogeneity. Both statistics are used to assess heterogeneity in a meta-analysis, but they have different interpretations and meanings. In meta-analysis, heterogeneity is assessed using statistical tests such as the Cochran's *Q*-statistic. The residual heterogeneity or unaccounted variability was obtained using I^2 as follows,

$$I^2 = (100\% * (Q–(k-1))/Q) \tag{5}$$

Where, Q is the Cochran's Q statistic and k is the number of studies included in the meta-analysis. This I^2 describes the percentage of variation across studies that is due to heterogeneity rather than chance (Higgins et al., 2003). So, both $\tau^2$and I^2 are used to assess the degree of heterogeneity in a meta-analysis, they differ in their interpretation and how they are used. τ^2 is used to estimate the amount of heterogeneity, while I^2 is used to quantify the proportion of heterogeneity in the total variation of effect sizes.

Cochran (1954) proposed the *Q*-statistic which is a weighted sum of squared deviations (the difference of each study effect size from the mean effect size of all the k studies considered). It was used to test the homogeneity of the effect sizes. So, the Q statistic is employed to distinguish the variability observed between studies into variability that could be due to random variation, and variability resulting from potential differences between the studies. It is calculated as follows,

$$Q = \sum w_i (y_i - \theta)^2 \tag{6}$$

where, $w_i = 1/\ v_i$ and $\theta = \sum w_i y_i / \sum w_i$ is the inverse-variance weighted estimate of θ under the assumption of homogeneity (all summations go from $I = 1$ to k throughout the paper unless otherwise noted). Also, the Q-statistic follows a chi-square distribution with n-1 degrees of freedom under the null hypothesis that the effect sizes are homogeneous. We equally followed Viechtbauer (2010) by employing Q_E as test for residual heterogeneity. In addition, he used the omnibus test (Q_M) to test for moderators. This omnibus test is a statistical test that tests for the significance of several parameters in a model at once. It test the null hypothesis that moderators are unrelated to the effect sizes. He further states; it can happen that the omnibus test is not significant, yet some of the individual coefficients are.

Estimated residual heterogeneity (τ^2) was determined using the Der Simonian and Laird (1986) estimator ($\widehat{\tau}^2$) which is computed as follows,

$$\widehat{\tau}^2 = \frac{Q-(k-1)}{c} \tag{7}$$

where,

$$c = \sum w_i - w_i^2 \Big/ \sum w_i$$

With the heterogeneity now gotten, the mean effect (μ) was estimated by,

$$\mu = \frac{\widetilde{w}_i y_i}{\sum \widetilde{w}_i} \tag{8}$$

where, $\widetilde{w}_i = 1/(v_i + \widehat{\tau}^2)$ are used as weights. The sampling variance (H^2)

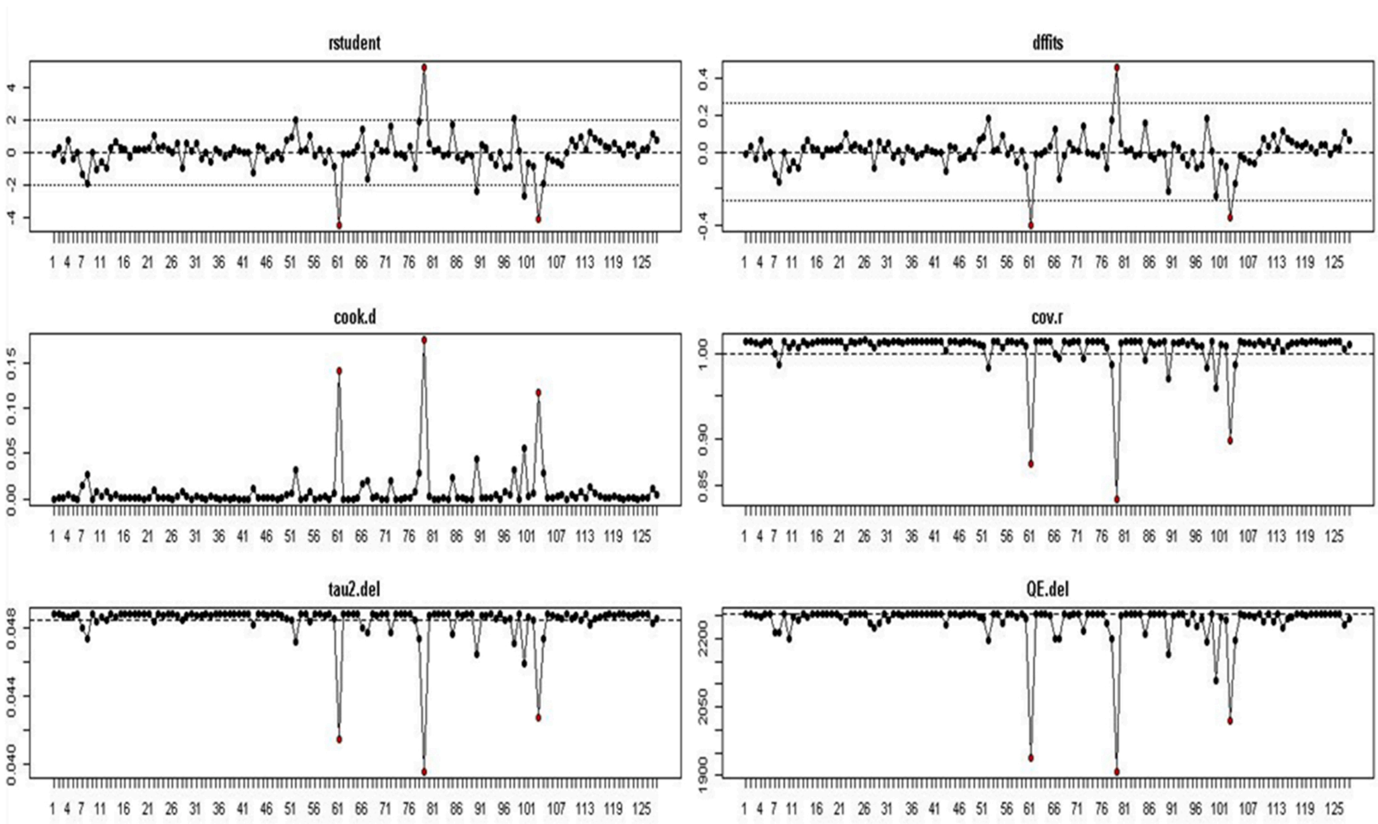

Fig. 4. Plot of the studentized deleted residuals (rstudent), DFFITS values (dffitts), Cook's distances (cook.d), COVRATIO values (cov.r), τ^2 (tau2.del), and test for heterogeneity (QE.del) showing three outlying studies.

was then calculated by,

$Var[\mu] = 1/\sum \widetilde{w}_i$. The 95% confidence intervals for μ were gotten by

$$\mu \pm 1.96\sqrt{Var[\mu]} \tag{9}$$

Two main diagnostics were used to check the influence of the outliers and included the Cook's distance and covariance ratio. The Cook's distance can be interpreted as the Mahalanobis distance between the entire set of predicted values once with the *i*th study included and once with the *i*th study excluded from the model fitting (Viechtbauer and Cheung, 2010). It is given by,

$$D_i = \sum \frac{(\mu_i - \mu_{i(-i)})^2}{v_i + \tau^2} \tag{10}$$

The covariance ratio was computed using the ratio of the generalized variances, which is the influence of the *i*th study examined by means of the change in the variance–covariance matrix of the parameter estimates when excluding the *i*th study from the model fitting (Viechtbauer and Cheung, 2010). This was performed on the results of a random effect model given as,

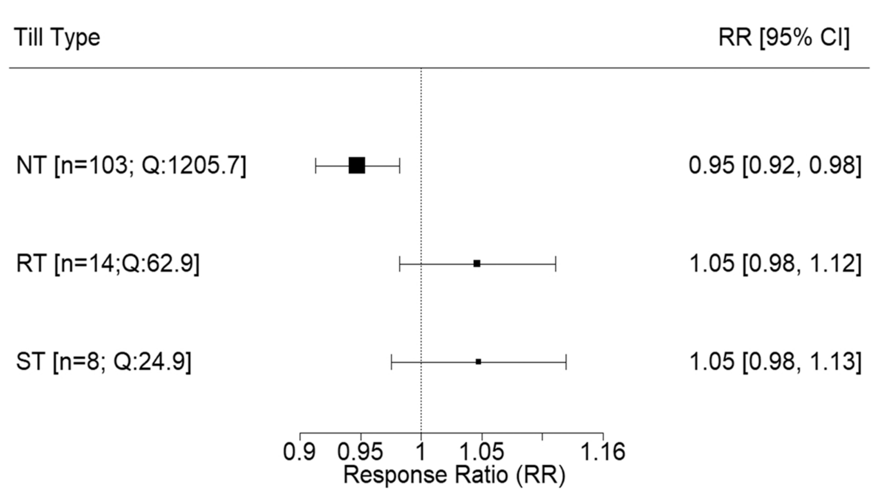

Fig. 5. Forest plot showing estimated average response ratios (RR) for No-tillage (NT), Ridge tillage (RT) and Strip tillage (ST) based on random effects model. n is sample size/number of studies for the specific tillage type. Q is the Cochran's Q-statistic. The error bars represent the 95% confidence intervals (CI). The square shapes represent the estimated RR and their size correspond to their weight (based on its sample size). The vertical broken line represents the line of no effect.

Table 2
Statistical data on no-tillage showing heterogeneity and accountability of the heterogeneity by single moderators.

	I^2 (%)	R^2 (%)	τ^2	H^2	Q_E (df)	p-value	QM (df)	p-value
Residue	94.84	2.30	0.04	19.39	2017.07(104)	< 0.0001	4.27(1)	0.04
Climate	94.95	0.34	0.04	19.78	2057.59(104)	< 0.0001	4.65(1)	0.03
Texture	94.00	16.57	0.04	16.67	1550.69(93)	< 0.0001	41.23(12)	< 0.0001
Rotation	94.98	0.00	0.04	19.93	2072.55(104)	< 0.0001	1.52(1)	0.22
Duration	95.01	0.00	0.04	20.03	2083.01(104)	< 0.0001	0.59(1)	0.44
Crop	95.13	0.00	0.05	20.53	2052.86(100)	< 0.0001	2.71(5)	0.74

I^2 (residual heterogeneity), τ^2 (estimated amount of residual heterogeneity), R^2 (amount of heterogeneity accounted for), Q_E (test of residual heterogeneity), H^2 (sampling variability) and QM (Omnibus test for moderators).

$$COVRATIO_i = \frac{Var[\mu_{(-i)}]}{Var[\mu]} \tag{11}$$

A COVRATIO value below 1(see Fig. 4), therefore, indicates that removal of the *i*th study actually yields more precise estimates of the model coefficients (or equivalently, that addition of the *i*th study actually reduces precision).

Also, large changes in the τ^2 estimate after exclusion of the *i*th study can signal the presence of outliers and/or influential cases (Viechtbauer and Cheung, 2010). Hedges and Olkin (1985) also suggested to examine changes in the Q (or Q_E) statistics when excluding a study in turn. Based on Viechtbauer and Cheung (2010), externally standardized residuals, the DFFITS value, Cook's distance, the covariance ratio, the leave-one-out amount of (residual) heterogeneity, and the leave-one-out test statistic for the test of (residual) heterogeneity were all used to check for outliers and their influence on the meta-analysis. Thus, three outlying studies were taken out before the analysis because the random model fitted best and gave a better estimate of the average effect without the outliers.

While presenting the results on forest plots, an attempt was made to include only the statistically significant (5% error margin) results as all data points were too many to successfully present on forest plots while allowing for readability. For the results not statistically significant, we thought it was still important to report the effect size, as it conveys the message about the magnitude of the moderator effect on RR, while controlling for the specific moderator. However, an appendix section with detailed results on the pairwise comparisons has been included for further clarifications. Likewise, even when a single moderator failed at explaining a significant heterogeneity, we went further at doing the pairwise comparison and explaining the moderator level effects as presented on the forest plots.

3. Results

3.1. Overall Results of Conservation Tillage

Initially, fitting a random model showed that there was variability among the studies. A random effect model for studies on NT resulted in a total of heterogeneity (I^2) of 92.62%. Based on the random model, a significant ($p = 0.02$) estimated RR of 0.949 with lower and upper confidence intervals at 0.916 and 0.982 respectively, was observed (see Fig. 5).

In a similar procedure, the impact of RT was investigated with a random model. This resulted in an estimated average response ratio of 1.05 with lower and upper boundary confidence interval (CI) of 0.98 and 1.12, respectively, a I^2 of 81.02% and, a $p = 0.16$ as seen in Fig. 5. Likewise, ST, analyzed separately with the random-effect model resulted in an estimated average response ratio of 1.05 with lower and upper CI at 0.98 and 1.13 respectively, I^2 of 70.00% and, $p = 0.19$. (Table 2).

3.2. Moderator Effect under No-Till

3.2.1. Effect of Crop Rotation under NT

Subgroup analysis with crop rotation using a mixed model showed that rotation did not account for much of the within study heterogeneity ($R^2 = 0.00$) when comparing NT with CT. This was also seen by the large amount of residual heterogeneity ($I^2 = 94.98\%$), with a very significant ($p < 0.0001$) Q_E (df = 104) = 2072.55. Moderator analysis (Fig. 5.) to estimate the average response ratio due to crop rotation showed that, for its implementation, there was a significant ($p = 0.005$) decrease (0.95) in RR. However, no rotation resulted in a similar decrease in RR (0.94; $p = 0.15$). Despite the low amount of heterogeneity accounted for, several two way combinations showed significant effects on RR as can be seen on Fig. 6. There was an increase in the RR of as much as 48% on combination of no rotation and sandy clay loam texture ($p = 0.02$). The remaining combinations all led to a decrease in RR. Other notable combinations included those of implementing rotation and cultivating

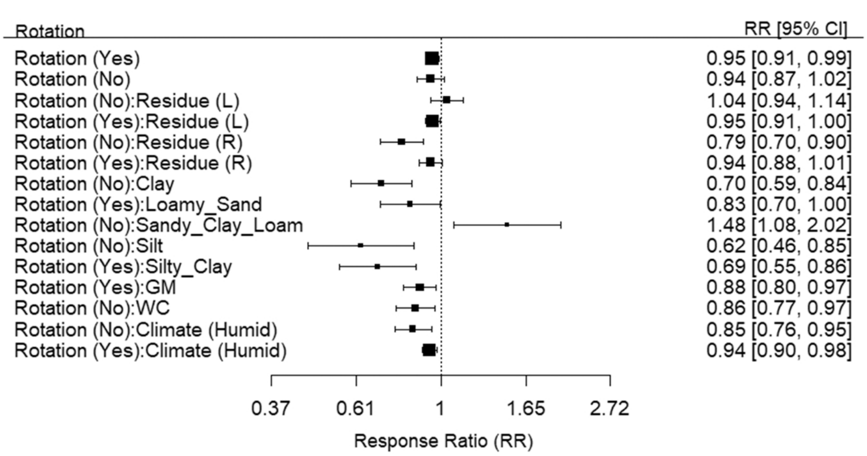

Fig. 6. Forest plot showing yield comparison (estimated average response ratios RR) between NT and conventional tillage as a function of crop rotation and interactions with other moderators: residue, soil texture, crop type, experimental duration, and climate. Symbol meanings: GM is grain maize, WC is winter cereal, R is soil surface without residue, L is soil surface with residue. The error bars represent the 95% confidence intervals CI. The square shapes represent the estimated RR and the size corresponds to its weight (based on its sample size). The vertical broken line represents the line of no effect.

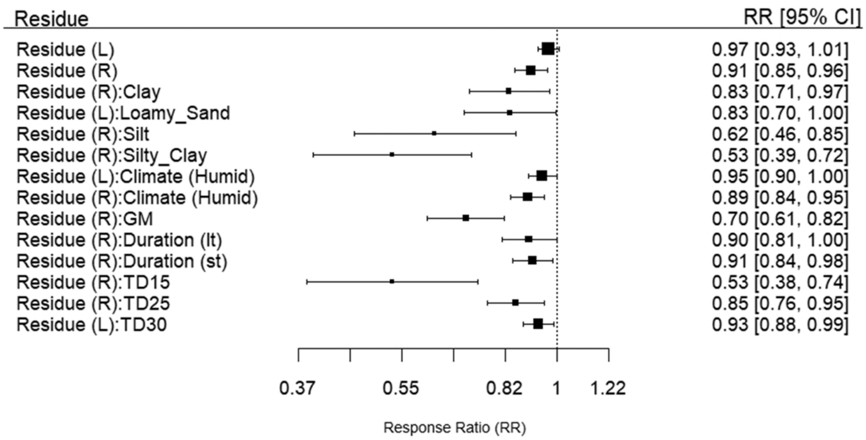

Fig. 7. Forest plot with yield comparisons (estimated average response ratios RR) between no-tillage (NT) and conventional tillage based on mainly residue and other moderators. lt is long term duration (>= 5 years), st is short term duration (< 5 years), GM is grain maize, residue left is L, residue removed R, TD is conventional tillage depth (at 15 cm, 25 cm and, 30 cm). The error bars represent the 95% confidence intervals (CI). The square shapes represent the estimated RR and the size corresponds to its weight (based on its sample size). The vertical broken line represents the line of no effect.

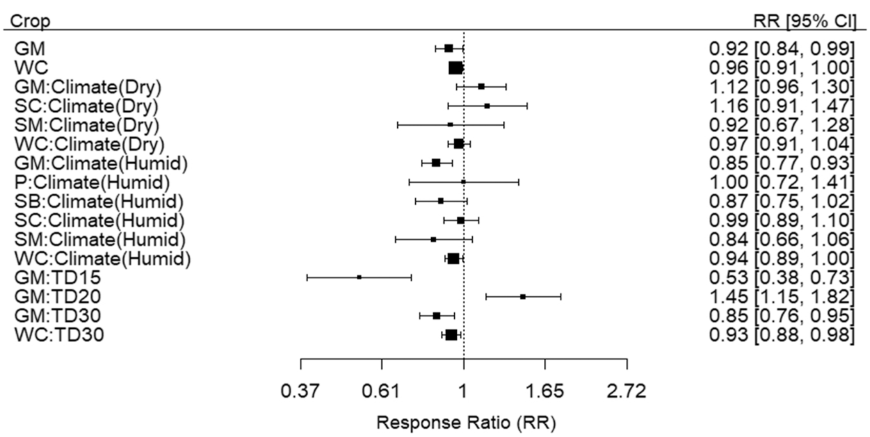

Fig. 8. Forest plot with yield comparisons (estimated average response ratios RR) between no-tillage (NT) and conventional tillage based on residue and other moderators. GM is grain maize, WC is winter cereal, P is Potato, SM is silage maize and SC is spring cereal, Climate(Humid) refers to a humid climatic area, Climate(Dry) refers to a dry climatic area. TD is conventional tillage depth (at 15 cm, 25 cm and, 30 cm). The error bars represent the 95% confidence intervals (CI). The square shapes represent the estimated RR and the size corresponds to its weight (based on its sample size). The vertical broken line represents the line of no effect.

grain maize which led to an estimated RR of 0.88 at $p = 0.01$. Also, when considering residue, it was found that whether residue was retained or not and whether combined with rotation or no rotation, there was mostly a decrease in RR. Taking a closer look, it can be seen that this decrease with residue was rather relative because, a situation with crop rotation and residue retained led to a response ratio 0.95 ($p = 0.03$) compared to 0.79 ($p = 0.0004$) when no rotation and no residue retained were implemented, while no rotation but residue retained resulted in a non-significantly difference in RR (1.04; $p = 0.14$). This trend of lower RR occurred again when considering the climate moderator. Other textures with a lowered RR were silt ($p = 0.003$) and clay (0.001) in combination with no rotation. When considering rotation in combination with tillage depth, the most significant conventional tillage depth was 30 cm with an estimated RR of 0.93 ($p = 0.006$).

3.2.2. *Effect of Residue Retention under NT*

The moderator Residue Retention alone accounted for 2.30% of the heterogeneity. This is a low value as the residual heterogeneity remained high ($I^2 = 94.84\%$) with an estimated amount of residual heterogeneity (τ^2) of 0.03 and significant ($p < 0.0001$), and Q ($df = 104$) = 2017.07. A reduction in RR was observed when residues were retained ($p = 0.11$). Also, removing residues from the surface led to reduction in RR ($p = 0.002$). Two-way interactions (Fig. 7) for residue and other moderators show that significant interactions all decreased RR under NT. Significant combinations were noted for texture, climate, duration and conventional tillage depth. On combination with texture, there was a reduction in RR with residue removed on clay ($p = 0.02$), silt ($p = 0.003$), and silty clay ($p < 0.0001$). The only exception for texture was the loamy sand ($p = 0.05$) showing a decrease in RR with residue retained. Another moderator was that of climate with reductions in RR when it was humid irrespective of whether residue was retained or removed with RR decreasing more when residue was removed ($p = 0.0005$). The only significant result with combination of residue and crop type was for grain maize ($p < 0.0001$) which showed a decrease in RR with residue removed. Irrespective of whether a study was long ($p = 0.05$) or short term ($p = 0.02$) the removal of residue under NT led to a decrease in RR.

3.2.3. *Crop Effects under NT*

As a single moderator in the mixed model, the crop moderator did not account for any heterogeneity in the study ($R^2 = 0.00\%$, $I^2 = 95.13\%$, $\tau^2 = 0.05$ and $Q_E = 2052.86$ with $p < 0.0001$). Crop type in combination with other moderators showed major effects for grain maize and winter cereals (Fig. 8). Main effects showed that grain maize ($p = 0.04$) and winter cereals ($p = 0.05$) had decreases in RR under NT. Crop effect combination with other moderators was also shown to have decreases in RR. When the climate moderator was considered, it was found that grain maize, spring cereals, silage maize and winter cereals in a dry climate have the potential to increase RR under NT though their effects were not significant. In a humid climate, grain maize clearly showed a reduction in RR under NT while potato seemed to show no effect. Other crops like sugar beet and silage maize all led to decreases in RR under NT.

3.2.4. *Textural Effects under NT*

The texture moderator accounted for some of the between study heterogeneity ($R^2 = 16.57\%$). So, the heterogeneity ($I^2 = 91.64\%$) was

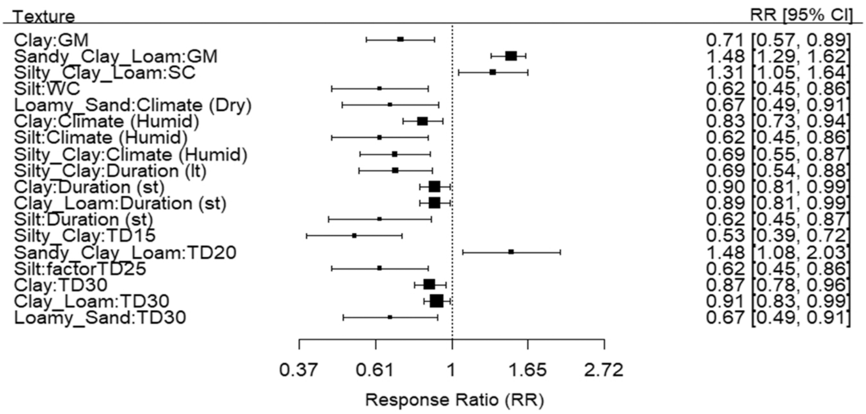

Fig. 9. Forest plot with yield comparisons (estimated average response ratios RR) between no-tillage (NT) and conventional tillage based on soil texture with other moderators in combination. lt is long term duration (>= 5 years)), st is short term duration (<5years), TD is conventional tillage depth (at 15 cm, 25 cm and, 30 cm). GM is grain maize, WC is winter cereal, and SC is spring cereal. The error bars represent the 95% confidence intervals CI. The square shapes represent the estimated RR and the size corresponds to its weight (based on its sample size). The vertical broken line represents the line of no effect.

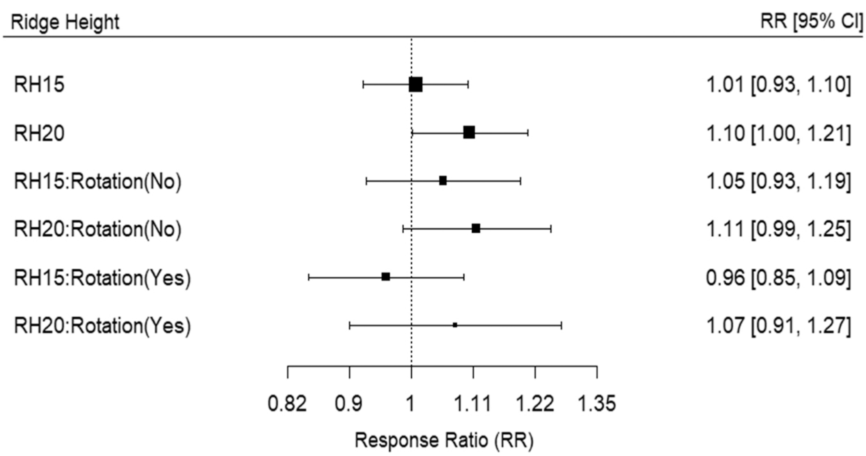

Fig. 10. Forest plot with yield comparisons (estimated average response ratios RR) between RT and conventional tillage based on ridge height with other moderators in combination. RH15 represents ridge heights below 20 cm and RH20 represents ridge heights above 20 cm. The error bars represent the 95% confidence intervals CI. The square shapes represent the estimated RR and the size corresponds to its weight (based on its sample size). The vertical broken line represents the line of no effect.

still high with Q_E (df = 93) of 1550.61 and a p-value < 0.0001. Also, the test of moderators proved significant ($p < 0.0001$) with QM (df =12) = 41.23. The only single soil textures having a significant effect (decrease) on RR are silt ($p = 0.01$) and silty clay ($p = 0.01$), as seen on Fig. 9. For combinations of textures and other moderators, most led to a decrease in RR under NT except for sandy clay loam ($p = 0.02$) with grain maize, silty clay loam ($P = 0.02$) with spring cereals and sandy clay loam with conventional tillage depth at 20 cm ($P = 0.02$). For the other combinations, silt and winter cereals together led to a decrease in RR under NT.

3.3. Effect of ridge tillage and moderator effects

With the random effect model showing a high amount of heterogeneity ($I^2 = 69.42$), it was found that some moderators introduced into the model with the aim of explaining this heterogeneity could not account for it (see Table 1). These moderators were residue and climate all having an $R^2 = 0.00$. A new moderator, ridge height (RH) was seen to account for some of the heterogeneity ($p = 0.001$) within the studies ($R^2 = 0.474$). It can also be seen that there was a low/reduced sampling variability for RH, texture and, crop type when compared to other

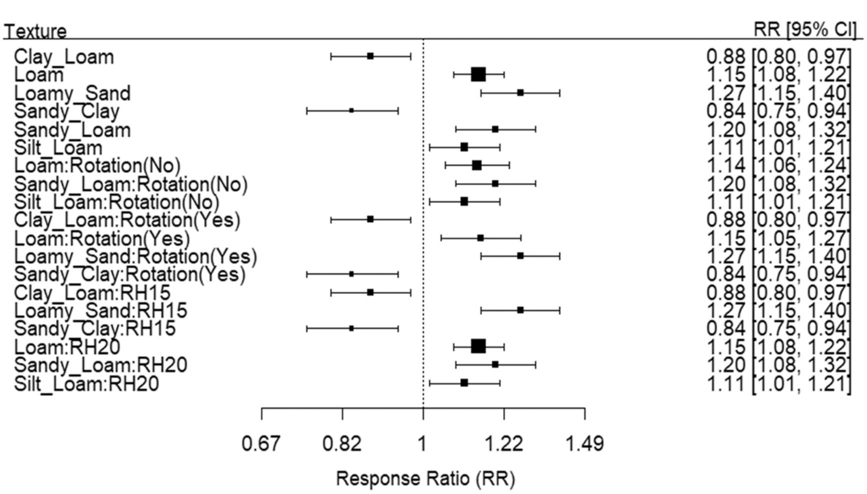

Fig. 11. Forest plot with yield comparisons (estimated average response ratios RR) between RT and conventional tillage based on texture with other moderators in combination. RH15 represents ridge heights below 20 cm and RH20 represents ridge heights above 20 cm. The error bars represent the 95% confidence intervals CI. The square shapes represent the estimated RR and the size corresponds to its weight (based on its sample size). The vertical broken line represents the line of no effect.

Table 3
Statistical data on ridge tillage showing heterogeneity and accountability of the heterogeneity by single moderators.

	I^2 (%)	R^2 (%)	τ^2	H^2	Q_E (df)	p-value	QM (df)	p-value
RH	69.42	47.4	0.008	3.31	25.5(8)	0.001	19.1(14)	0.12
Texture	0.00	100.00	0.00	1.00	3.51(6)	0.74	59.44(8)	< 0.0001
Rotation	80.56	2.58	0.01	5.14	58.6 (12)	< 0.0001	3.56(2)	0.22
Residue	82.48	0.00	0.01	5.71	62.62(12)	< 0.0001	0.0002(1)	0.99
Crop	56.53	64.40	0.04	2.30	11.65(5)	0.04	22.06(8)	0.005
Climate	82.66	0.00	0.01	5.77	62.73(12)	< 0.0001	0.01(1)	0.91

I^2 (residual heterogeneity), τ^2 (estimated amount of residual heterogeneity), R^2 (amount of heterogeneity accounted for), Ridge height (RH), Q_E (test of residual heterogeneity), H^2 (sampling variability) and QM (Omnibus test for moderators).

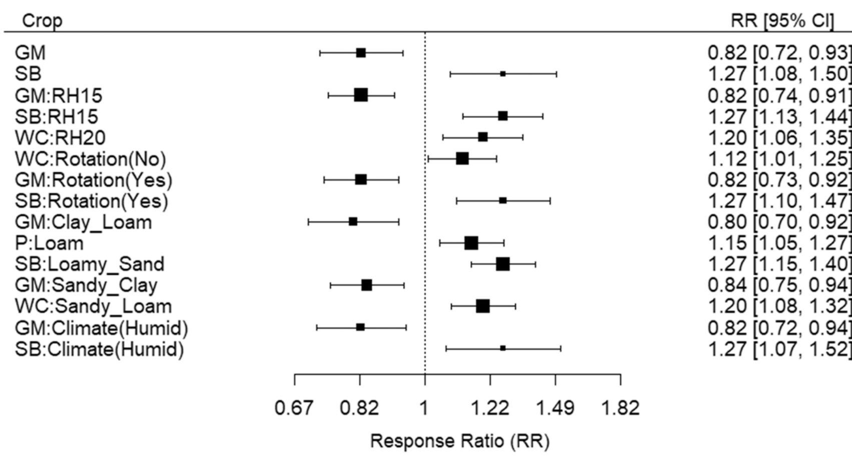

Fig. 12. Forest plot with yield comparisons (estimated average response ratios RR) between RT and conventional tillage based on crop type with other moderators in combination. RH15 represents ridge heights below 20 cm and RH20 represents ridge heights above 20 cm. SB is sugar beet SB, GM is grain maize, P is potato, and WC is winter cereals. The error bars represent the 95% confidence intervals CI. The square shapes represent the estimated RR and the size corresponds to its weight (based on its sample size). The vertical broken line represents the line of no effect.

moderators.

3.3.1. *Effects of ridge height under RT*

It was found that ridge height was a major source of between study variability. Under RT, the estimated average RR (1.05) was above the line of no effect. Main effects showed RR to significantly increase with ridge heights of at least 20 cm ($p = 0.05$) (Fig. 10). There are other interactions with and without crop rotation but all appearing to be non-significant.

3.3.2. *Textural Effects under RT*

Texture of ridge tilled soils was shown to be a major reason for some of the heterogeneity encountered. Two soil textures were found to cause a decrease in RR under RT when crop rotation was implemented. These were clay loam ($p = 0.009$) and sandy clay ($p = 0.002$) as seen in Fig. 11. Loam and a RH of at least 20 cm led to a significant increase in RR ($p < 0.0001$). Though loamy sand appeared to lead to a greater estimate of RR (1.27) at ridge heights less than 20 cm, it is important to look at the small sample size. Loamy texture had a larger sample size and thus, results are more accurate. It is also found that loamy sand with implementation of rotation ($p < 0.0001$) resulted in an increase in RR under RT.

3.3.3. *Crop type effects*

Crop type also accounted for part of the heterogeneity between the studies (see Table 3). Pair-wise comparisons showed that different crops did affect RR differently when ridge height changed as seen in Fig. 12. It was found that while grain maize on RT led to a significant ($p = 0.0002$) decrease in RR under ridge heights of less than 20 cm, sugar beet increased RR when in combination with a RH below 20 cm ($p = 0.0001$). Winter cereals ($p = 0.0003$) and potato ($p = 0.004$) were found to increase RR when in combination with a RH above 20 cm. On crop and rotation, it was seen that winter cereals without rotation led to an increase in RR ($p = 0.03$) as opposed to grain maize ($p = 0.0009$) with crop rotation where it caused a decrease in RR likewise did sugar beet ($p = 0.001$). With the texture moderator, a combination of grain maize on either clay loam ($p = 0.002$) or sandy clay ($p = 0.002$) led to a decrease in RR while sugar beet with loamy sand ($p < 0.0001$) and winter cereals with sandy loam ($p = 0.0003$) led to increases in RR.

3.4. *Effect of Strip Tillage and Moderator Effects*

The main moderators accounting for the heterogeneity in ST effects were strip till depth, texture and rotation (see Table 4). Depths of 8 cm ($p < 0.0001$) and 20 cm ($p = 0.009$) led to increases in RR, as shown in Fig. 13. The same increase in RR was found with a loamy sand texture ($p = 0.041$). Combinations showed that a strip depth of 20 cm and no

Table 4
Statistical data for strip tillage showing heterogeneity and accountability of the heterogeneity by single moderators.

	I^2 (%)	R^2 (%)	τ^2	H^2	Q_E (df)	p-value	QM (df)	p-value
SD	65.81	10.86	0.01	2.93	6.12(2)	0.05	5.12(5)	0.39
Texture	68.15	8.13	0.007	3.14	12.79(4)	0.01	3.12(3)	0.37
Rotation	69.53	0.23	0.007	3.28	21.22 (6)	0.002	0.9696(1)	0.32
Duration	73.05	0.00	0.009	3.71	24.45(6)	0.0004	0.0713(1)	0.79
Crop	75.36	0.00	0.01	5.06	17.81(4)	0.001	1.10(3)	0.78

I^2 (residual heterogeneity), τ^2 (estimated amount of residual heterogeneity), R^2 (amount of heterogeneity accounted for), Strip depth (SD), Q_E (test of residual heterogeneity), H^2 (sampling variability) and QM (Omnibus test for moderators).

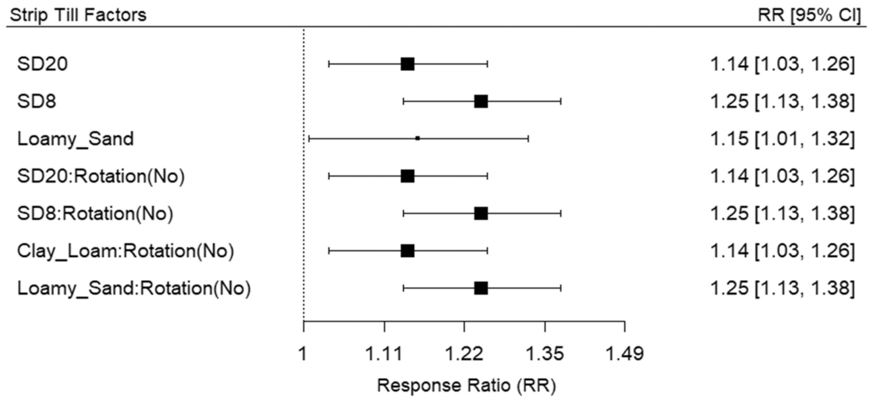

Fig. 13. Forest plot with yield comparisons (estimated average response ratios RR) between ST and conventional tillage based on significant moderator levels accounting for heterogeneity and effect on the estimated RR. SD is strip till depth at 8 cm and 20 cm. Error bars are the confidence intervals CI. The square shapes represent the estimated RR and the size corresponds to its weight (based on its sample size). The vertical broken line represents the line of no effect.

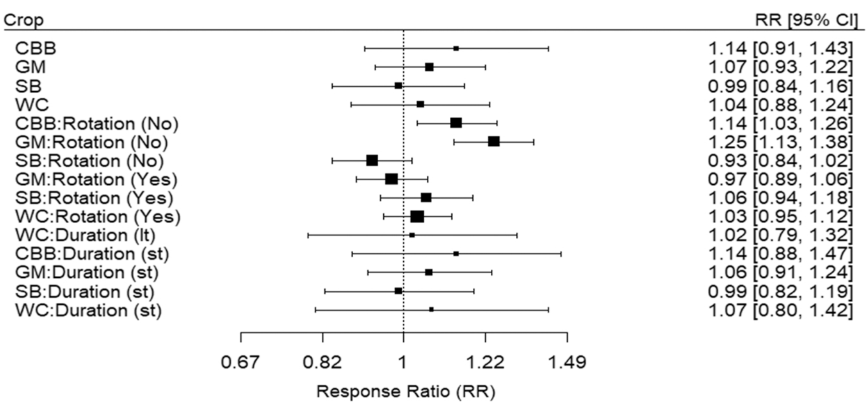

Fig. 14. Forest plot with yield comparisons (estimated average response ratios RR) between ST and conventional tillage based on crop moderator levels accounting for heterogeneity and effect on the estimated RR. CBB is cabbage, GM is grain maize, SB is sugar beet, WC is winter cereal, lt is long term duration (>= 5 years), st is short term duration (<5 years). Error bars are the confidence intervals CI. The square shapes represent the estimated RR and the size corresponds to its weight (based on its sample size). The vertical broken line represents the line of no effect.

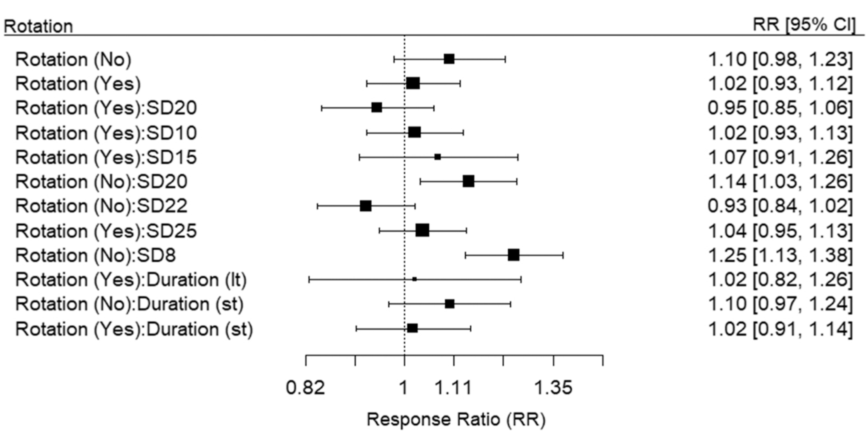

Fig. 15. Forest plot with yield comparisons (estimated average response ratios) between ST and conventional tillage based on rotation moderator levels accounting for heterogeneity and effect on the estimated RR.SD is strip till depth at 8 cm, 10 cm, 15 cm, 20 cm, 22 cm, and 25 cm, lt is long term duration (>= 5 years), st is short term duration (<5 years). Error bars are the confidence intervals CI. The square shapes represent the estimated RR and the size corresponds to its weight (based on its sample size). The vertical broken line represents the line of no effect.

rotation, resulted in an increase in RR ($p = 0.03$). Likewise, a strip till depth of 8 cm with no rotation led to higher RR ($p < 0.0001$). The texture moderator showed that clay loam ($p = 0.009$) and loamy sand ($p < 0.001$) in combination with no crop rotation led to increases in RR.

3.4.1. Crop Moderator

Main crop effects noted are those for cabbage ($p = 0.26$), grain maize ($p = 0.36$), sugar beet ($p = 0.88$) and winter cereals ($p = 0.63$), which all show an average estimated RR greater than 1 except for sugar beet as seen in Fig. 14. While checking for combinations, the only significant results were for cabbage ($p = 0.009$) and grain maize ($p < 0.0001$) in a no crop rotation scenario. All other possible combinations were not significant, with e.g. sugar beet leading to a non-significant decrease in RR ($p = 0.12$) in combination with no crop rotation and a non-significant increase in RR ($p = 0.33$) with crop rotation. Also grain maize without rotation led to a decrease in RR ($p = 0.51$). Winter cereals increased the RR when in combination with rotation ($p = 0.43$). All other interactions were for a short term duration, where sugar beet was the only crop with a decrease in RR ($p = 0.89$). So, cabbage ($p = 0.32$), grain maize ($p = 0.42$), and winter cereals ($p = 0.64$) all led to non-significant increases in RR.

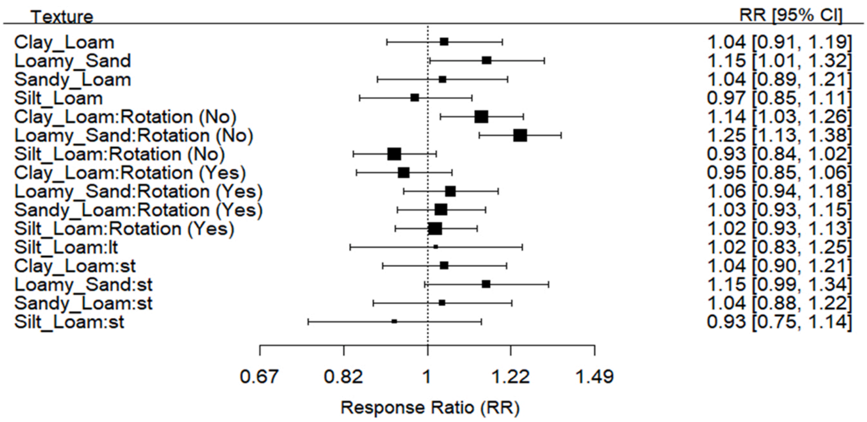

Fig. 16. Forest plot with yield comparisons (estimated average response ratios RR) between ST and conventional tillage based on texture moderator levels accounting for heterogeneity and effect on the estimated RR. lt is long term duration (>= 5 years), st is short term duration (<5 years). Error bars are the confidence intervals CI. The square shapes represent the estimated RR and the size corresponds to its weight (based on its sample size). The vertical broken line represents the line of no effect.

3.4.2. Rotation Moderator

Irrespective of whether there was a rotation or not, ST resulted in an increased RR, with no rotation ($p = 0.11$) leading to a higher increase than with rotation ($p = 0.71$), as can be seen in Fig. 15. With combinations, it was seen that rotation at a strip till depth of 20 cm and no rotation at a strip till depth of 22 cm led to decreases in RR with $p = 0.34$ and $p = 0.12$, respectively. The only depths with significant increases in RR without rotation were at 20 cm ($p = 0.009$) and 8 cm ($p < 0.0001$). With rotation and strip depth combinations, RR increased at 10 cm ($p = 0.68$), 15 cm ($p = 0.40$) and 25 cm ($p = 0.41$). An increase in RR is seen for a long term duration in combination with crop rotation. Likewise, a similar increase with/without rotation is observed with a short term duration.

3.4.3. Texture moderator

The main effects for texture with non-significant increases in RR were clay loam ($p = 0.55$) and sandy loam ($p = 0.65$), and with a significant increase loamy sand ($p = 0.04$), while for silt loam ($p = 0.68$), a non-significant decrease in RR was seen. All texture with rotation comparisons showed an increase in RR except for clay loam with rotation ($p = 0.12$) and silt loam with rotation ($p = 0.34$) where a decrease in RR was observed as seen in Fig. 16. So, clay loam ($p = 0.009$) and loamy sand ($p < 0.0001$) without rotation still led to increases in RR. Other relatively lower non-significant increases in RR came with implementation of rotation on loamy sand ($p = 0.33$), sandy loam ($p = 0.52$) and silt loam ($p = 0.68$). Textural comparison with duration showed a non-significant decrease in RR ($p = 0.46$) when silt loam was under a short term duration as opposed to long term where RR increases ($p = 0.84$). Clay loam ($p = 0.59$), loamy sand ($p = 0.06$) and sandy loam ($p = 0.66$) all showed increases in RR while in the short term duration.

3.4.4. Strip depth moderator

The SD moderator accounted for 10.86% of the heterogeneity. This is further seen with the I^2 value being less than 100%. With a test for moderators ($Q_M = 5.12$), we find no significant ($p = 0.39$) evidence that the differences between the various strip depth levels are simultaneously equal to zero. No results were obtained when performing pairwise comparisons with strip depth versus crop type and texture moderators because the numerous moderator levels relative to the small number of studies ($n = 8$).

4. Discussion

4.1. Crop yield and tillage type

To answer the first hypothesized question, it can be stated that yields are indeed not optimal under no till as opposed to RT and ST where yields are higher than those of conventional tillage. However, European crop yield varies substantially under both conservation (considering NT, RT and ST) and conventional tillage. This variation is attributable to crop type, crop rotation, residue retention, soil texture, tillage depth, climate and duration of tillage practice. However, given the limited data under RT and ST, it is found that reasonable moderators worth taking into account under RT are ridge height, texture and crop based on the model fit (Table 3) and strip depth for ST (Table 4). These results are generally consistent with the findings of Van den Putte et al. (2010) who found in Europe yield reductions (0–30%) under conservation tillage. They specifically found 8.5% average crop yield reductions under NT while this study found an average reduction of 5.1% under the same tillage type. Also, Pittelkow et al. (2015) in their global meta-analysis found that on average NT negatively impacted yields by 5.7%. A very promising major finding was the 5% yield increase for both RT and ST compared to CT.

Given the numerous advantages that come with crop rotation, it was still seen that yields are improving under RT and ST even without rotation. This can be explained partly by the fact that these techniques are usually implemented with residue left on the surface and receive a small amount of soil disturbance and helping to maximize soil temperature which helps in faster seed emergence (Licht and Al-Kaisi, 2005; Overstreet et al., 2010). Also, on the very nature of RT, there's decreased infiltration on ridges, leading to decreased leaching and erosion of minerals like nitrates and phosphates (Müller et al., 2009; Bargar et al., 1999) into deeper soil layers thus remain available for the plants. In addition, ridge heights of at least 20 cm lead to higher yields because in most cases, the elevated soil properties on ridged soil are responsible for better yield performance compared to an un-ridged soil (Benjamin et al., 1990; Sharratt et al., 1996; Hatfield et al., 1998; Krause et al., 2008). A major moderator is that higher ridge temperatures favouring quick emergence, soil aggregation and water use efficiency are being improved. This was shown to be very important especially to parts of northern Europe where ridging seems to be common practice in raising the temperature of the cultivated soil (Taivalmaa and Talvitie, 1997).

The greatest yields increases (25%) at 8 cm strip depth could be explained by the minimal shallow soil disturbance as a potential solution to problems like late seed emergence due to cold and wet soil conditions associated with methods such as NT (Licht and Al-Kaisi, 2005). The tilled zone enhances evaporation of water from soil and warming of the seedbed while total soil disturbance is minimized. These improved soil conditions work best on loamy sand texture leading to yield increases of up to 15% within Europe when considering only this texture type. Furthermore, Ren et al. (2019) showed an increase in maize biomass yield on a sandy loam under ST. Our study shows sandy loam to be resulting in an average of 4% yield increase with ST.

4.2. Crop yield and climate

Areas that are classified as dry tend to depend on rain fed or irrigated agriculture, thus, have been adapted to maximize the use of available water (Moret et al., 2007). Irrigation and water use efficiency was one of the major reasons for the yield gaps noticed between two of the three outlying studies. And this factor alone contributes to conservation tillage yields of semi-arid regions of Spain and most of Southern Europe to be

Table A1
Data table showing references and country of studies used in the meta-analysis.

No	Reference	Country	No	Reference	Country
1	Angas et al., (2006)	Spain	70	Moitzi et al., (2019)	Austria
2	Anken et al., (2004)	Switzerland	71	Moldovan et al., (2019)	Romania
3	Anken et al., (2004)	Switzerland	72	Morell et al., (2011)	Spain
4	Ali et al., (2019)	Italy	73	Moret et al. (2007)	Spain
5	Antmen (2019)	Turkey	74	Muhlbachova et al., (2015)	Czech
6	Avizienyte et al., (2013)	Lithuania	75	Neugschwandtner et al., (2015)	Austria
7	Biberdzic et al., (2019)	Serbia	76	Pagnani et al., (2019)	Italy
8	Borin and Sartori(1995)	Italy	77	Panasiewicz et al., (2020)	Poland
9	Borin and Sartori(1995)	Italy	78	Pareja-Sanchez et al., (2019)	Spain
10	Braim et al., (1992)	England	79	Plaza-Bonilla et al., (2017)	Spain
11	Buchi et al., (2017)	Switzerland	80	Plaza-Bonilla et al., (2018)	Spain
12	Buchi et al., (2017)	Switzerland	81	Rieger et al., (2008)	Switzerland
13	Calzarano et al., (2018)	Italy	82	Rusu et al., (2015)	Romania
14	Cantero-Martinez et al., (2007)	Spain	83	Rusu et al., (2015)	Romania
15	Carbonell-Bojollo et al., (2019)	Spain	84	Raus et al., (2016)	Romania
16	Casa and Lo Cascio (2008)	Italy	85	Ramos et al., (2019)	Spain
17	Castelli et al., (2019)	Italy	86	Ruisi et al., (2014)	Italy
18	Chetan et al., (2018)	Romania	87	Schlu?ter et al., (2019)	Germany
19	Chetan et al., (2019)	Romania	88	Schlu?ter et al., (2019)	Germany
20	Chetan et al., (2017)	Romania	89	Seddaiu et al., (2016)	Italy
21	Christian and Bacon(1990)	UK	90	Selvi et al., (2019)	Turkey
22	Celik et al., (2019)	Turkey	91	Sharma (1985)	UK
23	Cociu (2016)	Romania	92	Sieling et al., (1998)	Germany
24	Cociu (2016)	Romania	93	Stosic et al., (2017)	Croatia
25	Cociu and Alionte (2017)	Romania	94	Spiess et al., (2020)	Switzerland
26	Cociu and Alionte (2011)	Romania	95	Spiess et al., (2020)	Switzerland
27	Cociu and Alionte (2011)	Romania	96	Spiess et al., (2020)	Switzerland
28	Colecchia et al., (2015)	Italy	97	Struck et al., (2019)	Germany
29	Copec et al., (2015)	Croatia	98	Taner et al., (2016)	Turkey
30	Copec et al., (2015)	Croatia	99	Tellez-Rio (2017)	Spain
31	Dachraoui and Sombrero	Spain	100	Videnovic et al., (2011)	Serbia
32	De Vita et al., (2007)	Italy	101	Wozniak (2016)	Poland
33	Dekemati et al., (2019)	Hungary	102	Wozniak (2016)	Poland
34	Demir and Gozubuyuk (2020)	Turkey	103	Wozniak (2020)	Poland
35	Ekeberg and Riley(1996)	Norway	104	Wozniak (2020)	Poland
36	Faust et al., (2019)	Germany	105	Wozniak and Soroka (2018)	Poland
37	Faust et al., (2019)	Germany	106	Wozniak and Gos (2014)	Poland
38	Fernandez et al., (2007)	Spain	107	Casa & Cascio (2008)	Italy
39	Filipovic et al., (2006)	Croatia	108	Borin and Sartori (1995)	Italy
40	Filipovic et al., (2006)	Croatia	109	Borin and Sartori (1995)	Italy
41	Giambalvo et al., (2018)	Italy	110	Eggert et al., (2018)	Germany
42	Gozubuyuk et al., (2015)	Turkey	111	Gursoy et al., (2012)	Turkey
43	Habbib et al., (2017)	France	112	Henriksen et al., (2006)	Denmark
44	Hernanz et al., (2014)	Spain	113	Henriksen et al., (2007)	Denmark
45	Husnjak et al., (2002)	Croatia	114	Krause et al. (2009)	Germany
46	Jug et al., (2019)	Croatia	115	Klikocka and Sommer (2003)	Poland
47	Jug et al., (2019)	Croatia	116	Marcinek et al., (2013)	Poland
48	Jug et al., (2011)	Croatia	117	Mert et al., (2006)	Turkey
49	Koch et al., (2009)	Germany	118	Ozpinar and Isik (2004)	Turkey
50	K?hling et al., (2017)	Siberia	119	Stathakos et al., (2006)	Greece
51	Kovacev et al., (2018)	Crotia	120	Taivalmaa (1997)	Finland
52	Kovacev et al., (2018)	Croatia	121	Cociu and Alionte (2011)	Romania
53	Kosutic et al., (2001)	Croatia	122	Gaj et al., (2015)	Poland
54	Kosutic et al., (2001)	Croatia	123	Jaskulska et al., (2019)	Poland
55	Lampurlanes et al., (2016)	Spain	124	Laufer et al., (2016)	Germany
56	Lithourgidis (2005)	Greece	125	Ren et al. (2019)	Belgium
57	Lopez-Bellido et al., (2012)	Spain	126	Sommermann et al., (2018)	Germany
58	Lopez-Bellido et al., (2001)	Spain	127	Tauchnitz et al., (2018)	Germany
59	Lopez-Bellido et al., (1998)	Spain	128	Ubelhor et al., (2014)	Germany
60	Lopez-Fando et al., (2007)	Spain			
61	Lopez and Arrue., (1997)	Spain			
62	Malecka et al., (2015)	Poland			
63	Malecka et al., (2012)	Poland			
64	Martinez et al., (2016)	Switzerland			
65	Martinez et al., (2016)	Switzerland			
66	Martinez et al., (2016)	Switzerland			
67	Martin-Rueda et al., (2007)	Spain			
68	Mazzoncini et al., (2009)	Italy			
69	Mikanova et al., (2012)	Czech			

comparable or most often greater than those of CT. Van de Putte et al. (2010) also concluded that climate may affect relative yields under conservation tillage as it is suggested that conservation tillage has a significant effect on the overall water balance and the availability of soil water to the plants (Holland, 2004). Lower yields in humid European areas can be attributable to their temperature not favouring quick crop germination and the fact that soils might be too wet when farm activities are done hence resulting in more compaction which is not remediated by tillage. The results from this analysis seems to suggest that such a yield difference between dry and humid areas may hinge on a better water use efficiency in drier areas.

4.3. Crop yield and crop rotation

Though conservation tillage yields are seen to decrease with crop rotations, interactions with other moderator levels show some scenarios performing better than others. We noted in similar fashion as Pittelkow et al. (2015) that, yields under crop rotation and residue retention are 11% higher than with no rotation and no residue retention (Fig. 6). Despite this relative increase, the yields are still lower to those of CT. One explanation is the probable increase of pests and diseases (Turner, 2004; Cantero-Martínez, 2007) due to residue cover. However, incorporation of crop rotation helps counteract the negative effects (pest) of residue cover as rotations are used to prevent or at least partially control several pests and at the same time to reduce the farmer's reliance on chemical pesticides (Chaddad, 2016). Also, Turner (2004) suggested that decreases in yield with cereal rotations can in addition most probably be due to increased nitrogen loss. Though yields are less for conservation tillage compared to CT irrespective of rotation or residue retention, it was shown that there is a gain for combining the two methods. Rotation particularly leads to yield reductions when the crop involved is grain maize in a humid climate.

4.4. Crop type and texture on yield

When considering the effect of crop type as seen in Fig. 8, there exist variation among the crops and in their interactions with other moderator. It is seen that grain maize is the only single crop with significant lower yields on conservation tillage and it only shows yield increases with some interactions. There was an overall reduction in grain maize of 8% under NT and 18% under RT, while under ST a 7% increase was observed. Other crops like cereals begin to show reductions when other moderators like texture and duration of study come into play. Grain maize on clay, clay loam and silty clay textures all show yield reductions as opposed to cereals. Van den Putte et al. (2010) points out to the effects of conservation tillage practices on root development to be one of the moderators responsible for such yield reductions. In addition, RT with permanent beds brings about changes in the physical characteristics of the soil like increased penetration resistance and bulk density during the early growth stages (Katsvairo et al., 2002; Shi et al., 2012) as opposed to the tilled upper soil layers of CT. These physical characteristics also offer a resistance to root development under NT and RT (Katsvairo et al., 2002). Thus, this could be one of the reasons for the relatively better grain maize yield under ST. It has been identified that because of the higher soil temperatures, lower soil strength, and a uniform phosphorus availability under conventionally tilled soil, maize tends to perform better in such soils compared to a no tilled soil (Qin et al., 2006). They further pointed out to a better maize root system in soil under a CT system as one reason for its improved yield, whereas root development of other cereals under conservation tillage was not significantly affected it was for maize. Ren et al. (2019) also found maize root growth limitations under no-tilled strips and eventually demonstrated that lower aeration within these strips partly explained it. They, however, observed a marginal increase in maize yield of 1.3% under ST in comparison with CT. Thus, the consensus in the literature above seems that grain maize does well on tilled soils because of the lower soil strength and aeration the roots encounter. This same reason could be used to conclude that roots of cereals can withstand the rather higher soil strength (Qin et al., 2004) on NT and ST because they generally perform better than grain maize. Van den Putte et al. (2010) noted that the reasons for this crop root differences and or performance are not entirely understood but the tillage systems could be a possible explanation.

5. Conclusions

This study found that European crop yields under RT and ST can be expected to be similar to those under CT. For farmers, this may serve as an alternative to NT where yields are on average, are 5.1% lower than those of CT. It is seen that soil texture as a moderator among all the moderators in this study does better at explaining the between study variability among studies.

Based on results from the limited data for RT and ST data, one will be inclined to recommend these methods to farmers. It is worth noting that for farmers, these two practices are further beneficial in that several combinations as presented above offer them the possibility of choosing whether or not to use rotation, which crop type to choose that best meets their needs. For ST we see that a sandy loam texture is ideal while for RT we identify a loamy texture. One is further inclined to conclude that ST is an alternative for crops like grain maize with yield reductions under NT and RT. Winter cereals don't encounter the problem of tillage type like grain maize. This can be as a result of the immediate aerated top soil on tilled strips which provide less resistance to grain maize roots. This is also evident by their relative lack of yield reductions on clay, clay loam and silty clay textures. Thus, one will be inclined to recommend ST to farmers as an alternative to CT especially for maize cultivation. This study shows strip depth as key moderators worth considering.

Even with the lower crop yields under NT compared with CT, we show yield improvement with other moderators like improved grain maize yields under NT in a sandy clay loam textured soil. Our analysis particularly shows the benefits of always using residue retention and crop rotation together under NT.

Declaration of Competing Interest

We know of no conflicts of interest associated with this publication, and there has been no significant financial support for this work that could have influenced its outcome. As Corresponding Author, I confirm that the manuscript has been read and approved for submission by all the named authors. Thus, we declare that this manuscript is original, has not been published before and is not currently being considered for publication elsewhere.

Data Availability

I have shared the link to my data at the attach file step https://data.mendeley.com/datasets/j329rc9336.

Appendix

See Appendix Table A1.

Appendix A. Supporting information

Supplementary data associated with this article can be found in the online version at doi:10.1016/j.fcr.2023.108967.

References

Alaoui, A., Barao, L., Ferreira, C.S.S., Schwilch, G., Basch, G., Garcia-Orenes, F., Morugan, A., Mataix-Solera, J., Kosmas, C., Glavan, M., Szabó, B., Hermann, T., Vizitiu, O.P., Lipiec, J., Frąc, M., Reintam, E., Xu, M., Di, J., Fan, H., Sukkel, W., Lemesle, J., Geissen, V., Fleskens, L., 2020. Visual assessment of the impact of agricultural management practices on soil quality. Agron. J. https://doi.org/10.1002/agj2.20216.

Alvarez, R., Steinbach, H.S., 2009. A review of the effects of tillage systems on some soil physical properties, water content, nitrate availability and crops yield in the Argentine Pampas. Soil Tillage Res. 104, 1–15. https://doi.org/10.1016/j.still.2009.02.005.

Araya, T., Gebremedhin, A., Baudron, F., Hailemariam, M., Birhane, E., Nyssen, J., Cornelis, W., 2021. Influence of 9 years of permanent raised beds and contour furrowing on soil health in conservation agriculture based systems in Tigray region, Ethiopia. LAND Degrad. Dev. 32 (3), 1525–1539. https://doi.org/10.1002/ldr.3816.

Baker, C.J., Saxton, K.E., Ritchie, W.R., 2002. No-Tillage Seeding: Science and Practice, 2nd ed. CAB International, Oxford, UK.

Bargar, B., Swan, J., Jaynes, D., 1999. Soil water recharge under uncropped ridges and furrows. Soil Sci. Soc. Am. J. 63 (5), 1290–1299. https://doi.org/10.2136/sssaj1999.6351290x.

Benjamin, J.G., Blaylock, A.D., Brown, H.J., Cruse, R.M., 1990. Ridge tillage effects on simulated water and heat transport. Soil Tillage Res. 18, 167–180 https://doi.org/10.1016/0167-1987 (90)90057-K.

Cantero-Martínez, C., Angás, P., Lampurlanés, J., 2007. Long-term yield and water use efficiency under various tillage systems in Mediterranean rainfed conditions. Ann. Appl. Biol. 150 (3), 293–305. https://doi.org/10.1111/j.1744-7348.2007.00142.x.

Chaddad, Fabio, 2016. Econ. Organ. Braz. Agric. || Enabling Cond. 19–44 https://doi:10.1016/b978-0-12-801695-4.00002-1.

Debska, B., Jaskulska, I., Jaskulski, D., 2020. Method of tillage with the factor determining the quality of organic matter. Agronomy 10 (9), 1250. https://doi.org/10.3390/agronomy10091250.

FAO, 1993. Soil Tillage in Africa: Needs and Challenges. FAO Soils Bulletin 69. FAO, Rome. http://hdl.handle.net/10919/65448.

Fernández, F.G., Sorensen, B.A., Villamil, M.B., 2015. A comparison of soil properties after five years of no-till and strip-till. Agron. J. 107 (4), 1339–1346. https://doi.org/10.2134/agronj14.0549.

Hatfield, J.L., Allmaras, R.R., Rehm, G.W., Lowery, B., 1998. Ridge tillage for corn and soybean production: environmental quality impacts. Soil Tillage Res 48 (1998), 145–154. https://doi.org/10.1016/S0167-1987(98)00141-X.

Hedges, L.V., Olkin, I., 1985. Statistical Methods for Meta-Analysis. Academic Press, New York https://doi: 10.2307/1164953.

Hedges, L.V., Vevea, J.L., 1998. Fixed- and Random-Effects Models in Meta-Analysis. Psychol. Methods 3 (4), 486–504. https://doi.org/10.1037/1082-989X.3.4.486.

Higgins, J.P.T., Thompson, S.G., Deeks, J.J., Altman, D.G., 2003. Measuring inconsistency in meta-analyses. BMJ: Br. Med. J. 327 (7414), 557–560. https://doi.org/10.1136/bmj.327.7414.557.

Holland, J.M., 2004. The Environmental consequences of adopting conservation tillage in Europe: Reviewing the evidence. Agric., Ecosyst. Environ. 103, 1–25. https://doi.org/10.1016/j.agee.2003.12.018.

Jabro, J., Stevens, W., Iversen, W., Evans, R., 2011. Bulk density, water content, and hydraulic properties of a sandy loam soil following conventional or strip tillage. Appl. Eng. Agric. https://doi.org/10.13031/2013.39576. https://doi.

Jaskulska, I., Jaskulski, D., 2020. Strip-till one-pass technology in central and Eastern Europe: A mzuri pro-til hybrid machine case study. Agronomy 10 (7), 925. https://doi.org/10.3390/agronomy10070925.

Katsvairo, T., Cox, W.J., van Es, H., 2002. Tillage and rotation effects on soil physical characteristics. Agron. J. 94 (2), 299–304. https://doi.org/10.2489/jswc.72.6.553. https://.

Kertesz, A., Madarász, B., 2014. Conservation agriculture in Europe Int. Soil Water Conserv. Res. 2 (1), 91–96. https://doi.org/10.1016/S2095-6339(15)30016-2.

Krause, U., Koch, H.-J., Maerlaender, B., 2009. Soil properties effecting yield formation in sugar beet under ridge and flat cultivation. Eur. J. Agron. 31 (1), 20–28. https://doi.org/10.1016/j.eja.2009.02.002.

Lal, R., 1993. Tillage effects on soil degradation, soil resilience, soil quality, and sustainability. Soil Tillage Res. 27 (1), 1–8. https://doi.org/10.1016/0167-1987(93)90059-X.

Langhans, C., Diels, J., Clymans, W., Van den Putte, A., Govers, G., 2019. Scale effects of runoff generation under reduced and conventional tillage. Catena 176, 1–13. https://doi.org/10.1016/j.catena.2018.12.031. https://.

Laufer, D., Koch, H.J., 2017. Growth and yield formation of sugar beet (Beta vulgaris L.) under strip tillage compared to full width tillage on silt loam soil in Central Europe. Eur. J. Agron. 82, 182–189. https://doi.org/10.1016/j.eja.2016.10.017. https://doi.

Licht, M.A., Al-Kaisi, M., 2005. Strip-tillage effect on seedbed soil temperature and other soil physical properties. Soil Tillage Res. 80, 233–249. https://doi.org/10.1016/j.still.2004.03.017.

Light, R.J., Pillemer, D.B., 1984. Summing Up: The Science of Reviewing Research. Harvard University Press, Cambridge, MA. https://doi.org/10.3102/0013189X015008016.

Meurer, K.H.E., Haddaway, N.R., Bolinder, M.A., Kätterer, T., 2018. Tillage intensity affects total SOC stocks in boreo-temperate regions only in the topsoil—A systematic review using an ESM approach. Earth Sci. Rev. 177, 613–622. https://doi.org/10.1016/j.earscirev.2017.12.015. https://doi.

Meyer, L.D., Dabney, S.M., Murphree, C.E., Harmon, W.C., 1999. Crop production systems to control erosion and reduce runoff from upland silty soils. Trans. ASAE 42 (6), 1645–1652. https://doi.org/10.4141/S03-014. https://doi.

Moret, D., Arrúe, J.L., López, M.V., Gracia, R., 2007. Winter barley performance under different cropping and tillage systems in semiarid Aragon (NE Spain. Eur. J. Agron. 26 (1), 54–63. https://doi.org/10.1016/j.eja.2006.08.007.

Müller, E., Wildhagen, H., Quintern, M., Heß, J., Wichern, F., Joergensen, R.G., 2009. Spatial patterns of soil biological and physical properties in a ridge tilled and a ploughed luvisol. Soil Tillage Res. 105 (1), 88–95. https://doi.org/10.1016/j.still.2009.05.011. https://.

Nunes, M.R., Karlen, D.L., Moorman, T.B., 2020. Tillage intensity effects on soil structure indicators—a US meta-analysis. Sustainability 12, 2071. https://doi.org/10.3390/su12052071.

Overstreet, L.F. (2009). Strip tillage for sugarbeet production. Intern.Sugar J. 111: 292–304. 102 Journal of Sugar Beet Research Vol. 49 Nos. 3 & 4.

Panagos, P., Ballabio, C., Poesen, J., Lugato, E., Scarpa, S., Montanarella, L., Borrelli, P., 2020. A soil erosion indicator for supporting agricultural, environmental and climate policies in the European Union. Remote Sens. 12 (9), 1365. https://doi.org/10.3390/rs12091365.

Qin, R.J., Stamp, P., Richner, W., 2006. Impact of tillage on maize rooting in a cambisol and luvisol in Switzerland. Soil Tillage Res. 85, 50–61 https://doi:10.1016/j.still.2004.12.003.

Rasmussen, K.J., 1999. Impact of ploughless soil tillage on yield and soil quality: A Scandinavian review. Soil Tillage Res. 53, 3–14 https://doi.org/10.1016/S0167-1987 (99)00072-0.

Ren, L., Nest, T.V., Ruysschaert, G., D'Hose, T., Cornelis, W.M., 2019. Short-term effects of cover crops and tillage methods on soil physical properties and maize growth in a sandy loam soil. Soil and Tillage Research, 192, 76–86. https://doi.org/10.1016/j.still.2019.04.026.

Sharratt, B.S., 1996. Tillage and straw management for modifying physical properties of a subarctic soil. Soil Tillage Res 38 (3 – 4), 239–250 https://doi.org/10.1016/S0167-1987 (96)01035-5.

Shi, X., Yang, X., Drury, C., Reynolds, W., McLaughlin, N., Zhang, X., 2012. Impact of ridge tillage on soil organic carbon and selected physical properties of a clay loam in southwestern ontario. Soil Tillage Res. 120, 1–7. https://doi.org/10.1016/j.still.2012.01.003.

Soane, B., Ball, B., Arvidsson, J., Basch, G., Moreno, F., Roger-Estrade, J., 2012. No-till in northern, western and south-western Europe: A review of problems and opportunities for crop production and the environment. Soil Tillage Res 118, 66–68. https://doi.org/10.1016/j.still.2011.10.015.

Sterne, J.A.C., Egger, M., 2001. Funnel plots for detecting bias in meta-analysis: guidelines on choice of axis. J. Clin. Epidemiol. 54 (10), 1046–1055 https://doi:10.1016/s0895-4356 (01)00377-8.

Strauss,P., Swoboda, D., Blum, E.H. (2003). How effective is mulching and minimum tillage to control runoff and soil loss? – A literature review Proceedings of the Conference on 25 Years of Assessment of Erosions "Ghent, 22–26 September, pp. 545–550.

Taivalmaa, S.L., Talvitie, H., 1997. The effects of ridging, row-spacing and seeding rate on carrot yield. Sci. Agric. Soc. Finl. Volume: 6, 363–369. https://doi.org/10.23986/afsci.72799.

Trevini, M., Benincasa, P., Guiducci, M., 2013. Strip tillage effect on seedbed tilth and maize production in Northern Italy as case-study for the Southern Europe environment. Eur. J. Agron. 50–56 https://doi:10.1016/j.eja.2013.02.007.

Van den Putte, A., Govers, G., Diels, J., Gillijns, K., Demuzere, M., 2010. Assessing the effect of soil tillage on crop growth: a meta-regression analysis on European crop yields under conservation agriculture (May). Present., EGU, Vienna, Austria 2–7. https://doi.org/10.1016/j.eja.2010.05.008.

Viechtbauer, W., 2010. Conducting meta-analyses in R with the metafor package. J. Stat. Softw. 36 (3), 1–48. http://www.jstatsoft.org/v36/i03/.

Viechtbauer, W., Cheung, M.W., 2010. Outlier and influence diagnostics for meta-analysis. Res. Synth. Methods 1, 112–125. https://doi.org/10.1002/jrsm.11. https://doi.

Zuur, A.F., Ieno, E.N., Walker, N.J., Saveliev, A.A., Smith, G.M., 2009. Mixed Effects Model-Ling for Nested Data. In Mixed Effects Models and Extensions in Ecology with R. Berlin/Heidelberg, Germany. Springer,, pp. 101–142. https://doi.org/10.1086/64813.