Authors' responses to the comments of anonymous Reviewer 1

We would like to thank Reviewer 1 for the constructive criticisms and suggestions made to our manuscript. Below, we respond to the comments in a point-by-point manner.
All revisions are visible in the supplement document ("Revised manuscript_visible changes.doc").

Responses to Specific Comments

1. Description of two models should be a bit extended. It should include a description of spatial disaggregation schemes and routing schemes used in both models

We agree with the Reviewer and extended description of the models with respect to their spatial disaggregation and routing schemes. The following paragraphs have been added

> The ECOMAG model utilizes semi-distributed approach with the whole river basin interpreted as a number of sub-basins. It takes into consideration topography, soil and land cover characteristics of a particular sub-basin. For each sub-basin hydraulic properties of soil as well as land-cover properties are scaled taking into account sub-basin area (Motovilov et al. 1999a, b). Subsurface and groundwater routing is based on the Darcy law, while the surface runoff and channel flow are described by the kinematic wave equation.
>
> The SWAP model utilizes a regular spatial grid with a size of $1^o \times 1^o$. The cells are connected by channel's network. Streamflow transformation within network is calculated with the use of a linear model using the TRIP algorithm (Oki et al., 1999)

2. The model performance is described too shortly in section 3 (only references). In addition to references it would be good to describe shortly, in 2-3 sentences, how the model calibration/ validation was done for these large river basins (for multiple gauges?), and to list obtained criteria of fit for the calibration and validation periods.

We have included the following paragraph and additional Table in order to clarify these issues

> Both models have been applied earlier for simulating runoff hydrographs on the basis of multi-year hydrometeorological observations in the Lena and Northern Dvina River basins and demonstrated good performance of simulations (Motovilov and Gelfan 2013; Gusev et. al, 2011; ; Krylenko et al., 2014, Gusev et al., 2015). Both trial-and-error manual procedure and Shuffle Complex Evolution (SCE-UA) automatic algorithm were applied for calibration of ECOMAG and SWAP, respectively. The widely-used Split-Sample Test (Klemeš, 1986) was utilized for validation of the models. Both calibration and validation procedures were carried out against daily streamflow data measured at several gauges of these large basins. The Nash and Sutcliffe (1970) efficiency, NSE, and bias evaluation criteria were adopted to summarize the goodness of fit of the simulated and measured daily discharge series. As an example, the evaluation criteria calculated for the outlets of Lena and Northern Dvina River basins and adopted from (Motovilov and Gelfan 2013; Gusev et. al, 2011; ; Krylenko et al., 2014, Gusev et al., 2015) are shown in Table 1.

Table 1 The Nash and Sutcliffe efficiency, NSE, and bias evaluation criteria calculated from simulated and measured daily discharge at the outlets of Lena and Northern Dvina River basins

River (Gauge)	Period	NSE	Bias, %
ECOMAG (calibration period)			
Lena (Stolb)	2000-2009	0.90	-2.9
N. Dvina(Ust'-Penega)	2000-2009	0.88	1.4
ECOMAG (validation period)			
Lena (Stolb)	1987-1999	0.86	1.4
N. Dvina(Ust'-Penega)	1970-1999	0.81	2.0
SWAT (calibration period)			
Lena (Stolb)	1971-1977	0.82	-4.9
N. Dvina(Ust'-Penega)	1986-1990	0.86	-1.1
SWAT (validation period)			
Lena (Stolb)	1978-1999	0.80	-3.7
N. Dvina(Ust'-Penega)	1967-1985; 1991-1998	0.85	-0.6

3. Both models are assigned as the physically based tools. Most probably, major processes are parametrized using physically-based approaches. However, the question is: is it sufficient to assign them to the class of physically-based models? Are both models fully distributed (3-dimentional), and what is the grid size? Do they both include full surface and groundwater balances and energy balance? Do they both include ONLY physically-based equations, and no any empirical or semi-empirical ones? Do they correspond to criteria outlined in Freeze and Harlan (1969) for a "physically based digitally simulated hydrologic response model"? (see also K. Beven paper, http://eprints.lancs.ac.uk/4421/1/Blueprint.pdf). Maybe the applied models should be rather classified as models of intermediate complexity or process-based models?

There are many different ways of classifying models in watershed hydrology and the most prevalent is the discrimination between black-box, conceptual and physically based models (e.g. Grayson, Blöschl, 2000) that is founded on the relationship between *a priori* (theoretical) and *a posteriori* (based on data) information assimilated by the model. K. Beven (2000, page 41) used the term "process based model" as a synonym of "physically based model". In contrast, R. Grayson and G. Blöschl (2000, page 55) used this term as a synonym of "conceptual model". If the respected reviewer shares the last opinion, then we would like to note that conceptualization of the dominant hydrological processes in the ECOMAG and SWAP models is based mostly on fundamental equations of hydro- and thermodynamics (in integrated form) that offers some advantages over conceptual model. The most important among these advantages is that a large amount of a priori information is used in conceptualization and parameters. Utilization of such prior knowledge and experience that modeler has brought, among other things, to the parametrization process greatly reduces the space for physically realistic parameter values and, consequently, reduces uncertainty of the model response associated with the parameter uncertainty.

Of course, neither ECOMAG nor SWAP model include only physically-based equations without any empirical or semi-empirical ones. However, in our opinion, use of the empirical relationships (particularly for calculation of the parameters through the basin attributes) does not convert a physically-based model into conceptual one. Moreover, we do not know any physically-based model (and widely recognized as such model; SHE, for instance) which could work without the empirical relationships.

Thus, taking into account these notes, we prefer to keep the term "physically based model" in the paper.

Beven, K. (2000): Rainfall-Runoff Modelling – The Primer. John Wiley & Sons, Chichester.

Grayson, R. and Blöschl, G. (2000) Spatial Modelling of Catchment Dynamics. In: Rodger Grayson and Günter Blöschl, eds. Spatial Patterns in Catchment Hydrology: Observations and Modelling. p. 51-81.

Responses to Technical Corrections

1. All abbreviations should be written in full when first mentioned (e.g., 2305, l. 16)

Changed

2, 3 2306: why "artificial" scenarios? why hydrometeorological "impact" (if it is forcing)

The sentence has been changed as follows

> The second group includes the approaches that are based on hydrological models forced by assigned scenarios of hydrometeorological inputs.

4. 2306: why "development" of this approach (maybe rather "application"?).

Changed

5. 2306: differ within ! differ by

Changed

6. 2307, l.2: favors ! favor

Changed

7. 2307: measurement data ! measured data

Changed

8. 2307: , primarily,! primarily

Changed

9. 2311, l. 27: belong to ! occurs in

Changed

10. 2312, l. 11: successively ! successfully?

Changed

11. Fig. 2: two identical graphs for P, no graph for T, please exchange.

Figure 2 has been changed.

12. 2316: similar fields ! similar patterns

Changed

13. 2317: to explain more accurately: if monthly or daily water discharge, then other indices are needed, and not j = 1, 2, . . ., 34.

We clarify this misunderstanding as follows:

> X_{ij} can be either annual discharge for a specific year, or monthly discharge for a specific calendar month, or daily discharge for a specific calendar day, derived from *i*-th realization and related to *j*-th year. Thus, according to the experimental design, any variable, be it annual, monthly or daily, is considered as 45×34 matrix (for instance, matrix of January discharges or matrix of July 25 discharges).

14. 2321, p. 4: not only "require different input data", but also "are differently structured and parametrized".
Thank you for this comment. Changed.

15. All Figures: please increase size of font on axes and subtitles.
All sizes are increased

16. 2326, l. 19-25: not necessary to repeat this here.
We would prefer to keep this fragment.

17. 2327: point 1 could be subdivided into two.
We have reduced the corresponding paragraph.