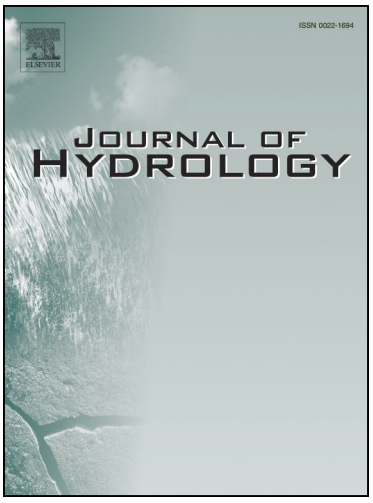
Journal Pre-proofs



Research papers

Opportunities and challenges in using catchment-scale storage estimates from cosmic ray neutron sensors for rainfall-runoff modelling

Katya Dimitrova-Petrova, Josie Geris, E. Mark Wilkinson, Rafael Rosolem, Lucile Verrot, Allan Lilly, Chris Soulsby

PII: S0022-1694(20)30338-3
DOI: https://doi.org/10.1016/j.jhydrol.2020.124878
Reference: HYDROL 124878

To appear in: *Journal of Hydrology*

Received Date: 6 January 2020
Revised Date: 26 February 2020
Accepted Date: 20 March 2020

Please cite this article as: Dimitrova-Petrova, K., Geris, J., Mark Wilkinson, E., Rosolem, R., Verrot, L., Lilly, A., Soulsby, C., Opportunities and challenges in using catchment-scale storage estimates from cosmic ray neutron sensors for rainfall-runoff modelling, *Journal of Hydrology* (2020), doi: https://doi.org/10.1016/j.jhydrol.2020.124878

Katya Dimitrova-Petrova (Orcid ID: 0000-0001-5119-7753)

Josie Geris (Orcid ID: 0000-0003-0159-0543)

Mark Wilkinson (Orcid ID: 0000-0001-5169-758X)

Rafael Rosolem (Orcid ID: 0000-0002-4914-692X)

Lucile Verrot (Orcid ID 0000-0001-9014-6908)

Allan Lilly (Orcid ID: 0000-0002-0142-9269)

Chris Soulsby (Orcid ID: 0000-0001-6910-2118)

Title page:

I. Key words:

Cosmic ray neutron sensor; rainfall-runoff modelling; storage-discharge relationship; managed landscapes; catchment hydrology

II. Title

Opportunities and challenges in using catchment-scale storage estimates from cosmic ray neutron sensors for rainfall-runoff modelling

III. The full name of the authors

Dimitrova-Petrova, Katya (1,2); Geris, Josie (1); Wilkinson, Mark, E. (2); Rosolem, Rafael (3); Verrot, Lucile (1); Lilly, Allan (2); Soulsby, Chris (1).

IV. The author's institutional affiliations

(1) Northern Rivers Institute, School of Geosciences, University of Aberdeen, Aberdeen, United Kingdom, (2) The James Hutton Institute, Craigiebuckler, Aberdeen, United Kingdom, (3) Department of Civil Engineering, University of Bristol, Bristol, United Kingdom

V. Corresponding Author’s name and contact details

Katya Dimitrova-Petrova k.dimitrovap@abdn.ac.uk

University of Aberdeen School of Geosciences St Mary's Building, Room B33 Elphinstone Road Aberdeen, AB24 3UF, United Kingdom

Opportunities and challenges in using catchment-scale storage estimates from cosmic ray neutron sensors for rainfall-runoff modelling

Authors: Dimitrova-Petrova, Geris, Wilkinson, Rosolem, Verrot, Lilly, Soulsby

Key words: Cosmic ray neutron sensor; rainfall-runoff modelling; storage-discharge relationship; managed landscapes; catchment hydrology

Abstract

Adequate characterization of catchment storage dynamics is crucial in hydrological models, yet scale-representative storage measurements are rare. Recent developments in Cosmic Ray Neutron Sensor (CRNS) technology and monitoring networks provide a powerful source of more scale-appropriate soil moisture data for many modelling applications. However, the potential in rainfall-runoff modelling is undeveloped. Here we present the first application of CRNS data in conceptual rainfall-runoff modelling and explore this potential in the context of a mixed-agricultural landscape in Scotland. We deployed and calibrated a CRNS in a heterogeneous soil-land use footprint over a ~3-year period. In this generally wet environment, the CRNS shallow sensing depth and relatively high neutron count uncertainty were identified as major challenges. However, given the better spatial coverage (up to 14 ha) and ease for maintenance, CRNS was thought to represent the simplest approach for long-term monitoring of managed mixed-agricultural sites. We used CRNS-derived, as well as single point-scale estimates, of near-surface soil storage (S_{NS}) to explore their characterisation of storage dynamics at the catchment-scale. Inter-comparison using linear regression showed that S_{NS} related well to catchment-scale storage dynamics, however this relationship was stronger for CRNS (R^2=0.91) compared to point-scale derived estimates (R^2=0.76). Based on this, we evaluated the effect of using

the CRNS and point scale derived S_{NS} data to constrain storage estimates controlling runoff generation in a common rainfall-runoff model (HBV-light). Including CRNS or point-scale field S_{NS} data alone in model calibration was especially useful for intermediate and wet periods. A combined model calibration using discharge and either S_{NS} storage estimates provided a better representation of catchment internal dynamics, additionally reducing uncertainty during low flows. In the context of mixed-agricultural landscapes in humid environments, this study showed the potential of using CRNS over point scale data (in terms of representativeness for single point data and practicality for point sensor networks) to characterise the catchment storage-discharge relationship and inform hydrological modelling.

1. Introduction

Water storage is one of the most important hydrological functions of catchments and strongly influences runoff generation processes (McNamara et al., 2011; Staudinger et al., 2017). Consequently, understanding and adequately characterising the catchment storage-discharge (S-Q) relationship is essential to identifying the dominant runoff processes (Spence, 2007). While we are usually unable to measure catchment-scale storage directly (Beven, 2016), there is a need for its better estimation in hydrological models (Birkel et al., 2015), which is especially challenging in landscapes with spatially variable properties (e.g. in soils and land use) (Blöschl et al., 2019; Beven, 2016; Birkel et al., 2015; Kormos et al., 2015; Zhang et al., 2014). Traditional methodologies for direct storage observations are often at point scale and limited to a single storage compartment (e.g. soils, groundwater, etc). Hence, they usually require rescaling to meet objectives of modelling exercises (Lievens et al., 2015; Loizu et al., 2018). Therefore, developing modelling frameworks that can incorporate new sources of data, (*e.g.* Bogena et al., 2015; Creutzfeldt et al., 2014; Dehaspe et al., 2018), and integrate larger catchment-scale characteristics and address heterogeneities, is crucial but challenging. In this context, Cosmic Ray Sensor (CRNS) data offer new opportunities, as they provide

continuous estimates of near-surface soil water content (SWC) at the intermediate field scale (up to 14 ha or 210 m sensing radius) (Schrön et al., 2017) (see Appendix I for all abbreviations).

Over the last decade, extensive research on the calibration of CRNS sensors has been conducted (Baroni et al., 2018; Franz et al., 2013; Heidbuchel et al., 2016; Iwema et al., 2015; Schrön et al., 2017; Zreda et al., 2008) to improve SWC estimates in a variety of environmental settings (Bogena et al., 2018; Duygu & Akyürek, 2019; Evans et al., 2016; Hawdon et al., 2014; Zhu et al., 2016). Research has also focussed on characterising soil moisture profiles (Baroni et al., 2018) and addressing spatial heterogeneities in soil and vegetation cover (*i.e.* land use)(Han et al., 2014a; Nguyen et al., 2017) with regards to CRNS observations. Increasing availability of CRNS networks worldwide (Baatz et al., 2014; Evans et al., 2016; Hawdon et al., 2014; Zreda et al., 2012) potentially offers a powerful data source for informing hydrological models. Indeed, CRNS data has already been used in some modelling applications, *e.g.* for inverse estimation of soil hydraulic properties (Brunetti et al., 2019; Foolad et al., 2017; Rivera Villarreyes et al., 2014), improving evapotranspiration estimates (Han et al., 2015) and calibration of land-surface models (Baatz et al., 2017; Iwema et al., 2017). Nevertheless, while CRNS applications are growing, they have mainly focused on physically-based modelling either at the plot or at large scale (i.e. for satellite products validation). In contrast, examples of CRNS applications in catchment-scale conceptual rainfall-runoff models, which rely heavily on S-Q relationships, are still lacking.

Success in using CRNS data to constrain catchment S-Q relationships will depend on the relative importance of soil moisture in modulating streamflow generation (Seyfried et al., 2009; Tromp-van Meerveld & McDonnell, 2006). Key challenges relate to potential scaling issues (Brocca et al., 2010), their relatively shallow sensing depth (Martin Schrön et al., 2018) as well as the representativeness of the near-surface storage to the dynamic storage, which controls streamflow dynamics (Staudinger et al., 2017). Nevertheless, in humid environments such as Scotland, where soils are relatively shallow and wet, near-surface soil moisture measurements obtained via CRNS have potential to be highly

informative for representing dynamic catchment storage. Hence, these can aid in informing models to adequately represent these storage dynamics. Here we explored the use of CRNS-based catchment-scale storage estimates in a conceptual rainfall-runoff modelling approach in the context of a mixed-agriculture humid landscape. More specifically, the objectives were to (i) evaluate the extent to which CRNS derived storage is linked to changes in observed catchment storage (S_{WB}), (ii) explore how CRNS derived storage can be used to evaluate catchment-scale storage in rainfall-runoff modelling and finally (iii) evaluate how (i) and (ii) are different if using more traditional point scale measurements instead of CRNS.

2. Methodology

2.1. Study site and instrumentation

This study was conducted in the Elsick catchment (~10 km^2), NE Scotland, UK (Figure 1). Altitude ranges from 90 to 165 m.a.s.l. Mean annual precipitation is approximately 800 mm and potential evapotranspiration 350 mm (Met Office, 2019). The catchment is dominated by metamorphic bedrock, overlain by glacial drift and relatively thin soils (0-1 m) (British Geological Survey, 2019; Soil Survey of Scotland Staff, 1981). The main soil types are poorly drained gleys (Gleysols, IUSS, 2015), covering 68% of the catchment (Figure 1A). The landscape is mainly covered by rotational crops (wheat, barley and rape seed) and grazed grassland (Figure 1C). To sustain this, the gley soils are subject to artificial drainage (Lilly et al., 2012). Freely draining podzols cover approximately 17% of the catchment and typically support grazed grassland and some rotational crops. Soils in the NE headwaters of Elsick consist of organic-rich peats and peaty podzols (Histosols and histic Podzols, IUSS, 2015), associated with woodland and moorland (Figure 1C). The catchment drainage network is comprised of the stem (MAIN), its tributary (TRIB) and several smaller artificial channels that drain into these (Figure 1; see also Dimitrova-Petrova et al., 2020). The total study period spanned from 14 November 2015 to 30 September 2018 (1052 days).

A Cosmic neutron ray sensor (CRNS -1000/B, Hydroinnova, New Mexico, USA) was used to obtain continuous estimates of near-surface average soil water content (SWC_{CRNS}). The CRNS station (Figure 1E) was installed in November 2015 at the intersection of three fields (latitude 57°02'25.58''N and longitude 2°11'18.84''W, elevation= 95 m.a.s.l.), which are representative for the most common soil-land use units in the catchment (Figure 1 B & D): Wheat-Gley (approximately 50% of the CRNS footprint), Barley-Gley (25%) and Pasture-Podzol (25%). The fields are separated by small grass boundary strips (< 2m), typical for this agricultural landscape (Figure 1E). The experimental set-up included a field-camera to monitor crop growth in the Wheat-Gley field, which was ultimately used to correct the CRNS signal for the influence of biomass water equivalent (BWE, in mm). Simultaneously, we measured volumetric soil water content SWC_{point}, every 30 min at four depths (100 mm, 200 mm, 300 mm, 400 mm) using a Profile-Probe PR2 (Delta-T Devices. Ltd), from 4 April 2017 onwards, at a single location. The point-scale probe was installed in the grass strip within a radius of <5 m next to the CRNS-weather station. In addition, we routinely measured topsoil (upper 0.06 m) volumetric SWC (ML2 soil moisture sensor, Delta T Devices Ltd.) every 7 to 14 days to explore spatiotemporal variability between fields. Topsoil SWC was estimated as the average of ten replicates randomly distributed both within each field and across the grass strip within the first 25 meters radius from the CRNS -weather station.

Precipitation was measured using an *Environmental Measurements* tipping bucket rain gauge (ARG100 gauge) at the CRNS station. At the same location, potential evapotranspiration (Penman-Monteith PET) rates were estimated using climatic variables measured every 30 min from 4 April 2017. Stream water levels (TD-Diver, Van Essen Instruments) were recorded at 15 min intervals at the outlets of the MAIN and TRIB catchments. Discharge gauging across the full range of observed levels were used to obtain continuous stream discharge from the stream water level data. Stream discharge for the Elsick catchment outlet Q_{obs} was calculated as the sum of observations at MAIN and TRIB. Daily precipitation data prior to local monitoring (January 2011 – January 2015), were obtained via distance weighted interpolation using 14 neighbouring gauges from a national monitoring network (Centre for

Environmental Data Analysis, Met Office, 2019b) within a 3 to 35 km radius of the catchment. This was also used to fill occasional data gaps in precipitation, while site corrected meteorological data from Dyce Aberdeen Airport (<25 km to the north of Elsick) were used to fill gaps in PET estimates (Met Office, 2019b). To provide an overview of the antecedent wetness conditions we calculated the Antecedent Precipitation Index for the last 7 days (API_7), using the Elsick daily precipitation data with a constant decay coefficient of 0.9 (Hooke, 1979).

2.2. General Methodology

The general methodological framework followed is shown in six main steps on Figure 2, while a description of these major methodological steps is detailed in the sections below. In summary, (I) first we calibrated the CRNS sensor. (II) We then transformed time series of near-surface SWC information, CRNS (SWC_{CRNS}) and point data (SWC_{point}), to produce a continuous time series of near-surface soil storage (S_{NS}) estimates for the upper 400 mm soil profile. (III) To evaluate the extent to which these are linked to catchment-scale storage dynamics, we calculated the storage component of the catchment water balance (S_{WB}) using observed precipitation, actual evapotranspiration estimates (AET) and measured discharge Q_{obs}. (IV) The S_{WB} and the two S_{NS} time series were then compared to understand the relationship between measured near-surface storage and that estimated at the catchment scale. Furthermore, we explored how shallow soil moisture information $S_{NS,}$ either obtained from CRNS or point data, can help better characterize catchment storage in a widely-used rainfall-runoff model (HBV-light). (V) After initial exploration of the model structure and simulated storage dynamics, we assumed that the near-surface storage could be conceptually related to the model dynamic storage (S_{dyn}). (VI) Ultimately, we evaluated how using Q_{obs} or S_{NS} data, individually or jointly, to constrain the model Q_{sim} and S_{dyn} would affect model performance and representation of catchment internal processes.

2.3. Cosmic ray neutron sensor (CRNS) neutron count calibration procedure for soil water content

Cosmic ray neutron sensor (CRNS) technology is based on the relationship between naturally occurring neutrons generated by cosmic rays and hydrogen present in the near-surface (Zreda et al., 2008). Hydrogen pools can be static (e.g. soil organic matter (SOM)) or dynamic (e.g. soil moisture, water contained in biomass) (Baatz et al., 2015; Rosolem et al., 2013; Zreda et al., 2012). The CRNS neutron count rate (in counts per hours, cph) is used to estimate a spatially-integrated near-surface SWC_{CRNS}, which involves continuous correction for atmospheric influences and a site-specific calibration based on field sampling campaigns. For this, we followed common procedures as presented in the literature. Time series of neutron counts rate [cph] were first corrected (N_{pih}) at a 30 min step for atmospheric pressure (p), incoming solar radiation (i) and relative humidity (h) (Evans et al., 2016; Zreda et al., 2012). Time aggregation to daily (24h) of the N_{pih} counts was then applied to meet the requirements of the study and reduce the statistical uncertainty associated with the relatively low counts at 30 min integration time. A daily vegetation correction factor F_{veg} (Baatz et al., 2015) as a function of biological water content (BWE, in mm) was applied to account for the effect of above-ground biomass on the signal to produce daily time series of N_{pihv} (Coopersmith et al. 2014; Jakobi et al., 2018; Tian et al., 2016). Biomass samples were collected from the three main fields and grass strip (n = 6 per land use) at six different times during the growing season (Table 1, Figure 3E) and processed for BWE following Franz et al., (2013); Tian et al. (2016). We then established relationships between these BWE data and crop height readings from a field camera with height gauge. These were used to estimate the BWE for each of the fields at the start of each month. Values of areal-weighted BWE within the CRNS footprint were then linearly interpolated to daily values, while also considering the timing of management activities (*e.g.* harvesting). The effect of snow was accounted for on a daily basis using field-camera records and information on snow depth [cm] from Dyce - Aberdeen Airport weather station. On days when snow was present and estimated SWC_{CRNS} was greater than total porosity (0.6), SWC_{CRNS} was assumed to be equal to the total porosity value.

Uncertainties in the daily SWC_{CRNS} estimate due to temporal heterogeneities (e.g. in bulk density) were further addressed with five soil moisture calibration field campaigns (Table 1). This CRNS calibration

procedure (Franz et al., 2013) was based on the benefits of employing multiple field sampling (Iwema et al., 2015) to also cover a range of soil wetness-vegetation conditions observed in the area of interest (Table 1). The five sampling campaigns involved: three throughout the 2017 growing season (April – November 2017); one in July 2018, to capture unusually dry summer conditions and one in September 2018 soon after harvesting of the Wheat-Gley and Barley-Gley fields. Soil samples were collected at 21 locations along 6 transects (2 in each field) at radial distances from the CRNS probe at 5, 25 and 75 m (Figure 2) using a soil corer and metal ring of known volume (95 cm^3) (*Eijkelkamp*, Giesbeek, the Netherlands). We selected these distances as they are most appropriate for the relatively small CRNS footprint in this generally wet environment (Schrön, et al., 2017). Additionally, 3 profiles on the grass strip where the CRNS was installed were sampled at distances 1, 2 and 5 meters. Samples were collected at three depths 0-5, 5-10, 10 -15 cm giving a total of 63 samples per campaign. We measured gravimetric SWC and dry bulk density (ρ_{bulk} in g cm^{-3}) of each oven-dried (24h/105°C) sample. The ρ_{bulk} was further corrected for the influence of stones (Hall et al., 1971), due to the stony nature of the soils (Table 1). For each sample, the proportion of soil organic matter (SOM) and soil lattice water (LW, i.e. the water contained in the soil minerals), were measured from 3 g sub-samples (loss-on-ignition, Davies, 1974). The effect of SOM and LW on the CRNS signal was evaluated jointly (Heidbuchel et al., 2016). We applied depth-distance weighting on the field calibration data, following Köhli et al. (2015) and also considered areal weighting taking into account the proportion of each field into the footprint (which remained constant with varying footprint size). Ultimately the daily SWC_{CRNS} was estimated as follows:

$$SWC_{CRNS} = \left(\frac{a_o}{\frac{N_{pihvv}}{N0} - a_1} - a_2 - (SOM + LW) \right) * \frac{\rho_{dry}}{\rho_w} \quad \text{(Eq 1)}$$

where SWC is in [cm^3 cm^{-3}], N_{pihv} are the biomass included corrected neutron counts integrated over 24 hours (in counts per day, cpd), N_0 is a theoretical site-specific value of the count rate over dry soil (here N_0=82800 cpd, optimized through field calibration, and a_i are fixed coefficients (a_0=0.0808, a_1=0.372 and a_2=0.115, Desilets et al., 2010). The ρ_w is the density of water (1 g cm^{-3}). While more

recent research by Schrön et al. (2017) proposed a more precise method for estimating the daily average CRNS effective sensing depth (z_{eff}), we followed the less complex original approach by Franz et al. (2012). We found that overall, the simpler approach caused an underestimation of z_{eff} but, most importantly, this choice had no impact on the moisture dynamics nor on the outcome of our near surface storage estimates used for the rainfall-runoff modelling, which was the main focus of this work (data not shown).

Uncertainty in SWC_{CRNS} associated with the ρ_{dry} was estimated by the difference in the SWC_{CRNS} time series computed using the mean ρ_{dry} and those using the minimum and maximum ρ_{dry} of each sampling (Table 1). We also assessed the effect of correcting the SWC_{CRNS} for BWE by subtracting the difference between the N_{pihv} derived SWC_{CRNS} and the non-vegetation corrected. However, since it had no detectable effects on the results it was not presented here.

2.4 Estimating near-surface storage (S_{NS}) from soil water content data

Time series of near-surface SWC data from the CRNS and point-scale sensors were transformed to near-surface storage (S_{NS}) estimates, being (S_{NS_CRNS}) and (S_{NS_point}), respectively. In this study, the S_{NS} is represented as the total soil water storage in a column of 400 mm. A total depth of 400 mm was chosen because (i) in the poorly draining gley soils (covering 75% of the footprint and ~68% of the catchment; Figure 2 A&B), a gley layer is present at depth of 300-400 mm, which marks the limit to the C soil horizon and parent material and above which most of the dynamic changes in soil moisture occur and; (ii) for the same reason, point-scale data were available for depths up to 400 mm.

While CRNS provides useful field-scale SWC information, issues such as the shallow and time-variable sensing depth need to be addressed to determine total soil storage across deeper profiles (here 400 mm). To account for that, we applied an exponential filter (as described by Peterson et al. (2016) for CRNS applications; and developed originally by Wagner et al. (1999) on daily SWC_{CRNS} [m^3 m^{-3}] to obtain depth-integrated SWC averages for the 400 mm soil column. For this the soil is divided into two

layers with variable depth. The depth of the upper soil layer is equal to the CRNS effective measurement depth (z_{eff}) with soil water content as SWC_{CRNS}. The deeper layer has depth 400 mm - z_{eff}, and soil water content SWC_{CRNS}^{SUB} , which is modelled according to:

$$(400\ mm - z_{eff}) * (d\ SWC_{CRNS}^{SUB}/dt) = C * (SWC_{CRNS} - SWC_{CRNS}^{SUB}) \quad \text{(Eq 2)}$$

where t is time and C is a proportionality constant. Following Peterson et al. (2016) Equation 2 can be reformulated as:

$$SWI_{CRNS\_SUB}(t) = SWI_{CRNS\_SUB}(t-1) * (1 - K_t) + SWI_{CRNS} * K_t \quad \text{(Eq 3)}$$

Where SWI_{CRNS}^{SUB} and SWI_{CRNS} are the soil water index of the modelled and monitored soil layers, respectively, t is the time index (in days) and K is the gain parameter. The K parameter was added by Stroud, (1999) to the original method proposed by Wagner et al., (1999), mainly to allow for missing SWI data (Albergel et al., 2008). The SWI represents the soil water content scaled from 0 to 1 with 0 the minimum (0.15 m^3m^{-3}) and 1 the maximum (0.6 m^3m^{-3}) SWC_{CRNS} recorded during the study. The K parameter, ranging from 0 to 1 is calculated according to:

$$K(t) = \frac{K(t-1)}{K(t-1)+\exp(-\frac{dt}{T})} \quad \text{(Eq 4)}$$

Where K (t-1) is the gain from the previous time, dt is the time step and T= (400 mm – z_{eff})/C is the characteristic time length. The characteristic time length T represents the timescale of soil moisture variation and depends on several factors such as thickness of the modelled soil layer, and soil properties influencing water transmission rates (such as hydraulic conductivity and soil texture). Ideally, T is calibrated based on an independent moisture sensor network within the footprint, which was not available here. However, in agreement with Peterson et al. (2016), an initial sensitivity test revealed that only the amplitude was slightly sensitive to T, but not the mean and general dynamics of S_{NS_CRNS}. For T we used a reference value of 15 days, as soils within the CRNS footprint drain relatively fast (Dimitrova-Petrova et al., 2020). The exponential filter was initialized with K_1=1 and SWI_{CRNS} =

SWI_{CRNS_SUB}. The intermediate step is necessary, as the SWI_{CRNS}^{SUB} was used to calculate SWC_{CRNS}^{SUB}. The near-surface soil water storage S_{NS_CRNS} was then calculated as the sum of storage in upper (sensed by CRNS) and lower soil layers:

$$S_{NS\_CRNS}(t) = SWC_{CRNS}(t) * z_{eff}(t) + SWC_{CRNS}^{SUB}(t) * (400 - z_{eff}) \quad \text{(Eq 5)}$$

for which S_{NS_CRNS} and z_{eff} are both in mm.

For S_{NS} estimates using point scale data (S_{NS_point} in mm), the total soil storage for the upper 400 mm was calculated from the soil profile probe volumetric SWC data at the four different depths (all in m^3 m^{-3}) assuming that measurement at each depth is representative for the 100 mm above that depth:

$$S_{NS\_point} = SWC_{(100\ mm)} * 100 + SWC_{(200\ mm)} * 100 + SWC_{(300\ mm)} * 100\ mm + SWC_{(400\ mm)} * 100 \quad \text{(Eq 6)}$$

2.5 Observed catchment storage comparison with near-surface storage estimates

We calculated the daily value of catchment-scale storage (S_{WB}) using a simple water balance approach. We used daily P, PET and Q_{obs} data collected for the Elsick catchment:

$$S_{WB}(t) = S_{WB}(t-1) + P(t) - Q_{obs(t)} - AET(t) \quad \text{(Eq 7)}$$

Whereby AET= m*PET(t) (Eq 8)

For which m is assumed to rise linearly from 0 at Q_{obs}= 0 to 1 at a Q_{obs} threshold (Carrer et al., 2019). The Q_{obs} threshold was set to 0.24 mm day^{-1} for Elsick, which coincides with the 20th percentile of the Q_{obs} data. The m coefficient constrains AET for when (soil) water availability is limited. In order to obtain an estimate of catchment water storage independent of data which we used to obtain S_{NS}, we used Q_{obs} rather than an index based any of the soil water content data to constrain AET. The water balance approach considers catchment storage S_{WB} as a lumped metric aggregating multiple reservoirs (*e.g.* near-surface soil storage, groundwater etc.). Underlying assumptions are that there is no leakage from the catchment and errors in the estimates of each water balance component are accounted for individually beforehand. The method assumes that precipitation constitutes the only input to the

catchment, AET and Q_{obs} are the only outputs of the system and on a daily basis the difference is reflected in the total storage S_{WB} of the catchment.

To evaluate the correspondence between derived near-surface storages (S_{NS_CRNS} and S_{NS_point}) and estimated catchment storage S_{WB}, we used the Pearson *r* correlation coefficient for the period where daily estimates of all three storages (S_{WB} and S_{NS}) were available, which was a total of 472 days (limited by the availability of the point-scale data). We also assessed if the S-Q relationship of discharge with the S_{WB} and the S_{NS} indicated threshold behaviour for the same 472-day period.

2.6 Near-surface storage estimates in conceptual rainfall-runoff catchment modelling

2.6.1 Model choice and storage representation

We explored the usefulness of S_{NS} estimates in evaluating catchment-scale storage derived from rainfall-runoff modelling. More specifically, this involved determining whether including S_{NS} data in the calibration process could achieve acceptable stream discharge predictions while also more accurately representing catchment internal processes. For this, we chose the widely used HBV-light model (Lindström et al., 1997; Seibert & Vis, 2012) , as it is simple, versatile (Geris et al., 2015; Motavita et al., 2019) and commonly used in humid environments (e.g. Soulsby et al., 2011; Staudinger et al., 2017). HBV-light is a lumped conceptual model that simulates stream discharge Q_{sim} using a minimal input time series of precipitation (P), air temperature (T) and potential evapotranspiration (PET). The model comprises four main components: a snow routine (snow accumulation and melt), a soil routine (groundwater recharge and AET), a response routine (computes run-off as function of storage) and routing routine (triangular weighting function for routing run-off to catchment outlet) (Seibert & Vis, 2012) (Figure 2, Panel V). The model was set up for Elsick at daily time steps using P, T and PET input from 1 January 2011 to 30 September 2018, allowing for a relatively long warm-up period to eliminate the effects of initial conditions, especially on storage. Q_{obs} was available from 12 January 2015 to the end of the study period. Based on initial testing, the simplest plausible model structure characterises

storage as three reservoirs, representing soil zone dynamics (SM), and upper (SUZ) and lower stores (SLZ) which roughly represent shallow and deeper run-off generating stores.

The S_{NS_CRNS} or S_{NS_point} represent the total storage for a certain physical depth as opposed to HBV dynamic storage S_{dyn}, which estimates the storage involved in the runoff generation response and is not bound to a specific depth (as for most conceptual rainfall-runoff models). Therefore, direct comparison of SWC estimates (S_{NS_CRNS} or S_{NS_point}) with modelled storage dynamics S_{dyn} is not straightforward and could be done in multiple ways. Here, we compared the total soil column storage 0-400 mm to the sum of SM and SUZ storage (Figure 2). The rationale being that SM is responsible for the partitioning of precipitation input to deeper storage and ET but does not produce runoff. While SUZ is the upper box that above a certain threshold (UZL parameter) generates Q. This combined function is considered equivalent to the role of the topsoil, where the CRNS and point profile probes sense.

2.6.2 Model parameterization, calibration and evaluation

We used a Monte Carlo approach (10,000 independently generated parameter sets) to calibrate the model. Initial parameter ranges were set based on literature values (Seibert & Vis, 2012; Tetzlaff et al., 2015) and exploratory model runs (Table 2). We fixed the parameters of the snow routine, as snow contribution to precipitation in the catchment is minor. Model performance was evaluated using Kling-Gupta efficiency (KGE) (Gupta et al., 2009), being a more balanced goodness-of-fit measure as compared to e.g. Nash–Sutcliffe (Nash & Sutcliffe, 1970), which can lead to underestimation of high and intermediate flows, or volumetric error (Criss & Winston, 2008), which has a bias towards storage at expense of the Q (Mizukami et al., 2019).

To account for model uncertainty, the best 100 parameter sets were used, according to single or combined optimization criteria (see Table 2). These included classic model calibration on measured discharge Q_{obs},; S_{dyn} calibrated on S_{NS} estimates, either S_{NS_CRNS} or S_{NS_point}, and a combined calibration

of discharge and storage, $Combo_{CRNS}$ and $Combo_{Point}$, for which we retained the 100 runs that simultaneously had efficiency of KGE_Q>0.5 and KGE_{NS}>0.5. The calibration period was 4 April 2017 to 30 September 2018 (~ 1.5 years) for all calibration targets, since this is the period for which all S_{NS_Point} and S_{NS_CRNS} data were available. Despite its relatively short length, the calibration period spanned a range of hydrometeorological conditions including a one in 25-year return-period flood (in March 2018) and a prolonged recession during the summer of 2018. Days with missing S_{NS_Point} data were not included in the calibration period.

3. Results

3.1 CRNS-derived soil water content (SWC) dynamics within a mixed-agriculture footprint

The monitoring period spanned ~ 3 years and covered a range of hydrometeorological conditions (Figures 3,4). The observations started during exceptionally wet conditions, characterised by high precipitation inputs and floods in winter 2015-16. This was followed by a period of more seasonal wetting-drying cycles (March 2016 – June 2017). Finally, a relatively wet spell August 2017 – April 2018, including a rainfall on snowmelt event in March 2018, was followed by a prolonged recession in summer 2018 (Figure 3).

During the study period, the N_{pihv} counts (Figure 3B) were on average 1834 cph and ranging between 1606 and 2195 cph. The derived SWC_{CRNS} averaged 0.37 m^3 m^{-3}, ranging between 0.13 m^3 m^{-3} in summer 2018 and reaching porosity (n=0.6) during two extreme precipitation events in December 2015-January 2016 and March 2018, while field sampling captured SWC between 0.22 and 0.48 m^3 m^{-3} (Figure 3D and Table 1) and a range of antecedent moisture conditions with API_7 on sampling days ranging between 0.2 to 53 mm (Table 1).

Soil properties measured within the CRNS footprint remained fairly similar across sampling events (Table 1) and variability in time was attributed to the combined effect of spatial heterogeneity, crop growth and associated soil management activities (*e.g.* ploughing). The dry bulk density (ρ_{dry}) varied

between 1.01 g cm^{-3} and 1.21 while measured stone content remained relatively high (13-18%, with variability assumed to be a random artefact of sampling) (Table 1). The SOM and LW values were also fairly stable and variability in time was associated with different stages of crop growth and post-harvest plant debris or litter in the soil. The relatively wet conditions at the study site translated in a generally shallow CRNS effective sensing depth (z_{eff}), which varied between 8 and 20 cm (mean=11 cm) (Figure 3F).

Overall, Biological water content (BWE) was relatively low, although we did observe variations both in space (between fields) and time (between growing seasons). BWE characterisation was based on field sampling of two distinct growing seasons. The first in 2017 was fairly normal in terms of crop production and BWE fluctuated between 0.1 mm, when the cropped (Wheat-Gley and Barley-Gley) fields were bare, and 5.4 mm during the peak of the growing season. During the extremely dry 2018 growing season BWE was much lower (max of 2.5 mm). In general, higher vegetation cover was associated with higher uncertainty in BWE (Figure 3C). Soil management activities, such as harvesting and ploughing, did not have more influence on variations in the CRNS signal than hydrometeorological conditions (*i.e.* precipitation) (Figure 3B).

Uncertainty in the SWC_{CRNS} estimates originating from ρ_{dry} were on average 0.03 $m^3\ m^{-3}$. As expected, differences were minimal in the dry summer months (0.01 $m^3\ m^{-3}$), while these were most pronounced during wet winter periods, reaching up to 0.05 $m^3\ m^{-3}$. Average difference in calculated SWC_{CRNS} time series using vegetation correction and non-corrected time series, was 0.015 $m^3\ m^{-3}$ (median 0.01 $m^3\ m^{-3}$), ranging from zero in periods of fallow land to 0.06 $m^3\ m^{-3}$ when crop growth was at its peak.

For the top 400mm of the soil, the CRNS estimated SWC_{CRNS} (Figure 3D) was generally lower than the SWC_{point} measured at the point-profile (Figure 3E), even during the dry summer of 2018. Additionally, more pronounced peaks in SWC_{point} during rainfall events could have reflected the influence of local conditions at the location of the point-scale probe. At this site, the upper 0 to 200 mm experienced the greatest variability in SWC during spring-autumn 2017 as well as drying in summer 2018. The lower

portion of the soil profile, below the root-zone, was generally wetter with smoother variations, although at 300 mm depth SWC was consistently greater than the 400 mm one. SWC at 400 mm showed the least variability in time and appeared less affected by the prolonged drying in summer 2018.

Consequently, near-surface storage dynamics obtained from CRNS and point scale measurements (S_{NS_CRNS} and S_{NS_point} , respectively), were of comparable magnitude, although S_{NS_point} was generally higher (Figure 4C). Near-surface storage S_{NS_CRNS} at the mixed-land use footprint ranged between 58 and 218 mm (mean 148±33, median 153 mm). For the period of storage intercomparison, 4 April 2017 – 30 September 2018, S_{NS_CRNS} storage varied less (range 58-194 mm) and was lower on average (137±35, median 144 mm) compared to the point-profile S_{NS_Point} estimates. The latter were higher on average 162±37 mm (median 164) and in range 79 to 221 mm. The consistently greater soil wetness at the point-profile location was explained, on one hand by the topsoil being rich in organic matter i.e. better at retaining soil moisture, and on the other hand, by being less affected by soil evaporation owing to the undisturbed grass cover. This relative difference was also detected in the synoptic topsoil measurements among fields (Figure 3E).

The SWC of the top 0-60 mm within the footprint was consistently different between the individual land use-units, though it followed the general near-surface SWC dynamics (Figure 3E). The topsoil of the (compacted) Pasture-Podzol and the grass strip were consistently wetter than the cropped gleys, which are systematically drained (Wheat and Barley). The grass strip was generally wetter than the Pasture-Podzol and similar to the SWC_{point} at 100 mm depth, however the pattern was inverted after the dry summer 2018 period and consequent rewetting phase in August 2018. The subtle differences among the gley units were associated with relative differences in crop type and growth stages.

3.2 Near-surface storage estimates and their relationship with catchment storage dynamics

Overall, in this wet mixed-agricultural landscape, near-surface soil moisture S_{NS} data both at the point and field scale appeared useful for characterising catchment-scale storage dynamics. Generally, all storage (S_{WB} and S_{NS}) dynamics closely tracked precipitation patterns and mostly the discharge dynamics (Figure 4 A,C). However, slight differences in the amplitude of their response to precipitation input possibly reflect the scale-dependence of the hydrological processes governing the dynamics of that storage compartment (Figure 4C). Figure 4D reveals the importance of SWC_{CRNS}^{SUB} to S_{NS_CRNS}. While both types of S_{NS} were closely correlated with S_{WB}, S_{NS_CRNS} showed a stronger linear relationship (R^2=0.91) than S_{NS_point} (R^2=0.76) (Fig 5 A and B, respectively), and effects of hysteresis (mainly clock-wise hysteresis loops) were more evident in the S_{NS_point} relationship than for S_{NS_CRNS}.

The S-Q relationship both at the catchment and near-surface scale was roughly described by an exponential equation (Figure 5 C & D). The R^2 of these relationships were 0.68 for S_{WB}–Q and 0.71 and 0.67 for S_{NS_CRNS} – Q and S_{NS_point}–Q, respectively). These data suggest that there might be a non-linear increase in flow once a storage threshold is exceeded, e.g. at approximately 150 mm of S_{NS_CRNS} and around 120 mm for S_{NS_point}.

3.3. Rainfall runoff modelling with near-surface storage estimates

The results of the HBV simulations using single or combined model calibration targets are shown in Figure 6. Efficiency ranges of the 100 best simulations and final parameter distributions are summarised in Table 2 and Supplementary Figures S1–S5, respectively. In all cases, the model was able to reproduce the timing and peak responses of the hydrograph reasonably well, though with varying degrees of uncertainty for the recession periods, depending on the calibration target (Figure 6 A.1 – E.1). As expected, single variable calibrations (Q_{obs}, S_{NS_CRNS} and S_{NS_point}) gave the best efficiencies (KGE, Table 2) for the variable in question and the median of the best 100 runs followed closely the time series of observed data (e.g. Figure 6 A1 for Q_{obs}; B2 for S_{NS_point} and S_{NS_CRNS}). For example, the classic Q_{obs} calibration produced a best model fit for discharge with and average KGE_Q of

0.72, ranging between 0.69 to 0.80 (Table 2). However, for these simulations S_{dyn} was largely unconstrained so that internal storage dynamics were not well represented.

Overall, the near-surface soil storage appeared to be a suitable indicator of catchment-scale storage controlling runoff dynamics (S_{dyn}), especially during intermediate and wet conditions. That is evident both from the single S_{NS} (Figure 6 B.1 & C.1) as well as the combined model calibrations (Figure 6 D.1 & E.1). Model performance for either of the single S_{NS} calibrations was reasonably good, although the S_{NS_CRNS} (median 0.79) was slightly greater than S_{NS_point} (median 0.74) (Table 2). Nevertheless, during dry conditions the near-surface soil storage S_{NS} (either derived from point scale or CRNS based data) was not indicative of the runoff-generating storage S_{dyn} (Figure 6 B.1 & C.1).

A combination of good model performance and better representation of catchment internal storage processes was achieved when the combined criteria were used (Table 2; Figure 6). The discharge and storage calibrated model simulated wet and intermediate flow conditions similarly well, as compared to the calibration on Q_{obs} alone and additionally reduced uncertainty during low flows and the prolonged recession in summer 2018 (Figure 6, D1 and E1). In addition, a more balanced constraint of simulated storage S_{dyn} was obtained, which captured the time series of observed S_{NS} well (Figure 6 D.2 and E.2). That contrasted with single model calibration targets, where S_{dyn} was either unconstrained (Figure 6 A.2) or too narrow to bracket measured storage variation and simulate flows adequately (Figure 6 B.2 and C.2). Combining calibration targets resulted in a trade-off in terms of overall model performance (Table 2) for both Q_{obs} (median KGE Q_{sim} 0.63 in both Combos) and S_{dyn} (KGE $Combo_{point}$ was 0.68 and $Combo_{CRNS}$ was 0.71, respectively). At the same time, we obtained a wider range of the 100 best runs of KGE Q_{sim} and KGE S_{dyn} as opposed to the single calibration targets.

Sensitive and identifiable parameters changed depending on the calibration target. For the flow calibrated model (Q_{obs}) the UZL (Figure S1), which controls the outflow from the upper storage box was identifiable, and the parameters K_0 and BETA, responsible for runoff, were better-constrained (Table 2). Consistent with model structure, the parameters controlling the soil routines (FC and LP)

became more sensitive when the S_{NS} was the model calibration target (Table 2, Figure S2 and S3). In line with the parameters for the respective single calibrations, for $Combo_{Point}$ or $Combo_{CRNS}$ model calibrations, both the UZL and FC parameters became clearly identifiable and most constrained (Table 2, Figure S4 and S5).

As the S_{NS_CRNS} and S_{NS_point} dynamics were very similar, differences in final parameter ranges for the S_{NS} and combined model calibrations were overall subtle. However, for both model calibrations that involved point-scale data we obtained a higher median (213 mm) and wider final FC parameter ranges as compared to the calibrations based on CRNS data (Table 2), where median FC for S_{NS_CRNS} and the $Combo_{CRNS}$ were 176 and 178 mm, respectively. These subtle differences between models are likely linked to the relative differences in estimated S_{NS} using either CRNS or point-scale profile measurements. The slightly greater S_{NS_point} values (Figure 4C) measured over the profile probe resulted in greater median FC values. The FC remained as constrained for the Combos, which likely underlines the important role of near-surface storage in streamflow generation. On the contrary, the UZL parameter was constrained to its lower boundary during the classic (Q_{obs}) model calibration and generally remained so, although parameter ranges for the Combos were wider for $Combo_{point}$ and especially so for $Combo_{CRNS}$.

4. Discussion

4.1. What are the advantages and limitations of CRNS-based soil water content monitoring in mixed agricultural landscapes in humid environments?

In this study we estimated near-surface SWC using a CRNS sensor in a humid mixed agricultural landscape over an approximately 3-year period. The main benefit of deploying CRNS is scale representation and overcoming heterogeneity in point-scale SWC measurements, especially in mixed land use areas (Figure 4C). Calibrating the sensor and correcting the signal in Elsick was mostly straightforward, and based on prior CRNS research (Franz et al., 2013; Köhli et al., 2015; Schrön, et al.,

2017; Zreda et al., 2008). However, site and climate specific challenges related to the generally wet soils, humid climate and spatial heterogeneities of soil and land use properties.

In humid environments, such as Scotland, CRNS technology used as a stand-alone method for near-surface SWC estimate faces issues like high neutron count uncertainty and relatively shallow (here 140 mm on average) sensing depth (Evans et al., 2016). To tackle these, we conducted field calibration at multiple days (Iwema et al., 2015), covering a range of hydroclimatic conditions, and used a relatively long 24 hours temporal aggregation of the CRNS signal (Martin Schrön et al., 2018). Independent, in-situ SWC measurements (Nguyen et al., 2019) were not available at high temporal resolution to obtain SWC information at greater depths (i.e. > z_{eff}). Instead, in-depth representativeness of the CRNS data was achieved by using an exponential filter approach (Peterson et al., 2016).

Information on the soil moisture profile is key to constraining CRNS signal uncertainty (Baroni et al., 2018). In fact, many intensely monitored research sites (*e.g.* Bogena et al., 2018) usually rely on a network of point-profile soil moisture measurements to complement the CRNS. Nevertheless, at an intensely managed agricultural catchment, such as that of Elsick, regular agricultural management (e.g. harvesting, cattle grazing) makes installing and maintaining such a network impractical (Vather et al., 2019). For comparative reasons only, the CRNS station in Elsick was equipped with a single profile probe, at a convenient location on undisturbed soil.

Compared to traditional distributed point scale measurements (Brunetti et al., 2019), CRNS sensors require little maintenance and are likely to be less influenced by the typical sources of uncertainty in hydraulic properties induced by soil management (*e.g.* ploughing). Moreover, in Elsick, we aimed at better representativeness of the CRNS data for the surrounding landscape, by placing the sensor between fields with rotational crops and grazed pasture, and two distinct soil types. As suggested by Evans et al. (2016), CRNS research in the UK would benefit from monitoring more sites representative for the soil types and land uses in the surrounding (~5km^2 area). The soil-vegetation units within the CRNS footprint represent approximately 85% of the Elsick catchment. However, it needs to be

recognised that the footprint of the CRNS is still considerably smaller than the total catchment area. Using a roving-type CRNS approach (e.g. Franz, 2018; Schrön et al., 2018; Vather et al., 2019), could be one way to fully explore this scale representativeness. Nevertheless, despite the remaining spatial difference between the total catchment area and the CRNS footprint, we found that the CRNS signal detected the dynamics that are relevant for improving model calibration at the catchment scale. Hence, using a more "realistically" heterogeneous footprint, as done in this study, potentially increases the value of the obtained SWC estimate for large scale modelling applications in catchments or for satellite validation. Therefore, the CRNS could offer a better and more flexible monitoring technique of spatially representative near-surface SWC in dynamically changing agricultural environments, despite some aspects requiring specialised effort (e.g. site-specific calibration, snow and vegetation correction), which are to be considered.

4.2. To what extent is near surface storage coupled to catchment-scale storage and discharge dynamics?

Soil moisture plays a critical role in modulating streamflow and hydrological connectivity between landscape units and the channel network (Tromp-van Meerveld & McDonnell, 2006). This is especially the case in humid environments, like Elsick, with relatively wet and temporarily waterlogged soils. At this site, comparison between simple catchment and near-surface storage data showed that these are closely related (Figure 5A and 5B). The better relationship between the S_{NS_CRNS} to S_{WB} as compared to S_{NS_point} was attributed to the overall better scale representativeness of the CRNS data i.e. overcoming spatial heterogeneities (Brunetti et al., 2019).

Additionally, all S-Q relationships in the catchment were described by an exponential equation, although hysteresis loops were more pronounced for S_{NS_point} (Figure 4 C, D). This difference can again be attributed to the CRNS data representing a more integrated signal than point scale measurements. Therefore, in the context of Elsick, the S_{NS_CRNS} seems to represent the soil storage dynamics better since it reflects the more dynamic wetting-drying trends of the managed soils in the catchment (Figure

4B). This is opposed to the S_{NS_point} data which would be more representative for the unmanaged soil-land use units or for soil moisture dynamics in the narrow riparian (often saturated) areas (Zuecco et al., 2018).

While characterising catchment scale storage and its relationship with discharge has been the focus of many empirical and modelling studies (e.g. Brauer et al., 2013; McNamara et al., 2011), fully characterising catchment scale storage remains a challenge. The latter relates to physical observations often being divergent from the conceptually defined rainfall-runoff model storage (Staudinger et al., 2017). In that sense, monitoring focused on key storage compartments largely connected to the streamflow generation (Spence et al., 2010), like the near-surface storage as was the focus here, could potentially benefit modelling efforts.

Despite small differences in transformation approaches, many studies have utilized soil moisture data of varying soil column depth for catchment storage modelling and comparison (e.g. Brauer et al., 2013; Nguyen et al., 2019). Here, we defined near-surface storage as the total water storage in the soil column of 400 mm depth, using CRNS and point-scale SWC estimates. For a different HBV modelling application, Seibert et al. (2011) transformed TDR soil moisture data of the first 300 mm and assumed it represented half of the total catchment storage. Other studies combined soil moisture and groundwater observations at the point (Carrer et al., 2019) or large (*i.e.* remote sensing) (Demirel et al., 2019; Nijzink et al., 2018) scale to characterise catchment-scale storage. Overall, our approach is complementary to these others as some sort of SWC data transformation specific to the study site remains a requirement for catchment storage characterisation.

4.3. How can shallow soil moisture information be used to help better characterize catchment storage in rainfall-runoff models and what are the implications for streamflow predictions?

In this study, a likely better internal process representation was achieved by including S_{NS_CRNS} in the model calibration (Figures 4,6). Calibration to near-surface storage estimates using point scale

measurements gave similar rainfall-runoff modelling results as CRNS derived ones in terms of fit, although with different parameter ranges so that internal processes were represented in a slightly different way. Additionally, median simulated discharge was often overestimated during dry periods using S_{NS} model calibration (Figure 6). However, the uncertainty bands of the 100 best runs using S_{NS_point} did not explain sufficiently the variability in Q_{obs}. Inherent to the conceptual rainfall-runoff model internal functioning, model uncertainty was always relatively high during recessions and consequent rewetting. However, the uncertainty in Q_{sim} in those periods was reduced when a Combo calibration was applied (Figure 6). Therefore, model simulations are more likely providing answers for the right reasons when including CRNS information than when using point scale measurements, and both than when only discharge data are used. However, the effect of including SWC estimates in model calibration in terms of improvement of the Q predictions and better representation of catchment internal processes can be relative and depends on several factors such as model structure (Zhuo & Han, 2016), scale representativeness of the data (Peters-Lidard et al., 2017) and the degree of connectivity of the monitored store to streamflow (Ali et al., 2014; Spence, 2010). Using SWC data for rainfall-runoff model calibration may lead to little (Brocca et al., 2010) or significant improvements in model structure identification and parameterization (Kuppel et al., 2018). The relative success mainly relates to the spatial mismatch and representativeness of the data, with limited SWC availability and a trade-off between spatial and temporal resolution, as well as the shallow monitoring depth. As also shown in this study, the issue of the relatively shallow CRNS sensing depth can be addressed by simple in-depth scaling approaches (Nguyen et al., 2019; Peterson et al., 2016). Moreover, the spatial heterogeneities characteristic of point scale sensors (as shown in Figure 3E) are largely overcome by the CRNS technology. In addition, because the hysteresis pattern in the storage-discharge relationship was less prominent when using CRNS data, this would particularly make CRNS-derived near-surface storage more amenable for use within conceptual rainfall-runoff models.

In the context of a mixed-agricultural humid environment, this study shows that CRNS data can be especially useful in wet and intermediate conditions (e.g. flood modelling applications) when near-

surface storage plays a dominant role in runoff generation. Additionally, landscape-average estimates of SWC_{CRNS} can be extremely valuable for agricultural management applications that reply on soil moisture status (Franz et al., 2016; Stevanato et al., 2019). These and other applications, such as flood warning systems, would therefore benefit significantly from the better internal representation of SWC that CRNS data can bring to rainfall-runoff models. However, during dry conditions, as the contribution of near-surface storage to streamflow diminishes (i.e. below the threshold in the storage-discharge relationship), CRNS data may become increasingly less informative for characterising catchment storage. A study using the HBV model by Demirel et al. (2019) drew similar general conclusions when using spatially averaged remotely-sensed soil moisture and groundwater observations for the Moselle basin (165,000 km^2). In agreement with our findings, when also exploring a range of different calibration objectives, they found that storage-only calibration failed to reproduce low flows and FC and LP model parameters only became sensitive when soil moisture information was used as a model calibration target. Nevertheless, for temperate climate catchments Orth et al., (2015) were able to consistently validate the HBV model successfully using SWC information. As such, the degree to which CRNS data can help to calibrate or constrain storage in models might be temporally and spatially variable, whereby uncertainty increases with drier conditions. Nevertheless, a quantitative evaluation of how this uncertainty varies should also consider that the effective sensing depth increases with dry periods.

4.4. Future opportunities for CRNS modelling approaches

The use of CRNS is increasing worldwide (Andreasen et al., 2017) covering an array of climatic conditions, soil types and land covers (Evans et al., 2016; Hawdon et al., 2014; Zreda et al., 2012). In heterogeneous managed environments, which represent a growing proportion of the Earth's surface, collected CRNS data can help address important questions of socioeconomic and environmental relevance (*e.g.* flood-warning systems, agricultural models). Moreover, while the SWC information of point-scale sensors is mostly related to soil texture characteristics, the CRNS signal captures additional

information on the influence of microtopography and vegetation cover (Jakobi et al., 2018) on soil moisture dynamics. This additional information value of CRNS may become important in modelling applications (Iwema et al., 2017), especially in mixed land use catchments like Elsick, where multiple vegetation covers are observed. Although using SWC_{CRNS} in conceptual rainfall-runoff modelling might not seem straightforward and requires adapting to local conditions (*e.g.* deeper soils; accounting for the effect of denser vegetation or important snow input on the CRNS signal), its use to help characterise catchment storage looks promising. Since storage conceptualization may vary between models, cross comparison with rainfall-runoff models with slightly different internal dynamics to assess the uncertainties from model selection is likely to be informative (Demirel et al., 2019).

The outcomes of this study show that CRNS data can be especially useful for soil moisture monitoring in flood warning modelling applications or at least to identify antecedent states when catchments may be vulnerable to the effects of high rainfall events. Moreover, while CRNS sensing depth can be shallow in wet environments such as the UK, it still provides information at a better resolution and at greater depths than many satellite products (typically only up to a few cm) and presents a valuable data source for validating these (Duygu & Akyürek, 2019; Kędzior & Zawadzki, 2016). Future research in mixed-land use environments could usefully focus on combining large scale remote sensing data (e.g., SMAP or Sentinel), which offer larger spatial coverage but with lower temporal frequency and CRNS data to cross compare modelling results. Despite the various advantages of the CRNS technology and its potential in rainfall-runoff modelling, a combination with traditional networks of point-scale measurements at greater depths is still encouraged (Baroni et al., 2018).

5. Conclusions

We deployed a cosmic ray neutron sensor (CRNS) probe to estimate near-surface soil storage dynamics in a mixed-agricultural landscape in a humid environment. While the method presents various challenges in wet climates, including high neutron count uncertainty and shallow sensing depth, its key benefit is that it largely overcomes the effect of spatial heterogeneities inherent in point-

scale sensors for obtaining landscape-average soil water content estimates. Moreover, in intensely managed agricultural catchments, where maintenance of profile sensors becomes logistically not feasible, CRNS may represent a better option for long-term high-resolution monitoring. Furthermore, we used CRNS along with point-scale soil moisture dynamics to explore the coherence between near-surface (S_{NS}) and catchment-scale (S_{WB}) storage using a water balance approach. Significant correlation of the either S_{NS} to the S_{WB} and discharge indicated close coupling with streamflow generation processes. CRNS storage estimates related better to water balance storage estimates and discharge dynamics and showed lesser effect of hysteresis, as compared to point-scale estimates, which is consistent with the better spatial representativeness of the CRNS-sensed S_{NS} storage. The close link between S_{NS} and S_{WB} provided a rationale for exploring the effect of using S_{NS} data to inform a conceptual rainfall runoff model. For that we calibrated HBV-light with either discharge or S_{NS} data or a combination of them. In the context of humid environments, CRNS data proved useful for model calibration, especially during wet and intermediate conditions as the soil storage is often connected to runoff generation. Better internal processes representation in the model was achieved when jointly calibrated using discharge and near-surface storage data. These findings imply the high potential of CRNS data, for example in flood-forecasting applications in humid mixed-land use environments. Dynamic land management and spatial heterogeneity create a challenge to catchment storage characterisation using traditional point scale soil moisture probes. In this sense, worldwide proliferation of CRNS monitoring and related research present exciting opportunities for the development of methodological frameworks that can inform modelling applications in heavily managed environments.

Acknowledgements

We thank the Macaulay Development Trust and School of Geosciences, University of Aberdeen for KDPs scholarship. JG would like to acknowledge funding from the Royal Society and the Carnegie Trust for the Universities of Scotland (project 70112). JG and LV acknowledge funding from the UK Natural

Environment Research Council (project NE/N007611/1 and CC13_080). MW was supported by the Rural & Environment Science & Analytical Services Division of the Scottish Government. RR received funding from the Natural Environment Research Council (projects NE/M003086/1, NE/R004897/1 and NE/T005645/1) and from the International Atomic Energy Agency of the United Nations (IAEA/UN) (project CRP D12014). Special thanks to Carol Taylor, Jessica Fennell, Alice Poli and many more for assistance with fieldwork. Finally, we would like to acknowledge Kenneth Loades for providing us with essential equipment for soil sampling and thank David Finlay and his team for enabling land access in Elsick.

Appendix I: Definitions of storage and related terms.

Abbreviation	Units	Definition	Reference
CRNS	[-]	Cosmic ray neutron sensor	Hydroinnova, New Mexico
N (pihv)	[cph]	Cosmic ray neutron intensity measured as neutron counts per hour [cph], inversely correlated to all hydrogen present in the upper decimetres of the subsurface and the first few hectometres of the atmosphere above the ground surface. The N signal is corrected for effects of atmospheric pressure (p), incoming neutron flux (i), air humidity (h) and optionally for the effect of the water contained in aboveground vegetation biomass (v).	(Baatz et al., 2015; Zreda et al., 2012; Zreda et al., 2008)
SWC_{CRNS}	[$m^3\ m^{-3}$]	Field average (~ 14 ha) soil water content based on calibrated Cosmic Ray Neutron Sensor data; integrated over a time-variable sensing depth z_{eff} (between 0.08 and 0.02 m)	(Schrön et al., 2017; Zreda et al., 2008)
SWC_{CRNS}^{sub}	[$m^3\ m^{-3}$]	Field average soil water content for depth 400 mm m-z_{eff}, derived from SWC_{CRNS} using an exponential filter.	This study, adapted from (Peterson et al., 2016)
SWC_{point}	[$m^3\ m^{-3}$]	Point scale soil water content based on Delta-T PR2 Sensor data; can be at different depths (100, 200, 300 and 400 mm)	(Delta T Devices Ltd.)
SWI	[-]	Soil water index. Ranges from 0 to 1, being the minimum and the maximum the SWC_{CRNS} at the site during the study period.	(Wagner et al., 1999)
S_{WB}	[mm]	Storage component based on water balance (P-Q-AET)	(Carrer et al., 2019; Pfister et al., 2017)
S_{NS}	[mm]	Near-surface soil storage for a defined depth (z=0.4 m)	This study
S_{NS_CRNS}	[mm]	Near-surface storage for a defined depth (z=0.4 m) determined as the sum of SWC_{CRNS} and SWC_{CRNS}^{sub}	This study
S_{NS_point}	[mm]	Near-surface storage for a defined depth (z=0.4 m) determined using point scale data (SWC_{point}) across different depths	This study
S_{dyn}	[mm]	Dynamic Storage (catchment scale), considered to control the majority of streamflow response. In the selected model structure	This study

		set-up, it is the sum of the storage in the SM (soil moisture) and the SUZ (upper groundwater zone) box in the HBV-light model.	
SM	[mm]	HBV model: Soil moisture box with its largest value equal to FC (field capacity). Partitioning of rainfall in soil water content and groundwater recharge. Does not produce runoff	(Seibert, 2005)
SUZ	[mm]	HBV model: Upper groundwater box, recharged by the SM box. Faster runoff (Q_0) of the SUZ box depends on the UZL (upper zone limit) parameter which acts as a threshold above which runoff is produced. Slower runoff Q_1 from this box depends on K1 recession constant.	(Seibert, 2005)
SLZ	[mm]	Lower groundwater box (PERC in mm day^{-1} defines the max percolation rate from the upper to the lower groundwater box)	(Seibert, 2005)

References

AHDB Cereals & Oilseeds (2018a). Barley growth guide. Kenilworth. UK https://projectblue.blob.core.windows.net/media/Default/Imported%20Publication%20Docs/Barley%20growth%20guide%20130718.pdf

AHDB Cereals & Oilseeds (2018b). Wheat growth guide. Kenilworth. UK https://projectblue.blob.core.windows.net/media/Default/Imported%20Publication%20Docs/Wheat%20growth%20guide.pdf

Albergel, C., Rüdiger, C., Pellarin, T., Calvet, J. C., Fritz, N., Froissard, F., … Martin, E. (2008). From near-surface to root-zone soil moisture using an exponential filter: An assessment of the method based on in-situ observations and model simulations. Hydrology and Earth System Sciences, 12(6), 1323–1337. https://doi.org/10.5194/hess-12-1323-2008

Andreasen, M., Jensen, K. H., Desilets, D., Franz, T. E., Zreda, M., Bogena, H. R., & Looms, M. C. (2017). Status and Perspectives on the Cosmic-Ray Neutron Method for Soil Moisture Estimation and Other Environmental Science Applications. Vadose Zone Journal, 16(8), 0. https://doi.org/10.2136/vzj2017.04.0086

Baatz, R., Bogena, H. R., Hendricks Franssen, H. J., Huisman, J. A., Montzka, C., & Vereecken, H. (2015). An empirical vegetation correction for soil water content quantification using cosmic ray probes. Water Resources Research, 51, 2030–2046. https://doi.org/10.1002/2014WR015432.Received

Baatz, R., Bogena, H. R., Hendricks Franssen, H. J., Huisman, J. A., Qu, W., Montzka, C., & Vereecken, H. (2014). Calibration of a catchment scale cosmic-ray probe network: A comparison of three parameterization methods. Journal of Hydrology, 516, 231–244. https://doi.org/10.1016/j.jhydrol.2014.02.026

Baatz, R., Franssen, H. J. H., Han, X., Hoar, T., Reemt Bogena, H., & Vereecken, H. (2017). Evaluation of a cosmic-ray neutron sensor network for improved land surface model prediction. Hydrology and Earth System Sciences, 21(5), 2509–2530. https://doi.org/10.5194/hess-21-2509-2017

Baroni, G., Schei, L. M., Schrön, M., Ingwersen, J., & Oswald, S. E. (2018). Uncertainty , sensitivity and improvements in soil moisture estimation with cosmic-ray neutron sensing, 564(July 2017), 873–887. https://doi.org/10.1016/j.jhydrol.2018.07.053

Beven, K. (2016). Advice to a young hydrologist. Hydrological Processes, 30(20). https://doi.org/10.1002/hyp.10879

Birkel, C., Soulsby, C., & Tetzlaff, D. (2015). Conceptual modelling to assess how the interplay of hydrological connectivity, catchment storage and tracer dynamics controls nonstationary water age estimates. Hydrological Processes, 29(13), 2956–2969. https://doi.org/10.1002/hyp.10414

Blöschl, G., Bierkens, M. F. P., Chambel, A., Cudennec, C., Destouni, G., Fiori, A., … Zhang, Y. (2019). Twenty-three unsolved problems in hydrology (UPH) – a community perspective. Hydrological Sciences Journal, 6667, 1–18. https://doi.org/10.1080/02626667.2019.1620507

Bogena, H.R., Montzka, C., Huisman, J. A., Graf, A., Schmidt, M., Stockinger, M., … Vereecken, H. (2018). The TERENO-Rur Hydrological Observatory: A Multiscale Multi-Compartment Research Platform for the Advancement of Hydrological Science. Vadose Zone Journal, 17(1), 0. https://doi.org/10.2136/vzj2018.03.0055

Bogena, H. R, Huisman, J. A., Güntner, A., Kusche, J., Jonard, F., & Vey, S. (2015). Emerging methods for noninvasive sensing of soil moisture dynamics from field to catchment scale : a review, 2(December), 635–647. https://doi.org/https://doi.org/10.1002/wat2.1097

Brauer, C. C., Teuling, A. J., Torfs, P. J. J. F., & Uijlenhoet, R. (2013). Investigating storage-discharge relations in a lowland catchment using hydrograph fitting, recession analysis, and soil moisture data. Water Resources Research, 49(7), 4257–4264. https://doi.org/10.1002/wrcr.20320

Brocca, L., Melone, F., Moramarco, T., Wagner, W., Naeimi, V., Bartalis, Z., & Hasenauer, S. (2010). Improving runoff prediction through the assimilation of the ASCAT soil moisture product. Hydrology and Earth System Sciences, 14(10), 1881–1893. https://doi.org/10.5194/hess-14-1881-2010

Brunetti, G., Šimůnek, J., Bogena, H., Baatz, R., Huisman, J. A., Dahlke, H., & Vereecken, H. (2019). On the Information Content of Cosmic-Ray Neutron Data in the Inverse Estimation of Soil Hydraulic Properties. Vadose Zone Journal, 18(1). https://doi.org/10.2136/vzj2018.06.0123

Carrer, G. E., Klaus, J., & Pfister, L. (2019). Assessing the Catchment Storage Function Through a Dual-Storage Concept. Water Resources Research, 55(1), 476–494. https://doi.org/10.1029/2018WR022856

Coopersmith, E. J., Cosh, M. H., & Daughtry, C. S. T. (2014). Field-scale moisture estimates using COSMOS sensors: A validation study with temporary networks and Leaf-Area-Indices. Journal of Hydrology, 519(PA). https://doi.org/10.1016/j.jhydrol.2014.07.060

Creutzfeldt, B., Troch, P. A., Güntner, A., Ferré, T. P. A., Graeff, T., & Merz, B. (2014). Storage-discharge relationships at different catchment scales based on local high-precision gravimetry. Hydrological Processes, 28(3), 1465–1475. https://doi.org/10.1002/hyp.9689

Criss, R. E., & Winston, W. E. (2008). Do Nash values have value? Discussion and alternate proposals. Hydrological Processes, 22(14), 2723–2725. https://doi.org/10.1002/hyp.7072

Dehaspe, J., Birkel, C., Tetzlaff, D., Sánchez-Murillo, R., Durán-Quesada, A. M., & Soulsby, C. (2018). Spatially distributed tracer-aided modelling to explore water and isotope transport, storage and mixing in a pristine, humid tropical catchment. Hydrological Processes, 32(21), 3206–3224. https://doi.org/10.1002/hyp.13258

Demirel, M. C., Özen, A., Orta, S., Toker, E., Demir, H.K. Ekmekcioğlu, O., … Booij, M.J. (2019). Additional Value of Using Satellite-Based Soil Moisture and Two Sources of Groundwater Data for Hydrological Model Calibration. Water, 11(10), 2083. https://doi.org/10.3390/w11102083

Desilets, D., Zreda, M., & Ferré, T. P. A. (2010). Nature's neutron probe: Land surface hydrology at an

elusive scale with cosmic rays. Water Resources Research, 46(11), 1–7. https://doi.org/10.1029/2009WR008726

Dimitrova-Petrova, K., Geris, J., Wilkinson, M. E., Lilly, A., & Soulsby, C. (2020). Using isotopes to understand the evolution of water ages in disturbed mixed land-use catchments. Hydrological Processes, hyp.13627. https://doi.org/https://doi.org/10.1002/hyp.13627

Duygu, M. B., & Akyürek, Z. (2019). Using Cosmic-Ray Neutron Probes in Validating Satellite Soil Moisture Products and Land Surface Models. Water, 11(7), 1362. https://doi.org/10.3390/w11071362

Evans, J. G., Ward, H. C., Blake, J. R., Hewitt, E. J., Morrison, R., Fry, M., … Jenkins, A. (2016). Soil water content in southern England derived from a cosmic-ray soil moisture observing system - COSMOS-UK. Hydrological Processes, 4999(August), 4987–4999. https://doi.org/10.1002/hyp.10929

Foolad, F., Franz, T. E., Wang, T., Gibson, J., Kilic, A., Allen, R. G., & Suyker, A. (2017). Feasibility analysis of using inverse modeling for estimating field-scale evapotranspiration in maize and soybean fields from soil water content monitoring networks, 1263–1277. https://doi.org/10.5194/hess-21-1263-2017

Franz, T. E. (2018). Soil moisture Mapping with a POrtable Cosmic Ray Neutron Sensor IAEA-TECDOC-1845 Soil.

Franz, T. E., Zreda, M., Rosolem, R., & Ferre, T. P. A. (2013). A universal calibration function for determination of soil moisture with cosmic-ray neutrons. Hydrology and Earth System Sciences, 17(2), 453–460. https://doi.org/10.5194/hess-17-453-2013

Franz, T. E., Wahbi, A., Vreugdenhil, M., Weltin, G., Heng, L., Oismueller, M., … Desilets, D. (2016). Using cosmic-ray neutron probes to monitor landscape scale soil water content in mixed land use agricultural systems. Applied and Environmental Soil Science, 2016. https://doi.org/10.1155/2016/4323742

Franz, T. E., Zreda, M., Ferre, T. P. A., Rosolem, R., Zweck, C., Stillman, S., … Shuttleworth, W. J. (2012). Measurement depth of the cosmic ray soil moisture probe affected by hydrogen from various sources. Water Resources Research, 48(8), 1–9. https://doi.org/10.1029/2012WR011871

Geris, J., Tetzlaff, D., Seibert, J., Vis, M., & Soulsby, C. (2015). Conceptual Modelling to Assess Hydrological Impacts and Evaluate Environmental Flow Scenarios in Montane River Systems Regulated for Hydropower. River Research and Applications, 31(9), 1066–1081. https://doi.org/10.1002/rra.2813

Gupta, H. V., Kling, H., Yilmaz, K. K., & Martinez, G. F. (2009). Decomposition of the mean squared error and NSE performance criteria: Implications for improving hydrological modelling. Journal of Hydrology, 377(1–2), 80–91. https://doi.org/10.1016/j.jhydrol.2009.08.003

Han, X., Jin, R., Li, X., & Wang, S. (2014). Soil moisture estimation using cosmic-ray soil moisture sensing at heterogeneous farmland. IEEE Geoscience and Remote Sensing Letters, 11(9), 1659–1663. https://doi.org/10.1109/LGRS.2014.2314535

Hawdon, A., McJannet, D., & Wallace, J. (2014). Calibration and correction procedures for cosmic-ray neutron soil moisture probes located across Australia. Water Resources Research, 50(6), 5029–5043. https://doi.org/10.1002/2013WR015138

Heidbuchel, I., Guntner, A., & Blume, T. (2016). Use of cosmic-ray neutron sensors for soil moisture monitoring in forests. Hydrology and Earth System Sciences, 20(3), 1269–1288. https://doi.org/10.5194/hess-20-1269-2016

Hooke, J. M. (1979) An analysis of the processes of river bank erosion, Journal of Hydrology, 42, 39–62.

Iwema, J., Rosolem, R., Baatz, R., Wagener, T., & Bogena, H. R. (2015). Investigating temporal field sampling strategies for site-specific calibration of three soil moisture-neutron intensity parameterisation methods. Hydrology and Earth System Sciences, 19(7), 3203–3216. https://doi.org/10.5194/hess-19-3203-2015

Iwema, J., Rosolem, R., Rahman, M., Blyth, E., & Wagener, T. (2017). Land surface model performance using cosmic-ray and point-scale soil moisture measurements for calibration. Hydrology and Earth System Sciences, 21(6), 2843–2861. https://doi.org/10.5194/hess-21-2843-2017

Jakobi, J., Huisman, J. A., Vereecken, H., Diekkrüger, B., & Bogena, H. R. (2018). Cosmic Ray Neutron Sensing for Simultaneous Soil Water Content and Biomass Quantification in Drought Conditions. Water Resources Research, 54(10), 7383–7402. https://doi.org/10.1029/2018WR022692

Kędzior, M., & Zawadzki, J. (2016). Comparative study of soil moisture estimations from SMOS satellite mission, GLDAS database, and cosmic-ray neutrons measurements at COSMOS station in Eastern Poland. Geoderma, 283, 21–31. https://doi.org/10.1016/j.geoderma.2016.07.023

Köhli, M., Schrön, M., Zreda, M., Schmidt, U., Dietrich, P., & Zacharias, S. (2015). Footprint characteristics revised for field-scale soil moisture monitoring with cosmic-ray neutrons. Water Resources Research, 51(7), 5772–5790. https://doi.org/10.1002/2015WR017169

Kormos, P. R., McNamara, J. P., Seyfried, M. S., Marshall, H. P., Marks, D., & Flores, A. N. (2015). Bedrock infiltration estimates from a catchment water storage-based modeling approach in the rain snow transition zone. Journal of Hydrology, 525, 231–248. https://doi.org/10.1016/j.jhydrol.2015.03.032

Kuppel, S., Tetzlaff, D., Maneta, M. P., & Soulsby, C. (2018). What can we learn from multi-data calibration of a process-based ecohydrological model? Environmental Modelling and Software, 101, 301–316. https://doi.org/10.1016/j.envsoft.2018.01.001

Lievens, H., Tomer, S. K., Al Bitar, A., De Lannoy, G. J. M., Drusch, M., Dumedah, G., … Pauwels, V. R. N. (2015). SMOS soil moisture assimilation for improved hydrologic simulation in the Murray Darling Basin, Australia. Remote Sensing of Environment, 168, 146–162. https://doi.org/10.1016/j.rse.2015.06.025

Lindström, G., Johansson, B., Persson, M., Gardelin, M., & Bergström, S. (1997). Development and test of the distributed HBV-96 hydrological model. Journal of Hydrology, 201(1–4), 272–288. https://doi.org/10.1016/S0022-1694(97)00041-3

Loizu, J., Massari, C., Álvarez-Mozos, J., Tarpanelli, A., Brocca, L., & Casalí, J. (2018). On the assimilation set-up of ASCAT soil moisture data for improving streamflow catchment simulation. Advances in Water Resources, 111(December 2016), 86–104. https://doi.org/10.1016/j.advwatres.2017.10.034

McNamara, J. P., Tetzlaff, D., Bishop, K., Soulsby, C., Seyfried, M., Peters, N. E., … Hooper, R. (2011). Storage as a Metric of Catchment Comparison. Hydrological Processes, 25(21), 3364–3371.

https://doi.org/10.1002/hyp.8113

Met Office. (2019a): MIDAS: UK Daily Rainfall Data. NCAS British Atmo- spheric Data Centre. 855 http://catalogue.ceda.ac.uk/uuid/c732716511 d3442f05cdeccbe99b8f90 856

Met Office. (2019b): MIDAS: UK Daily Temperature Data. NCAS British Atmospheric Data Centre. 857 http://catalogue.ceda.ac.uk/uuid/ 1bb479d3b1e38c339adb9c82c15579d8

Motavita, D. F., Chow, R., Guthke, A., & Nowak, W. (2019). The comprehensive differential split-sample test: A stress-test for hydrological model robustness under climate variability. Journal of Hydrology, 573(March), 501–515. https://doi.org/10.1016/j.jhydrol.2019.03.054

Nash, J., & Sutcliffe, J. (1970). River flow forecasting through conceptual models Part I - A discussion of principles. Journal of Hydrology, 10, 282–290. Retrieved from https://www.sciencedirect.com/science/article/pii/0022169470902556

Nguyen, H. H., Jeong, J., & Choi, M. (2019). Extension of cosmic-ray neutron probe measurement depth for improving field scale root-zone soil moisture estimation by coupling with representative in-situ sensors. Journal of Hydrology, 571, 679–696. https://doi.org/10.1016/j.jhydrol.2019.02.018

Nguyen, H. H., Kim, H., & Choi, M. (2017). Evaluation of the soil water content using cosmic-ray neutron probe in a heterogeneous monsoon climate-dominated region. Advances in Water Resources, 108, 125–138. https://doi.org/10.1016/j.advwatres.2017.07.020

Nijzink, R. C., Almeida, S., Pechlivanidis, I. G., Capell, R., Gustafssons, D., Arheimer, B., … Hrachowitz, M. (2018). Constraining Conceptual Hydrological Models With Multiple Information Sources. Water Resources Research, 54(10), 8332–8362. https://doi.org/10.1029/2017WR021895

Orth, R., Staudinger, M., Seneviratne, S. I., Seibert, J., & Zappa, M. (2015). Does model performance improve with complexity? A case study with three hydrological models. Journal of Hydrology, 523, 147–159. https://doi.org/10.1016/j.jhydrol.2015.01.044

Peterson, A. M., Helgason, W. D., & Ireson, A. M. (2016). Estimating field-scale root zone soil moisture using the cosmic-ray neutron probe. Hydrology and Earth System Sciences, 20(4), 1373–1385. https://doi.org/10.5194/hess-20-1373-2016

Pfister, L., Martínez-Carreras, N., Hissler, C., Klaus, J., Carrer, G. E., Stewart, M. K., & McDonnell, J. J. (2017). Bedrock geology controls on catchment storage, mixing, and release: A comparative analysis of 16 nested catchments. Hydrological Processes, 31(10), 1828–1845. https://doi.org/10.1002/hyp.11134

Rivera Villarreyes, C. A., Baroni, G., & Oswald, S. E. (2014). Inverse modelling of cosmic-ray soil moisture for field-scale soil hydraulic parameters. European Journal of Soil Science. https://doi.org/10.1111/ejss.12162

Rosolem, R., Shuttleworth, W. J., Zreda, M., Franz, T. E., Zeng, X., & Kurc, S. A. (2013). The Effect of Atmospheric Water Vapor on Neutron Count in the Cosmic-Ray Soil Moisture Observing System. Journal of Hydrometeorology, 14(5), 1659–1671. https://doi.org/10.1175/JHM-D-12-0120.1

Schrön, M., Rosolem, R., Köhli, M., Piussi, L., Schröter, I., Iwema, J., … Zacharias, S. (2018). Cosmic-ray Neutron Rover Surveys of Field Soil Moisture and the Influence of Roads. Water Resources Research, 54(9), 6441–6459. https://doi.org/10.1029/2017WR021719

Schrön, M., Köhli, M. O., Schrön, M., Köhli, M., Scheiffele, L., Iwema, J., … Lv, L. (2017). Improving calibration and validation of cosmic-ray neutron sensors in the light of spatial sensitivity Improving Calibration and Validation of Cosmic-Ray Neutron Sensors in the Light of Spatial Sensitivity – Theory and Evidence, Hydrology and Earth System Sciences Discussions. https://doi.org/10.5194/hess-2017-148

Schrön, M., Zacharias, S., Womack, G., Köhli, M., Desilets, D., Oswald, S. E., … Dietrich, P. (2018). Intercomparison of cosmic-ray neutron sensors and water balance monitoring in an urban environment. Geoscientific Instrumentation, Methods and Data Systems, 7(1), 83–99. https://doi.org/10.5194/gi-7-83-2018

Seibert, J., & Vis, M. J. P. (2012). Teaching hydrological modeling with a user-friendly catchment-runoff-model software package. Hydrology and Earth System Sciences, 16(9), 3315–3325. https://doi.org/10.5194/hess-16-3315-2012

Seibert, J., Bishop, K., Nyberg, L., & Rodhe, A. (2011). Water storage in a till catchment. I: Distributed modelling and relationship to runoff. Hydrological Processes, 25(25), 3937–3949. https://doi.org/10.1002/hyp.8309

Seyfried, M. S., Grant, L. E., Marks, D., Winstral, A., & McNamara, J. (2009). Simulated soil water storage effects on streamflow generation in a mountainous snowmelt environment, Idaho, USA. Hydrological Processes, 23(6), 858–873. https://doi.org/10.1002/hyp.7211

Soulsby, C., Piegat, K., Seibert, J., & Tetzlaff, D. (2011). Catchment-scale estimates of flow path partitioning and water storage based on transit time and runoff modelling. Hydrological Processes, 25(25), 3960–3976. https://doi.org/10.1002/hyp.8324

Spence, C. (2007). On the relation between dynamic storage and runoff: A discussion on thresholds, efficiency, and function. Water Resources Research. https://doi.org/10.1029/2006WR005645

Spence, C., Guan, X. J., Phillips, R., Hedstrom, N., Granger, R., & Reid, B. (2010). Storage dynamics and streamflow in a catchment with a variable contributing area. Hydrological Processes. https://doi.org/10.1002/hyp.7492

Staudinger, M., Stoelzle, M., Seeger, S., Seibert, J., Weiler, M., & Stahl, K. (2017). Catchment water storage variation with elevation. Hydrological Processes, 31(11), 2000–2015. https://doi.org/10.1002/hyp.11158

Stevanato, L., Baroni, G., Cohen, Y., Lino, F. C., Gatto, S., Lunardon, M., … Morselli, L. (2019). A novel cosmic-ray neutron sensor for soil moisture estimation over large areas. Agriculture (Switzerland), 9(9). https://doi.org/10.3390/agriculture9090202

Stroud, P. D. (1999). A Recursive Exponential Filter For Time-Sensitive Data, 1–8.

Tetzlaff, D., Buttle, J., Carey, S. K., van Huijgevoort, M. H. J., Laudon, H., Mcnamara, J. P., … Soulsby, C. (2015). A preliminary assessment of water partitioning and ecohydrological coupling in northern headwaters using stable isotopes and conceptual runoff models. Hydrological Processes, 29(25), 5153–5173. https://doi.org/10.1002/hyp.10515

Tian, Z., Li, Z., Liu, G., Li, B., & Ren, T. (2016). Soil water content determination with cosmic-ray neutron sensor: Correcting aboveground hydrogen effects with thermal/fast neutron ratio. Journal of Hydrology, 540, 923–933. https://doi.org/10.1016/j.jhydrol.2016.07.004

Tromp-van Meerveld, H. J., & McDonnell, J. J. (2006). Threshold relations in subsurface stormflow: 2. The fill and spill hypothesis. Water Resources Research, 42(2), 1–11.

https://doi.org/10.1029/2004WR003800

Vather, T., Everson, C., & Franz, T. E. (2019). Calibration and Validation of the Cosmic Ray Neutron Rover for Soil Water Mapping within Two South African Land Classes. Hydrology, 6(3), 65. https://doi.org/10.3390/hydrology6030065

Wagner, W., Lemoine, G., & Rott, H. (1999). A method for estimating soil moisture from ERS Scatterometer and soil data. Remote Sensing of Environment, 70(2), 191–207. https://doi.org/10.1016/S0034-4257(99)00036-X

Zhang, L., Brutsaert, W., Crosbie, R., & Potter, N. (2014). Long-term annual groundwater storage trends in Australian catchments. Advances in Water Resources, 74, 156–165. https://doi.org/10.1016/j.advwatres.2014.09.001

Zhu, X., Shao, M., Zeng, C., Jia, X., Huang, L., Zhang, Y., & Zhu, J. (2016). Application of cosmic-ray neutron sensing to monitor soil water content in an alpine meadow ecosystem on the northern Tibetan Plateau. Journal of Hydrology, 536. https://doi.org/10.1016/j.jhydrol.2016.02.038

Zreda, M., Shuttleworth, W. J., Zeng, X., Zweck, C., Desilets, D., Franz, T., & Rosolem, R. (2012). COSMOS: The cosmic-ray soil moisture observing system. Hydrology and Earth System Sciences, 16(11), 4079–4099. https://doi.org/10.5194/hess-16-4079-2012

Zreda, M., Desilets, D., Ferré, T. P. A., & Scott, R. L. (2008). Measuring soil moisture content non-invasively at intermediate spatial scale using cosmic-ray neutrons. Geophysical Research Letters, 35(21), 1–5. https://doi.org/10.1029/2008GL035655

Zuecco, G., Penna, D., & Borga, M. (2018). Runoff generation in mountain catchments: long-term hydrological monitoring in the Rio Vauz Catchment, Italy. Cuadernos de Investigación Geográfica, 44(2), 397. https://doi.org/10.18172/cig.3327

Table 1: Summary of the outcomes of the calibration sampling campaigns, including corrected N counts (N_{pihv}); depth-distance and areal weighted averages of volumetric SWC from field samples; sensing depth z_{eff}, bulk density; Stone content, measured as the particles of >2mm and dry bulk density was corrected for it; biomass water equivalent as well as hydrometeorological conditions for each sampling date. No rainfall fell in the duration of the sampling itself (i.e. SWC stable).

		Sampling 1	**Sampling 2**	**Sampling 3**	**Sampling 4**	**Sampling 5**
	Units/Date	18/04/2017	27/07/2017	29/11/2017	25/07/2018	26/09/2018
N_{pihv}	[counts h^{-1}]	1893	1763	1737	2099	1952
Field SWC	[m^3 m^{-3}]	0.36	0.44	0.48	0.24	0.22
ρ_{dry}	[g cm^{-3}]	1.13	1.21	1.10	1.01	1.03
Stone content	[%] ± stdev	13±6	18±9	15±6	18±5	14±3
Vol SOM+WL	[m^3 m^{-3}]	0.09	0.08	0.11	0.05	0.08
Z_{eff}	[cm]	12	10	8	16	15
BWE	[mm]	1.7 ±0.5 (1)	4.3± 1.3	0.1 ±0.1	1.7 ±0.8	0.1 ±0.1
Precip	[mm day^{-1}]	0	3	4.6	0	0
API_7	[mm]	0.2	53	24.8	10.6	11
Long term Precip (2)	[mm $month^{-1}$]	58.8	61.6	93.3	61.6	67.9
Q_{obs}	[mm day^{-1}]	0.16	1.7	3.2	0.06	0.21
PET	[mm day^{-1}]	1.8	2.6	0.2	4.3	1.7
Long term PET (3)	[mm $month^{-1}$]	28	40	26	40	37

(1) Estimates based on field observations, crop camera and crop growth cycle estimate, but not actual measurement.

(2) Long term averages (1981-2010) retrieved from Dyce Airport Station, Aberdeen (<25 km from Elsick), MetOffice, 2019a.

(3) Long term PET averages (1981-2010) using Thortwaite and long-term minimum and maximum temperature record from from Dyce Airport Station, Aberdeen (<25 km from Elsick), MetOffice, 2019b.

Table 2. Left hand side: initial parameter ranges for the HBV model. Parameters Pcorr=1, TT=0, CFMAX=1, CET=1. CFR=0.05, CWH=0.1 were fixed. Right hand side: final parameter ranges (median [min max] reported) for the 100 best runs using three single (Q_{obs}, S_{NS_point}, S_{NS_CRNS}) and two combined ($Combo_{Point}$ *and* $Combo_{CRNS}$*) calibration targets. In bold the parameters that are either sensitive or constrained in at least one of the model calibrations. KGE used as goodness of fit measure. KGE range of the 100 best runs is shown for individual calibration targets and KGE for both Q_{sim} and S_{dyn} are reported in the combined calibrations. Final parameter distribution graphically shown in Figures S1-S4 (supplementary material).*

Initial parameter ranges			Final parameter ranges Median [Min Max] for each calibration target				
Parameter	Lower limit	Upper limit	Q_{obs}	S_{NS_point}	S_{NS_CRNS}	$Combo_{point}$	$Combo_{CRNS}$
BETA	1	6	***4.40 [1.25 – 5.98]***	***3.08 [1.15-6.00]***	***5.12 [3.41-5.96]***	***4.01[1.18-5.98]***	***3.85 [1.25-5.98]***
FC	10	250	***73 [16-243]***	***213 [191-249]***	***176 [161-190]***	***213 [155-249]***	***178 [127-243]***
K0	0.1	0.5	0.43 [0.19-0.50]	0.29 [0.10-0.50]	0.27 [0.11-0.49]	0.42 [0.12-0.50]	0.41 [0.11-0.50]
K1	0.05	0.2	0.14 [0.05-0.20]	0.11 [0.05-0.20]	0.13 [0.05-0.20]	0.14 [0.05-0.20]	0.14 [0.05-0.2]
K2	0.001	0.01	0.01 [0.001-0.01]	0.01 [0.001-0.01]	0.01 [0.001-0.01]	0.01 [0.002-0.01]	0.01 [0.001-0.01]
LP	0.3	1	0.68 [0.31-1.00]	0.70 [0.57-0.93]	0.61 [0.43-0.70]	0.72 [0.31-0.98]	0.72 [0.31-0.98]
MAXBAS	1	2.5	1.39 [1.00-2.50]	1.68 [1.01-2.46]	1.69 [1.00-2.50]	1.55 [1.01-2.49]	1.55 [1.01-2.50]
PERC	0	4	0 [0-3]	1 [0-4]	2 [0-4]	0 [0-2]	0 [0-2]
UZL	0	70	***4 [0-22]***	***40 [2-69]***	***33 [1-70]***	***5 [0-27]***	***5 [0-57]***
		KGE Q_{sim}	0.72 [0.69-0.80]	-0.25 [-0.99 – -0.13]	-0.33 [-0.64 – -0.02]	0.63 [0.50-0.80]	0.63 [0.50-0.80]
		KGE S_{dyn}	For S_{NS_point} 0.19 [-0.24 – 0.77] For S_{NS_CRNS} 0.26 [-0.2 – 0.56]	0.74 [0.74-0.76]	0.79[0.78-0.81]	0.68 [0.55-0.74]	0.71 [0.52-0.79]

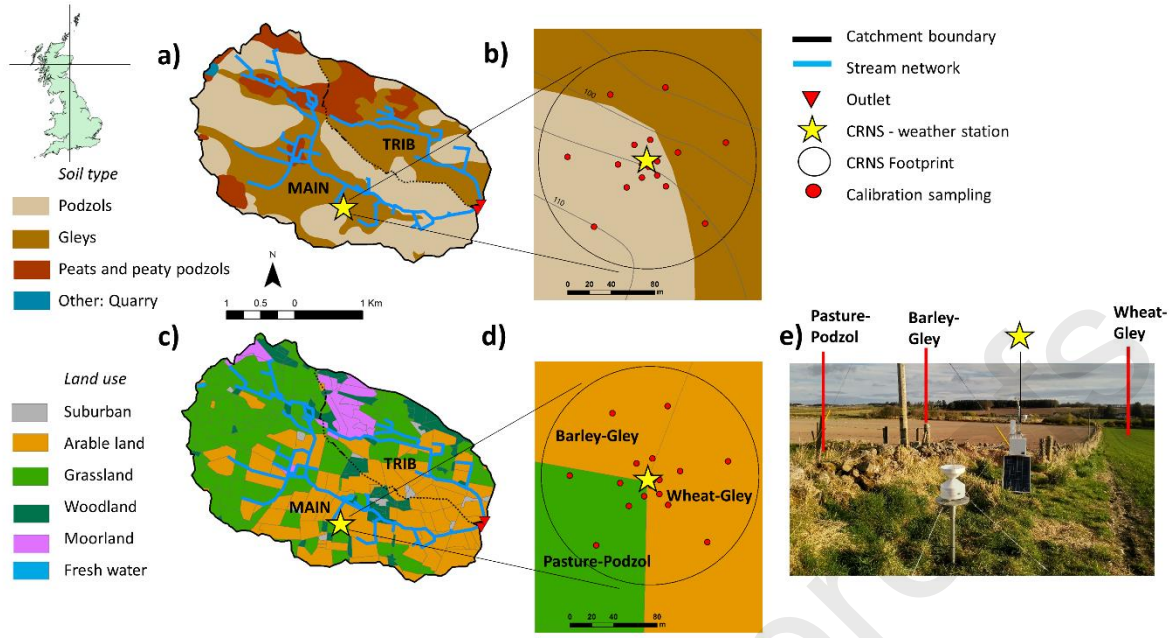
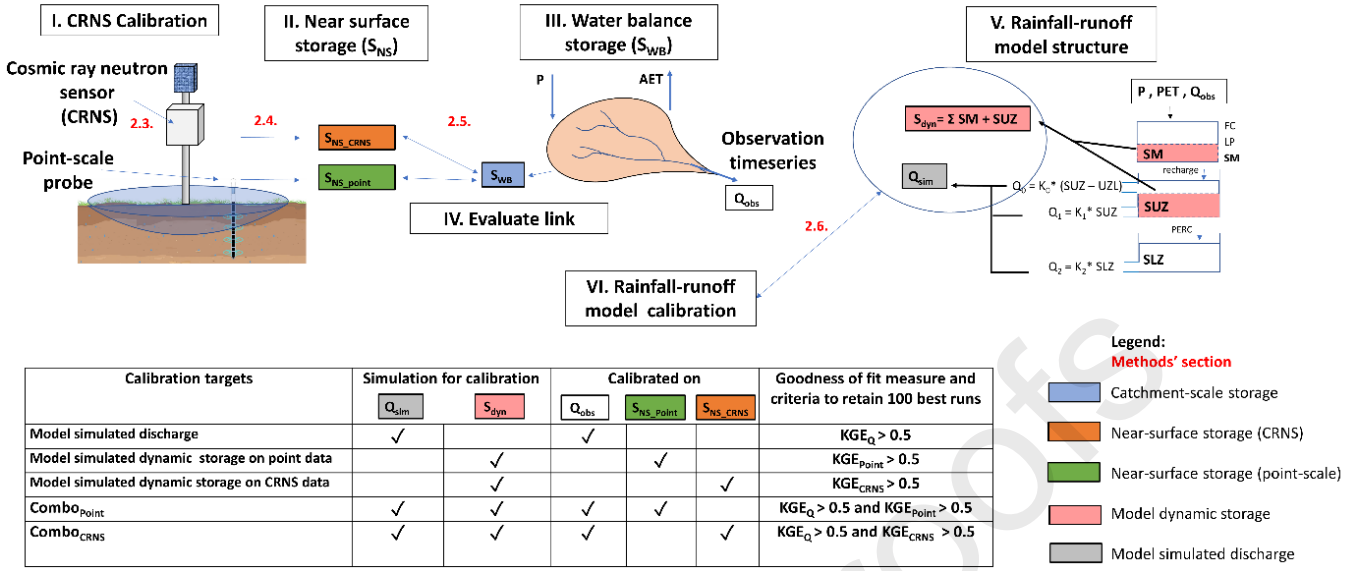


Figure 1: The Elsick catchment study site and instrumentation, showing (a) the soil type distribution and locations of the CRNS-weather station and gauging stations (b) zoom of the CRNS footprint covering two soil types and sampling locations (red dots); (c) Land use classes and (d) location of the CRNS between three different vegetation covers (rotating crops and pasture). (e) field setting showing the CRNS-weather station and rain gauge. Map data 1:25 000 Soil Map, The James Hutton Institute/ Land Cover Map 2015, NERC Environmental Information Data Centre.



Calibration targets	Simulation for calibration		Calibrated on			Goodness of fit measure and criteria to retain 100 best runs
	Q_{sim}	S_{dyn}	Q_{obs}	S_{NS_Point}	S_{NS_CRNS}	
Model simulated discharge	✓		✓			$KGE_Q > 0.5$
Model simulated dynamic storage on point data		✓		✓		$KGE_{Point} > 0.5$
Model simulated dynamic storage on CRNS data		✓			✓	$KGE_{CRNS} > 0.5$
$Combo_{Point}$	✓	✓	✓	✓		$KGE_Q > 0.5$ and $KGE_{Point} > 0.5$
$Combo_{CRNS}$	✓	✓	✓		✓	$KGE_Q > 0.5$ and $KGE_{CRNS} > 0.5$

Figure 2. Step by step procedure followed in using near-surface soil storage data (S_{NS}) to evaluate catchment-scale dynamic storage (S_{dyn}) in HBV-light rainfall-runoff model. Workflow is shown using blue arrows and red numbering corresponding to methods section describing it. Model flow is depicted by black arrows. Rainfall-runoff model outcomes and different storages used for comparison are shown in distinct colours (see legend).

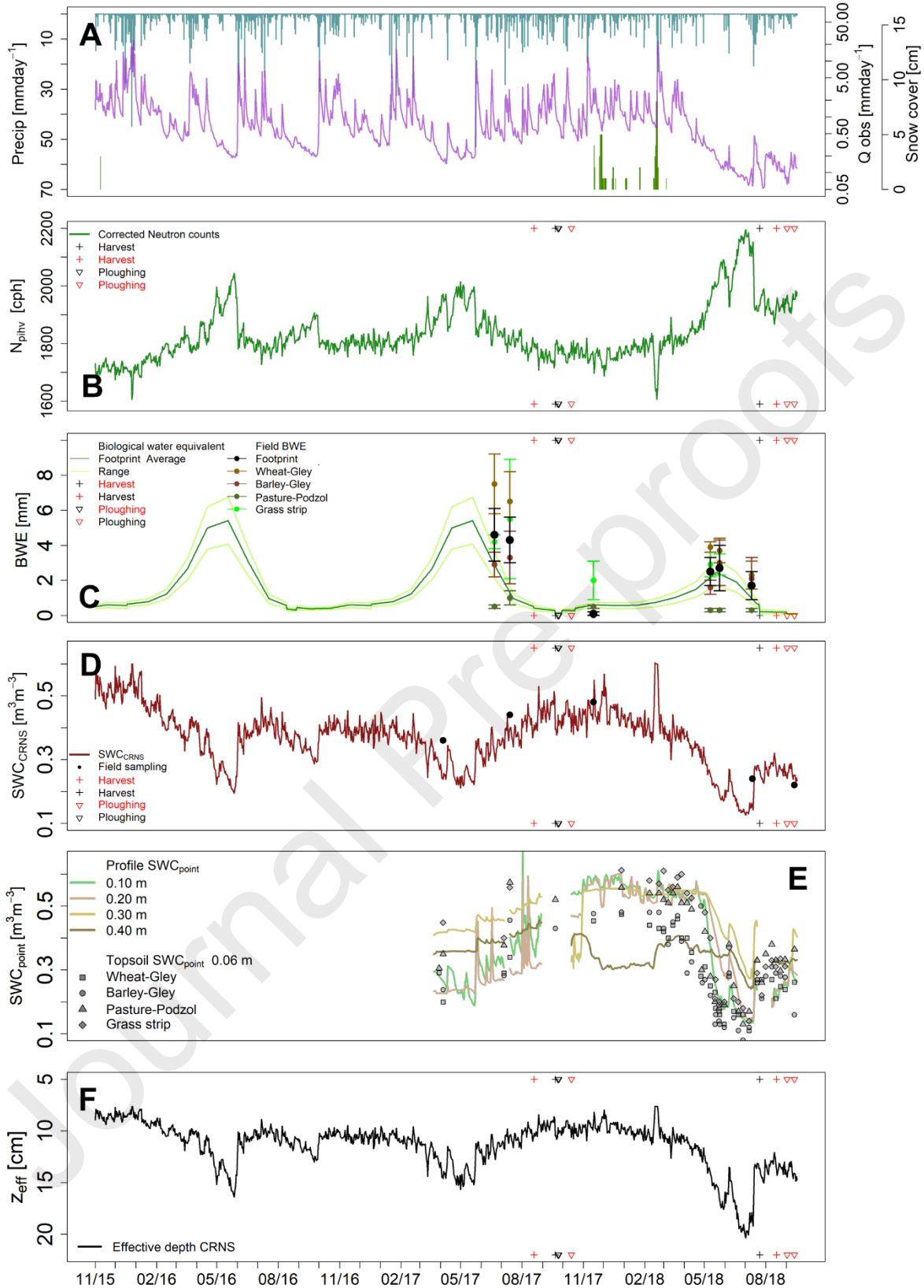


Figure 3. (A) Temporal dynamics of daily precipitation, stream discharge [mm day^{-1}] and snow depth [cm] during the period 14 November 2015 – 30 September 2018. (B) Daily average N_{pihv} counts, scaled

to hourly values [cph] corrected for influence of atmospheric pressure, humidity and incoming radiation as well as vegetation; (C) BWE [mm] measured at each field on sampling dates, error bars = standard deviation, continuous time series represents estimated BWE scenario at the footprint. (D) Estimated SWC_{CRNS} [m^3 m^{-3}] and field depth-distance and areal weighed SWC on the five sampling days; (E) Point scale SWC data available for the site. Lines represent time series of SWC_{point} at different depths at and dots represent arithmetic average SWC for each field and the grass patch; (F) Effective sensing depth of the CRNS sensor z_{eff} [cm]. Where available, information on dates of agricultural management activities, ploughing (as a plus) and harvesting (inverted triangle) indicated in black for the Wheat-Gley and in red for the Barley- Gley.

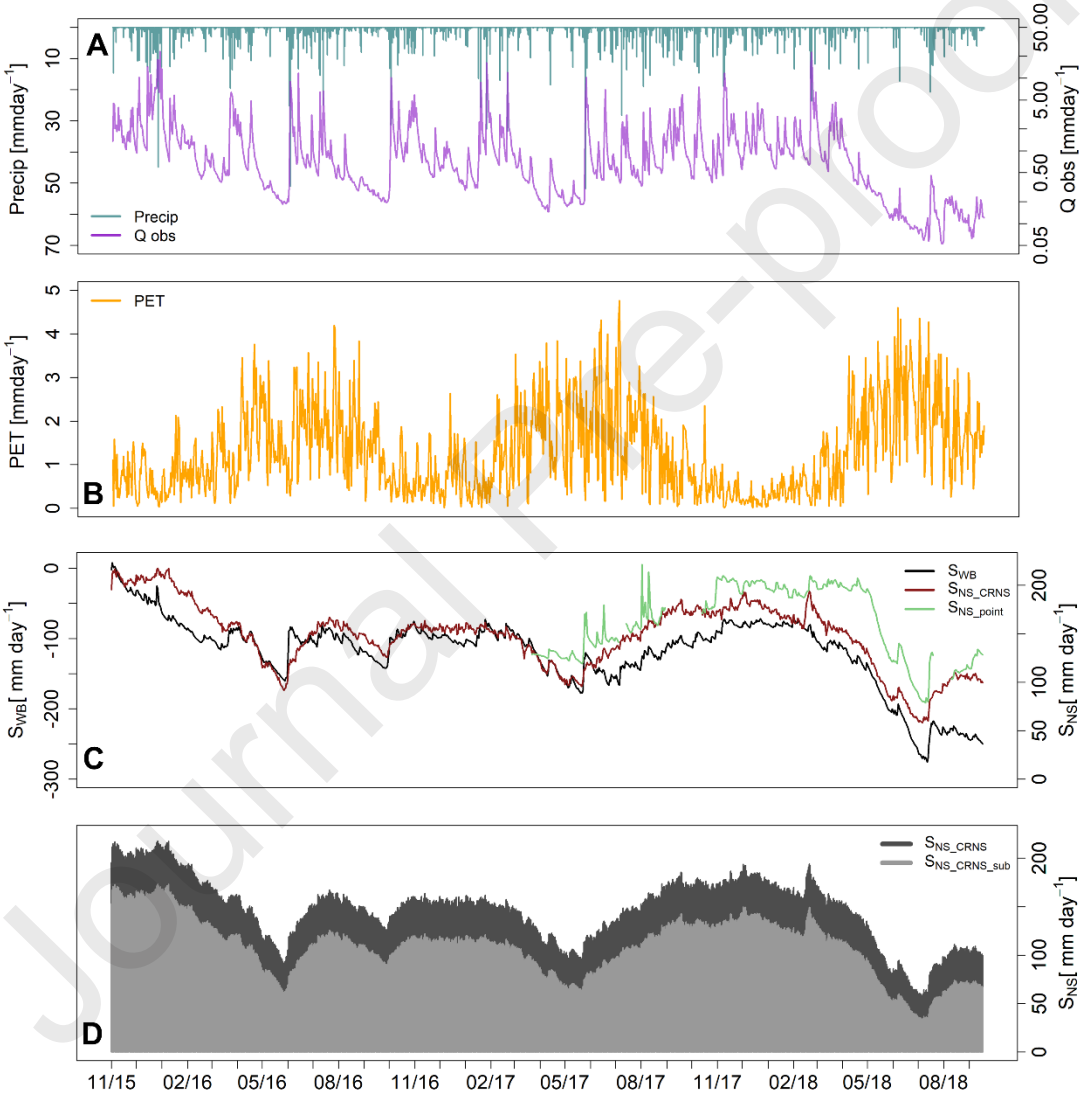


Figure 4. (A) Daily time series of precipitation and discharge [mm day^{-1}] (B) and PET [mm day^{-1}] for the period 14 November 2015 to 30 September 2018 (C) Time series of catchment storage (S_{wb}) dynamics [mm day^{-1}] and near-surface storage estimates based on point-scale sensor estimates (S_{NS_point}). (D) the

CRNS-based storage estimate (S_{NS_CRNS}), additionally showing the component of S_{NS_CRNS} estimated using exponential filter ($S_{NS\_CRNS\_sub}$).

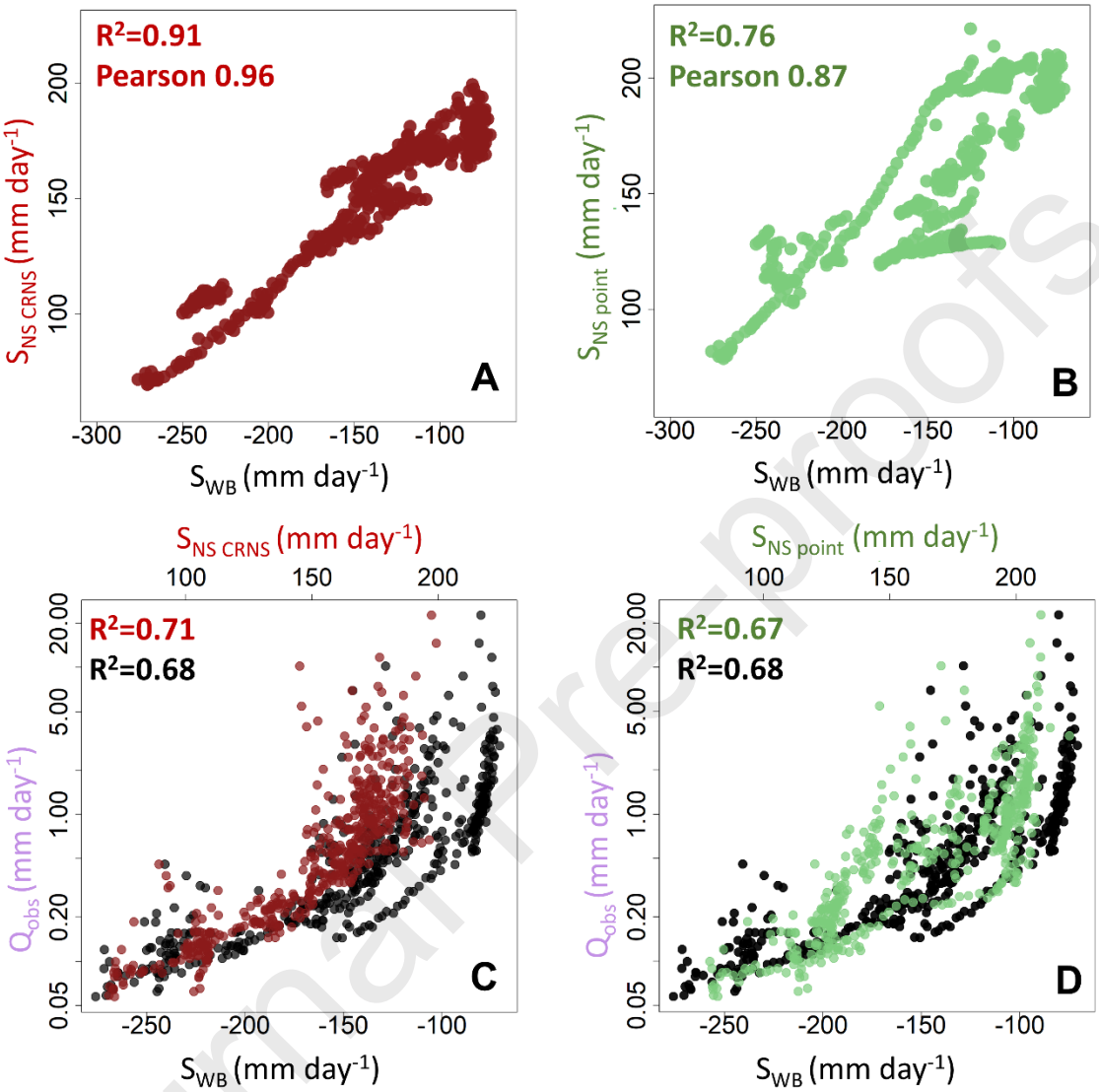


Figure 5: Relationship between catchment scale S_{wb} and near-surface storage estimates (A) S_{NS_CRNS} as well as (B) S_{NS_point}; (C) Relationship between discharge and different storage components: S_{wb} (black), S_{NS_CRNS} (dark red) and (D) S_{NS_point} (green) . Relationships (linear in panel A and B; exponential in C and D) were derived for days where information on the three storages was available (471 days).

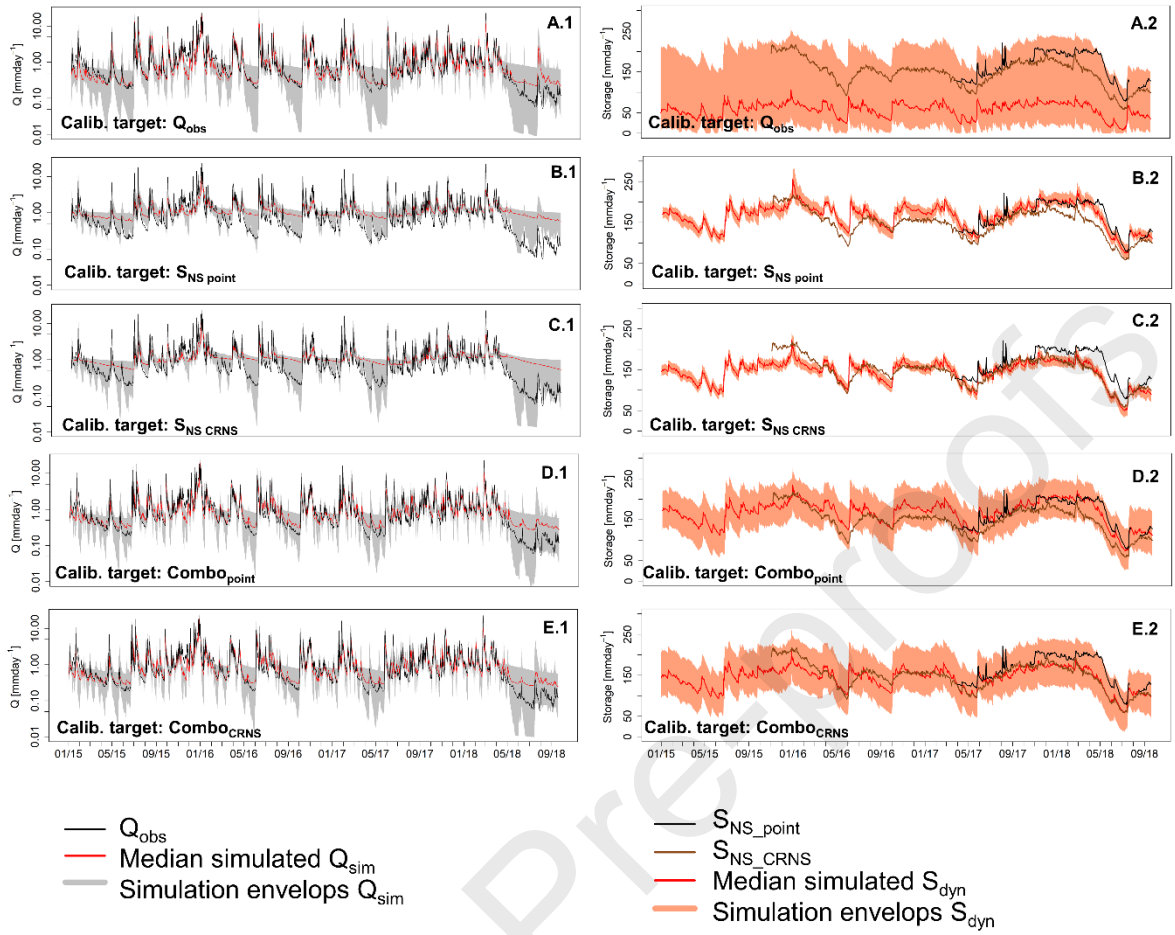


Figure 6. Left panels (A1 to E1) show the time series of observed Q, median and uncertainty bands of Q_{sim} of the best 100 runs using the five calibration criteria and right panels (A2 to E2) show observed S_{NS_CRNS} and S_{NS_point} with median simulated and uncertainty bands of the modelled dynamic storage of the best 100 runs using the following calibration targets: (A) Q_{obs}, (B) S_{NS_point}, (C) S_{NS_CRNS} , (D) Combo S_{NS_point}, (E) Combo S_{NS_CRNS}). KGE used as a goodness of fit measure.

Supplementary Material

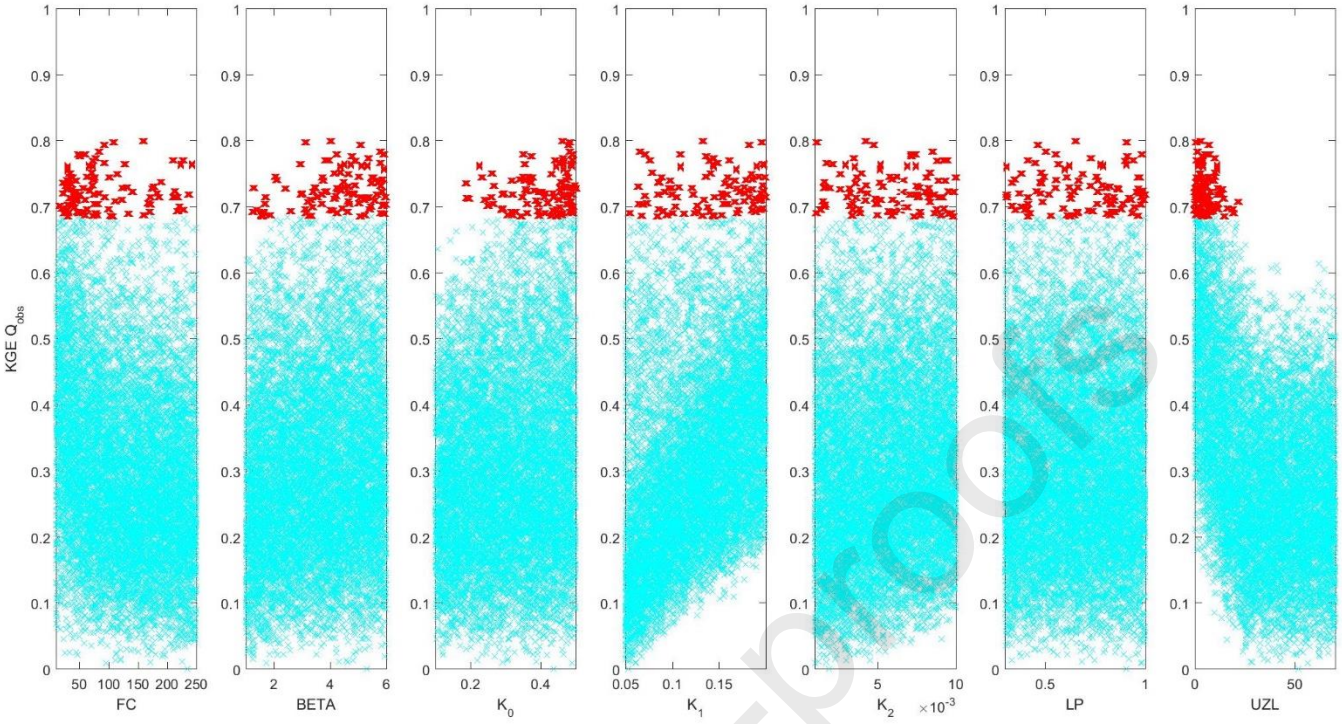


Figure S1: Calibration target Q_{obs}: Final parameter range of selected parameters (either sensitive or identifiable in some of the calibrations) of best 100 runs (in red), KGE used as the Goodness of fit criteria.

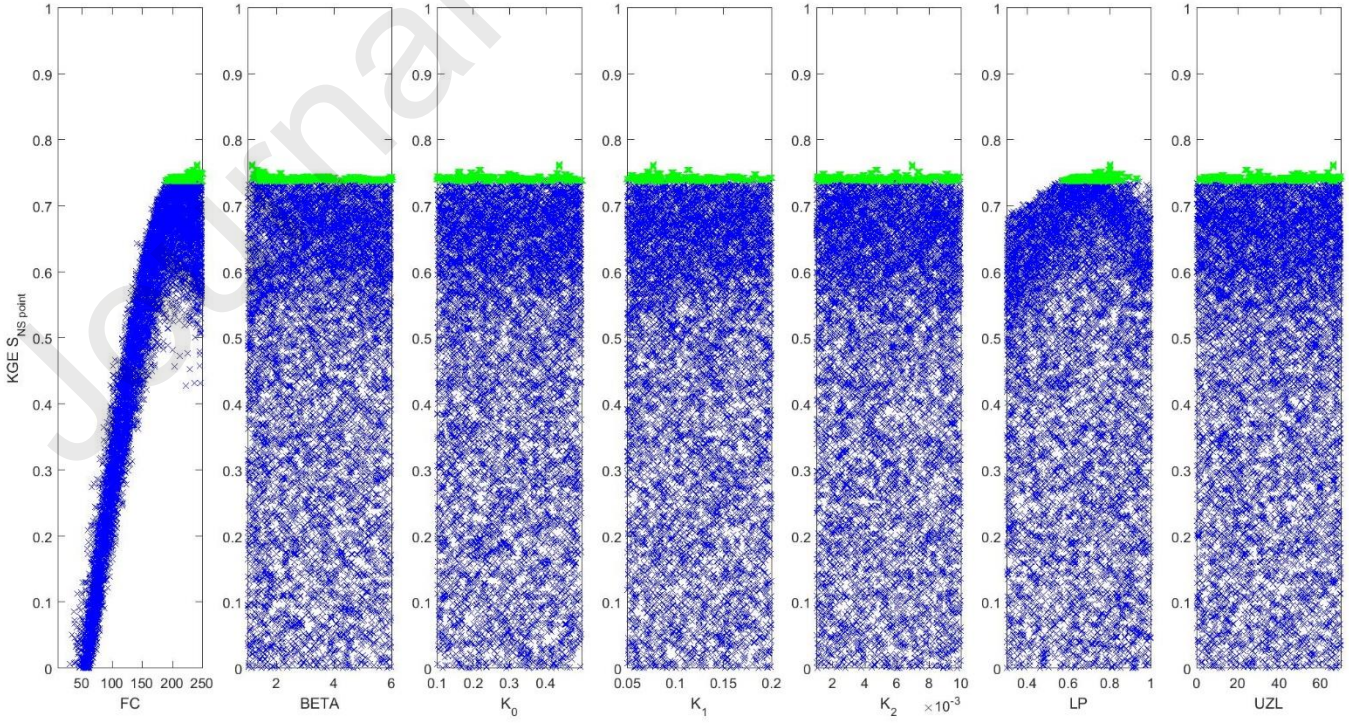


Figure S2: Calibration target S_{NS_CRNS}: Final parameter range of selected parameters (either sensitive or

identifiable in some of the calibrations) of best 100 runs (in green), KGE used as the Goodness of Fit criteria.

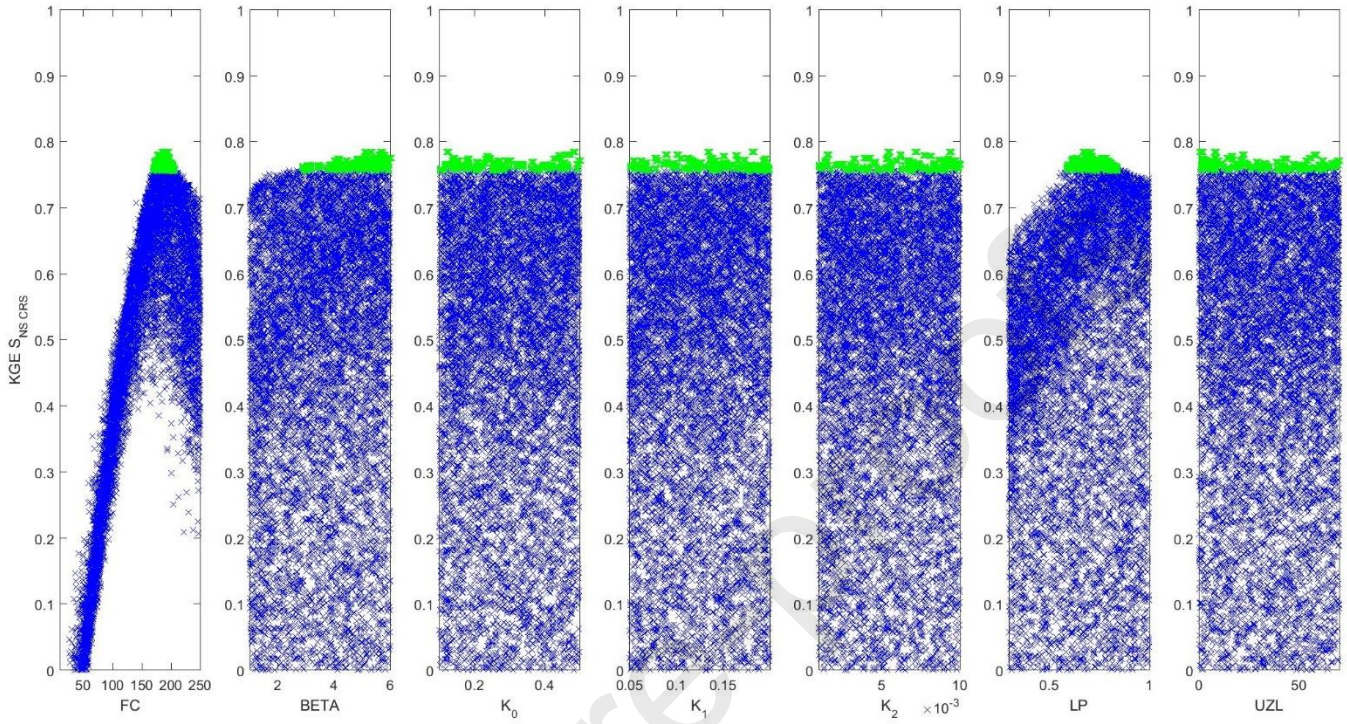


Figure S3: Calibration target S_{NS_point} : Final parameter range of selected parameters (either sensitive or identifiable in some of the calibrations) of best 100 runs (in green), KGE used as the Goodness of Fit criteria.

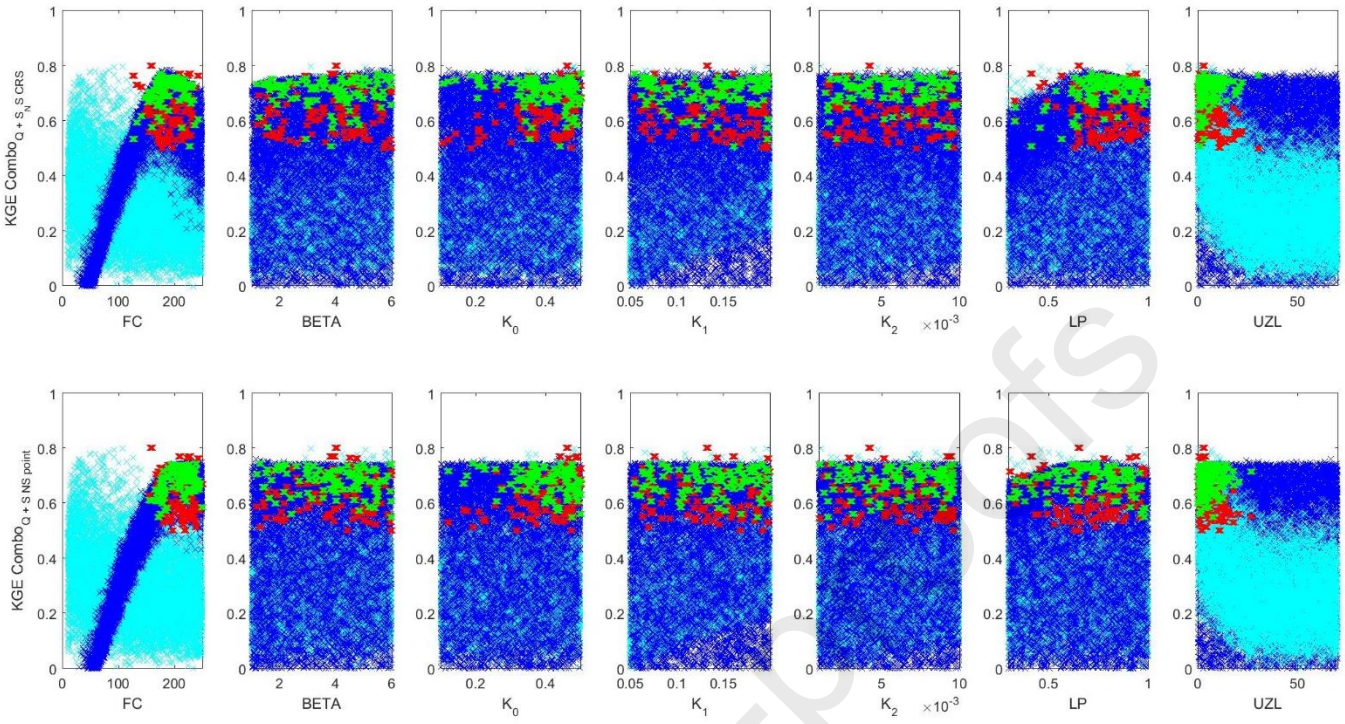


Figure S4: Calibration target is Combo Q + S_{NS}. Top panel Combo with S_{NS_CRNS}; Bottom panel Combo with S_{NS_point} : Final parameter range of selected parameters (either sensitive or identifiable in some of the calibrations) of best 100 runs (discharge in red, S_{NS} in green), KGE used as the Goodness of Fit criteria.

KD, JG, CS and MW developed the objectives of the study, with input from RR and LV on the use of the COSMOS probe. KD undertook the field work and modelling study with supervision from the co-authors. All authors discussed the data and interpretation of the results. KD drafted the manuscript with inputs and edits from all co-authors.

- CRS gives reliable soil moisture measurements at wet agricultural sites
- Measured soil water storage correlates with catchment storage and discharge
- Rainfall-runoff model calibration on CRS data was most useful in wetter periods
- Calibration on discharge and CRS effectively captured catchment internal dynamics
- CRS potential for long-term monitoring and modelling at wet agricultural sites